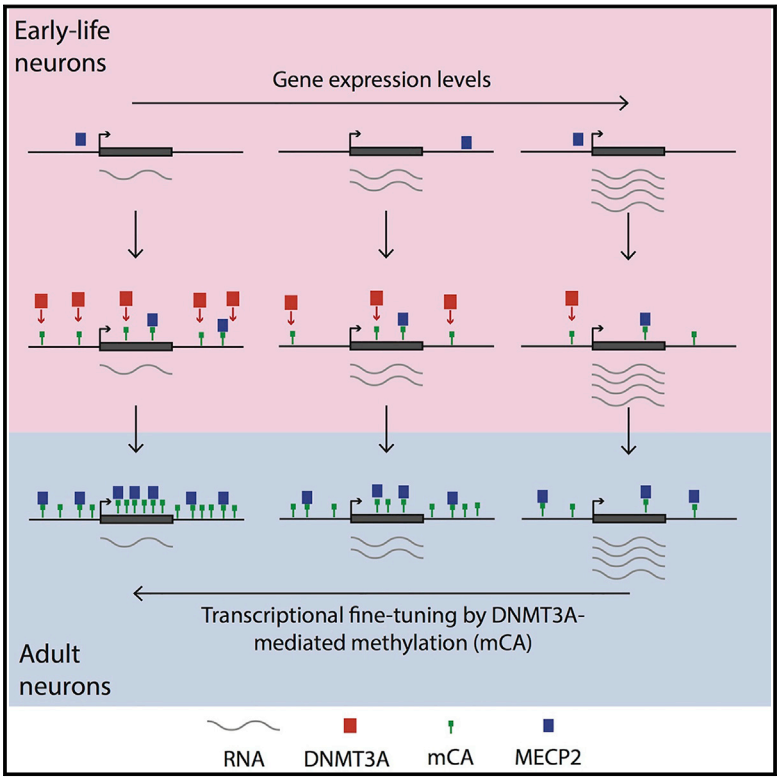
Article

Cell

Early-Life Gene Expression in Neurons Modulates Lasting Epigenetic States

Graphical Abstract



Highlights

- In the brain, DNMT3A binds the genome during early life to specify CA methylation
- DNMT3A preferentially binds across transcribed regions of lowly expressed genes
- DNMT3A binding across genes is modulated by the transcription states of genes
- mCA recruits MECP2 and fine-tunes gene expression in the adult brain

Authors

Hume Stroud, Susan C. Su, Sinisa Hrvatin, ..., Benyam Kinde, Harrison W. Gabel, Michael E. Greenberg

Correspondence

michael_greenberg@hms.harvard.edu

In Brief

The deposition of repressive mCA marks by the methyltransferase DNMT3A across specific brain genes during early postnatal life is important for their regulation throughout life.

Stroud et al., 2017, Cell *171*, 1151–1164
November 16, 2017
https://doi.org/10.1016/j.cell.2017.09.047

CellPress

Article

Cell

Early-Life Gene Expression in Neurons Modulates Lasting Epigenetic States

Hume Stroud,[1] Susan C. Su,[1] Sinisa Hrvatin,[1] Alexander W. Greben,[1] William Renthal,[1] Lisa D. Boxer,[1] M. Aurel Nagy,[1] Daniel R. Hochbaum,[1] Benyam Kinde,[1] Harrison W. Gabel,[1] and Michael E. Greenberg[1,2,*]

[1]Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA

[2]Lead Contact

*Correspondence: michael_greenberg@hms.harvard.edu

https://doi.org/10.1016/j.cell.2017.09.047

SUMMARY

In mammals, the environment plays a critical role in promoting the final steps in neuronal development during the early postnatal period. While epigenetic factors are thought to contribute to this process, the underlying molecular mechanisms remain poorly understood. Here, we show that in the brain during early life, the DNA methyltransferase DNMT3A transiently binds across transcribed regions of lowly expressed genes, and its binding specifies the pattern of DNA methylation at CA sequences (mCA) within these genes. We find that DNMT3A occupancy and mCA deposition within the transcribed regions of genes is negatively regulated by gene transcription and may be modified by early-life experience. Once deposited, mCA is bound by the methyl-DNA-binding protein MECP2 and functions in a rheostat-like manner to fine-tune the cell-type-specific transcription of genes that are critical for brain function.

INTRODUCTION

The methylation of cytosine is a classical epigenetic modification that plays critical roles in regulating gene expression and maintaining cellular identity (Bird, 2002; Jaenisch and Bird, 2003). The *de novo* methyltransferases DNA methyltransferase (DNMT)3A and DNMT3B initiate methylation of cytosines in DNA at the 5′ position of the pyrimidine ring, and DNMT1 maintains methylation of these sites during DNA replication.

In most cell types, DNA methylation occurs at CG sequences (mCG). The mCG mark recruits the methyl-CpG-binding domain (MBD) proteins MBD1, MBD2, MBD4, and MECP2, together with co-repressors that mediate histone deacetylation and transcriptional repression, thus facilitating diverse biological processes such as genomic imprinting, X chromosome inactivation, and cellular differentiation (Bird, 2002). Recent studies have revealed the presence of methylation at non-CG sequences, primarily occurring in CA sequences (mCA), in many tissues and cell types, with neurons appearing to be unique in that they accumulate relatively high levels of mCA specifically during postnatal development (He and Ecker, 2015). While, in a given population of neurons, specific CA sites are methylated at a lower frequency than CG sequences are across the genome, there are many more CA sites than CG sites so that ultimately, CA and CG methylation reach similar levels in adult neurons. The timing of mCA accumulation coincides with an increase in DNMT3A expression in the developing postnatal brain (Lister et al., 2013), and deletion of DNMT3A in the brain results in reduction in mCA across the genome (Gabel et al., 2015; Guo et al., 2014), suggesting that DNMT3A is the primary enzyme that mediates mCA in neurons.

The functional significance of mCA in neurons is underscored by the fact that brain-specific deletions of DNMT3A, or the mCA-binding protein MECP2, give rise to significant neuronal and behavioral deficits (Chahrour and Zoghbi, 2007; Lister and Mukamel, 2015; Nguyen et al., 2007; Sztainberg and Zoghbi, 2016). Consistent with these findings, in humans, DNMT3A is mutated in some individuals with autism spectrum disorder (ASD) (Sanders et al., 2015), as well as those with intellectual disabilities (Tatton-Brown et al., 2014). In addition, loss-of-function mutations in MECP2 give rise to Rett syndrome (RTT), an ASD that is one the most common causes of severe cognitive impairments in females (Chahrour and Zoghbi, 2007).

To better understand the specific functions of mCA during nervous system development and how its dysregulation contributes to neurological disorders such as RTT, we investigated the mechanisms regulating the deposition of the mCA mark and how mCA contributes to the control of gene expression as the brain matures. Previous whole-genome methylation analyses have shown that the presence of mCA is correlated with decreased gene expression in the brain and in neurons (He and Ecker, 2015), suggesting that mCA might recruit transcription factors, which function to negatively regulate gene expression, to the DNA. Consistent with this idea, in the cerebellum, the absence of *Dnmt3a* leads to the subtle upregulation of many neuronal genes, with long genes that have a high density of mCA across their transcribed regions being affected the most (Gabel et al., 2015). These findings suggest that DNMT3A, by mediating the deposition of mCA within the transcribed regions of genes, functions as a repressor that dials down the level of gene expression.

Here, we show that in the mouse brain during early life, the DNA methyltransferase DNMT3A preferentially binds transiently to intergenic regions and across transcribed regions of lowly expressed genes and that this binding primarily determines the pattern of mCA in the adult brain. We find that at sites where

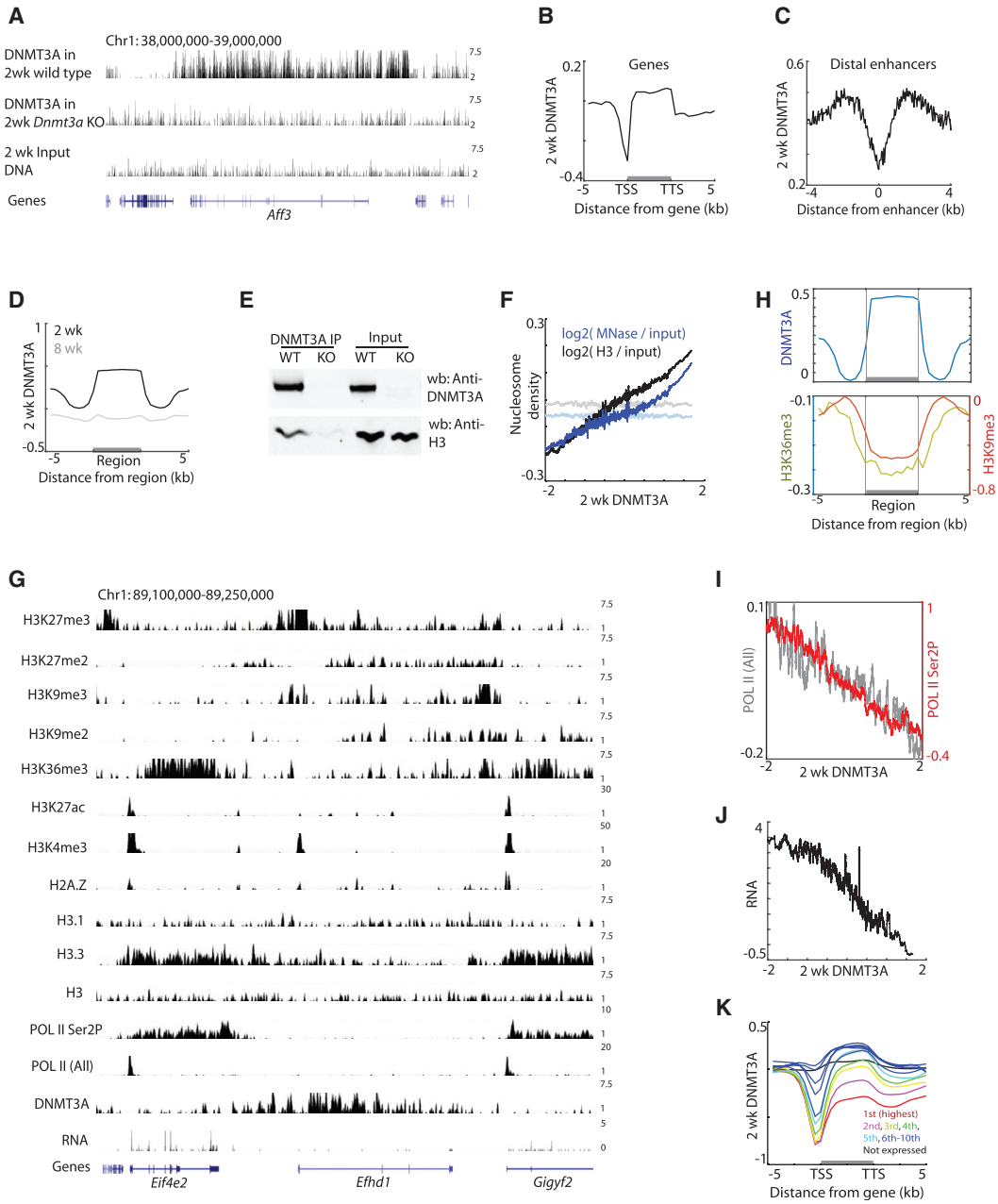


Figure 1. Genomic Binding of DNMT3A in the Brain during Early Life

(A) Genome browser views of DNMT3A ChIP-seq data in wild-type and *Dnmt3a* cKO cortex.

(B) Average DNMT3A distribution across all genes. DNMT3A signal was normalized to input DNA. TSS, transcription start site. TTS, transcription termination site.

(C) Average DNMT3A distribution across putative distal enhancers. H3K27ac data from the forebrain over multiple developmental points was used (Nord et al., 2013), and H3K27ac peaks more than 1 kb apart from annotated TSS were analyzed (n = 38,620).

(D) Average distribution of 2-week and 8-week DNMT3A across DNMT3A-enriched regions in 2-week cortex (n = 22,223).

(E) Immunoprecipitation of DNMT3A under non-denaturing conditions from 2-week cortical extracts from wild-type and *Dnmt3a* cKO (KO) mice.

(F) DNMT3A binding and nucleosome density in 2-week cortex across the genome in 5-kb tiles. Both MNase-seq and H3 ChIP-seq reads were normalized to sonicated input DNA (log2 ratio). The average nucleosome density was binned according to DNMT3A enrichment relative to input DNA. Nucleosome densities after random grouping of genomic tiles are shown as controls in faded lines ($p < 0.001$, permutation test).

(G) Genome browser view of ChIP-seq data in the 2-week cortex.

(H) Average distribution of H3K9me3 and H3K36me3 across defined DNMT3A-enriched regions.

(I) Correlation between DNMT3A binding and RNA-POL II occupancy across gene bodies in the 2-week cortex. The average POL II (all) or POL II (Ser2P) across gene bodies were binned according to DNMT3A density. Promoter regions were excluded from the analyses.

(legend continued on next page)

mCA is deposited in early life, this methyl mark recruits MECP2, which then functions to restrain the transcription of genes in the maturing brain in a neuronal subtype-specific manner. These findings suggest a rheostat model of gene regulation in which DNMT3A, mCA, and MECP2 collaborate to fine-control the genes whose precisely tuned expression may be critical for normal brain development and function.

RESULTS

DNMT3A Transiently Interacts with the Genome during Early Life

A previous study has shown that the patterns of mCA are at least partly conserved through evolution (Lister et al., 2013) and are neuronal subtype specific (Mo et al., 2015), suggesting that there must be a conserved mechanism by which DNMT3A is recruited to particular sites in the genome. To identify this mechanism, we investigated where across the genome DNMT3A binds, speculating that the sites of DNMT3A binding might provide insight into the mechanism by which DNMT3A is recruited to DNA and ultimately how DNMT3A regulates mCA to control gene expression. Given that in neurons the increased expression of the *Dnmt3a* mRNA and protein peaks 2 weeks after birth (Feng et al., 2005; Lister et al., 2013), we assessed DNMT3A binding to DNA in 2-week-old mouse brain by chromatin immunoprecipitation followed by sequencing (ChIP-seq) using an anti-DNMT3A-specific antibody to enrich for sites of DNMT3A binding (Figures 1A, S1A, and S1B). The specificity of the DNMT3A ChIP-seq analysis was validated using cortical and hippocampal extracts from 2-week-old mice in which *Dnmt3a* expression was conditionally disrupted in the brain (*Nestin-cre*; *Dnmt3a*$^{fl/fl}$, hereafter designated *Dnmt3a* cKO) (Figures 1A and S1A–S1C). These DNMT3A ChIP-seq experiments revealed that in both brain regions, DNMT3A binds broadly across the genome and that DNMT3A binding is significantly decreased at gene promoters and enhancers (Figures 1B, 1C, S1D, and S1E). Moreover, consistent with the reduction in DNMT3A expression following early postnatal development (Figure S1F), DNMT3A ChIP-seq using cortical extracts prepared from 8-week-old adult cortex showed that DNMT3A binding across the genome is significantly reduced relative to that in 2-week-old mice (Figures 1D and S1G). Thus, we conclude that during the early postnatal period, DNMT3A is transiently recruited to and broadly coats the neuronal genome.

DNMT3A Binds across the Transcribed Regions of Lowly Expressed Genes

We next investigated what features of the neuronal genome determine where DNMT3A will bind. Previously, the mechanisms of DNMT-protein targeting to the genome *in vivo* have been challenging to characterize, and most advances in our understanding of this process have been made using *in vitro* approaches. Prior work in other cell types has suggested that DNMT proteins are targeted to the genome through their interaction with specific accessory factors that associate with nucleosomes by recognizing specific histone modifications or by binding specific genomic sequences (Rose and Klose, 2014). On the basis of these findings, we hypothesized that in neurons, DNMT3A recruitment might also be dictated by specific features of chromatin that are present during early postnatal development.

We asked whether DNMT3A binds to nucleosomes. Consistent with this possibility, we found that DNMT3A co-precipitates with histones (Figures 1E and S1H). In addition, micrococcal nuclease digestion of chromatin isolated from the cortex followed by sequencing (MNase-seq), as well as histone H3 ChIP-seq, revealed that DNMT3A binding positively correlates with the density of nucleosomes across the genome (Figure 1F). These results suggest that in the brain during early postnatal life, DNMT3A interacts with histones within nucleosomes.

To determine if DNMT3A recruitment is correlated with specific features of nucleosomes, we employed ChIP-seq analysis to generate genome-wide maps of a variety of chromatin modifications in mouse cortex at 2 weeks old, a developmental time point during which the brain epigenome has not previously been well characterized. Our analysis included an assessment of seven histone marks associated with active transcription (H3K4me3, H3K36me3), active enhancers (H3K27ac), heterochromatin (H3K9me2, H3K9me3), and Polycomb targeting (H3K27me2, H3K27me3), as well as binding data for three well-characterized histone variants (H3.1, H3.3, and H2A.Z) (Figures 1G and S1I). We also carried out ChIP-seq on total RNA polymerase II (POL II) and elongating forms of POL II (phosphorylated in Ser2 of POL II), as well as RNA sequencing (RNA-seq) analysis to identify the regions of the neuronal genome that are transcribed 2 weeks postnatally and to assess the level of transcription of each gene.

While no single chromatin feature was sufficient to predict the sites of DNMT3A binding, several broad trends were readily apparent (Figure S1J). DNMT3A is for the most part excluded from genomic regions marked by chromatin features associated with active regulatory elements (e.g., active enhancers and promoters). The exclusion of DNMT3A from these sites is in part consistent with *in vitro* evidence suggesting that H3K4me3 blocks DNMT3A binding (Ooi et al., 2007). In addition, somewhat surprisingly given the finding that the homologous protein DNMT3B interacts with H3K36me3 in embryonic stem cells (Baubec et al., 2015), DNMT3A was found to be reduced in regions of chromatin that are associated with H3K36me3—a mark that is typically present within the gene bodies of actively transcribed genes (Figures 1H and S1K–S1M).

We found that DNMT3A is largely excluded from the bodies of genes that are transcribed at a high level and is also absent from genes that are silent due to their inclusion in H3K9 methylation-marked heterochromatin (Figures 1H–J, S1L, and S1N). Instead, DNMT3A appears to bind to most other regions of the genome

(J) Correlation between DNMT3A binding and gene expression in the 2-week cortex. The average gene expression level was binned according to DNMT3A density.

(K) Average DNMT3A distribution over genes of different expression levels in the wild-type cortex.

See also Figure S1.

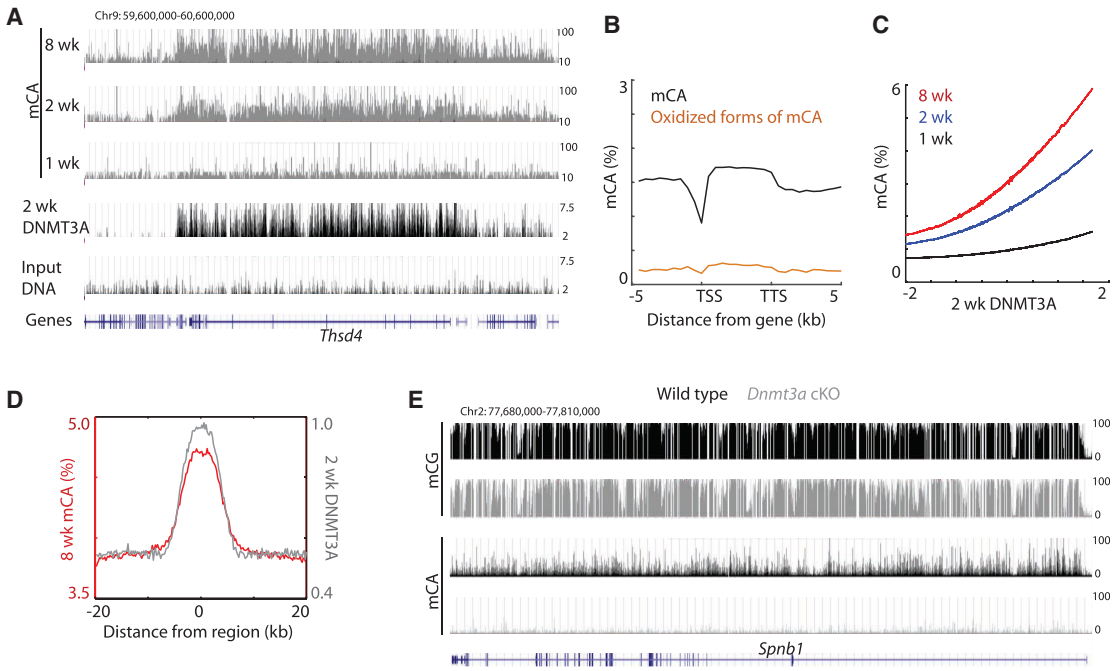


Figure 2. DNMT3A Binding Specifies mCA Patterns

(A) Genome browser view of mCA and DNMT3A ChIP-seq data.

(B) Average distribution of mCA and oxidized forms of mCA over genes.

(C) Correlation between 2-week DNMT3A binding and mCA in the cortex. mCA was binned according to DNMT3A densities within 5-kb tiles across the genome.

(D) Average distribution of 8-week mCA levels over 2-week DNMT3A-enriched regions. Distribution of DNMT3A is also shown.

(E) Genome browser view of DNA methylation in wild-type and *Dnmt3a* cKO cortex.

See also Figure S2.

and, perhaps most importantly, within the transcribed regions of neuronal genes that are expressed at a relatively low level (Figures 1I–1K) such that the more lowly expressed the gene, the higher the level of DNMT3A binding across the gene. Taken together, these findings indicate that DNMT3A associates with nucleosomes and binds broadly across the genome but is excluded from highly transcribed genes and silenced genes that reside in heterochromatin. Thus, DNMT3A displays preferential binding across genes that are transcribed at a low but detectable level.

Genomic DNMT3A Targeting Specifies mCA Patterns

We next asked if DNMT3A binding across the genome determines the sites of mCA deposition. It is unclear, for example, if the sites of mCA deposition are determined primarily by where across the genome DNMT3A is bound, or if once bound to DNA, DNMT3A's enzymatic activity is regulated such that only a subset of DNMT3A-binding sites become methylated (Guo et al., 2015; Li et al., 2011). Alternatively, mCA may be catalyzed by DNMT3A at all sites to which DNMT3A binds, but the mCA mark might be actively removed at specific sites. Having first confirmed the ability of DNMT3A to directly catalyze mCA *in vitro* (Figure S2A), we performed whole-genome bisulfite sequencing (WGBS) to generate base-resolution maps of DNA methylation in both the mouse cortex and the mouse hippocampus at 1, 2, and 8 weeks of age to determine if the genomic sites that DNMT3A preferentially binds become methylated. We observed a marked genome-wide increase in mCA in the cortex and hippocampus between 1 week and 8 weeks after birth, consistent with a previous study in the frontal cortex (Lister et al., 2013) (Figures 2A and S2B–S2D).

WGBS does not distinguish between cytosine methylation and oxidized forms of methylation at cytosine that are known to be present at high levels in the brain (Kriaucionis and Heintz, 2009). Because oxidized forms of methylation have been suggested to form as intermediates during the active removal of methyl groups in the CG context (Pastor et al., 2013), we employed oxidative bisulfite sequencing (oxBS-seq) (Booth et al., 2012), a method that enables direct measurement of DNA methylation, to accurately assess the level of mCA (Figure S2E). This analysis confirmed that in the 12-week cortex, the percentage of mCA that is oxidized is relatively low, suggesting that most CA sites that become methylated on cortical DNA in early postnatal life are not oxidized later in life (Figures 2B and S2F). This result, together with the finding that DNMT3A primarily binds the genome during early development (Figure 1) and the fact that genome-wide mCA levels in the adolescent mouse and human brains appear to be similar to those in mature brains (Lister et al., 2013), suggests that once CA methylation has occurred in the brain, a large proportion of the mCA mark is likely stable and persists throughout the lifetime of the animal.

Comparison of the DNMT3A genomic binding profile with the pattern of mCA revealed a strong correlation between DNMT3A binding to chromatin at 2 weeks after birth and the mCA pattern later in life (Figures 2A, 2C, S2C, and S2G). Notably, DNMT3A-bound sites are highly enriched for mCA (Figures 2D and S2H). Moreover, genomic regions that differed in their level of DNMT3A binding between the cortex and hippocampus also display differential patterns of mCA such that regions that had a high density of DNMT3A become highly methylated, and those that bind low amounts of DNMT3A are lowly methylated (Figure S2I). Thus, the loci and density of DNMT3A binding are highly correlated with the pattern of mCA with respect to position and density across the neuronal genome. DNMT3A appears to be the sole enzyme responsible for postnatal CA methylation. By WGBS, we found that mCA is essentially eliminated in the cortex of *Dnmt3a* cKO mice (Figures 2E, S2J, and S2K). By contrast, mCG is partially reduced across the neuronal genome in *Dnmt3a* cKO mice compared to wild-type mice, specifically at sites that are bound by DNMT3A in the wild-type cortex at 2 weeks after birth (Figure S2L). This suggests that DNMT3A controls the genome-wide postnatal increase in mCG observed between young and adult brains. However, we note that since the majority of mCG is deposited early in brain development, prior to when DNMT3A expression was disrupted, our experiments do not identify which methyltransferases catalyze this early wave of mCG. Collectively, our results indicate that in the early postnatal period, a transient physical interaction of DNMT3A with chromatin primarily determines the pattern of mCA and, in part, the mCG pattern in the adult brain.

Experience-Dependent Transcriptional Activation Disrupts Local Binding of DNMT3A and mCA Deposition

Given that DNMT3A binds across the neuronal genome during the time postnatally when sensory experience first affects neuronal gene expression, we asked if sensory experience and synaptic activation have an effect on DNMT3A binding to DNA or on deposition of the mCA mark, which, if affected, might ultimately lead to a change in neuronal gene expression.

To investigate this possibility, we administered 2-week-old mice the glutamate receptor agonist kainic acid (KA), a treatment that triggers seizure activity and is accompanied by a large increase in activity-dependent gene transcription throughout the hippocampus (Bloodgood et al., 2013). Hippocampi were then isolated for DNMT3A ChIP-seq. While compared to vehicle control, KA treatment had no obvious effect on the average amount of DNMT3A across the genome, we observed a significant reduction in DNMT3A binding in KA-treated mice compared to control mice across the transcribed region of genes whose expression is induced by KA treatment, including key immediate-early genes such as *Fos* and *Egr1* (Figures 3A, 3B, S3A, and S3B). A more modest but significant decrease in the binding of DNMT3A within the transcribed regions of activity-induced genes was also observed in the visual cortex in response to a more natural stimulation paradigm in which dark-reared mice were acutely exposed to light (Figures S3D–S3F). Importantly, deletion of DNMT3A had no effect on the level of expression of activity-regulated genes (Figure S3C), suggesting that rather than facilitating activity-dependent gene transcription, the decrease in DNMT3A binding and mCA deposition in KA-treated relative to untreated neurons is a consequence of increased expression of the activity-regulated genes upon KA treatment (or light exposure).

To determine if these changes in DNMT3A binding alter the adult mCA pattern at the relevant genes, we exposed mice to a daily injection of KA for 10 days, beginning on day 10 after birth, and then allowed the mice to mature into adults (8 weeks) (Figure 3C). While this treatment did not cause a persistent induction of activity-dependent gene transcription over the 10-day period, we reasoned that the daily exposure would induce transient increases in activity-dependent gene expression over a period of several hours each day (Figure S3G), resembling the effects of persistent seizures on the developing postnatal brain observed in children with epilepsy. WGBS on DNA extracted from the adult hippocampi of KA-treated or untreated mice revealed a modest but statistically significant decrease in mCA, specifically across the bodies of activity-regulated genes in the KA-treated mice (Figure 3D), suggesting that perturbing gene expression during early postnatal brain development has the potential to alter the distribution of mCA across the transcribed regions of genes later in life and, in this way, may have a long-lasting effect on gene expression in adult mice (Figure S3H).

Genetic Mutation-Driven Transcriptional Activation Disrupts DNMT3A Binding and mCA Deposition

To assess further whether changes in the transcription of genes during early life can alter the neuronal methylome, we next sought to test whether genetic mutations that disrupt the function of a transcriptional repressor in early life and thus lead to aberrant upregulation of specific genes during early development might lead to changes in mCA levels across the misregulated genes. Such a mechanism of gene misregulation could, in principle, contribute to the etiology of neurological disorders that are known to be due to the mutation of a transcriptional repressor (Vissers et al., 2016). One such disorder is Weaver syndrome, a congenital overgrowth disorder associated with intellectual disabilities that is due to mutation of EZH2, one of the two catalytic subunits of the Polycomb repressive complex 2 (PRC2) (Gibson et al., 2012). EZH2 catalyzes H3K27 methylation and functions as a transcriptional repressor in many cell types (Margueron and Reinberg, 2011), including neurons (von Schimmelmann et al., 2016).

To investigate the possible effects of EZH2 loss on neuronal mCA patterning, we crossed *Nestin*-cre mice to *Ezh2*$^{fl/fl}$ mice (Shen et al., 2008) to delete *Ezh2* in the brain (*Ezh2* cKO). As might be predicted, the disruption of EZH2 function in the brain led to severe growth retardation (data not shown) (Zhang et al., 2014), and ChIP-seq analyses of brain extracts from these mice at 2 weeks of age indicated a genome-wide reduction in H3K27me3, a chromatin signature of gene repression (Figure S4A). By RNA-seq, we found that the genome-wide decreased presence of the H3K27me3 mark in *Ezh2* cKO brains was accompanied by an upregulation in expression of small subsets of both neuronal and non-neuronal genes.

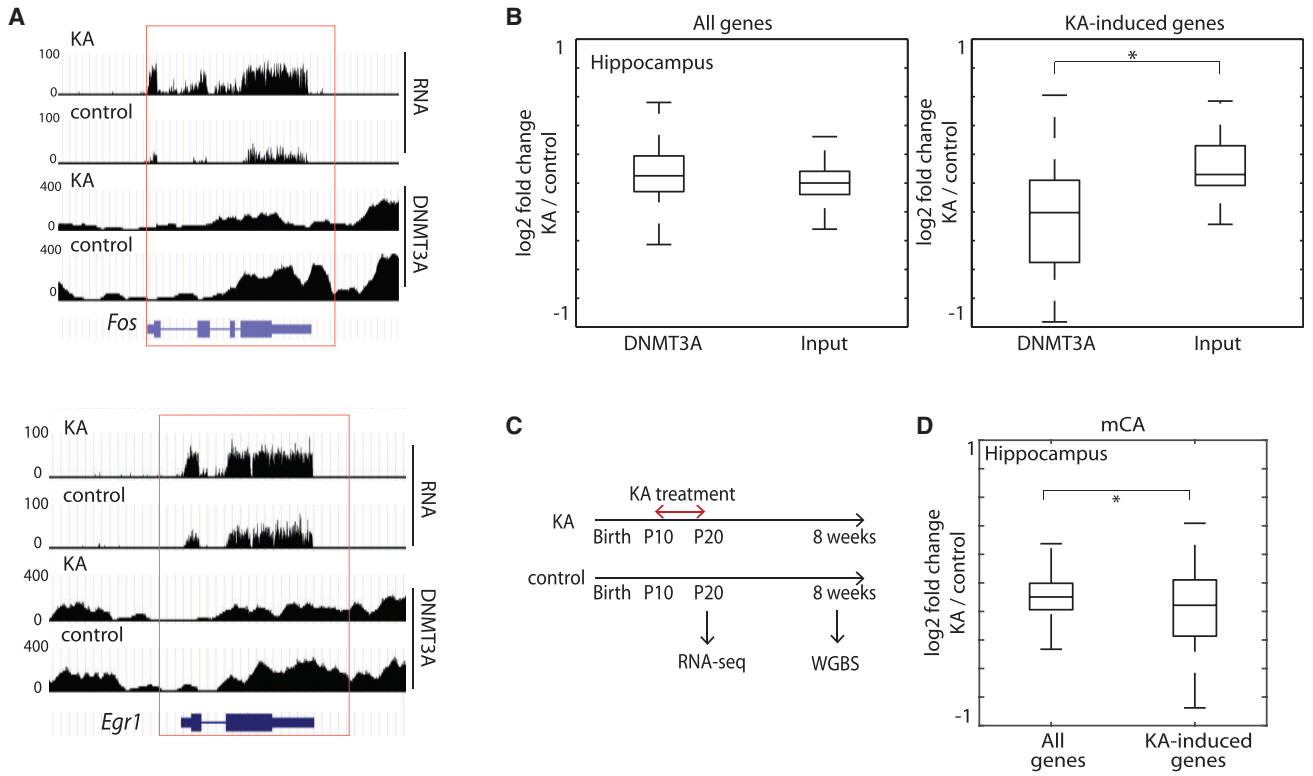


Figure 3. Experience-Driven Gene Transcriptional Induction Disrupts Local DNMT3A Binding

(A) Genome browser views of DNMT3A binding and gene expression data in KA-treated and PBS-treated (control) 2-week hippocampi.

(B) Effect of KA treatment on DNMT3A binding over all genes and KA-induced genes (n = 51). DNMT3A-ChIP-seq- and input-DNA-read densities of 2-week hippocampi within gene bodies were calculated, and log2 ratios between KA and control samples are shown. *p < 0.0001, Wilcoxon rank-sum test.

(C) Scheme illustrating the KA treatment from P10 to P20.

(D) Boxplots of the difference between 8-week mCA in KA-treated and control hippocampi within all genes and KA-induced genes. *p < 0.05, Wilcoxon rank-sum test.

See also Figure S3.

To determine whether DNMT3A binding and mCA patterns are altered across transcribed regions of the genes whose expression is upregulated in the *Ezh2* cKO cortex, we performed DNMT3A ChIP-seq and WGBS. When all genes were analyzed in aggregate, the disruption of EZH2 function was found to have no substantial effect on DNMT3A binding to DNA or mCA levels when *Ezh2* cKO and wild-type mice were compared. This suggests that EZH2 function is not required for DNMT3A binding to chromatin per se (Figures 4A, 4B, and S4A–S4E). However, when we focused our analyses in the vicinity of genes whose expression is selectively upregulated in *Ezh2* cKO cortex, we found that DNMT3A binding and mCA levels were significantly reduced within the transcribed regions of these genes in the *Ezh2* cKO mice compared to wild-type mice (Figures 4A and S4B).

Importantly, the disruption of DNMT3A function has no effect on the expression of the genes whose transcription is upregulated in the *Ezh2* cKO mice (Figure 4C). This suggests that the misregulation of gene expression in *Ezh2* cKO mice is not due to a reduction in mCA across the transcribed region of these genes. We conclude that, rather than facilitating the initial activation of gene transcription, the decrease in DNMT3A binding and mCA is likely a consequence of increased expression of these genes due to the *Ezh2* mutation.

Cell-Type-Specific mCA Patterns Are Shaped by Early-Life Gene Expression

Several recent studies have suggested that the DNA methylation in the brain varies across different neuronal subtypes (Lister et al., 2013; Mo et al., 2015). However, it is not known how these different methylation patterns are established during postnatal brain development and whether the pattern of methylation determines the level of neuronal subtype-specific gene expression or vice versa. To address these issues, we compared the developmental DNA methylome and transcriptome dynamics in two specific neuronal subtypes: parvalbumin (*Pv*)- and vasoactive-intestinal-peptide (*Vip*)-expressing interneurons.

Pv and *Vip* neurons are well-defined, relatively homogeneous, sparse cell types (< 5% of total cells in the brain) that function in distinct ways to inhibit cortical circuit function (Markram et al., 2004). *Pv* neurons mediate somatic inhibition of excitatory pyramidal neurons, whereas *Vip* neurons have been shown to inhibit somatostatin interneurons, which in turn mediate dendritic

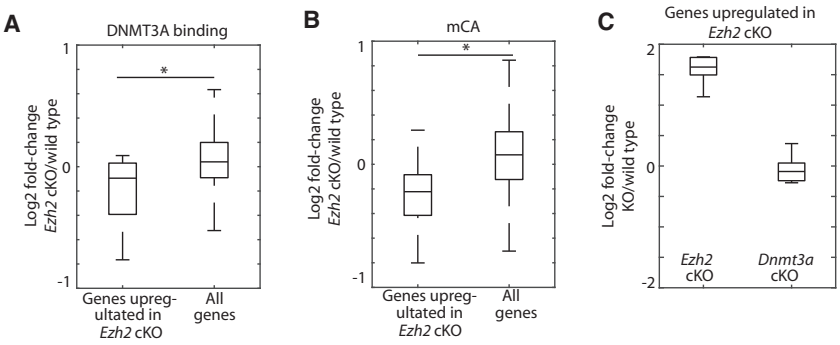


Figure 4. Genetic Mutation-Driven Gene Transcriptional Induction Disrupts Local DNMT3A Binding

(A) Effect of *Ezh2* mutation on DNMT3A binding over all genes and genes upregulated by at least 2-fold in *Ezh2* cKO (n = 11). DNMT3A- and H3-ChIP-seq-read densities of 2-week cortices within gene bodies were calculated, and log2 ratios between *Ezh2* cKO and wild-type control samples are shown. *p < 0.05, Wilcoxon rank-sum test.

(B) Boxplots of the difference between 2-week mCA in *Ezh2* cKO and wild-type control cortices within all genes and genes upregulated in *Ezh2* cKO. *p < 0.05, Wilcoxon rank-sum test.

(C) Gene expression levels of genes defined as genes upregulated in *Ezh2* cKO. RNA-seq data for 2-week *Dnmt3a* cKO and respective littermate wild-type controls are also shown.

See also Figure S4.

inhibition of pyramidal neurons. The distinct functions of *Pv* and *Vip* neurons may in part reflect the fact that these two neuronal subtypes display distinct patterns of mCA and mRNA expression (Mo et al., 2015). It is noteworthy that *Pv* neurons have substantially higher genome-wide levels of mCA compared to excitatory neurons (Mo et al., 2015), suggesting that mCA might play a particularly important role in this inhibitory neuronal subtype. By contrast, the level of mCA across the genome of *Vip* neurons is similar to that of excitatory neurons. The differences in the level of mCA across the genome between these two inhibitory neuron populations made them attractive choices for this analysis.

To determine if differential patterns of gene expression between *Pv* and *Vip* neurons in the early postnatal period lead to differential patterns of mCA in these neurons later in life, we used the INTACT (isolation of nuclei tagged in specific cell types) method (STAR Methods) to compare the mCA landscapes and gene expression profiles of *Pv*- and *Vip*-expressing neurons in the mouse cortex at various stages of postnatal neuronal development. *Pv* or *Vip* nuclei were purified at postnatal weeks one, three, and eight (Figures 5A and 5B), and nuclear RNA-seq data were obtained and compared to ribosome-bound mRNA-seq data for *Pv*- and *Vip*-expressing neurons to verify that the two approaches yielded similar results (Figure S5A). In addition, we performed WGBS on *Pv*- and *Vip*-purified nuclei to generate base-resolution maps of DNA methylation in the two cell types at each developmental time point.

In the 1-week-old mouse cortex, *Pv* and *Vip* neurons exhibited relatively similar methylomes that lack detectable mCA and possess similar, though not identical, genomic patterns of mCG (Figures 5C, S5B, and S5C). Importantly, we found that even before mCA accumulation, *Pv* and *Vip* neurons at 1 week of age exhibit pronounced differences in their gene expression profiles (Figures 5D and S5D). For example, young *Pv* neurons more highly express neuronal genes such as *Cacna2d2* and *Hapln4* when compared to *Vip* neurons, and *Vip* neurons more highly express genes such as *Kcnq4* and *Accn4* when compared to *Pv* neurons (Figure S5D).

We found that the density of mCA increases across the genome in both *Pv* and *Vip* neurons between postnatal week one and three (Figures 5C and S5B), and that mCA is preferentially associated with lowly transcribed genes in both neuronal subtypes (Figure S5E). Specifically, by postnatal week three, genes that are transcribed at a high level in *Pv* neurons, but not in *Vip* neurons at week one, have become highly methylated at CA sequences within their transcribed regions in *Vip*, but not *Pv*, neurons (Figures 5C, 5E, and S5F–S5I). A similar phenomenon was found to be the case for genes that are specifically transcribed at a high level in *Vip* neurons and a low level in *Pv* neurons; the lowly expressed genes in *Pv* neurons became highly methylated in *Pv* neurons, but not *Vip* neurons. Thus, the decision to express a gene in a neuronal subtype-specific manner appears to be made prior to methylation at CA sequences. Once a gene has been specified to be lowly expressed, the gene likely binds DNMT3A within its transcribed region and becomes methylated at CA sequences.

These findings suggest that neuronal subtype-specific patterns of gene expression are determined early during development and that the deposition of mCA occurs subsequently in a neuronal subtype-specific manner. In addition, these observations provide further support for the conclusion that highly expressed genes are refractory to DNMT3A binding and mCA deposition, as these genes only become methylated at CA sequences in neuronal subtypes where they are lowly expressed, not in subtypes where the genes are highly expressed.

mCA Mediates Cell-Type-Specific Fine-Tuning of Transcription

The function of intragenic mCA is not known, and several possible functions can be envisioned. Methylation within the transcribed regions of lowly expressed genes could, in principle, modulate the expression of these genes either positively or negatively. Alternatively, mCA within gene bodies might preserve transcriptional fidelity by slowing transcriptional elongation, inhibiting spurious transcriptional initiation, or preventing POL II from falling off the DNA. To begin to distinguish among these possibilities, we blocked the deposition of mCA within the transcribed regions of neuronal genes by disrupting DNMT3A function as described above and then assessed the effect on neuronal subtype-specific gene expression.

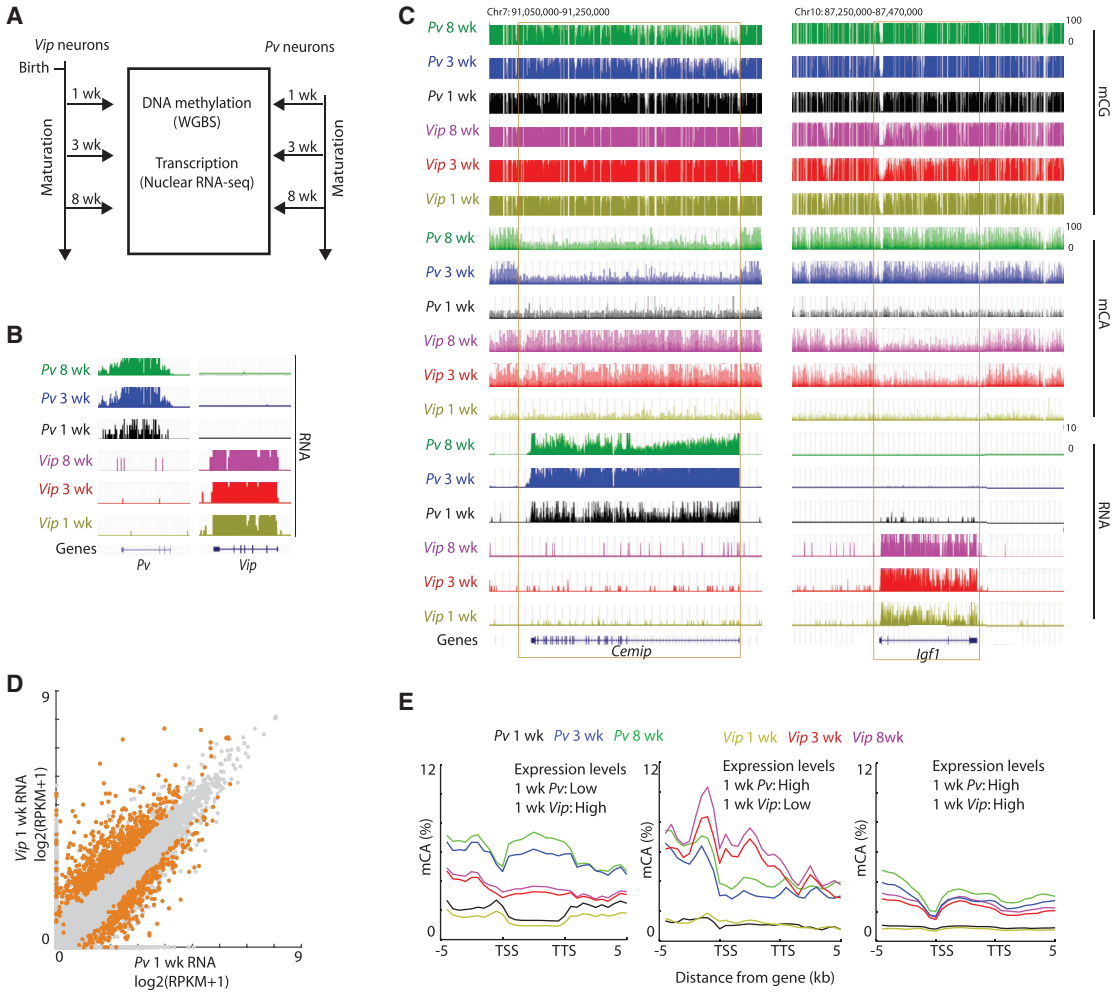


Figure 5. Cell-Type-Specific Transcription during Early Life Shapes mCA Patterns across Genes

(A) The scheme of experimental design.

(B) Genome browser views confirming specific enrichment of *Pv* and *Vip* transcripts in INTACT-isolated *Pv* and *Vip* nuclei, respectively.

(C) Genome browser views of DNA methylation and transcription in cortical *Pv* and *Vip* neurons at the indicated developmental time points.

(D) Scatterplot comparing 1-week *Pv* and *Vip* transcript levels across all genes. Genes with differential gene expression (false discovery rate [FDR] < 0.05) are highlighted in orange. *Pv* > *Vip* genes, n = 363; *Vip* > *Pv* genes, n = 698.

(E) Average distribution of mCA over genes that were classified as lowly transcribed in one cell type and highly transcribed in the other, as well as genes classified as highly expressed in both cell types (see STAR Methods).

See also Figure S5.

To assess the effect of DNMT3A disruption on gene expression in different neuronal subtypes, we adapted a protocol for high-throughput droplet-based microfluidic single-cell RNA-seq (Klein et al., 2015) so that the protocol could be applied to isolated nuclei. We reasoned that by measuring the level of nuclear RNA, we would obtain a more accurate estimate of transcription than by measuring the level of cytoplasmic RNA. We performed single-nuclei RNA-seq using nuclei isolated from the cortices of *Dnmt3a* cKO mice or wild-type mice and analyzed the transcriptomes of a total of 26,063 nuclei from the *Dnmt3a* cKO mice and 23,677 nuclei from wild-type controls (Figures 6A and S6A). Importantly, using this approach, we found that we can efficiently distinguish nuclei of different inhibitory neuron subtypes (Figures 6B and S6B).

We found by single-nuclei RNA-seq that wild-type *Pv* and *Vip* neurons exhibit gene profiles that are similar to those observed in the INTACT experiments, indicating that single-nuclei RNA-seq enables accurate cell-type-specific analyses of nuclear RNA (Figure S6C). In wild-type *Pv* and *Vip* neurons, genes with a high density of mCA within their transcribed regions were lowly expressed, thus confirming our findings using the INTACT approach (see above) (Figure 6C). When we compared *Dnmt3a* cKO mice to wild-type mice, we found that genes that are normally lowly expressed and have a high density of mCA across their gene bodies are expressed on average at a higher level in

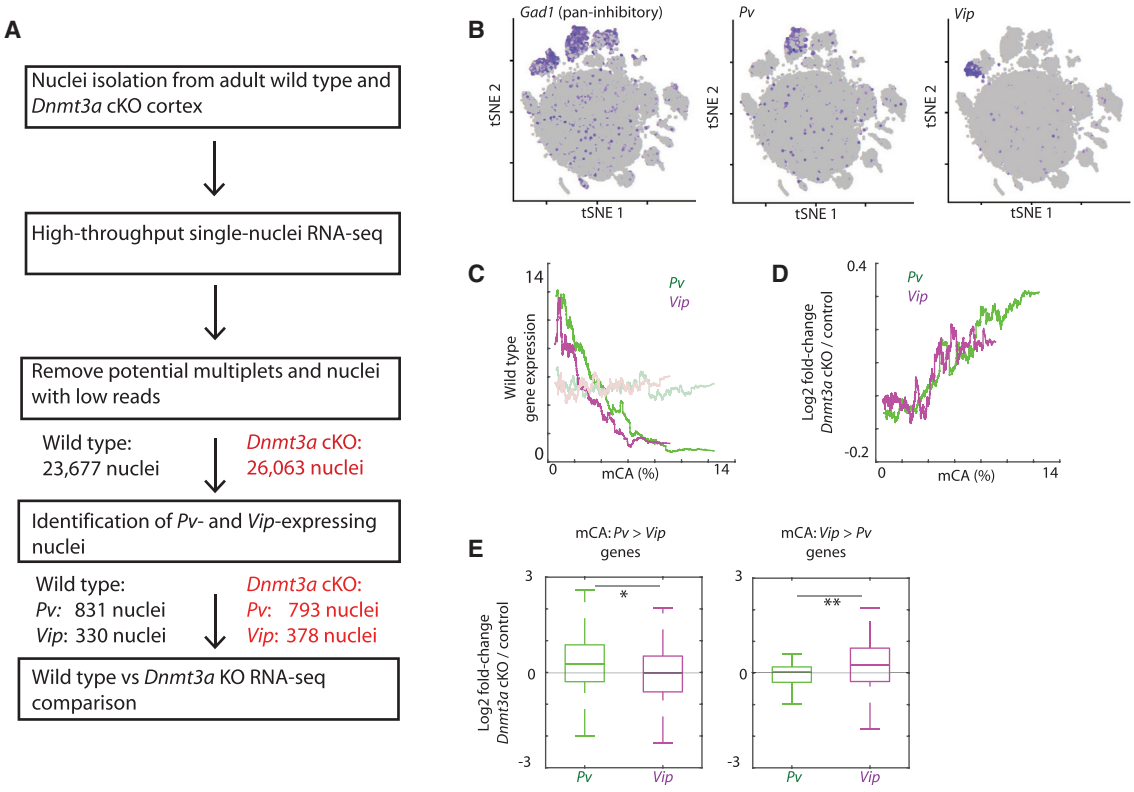


Figure 6. Single-Nuclei Sequencing Reveals Cell-Type-Specific Gene Expression Fine-Tuning by mCA

(A) The scheme of experimental and analysis design.

(B) Seurat t-distributed stochastic neighbor embedding (t-SNE) plots of all nuclei in samples. Nuclei expressing indicated marker genes are depicted in purple.

(C) Correlation between wild-type mCA densities within the gene bodies and the expression levels determined by single-nuclei RNA-seq. Gene expression levels after random groupings of genes are shown as controls in faded lines.

(D) Correlation between wild-type mCA densities within the gene bodies and transcriptional defects in *Dnmt3a* cKO cortex.

(E) Cell-type-specific gene repression defects in *Dnmt3a* KO neurons. Genes were separated based on the mCA densities within the gene bodies in *Pv* and *Vip* neurons. *p = 3.6e−4, **p = 0.001, Kolmogorov-Smirnov test.

See also Figure S6.

Dnmt3a KO neurons compared to wild-type neurons (Figures 6D and S6D). Importantly, while the magnitude of the increases in transcription observed in *Dnmt3a* KO neurons are relatively modest, a large number of neuronal genes are affected, suggesting that mCA likely functions at a genome-wide level to restrain the transcription of lowly expressed neuronal genes. Notably, the more lowly expressed a gene is, the higher the density of mCA and the greater the degree of misregulation when DNMT3A function is disrupted.

We next asked if the disruption of DNMT3A function differentially affects gene expression in *Pv* and *Vip* neurons. Remarkably, we observed substantial differences in gene misregulation in the *Dnmt3a* cKO cortex between neuronal subtypes such that genes with a high density of mCA in *Pv* neurons, but not in *Vip* neurons, were preferentially upregulated in *Pv* neurons when DNMT3A function was disrupted (Figure 6E). Likewise, when DNMT3A function is disrupted, genes with a relatively high density of mCA in *Vip* neurons are selectively upregulated in *Vip* neurons, but not in *Pv* neurons. These findings collectively suggest that DNMT3A, by catalyzing the methylation of CA sequences within the transcribed regions of lowly expressed genes, dampens the transcription of these genes in a neuronal subtype-specific manner. Given that the pattern of neuronal subtype-specific gene expression is determined prior to DNMT3A binding and the deposition of mCA, these findings are most consistent with a model in which mCA functions to restrain the expression of lowly expressed genes in the same manner across various neuronal subtypes, with the differences in gene expression between neuronal subtypes being determined prior to DNMT3A binding and mCA deposition.

mCA Controls Cell-Type-Specific Gene Expression by Recruiting MECP2 to Gene Bodies

To begin to investigate how DNMT3A catalyzed mCA functions to restrain the expression of lowly expressed genes, we next investigated the neuronal subtype-specific binding patterns of MECP2 and the effect of disrupting MECP2 function on neuronal subtype-specific gene transcription. Mutations in *MECP2* cause RTT, and this disorder has been suggested to be due to a loss of binding of MECP2 to mCG and mCA sequences across the neuronal genome (Chen et al., 2015; Gabel et al., 2015; Guo

et al., 2014; Lagger et al., 2017). However, whether the selective binding of MECP2 to mCA sequences in different neuronal subtypes accounts for neuronal subtype-specific gene misregulation in RTT is not known. In mixed neuronal populations, MECP2 has been shown to bind ubiquitously across the genome, accumulating at near histone-octamer levels (Skene et al., 2010), and studies have suggested that MECP2 might bind to DNA to some extent, regardless of the DNA methylation status (Baubec et al., 2013; Rube et al., 2016). Because accumulating evidence indicates that MECP2 binds avidly to mCA sequences in neurons to recruit the transcriptional corepressor NCOR and thereby temper gene transcription, we asked whether MECP2 might function to enforce the neuronal subtype-specific dampening of the transcription of mCA-marked genes.

To investigate this possibility, we performed MECP2 ChIP-seq using anti-MECP2-specific antibodies to map MECP2 binding across the genomes of *Pv* and *Vip* neurons that we isolated from the adult cortex. By aligning the *Pv* and *Vip* MECP2 ChIP-seq data with the *Pv* and *Vip* mCA data, respectively (Figure 7A), we found that MECP2 binding within the transcribed regions of genes is significantly correlated with the mCA distribution in both *Pv* and *Vip* neurons, providing further evidence that MECP2 binds to mCA. Moreover, we found that genes enriched for mCA specifically in *Pv* neurons versus *Vip* neurons are likewise enriched for MECP2 binding and vice versa (Figures 7B–7D, S7A, and S7B). We conclude that the binding pattern of MECP2 is at least in part neuronal subtype specific and that the subtype-specific pattern of MECP2 binding is associated with cell-type-specific differences in intragenic mCA patterns.

To test whether the differences in MECP2 binding between neuronal subtypes might explain the distinct patterns of gene misregulation observed in *Pv* and *Vip* neurons from the *Dnmt3a* cKO cortex, we performed high-throughput single-nuclei sequencing of adult *Mecp2* KO and wild-type control cortices (Figure S7C). By comparing the patterns of gene misregulation in *Pv*- and *Vip*-expressing nuclei from wild-type and *Mecp2* KO mice, we found that *Mecp2* KO neurons exhibit defects in gene expression such that genes with a high density of mCA within their transcribed region (and thus a high density of bound MECP2 in the same region) were preferentially upregulated (Figures 7E, 7F, and S7D), with the pattern of gene misregulation being quite similar to that detected in the same subtype of neurons obtained from *Dnmt3a* cKO mice (Figure 6D). These findings collectively suggest that mCA contributes to the fine-tuning of genes, including those with critical neuronal functions, in a neuronal subtype-specific manner at least in part by differentially recruiting MECP2 to neuronal gene bodies. Once bound to mCA, MECP2 appears to restrain gene transcription to a level of expression that is directly correlated with the number of mCA marks and MECP2 binding sites per gene, thus preferentially regulating some of the longest genes in the genome (Kinde et al., 2016; Lagger et al., 2017; Sugino et al., 2014). We note that the long genes that are upregulated when MeCP2 function is disrupted are typically lowly expressed in wild-type neurons (Figure S7E). These results suggest that transcriptional dysregulation of lowly expressed genes might contribute to RTT and possibly other neurodevelopmental disorders.

DISCUSSION

Our findings suggest that in the mammalian nervous system, DNMT3A binds transiently in a neuronal subtype-specific manner across the genome during early life, thereby catalyzing the methylation of nearby CA sequences and thus laying down a pattern of neuronal mCA that potentially lasts the lifetime of the animal. The probability that a given cytosine within a gene is methylated by DNMT3A is, in part, dependent on the level of expression of the gene such that the less the gene is expressed, the more likely cytosines within its transcribed region are to become methylated. While this appears to be a general trend, we note that a small number of high-mCA genes are highly expressed (Figure S7F).

mCA within a gene recruits MECP2 to fine-tune the expression of the gene such that the more mCA and MECP2 are recruited to the gene body, the more restrained expression of the gene becomes (Figure S7G). In contrast to enhancer-based mechanisms that provide coarse control of gene expression, thus augmenting the expression of actively transcribed genes over a wide range of expression levels anywhere between 2- and 1,000-fold, the fine control of gene expression by DNA methylation within gene bodies appears to provide precise regulation of lowly expressed genes with expression levels ranging from 1- to 2-fold at most.

The precise control of transcription is critical for proper neuronal function inasmuch as aberrant transcription of lowly expressed genes, when either DNMT3A or MECP2 are mutated, leads to significant defects in brain development and function. One possibility is that mCA- and MECP2-dependent fine control of gene expression has evolved to control the expression of genes whose transcription is sufficiently low that typical transcription-factor-bound regulatory elements, which are generally capable of coarse control of transcription, would not be able to achieve the precise level of gene expression required for normal neuronal function.

Previous studies have provided evidence that dynamic changes in mCG also occur in response to neuronal activity at specific loci in the genome, although the mechanism and functional significance of the changes remain elusive (Lister and Mukamel, 2015). These findings appear to be consistent with the observation that mCG is prone to oxidation by ten-eleven translocation (TET) proteins and thus is prone to active removal (Pastor et al., 2013). Here, we suggest a mechanism by which the deposition of mCA can potentially be affected by neuronal activity at a time during development when sensory input plays a key role in neuronal maturation. Given that once the mCA mark is deposited during this critical period in brain development, it appears to largely remain stable, our findings suggest that patterns of activity in the brain—specifically early in life—may trigger lasting changes in the epigenetic states of genes. This then may have a long-term effect on gene expression and neuronal function much later in life. We note that KA treatment during early life results in subtle changes in mCA, and by single-nuclei sequencing, we observe cell-type-specific changes in expression of a few thousand genes in the adult hippocampus

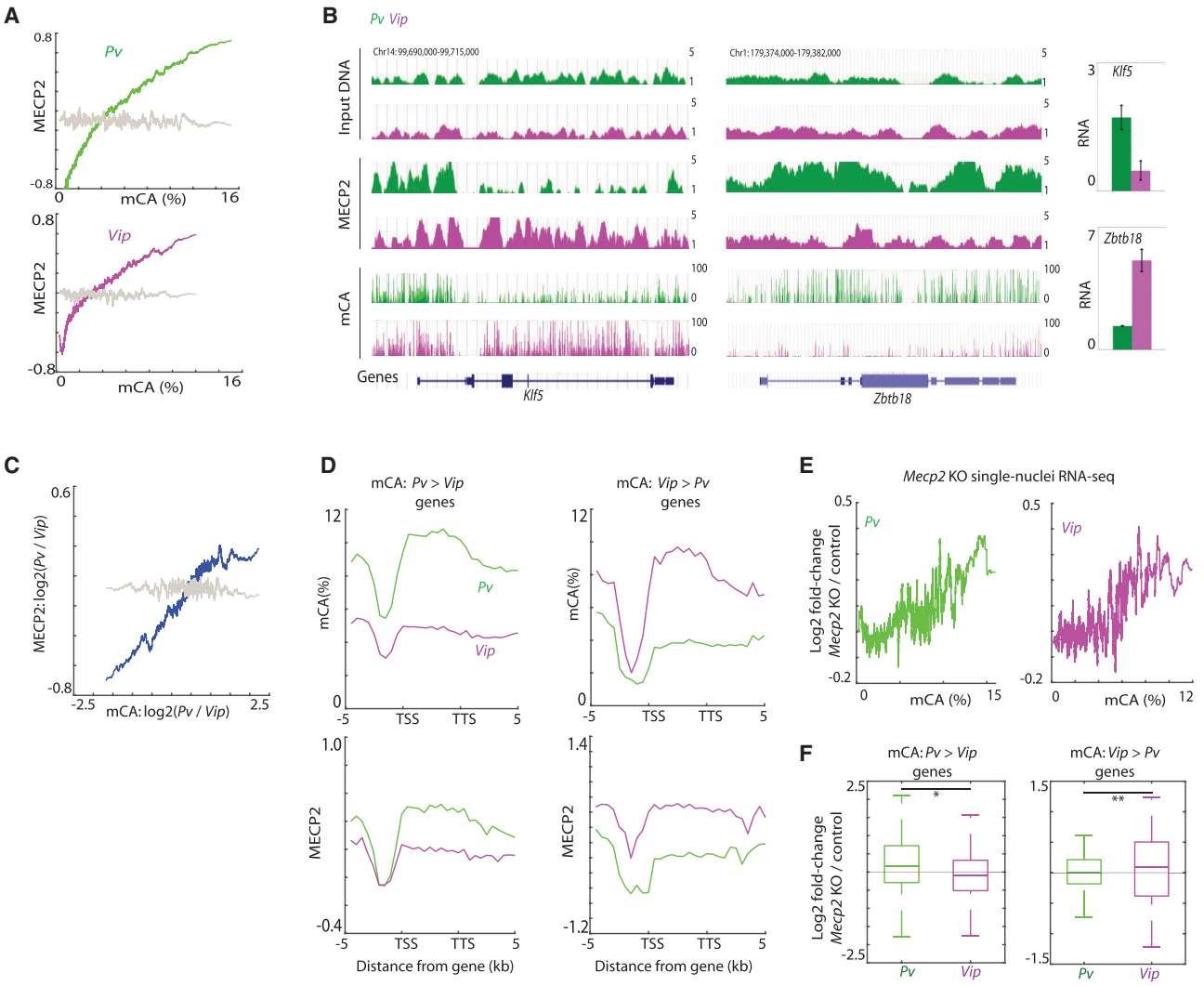


Figure 7. Cell-Type-Specific mCA Directs MECP2-Mediated Gene Repression

(A) Correlation between mCA and MECP2 binding in *Pv* and *Vip* neurons in the 8-week cortex within genes. The average MECP2 density relative to input DNA was binned according to mCA density within genes. MECP2 densities after random grouping of genes are shown as controls in gray. $p < 0.001$, permutation test.

(B) Genome browser views of mCA and MECP2 profiles in *Pv* and *Vip* neurons. Gene expression level determined by INTACT-purified RNA-seq is shown on right. Error bars represent SD between two biological replicates.

(C) Correlation between the difference in mCA between *Pv* and *Vip* neurons and MECP2 binding. MECP2 densities after random grouping of genes are shown as controls in gray $p < 0.001$, permutation test.

(D) Average distribution of MECP2 over genes differentially CA-methylated. (mCA *Pv* > *Vip* genes, n = 872; *Vip* > *Pv* genes, n = 84). Distributions of mCA over the genes are also shown.

(E) Correlation between mCA densities within the gene bodies and transcriptional defects in *Mecp2* KO cortex. $p < 0.001$ for *Pv* and $p = 0.001$ for *Vip*, permutation test.

(F) Cell-type-specific gene repression defects in *Mecp2* KO neurons. Genes were separated based on the mCA densities within the gene bodies in *Pv* and *Vip* neurons. *p = 3.6e−5, **p = 0.04, Kolmogorov-Smirnov test.

See also Figure S7.

(not shown). While at this stage we cannot directly assess the causality between the changes in mCA and gene expression, future studies should elucidate the overall role of experience-dependent changes in mCA on gene expression and neuronal function.

Consistent with the findings that DNMT3A catalyzes the deposition of mCA and MECP2 binds mCA and serves as a reader of mCA, humans and mice that have mutations in either DNMT3A or MECP2 appear to exhibit somewhat similar neurological deficits, including intellectual disability in humans and learning and memory deficits in mice (Chahrour and Zoghbi, 2007; Lister and Mukamel, 2015; Tatton-Brown et al., 2014). In addition, both adult *Mecp2* KO and *Dnmt3a* cKO mice are less active and have weaker grip strength and reduced motor

coordination (Figure S7H) (Chahrour and Zoghbi, 2007; Nguyen et al., 2007). In both *Dnmt3a* cKO and *Mecp2* KO mice, aberrant phenotypes are less apparent at birth and become more obvious as the mice mature (Figure S7I) (Chahrour and Zoghbi, 2007). Notably, *Mecp2* KO mice have been more thoroughly characterized than *Dnmt3a* cKO mice to date. In the *Mecp2* KO mice, neurological symptoms appear around the time during the postnatal period when DNMT3A begins to catalyze mCA across the transcribed regions of lowly expressed genes, suggesting that both RTT and the DNMT3A deficiency syndrome may be a consequence of the disruption of mCA-dependent gene regulation in neurons.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENTS AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mice
- METHOD DETAILS
  - Kainic Acid (KA) Treatment
  - Dark-Rearing/ Light Exposure Experiments
  - DNMT3A Immunoprecipitation
  - Chromatin Immunoprecipitation Sequencing (ChIP-Seq)
  - Immunoblotting
  - Micrococcal Nuclease Sequencing (MNase-Seq)
  - Whole Genome Bisulfite-Sequencing (WGBS)
  - Oxidative Bisulfite Sequencing (oxBS-Seq)
  - In Vitro DNA methyltransferase activity assay
  - RNA Sequencing (RNA-Seq)
  - INTACT Nuclei Isolation
  - Nuclear RNA-Seq
  - MECP2 ChIP-seq on SUN1 purified nuclei
  - Single-nuclei RNA sequencing (inDrops)
- QUANTIFICATION AND STATISTICAL AVAILABILITY
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental information includes seven figures and can be found with this article online at https://doi.org/10.1016/j.cell.2017.09.047.

AUTHOR CONTRIBUTIONS

H.S. designed the research and performed most of the experiments. H.S. and S.C.S. performed nuclei isolations. S.H., M.A.N., and D.R.H. established the computational pipeline for inDrops RNA-seq analyses. A.W.G. performed DNMT3A IP experiments. W.R. and L.D.B. generated the *Mecp2* KO inDrops data. B.K. and H.W.G. provided *Dnmt3a*$^{fl/fl}$ mice. M.E.G. advised on all aspects of the study. H.S. and M.E.G. wrote the manuscript.

ACKNOWLEDGMENTS

We thank J. Nathans and A. Mo for SUN1 mice and E. Griffith, E. Pollina, and T. Cherry for editorial assistance; P. Zhang, J. Wang, and K. Fazioli for assistance with the mouse colony; and G. Mandel, A. Bird, and M. Morselli for helpful discussions. This work was supported by grants from the Rett Syndrome Research Trust and the NIH (RO1NS048276) to M.E.G. and the NARSAD and Hearst funds to H.S. S.C.S. was supported by the NRSA training grant and fellowship (T32NS007484 and F32NS089186). H.S. is an HHMI Fellow of the Damon Runyon Cancer Research Foundation (DRG-2194-14).

Received: February 21, 2017
Revised: July 17, 2017
Accepted: September 25, 2017
Published: October 19, 2017

REFERENCES

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Baubec, T., Ivánek, R., Lienert, F., and Schübeler, D. (2013). Methylation-dependent and -independent genomic targeting principles of the MBD protein family. Cell *153*, 480–492.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schübeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature *520*, 243–247.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. *16*, 6–21.

Bloodgood, B.L., Sharma, N., Browne, H.A., Trepman, A.Z., and Greenberg, M.E. (2013). The activity-dependent transcription factor NPAS4 regulates domain-specific inhibition. Nature *503*, 121–125.

Booth, M.J., Branco, M.R., Ficz, G., Oxley, D., Krueger, F., Reik, W., and Balasubramanian, S. (2012). Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science *336*, 934–937.

Chahrour, M., and Zoghbi, H.Y. (2007). The story of Rett syndrome: from clinic to neurobiology. Neuron *56*, 422–437.

Chen, W.G., Chang, Q., Lin, Y., Meissner, A., West, A.E., Griffith, E.C., Jaenisch, R., and Greenberg, M.E. (2003). Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. Science *302*, 885–889.

Chen, L., Chen, K., Lavery, L.A., Baker, S.A., Shaw, C.A., Li, W., and Zoghbi, H.Y. (2015). MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. Proc. Natl. Acad. Sci. USA *112*, 5509–5514.

Feng, J., Chang, H., Li, E., and Fan, G. (2005). Dynamic expression of de novo DNA methyltransferases Dnmt3a and Dnmt3b in the central nervous system. J. Neurosci. Res. *79*, 734–746.

Gabel, H.W., Kinde, B., Stroud, H., Gilbert, C.S., Harmin, D.A., Kastan, N.R., Hemberg, M., Ebert, D.H., and Greenberg, M.E. (2015). Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. Nature *522*, 89–93.

Gibson, W.T., Hood, R.L., Zhan, S.H., Bulman, D.E., Fejes, A.P., Moore, R., Mungall, A.J., Eydoux, P., Babul-Hirji, R., An, J., et al.; FORGE Canada Consortium (2012). Mutations in EZH2 cause Weaver syndrome. Am. J. Hum. Genet. *90*, 110–118.

Gunnersen, J.M., Kim, M.H., Fuller, S.J., De Silva, M., Britto, J.M., Hammond, V.E., Davies, P.J., Petrou, S., Faber, E.S., Sah, P., and Tan, S.S. (2007). Sez-6 proteins affect dendritic arborization patterns and excitability of cortical pyramidal neurons. Neuron *56*, 621–639.

Guo, J.U., Su, Y., Shin, J.H., Shin, J., Li, H., Xie, B., Zhong, C., Hu, S., Le, T., Fan, G., et al. (2014). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. Nat. Neurosci. *17*, 215–222.

Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G., et al. (2015). Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. Nature *517*, 640–644.

Guy, J., Hendrich, B., Holmes, M., Martin, J.E., and Bird, A. (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat. Genet. *27*, 322–326.

He, Y., and Ecker, J.R. (2015). Non-CG Methylation in the Human Genome. Annu. Rev. Genomics Hum. Genet. *16*, 55–77.

Hippenmeyer, S., Vrieseling, E., Sigrist, M., Portmann, T., Laengle, C., Ladle, D.R., and Arber, S. (2005). A developmental switch in the response of DRG neurons to ETS transcription factor signaling. PLoS Biol. *3*, e159.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. *33* (*Suppl*), 245–254.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900–903.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Kinde, B., Wu, D.Y., Greenberg, M.E., and Gabel, H.W. (2016). DNA methylation in the gene body influences MeCP2-mediated gene repression. Proc. Natl. Acad. Sci. USA *113*, 15114–15119.

Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell *161*, 1187–1201.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929–930.

Lagger, S., Connelly, J.C., Schweikert, G., Webb, S., Selfridge, J., Ramsahoye, B.H., Yu, M., He, C., Sanguinetti, G., Sowers, L.C., et al. (2017). MeCP2 recognizes cytosine methylated tri-nucleotide and di-nucleotide sequences to tune transcription in the mammalian brain. PLoS Genet. *13*, e1006793.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Li, B.Z., Huang, Z., Cui, Q.Y., Song, X.H., Du, L., Jeltsch, A., Chen, P., Li, G., Li, E., and Xu, G.L. (2011). Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. Cell Res. *21*, 1172–1181.

Lister, R., and Mukamel, E.A. (2015). Turning over DNA methylation in the mind. Front. Neurosci. *9*, 252.

Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D., et al. (2013). Global epigenomic reconfiguration during mammalian brain development. Science *341*, 1237905.

Mamiya, P.C., Hennesy, Z., Zhou, R., and Wagner, G.C. (2008). Changes in attack behavior and activity in EphA5 knockout mice. Brain Res. *1205*, 91–99.

Mardinly, A.R., Spiegel, I., Patrizi, A., Centofante, E., Bazinet, J.E., Tzeng, C.P., Mandel-Brehm, C., Harmin, D.A., Adesnik, H., Fagiolini, M., and Greenberg, M.E. (2016). Sensory experience regulates cortical inhibition by inducing IGF1 in VIP neurons. Nature *531*, 371–375.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343–349.

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneurons of the neocortical inhibitory system. Nat. Rev. Neurosci. *5*, 793–807.

Mo, A., Mukamel, E.A., Davis, F.P., Luo, C., Henry, G.L., Picard, S., Urich, M.A., Nery, J.R., Sejnowski, T.J., Lister, R., et al. (2015). Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain. Neuron *86*, 1369–1384.

Nguyen, S., Meletis, K., Fu, D., Jhaveri, S., and Jaenisch, R. (2007). Ablation of de novo DNA methyltransferase Dnmt3a in the nervous system leads to neuromuscular defects and shortened lifespan. Dev. Dyn. *236*, 1663–1676.

Nord, A.S., Blow, M.J., Attanasio, C., Akiyama, J.A., Holt, A., Hosseini, R., Phouanenavong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., et al. (2013). Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. Cell *155*, 1521–1531.

Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D., et al. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714–717.

Pastor, W.A., Aravind, L., and Rao, A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. Nat. Rev. Mol. Cell Biol. *14*, 341–356.

Rose, N.R., and Klose, R.J. (2014). Understanding the relationship between DNA methylation and histone lysine methylation. Biochim. Biophys. Acta *1839*, 1362–1372.

Rube, H.T., Lee, W., Hejna, M., Chen, H., Yasui, D.H., Hess, J.F., LaSalle, J.M., Song, J.S., and Gong, Q. (2016). Sequence features accurately predict genome-wide MeCP2 binding in vivo. Nat. Commun. *7*, 11025.

Sanders, S.J., He, X., Willsey, A.J., Ercan-Sencicek, A.G., Samocha, K.E., Cicek, A.E., Murtha, M.T., Bal, V.H., Bishop, S.L., Dong, S., et al.; Autism Sequencing Consortium (2015). Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. Neuron *87*, 1215–1233.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502.

Shen, X., Liu, Y., Hsu, Y.J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.C., and Orkin, S.H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol. Cell *32*, 491–502.

Skene, P.J., Illingworth, R.S., Webb, S., Kerr, A.R., James, K.D., Turner, D.J., Andrews, R., and Bird, A.P. (2010). Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. Mol. Cell *37*, 457–468.

Song, C.X., Yi, C., and He, C. (2012). Mapping recently identified nucleotide variants in the genome and transcriptome. Nat. Biotechnol. *30*, 1107–1116.

Sugino, K., Hempel, C.M., Okaty, B.W., Arnson, H.A., Kato, S., Dani, V.S., and Nelson, S.B. (2014). Cell-type-specific repression by methyl-CpG-binding protein 2 is biased toward long genes. J. Neurosci. *34*, 12877–12883.

Sztainberg, Y., and Zoghbi, H.Y. (2016). Lessons learned from studying syndromic autism spectrum disorders. Nat. Neurosci. *19*, 1408–1417.

Taniguchi, H., He, M., Wu, P., Kim, S., Paik, R., Sugino, K., Kvitsiani, D., Fu, Y., Lu, J., Lin, Y., et al. (2011). A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. Neuron *71*, 995–1013.

Tatton-Brown, K., Seal, S., Ruark, E., Harmer, J., Ramsay, E., Del Vecchio Duarte, S., Zachariou, A., Hanks, S., O'Brien, E., Aksglaede, L., et al.; Childhood Overgrowth Consortium (2014). Mutations in the DNA methyltransferase gene DNMT3A cause an overgrowth syndrome with intellectual disability. Nat. Genet. *46*, 385–388.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. *32*, 381–386.

Tronche, F., Kellendonk, C., Kretz, O., Gass, P., Anlag, K., Orban, P.C., Bock, R., Klein, R., and Schütz, G. (1999). Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety. Nat. Genet. *23*, 99–103.

Vissers, L.E., Gilissen, C., and Veltman, J.A. (2016). Genetic studies in intellectual disability and related disorders. Nat. Rev. Genet. *17*, 9–18.

von Schimmelmann, M., Feinberg, P.A., Sullivan, J.M., Ku, S.M., Badimon, A., Duff, M.K., Wang, Z., Lachmann, A., Dewell, S., Ma'ayan, A., et al. (2016). Polycomb repressive complex 2 (PRC2) silences genes responsible for neurodegeneration. Nat. Neurosci. *19*, 1321–1330.

Wilkinson, D.G. (2001). Multiple roles of EPH receptors and ephrins in neural development. Nat. Rev. Neurosci. *2*, 155–164.

Xi, Y., and Li, W. (2009). BSMAP: whole genome bisulfite sequence MAPping program. BMC Bioinformatics *10*, 232.

Yu, Z.L., Jiang, J.M., Wu, D.H., Xie, H.J., Jiang, J.J., Zhou, L., Peng, L., and Bao, G.S. (2007). Febrile seizures are associated with mutation of seizure-related (SEZ) 6, a brain-specific gene. J. Neurosci. Res. *85*, 166–172.

Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics *25*, 1952–1958.

Zhang, J., Ji, F., Liu, Y., Lei, X., Li, H., Ji, G., Yuan, Z., and Jiao, J. (2014). Ezh2 regulates adult hippocampal neurogenesis and memory. J. Neurosci. *34*, 5184–5199.

Zilionis, R., Nainys, J., Veres, A., Savova, V., Zemmour, D., Klein, A.M., and Mazutis, L. (2017). Single-cell barcoding and sequencing using droplet microfluidics. Nat. Protoc. *12*, 44–73.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-DNMT3A	Santa Cruz	Cat# 20703; RRID: AB_2093990
Rabbit anti-H3.1/H3.2	EMD Millipore	Cat# ABE154
Rabbit anti-H3.3	EMD Millipore	Cat# 09-838; RRID: AB_10845793
Rabbit anti-H2A.Z	EMD Millipore	Cat# 07-594; RRID: AB_390152
Rabbit anti-H3K4me3	EMD Millipore	Cat# 07-473; RRID: AB_1977252
Mouse anti-H3K9me2	Abcam	Cat# ab1220; RRID: AB_449854
Rabbit anti- H3K9me3	Abcam	Cat# ab8898; RRID: AB_306848
Rabbit anti-H3K27me2	Cell Signaling	Cat# 9728; RRID: AB_1281338
Rabbit anti-H3K27me3	EMD Millipore	Cat# 07-449; RRID: AB_310624
Rabbit anti-H3K36me3	Abcam	Cat# ab9050; RRID: AB_306966
Rabbit anti-H3	Abcam	Cat# ab1791; RRID: AB_302613
Rabbit anti-MECP2	Chen et al., 2003	N/A
Rabbit anti-H3K27ac	Abcam	Cat# ab4729; RRID: AB_2118291
Mouse anti-RNA Pol II	Abcam	Cat# ab817; RRID: AB_306327
Rabbit anti-RNA Pol II Phospho-Ser2	Abcam	Cat# ab5095; RRID: AB_304749
Mouse anti-DNMT3A	Abcam	Cat# ab13888; RRID: AB_300714
Rabbit anti-GFP	Thermo Fisher	Cat# G10362; RRID: AB_2536526
Biological Samples		
Unmethylated lambda DNA	Promega	Cat# D1521
Chemicals, Peptides, and Recombinant Proteins		
Kanic acid	Sigma Aldrich	Cat# K0250
Dynabeads Protein G	Thermo Fisher	Cat# 10004D
Dynabeads Protein A	Thermo Fisher	Cat# 10002D
NuPAGE Novex 4-12% Bis-Tris gel	Thermo Fisher	Cat# NP0335BOX
Biotinylated histone H3	Epicypher	Cat# 12-0001
Human DNMT3A	Abcam	Cat# 170408
Agencourt AMPure XP beads	Beckman Coulter	Cat# A63881
Formaldehyde	Sigma Aldrich	Cat# F8775
Critical Commercial Assays		
Ovation Ultralow Library System	Nugen	Cat# 0330, 0331, 0344
TruSeq Strand Specific RNA-Sequencing kit	Illumina	Cat# RS-122-2101
NEBNext Ultra Directional RNA Library Prep Kits	New England Biolabs	Cat# E7420L
TrueMethylSeq Kit	Cambridge Epigenetix	N/A
NEBNext Multiplex Oligos for Illumina	New England Biolabs	Cat# E7335L
GeneRead rRNA Depletion Kit	QIAGEN	Cat# 180211
NEBNext rRNA Depletion Kit	New England Biolabs	Cat# E6310X
Qubit dsDNA HS Assay Kit	Thermo Fisher	Cat# Q32854
Deposited Data		
Raw data for sequencing	NCBI Gene Expression Omnibus	GEO: GSE103214
Experimental Models: Organisms/Strains		
Mouse: C57BL/6 WT	The Jackson Laboratory	N/A
Mouse: C57BL/6 WT	Charles River	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse: *Nestin-cre*	The Jackson Laboratory	Cat# 003771
Mouse: *Dnmt3a*$^{fl/fl}$	Kaneda et al., 2004	N/A
Mouse: *Ezh2* $^{fl/fl}$	The Jackson Laboratory	Cat# 022616
Mouse: SUN1-2xsfGFP-6xMYC	The Jackson Laboratory	Cat# 021039
Mouse: *Pv-cre*	The Jackson Laboratory	Cat# 017320
Mouse: *Vip-cre*	The Jackson Laboratory	Cat# 010908
Mouse: *Mecp2* KO	The Jackson Laboratory	Cat# 003890
Oligonucleotides		
Random hexamer primers: TCGGCATTCCTGCTGAA CCGCTCTTCCGATCTNNNNNN	Klein et al., 2015; Zilionis et al., 2017	N/A
PE1 primer: AATGATACGGCGACCACCGAGATCTAC ACTCTTTCCCTACACGA	Klein et al., 2015; Zilionis et al., 2017	N/A
PE2 primer: CAAGCAGAAGACGGCATACGAGATCGGTCTCGGC ATTCCTGCTGAAC	Klein et al., 2015; Zilionis et al., 2017	N/A
Software and Algorithms		
Bowtie V1	Langmead et al., 2009	http://bowtie-bio.sourceforge.net/index.shtml
SICER	Zang et al., 2009	N/A
BSmap V2.74	Xi and Li, 2009	https://code.google.com/archive/p/bsmap/
DAVID V6.8	Huang da et al., 2009	https://david.ncifcrf.gov/
Tophat2	Kim et al., 2013	https://ccb.jhu.edu/software/tophat/index.shtml
HTseq	Anders et al., 2015	https://pypi.python.org/pypi/HTSeq
DEseq	Anders et al., 2015	http://bioconductor.org/packages/release/bioc/html/DESeq.html
Indrops mapping	Klein et al., 2015	https://github.com/indrops/indrops
Seurat	Satija et al., 2015	http://satijalab.org/seurat/
Monocle	Trapnell et al., 2014	http://cole-trapnell-lab.github.io/monocle-release/

CONTACT FOR REAGENTS AND RESOURCE SHARING

Further information and requests for reagents should be directed to the Lead Contact Michael E. Greenberg (michael_greenberg@hms.harvard.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

All wild-type mice used in this study were in the C57BL/6J background. For conditional *Dnmt3a* knockout experiments, *Nestin-cre* mice (Tronche et al., 1999) were crossed to *Dnmt3a*$^{fl/fl}$ mice (Kaneda et al., 2004), where *Nestin-cre* was always introduced from the male mouse. Littermate controls were used for all comparisons between wild-type (*Dnmt3a*$^{fl/fl}$) and *Dnmt3a* cKO (*Nestin-cre*; *Dnmt3a*$^{fl/fl}$) mice. For conditional *Ezh2* knockout experiments, similarly *Nestin-cre* mice were crossed to *Ezh2* $^{fl/fl}$ mice (Shen et al., 2008), and littermate controls were used for all comparisons between wild-type (*Ezh2* $^{fl/fl}$) and *Ezh2* cKO (*Nestin-cre*; *Ezh2* $^{fl/fl}$) mice. Homozygous SUN1-2xsfGFP-6xMYC mice were generated in a previous study (Mo et al., 2015) and were crossed with homozygous *Pv-cre* (Hippenmeyer et al., 2005) and *Vip-cre* mice (Taniguchi et al., 2011). *Mecp2* KO mice have been previously described (Guy et al., 2001).

METHOD DETAILS

Kainic Acid (KA) Treatment

Postnatal day 14 (P14) mice were injected with either kanic acid (Sigma Aldrich K0250) (5mg/kg) or 1X PBS for controls. KA-injected animals began to seize ~10 minutes after injection, and hippocampi were dissected an hour post-seizure. Hippocampi of treated and untreated mice from the same litter were pooled for ChIP and RNA-seq experiments (two to three mice each). For daily KA administration, we exposed mice to daily injections of KA for ten days beginning day ten postnatally, and then allowed the mice to mature into adults without any injections (eight weeks). Littermate mice were used for all comparisons between KA-treated and control mice.

Dark-Rearing/ Light Exposure Experiments

Mice were housed in a standard light cycle until P10 (before eye opening). The mice were then dark-housed until P17. For light stimulation, animals were transferred to the light chamber for one hour prior to dissection of visual cortices. The remaining mice were sacrificed in the dark and not exposed to light (dark-housed controls). The eyes of all animals were enucleated before dissecting the visual cortices. Littermate mice were used for all comparisons between light-exposed and dark-reared mice.

DNMT3A Immunoprecipitation

Cortices from two-week-old wild-type mice or *Dnmt3a* cKO mice were homogenized in 8 mL buffer A (10 mM HEPES pH 8.0, 25 mM KCl, 1 mM EDTA, 0.5 mM EGTA, 2M sucrose, 10% glycerol, 1mM DTT, 0.15 mM spermine, 0.5 mM spermidine, Roche protease inhibitor tablets, phosphatase inhibitors 2/3) using a Dounce homogenizer with the "Tight" pestle. Homogenized cortices were rotated at 4°C for 10 min then layered on top of a 2 mL cushion of buffer A and spun at 24,000 rpm for 40 min at 4°C. The nuclei pellet was resuspended in 1 mL buffer NE1 (20 mM HEPES pH 8.0, 10 mM KCl, 1 mM MgCl2, 0.1% Triton X-100, 1 mM DTT, protease inhibitors, phosphatase inhibitors), and was spun at 800 g for 5 min at 4°C. The pellet was resuspended in an equal volume of NE1 (20-30 uL). Three uL of benzonase (Sigma) were added, and the sample incubated for one hour at 4°C. The volume of NE1 was raised to 500 uL, and the concentration of Triton X-100 raised to 0.5%. NaCl was added to a final concentration of 150 mM. The sample was rotated at 4°C for 20 min, then spun at 16,000 g for 20 min. The supernatant was pre-cleared by incubating with 40 uL Dynabeads protein G (Thermo) for 1 hr, then transferred to a tube containing 40 ul Dynabeads protein G pre-bound with anti-DNMT3A antibodies (Abcam ab13888), and rotated for one hour at 4°C. After incubation, the beads were washed four times with NE1+150 mM NaCl+0.5% Triton X-100, and finally eluted off the beads by adding 50 ul 2X LDS buffer (Thermo) and incubating at 95°C for 10 min.

Chromatin Immunoprecipitation Sequencing (ChIP-Seq)

Tissue was crosslinked 1% formaldehyde. The reaction was stopped with 0.125M glycine. After one wash with PBS, samples were lysed twice with L1 buffer (50mM HEPES pH7.5, 140mM NaCl, 1mM EDTA, 1mM EGTA, 0.25% Triton X-100, 0.5% NP40, 10% Glycerol, protease inhibitors). The samples were incubated in L2 buffer (10mM Tris-HCl pH8.0, 200mM NaCl), then lysed in L3 buffer (1M Tris-HCl pH8.0, 5M NaCl, 0.5 M EDTA, 0.5 M EGTA, 10% Na-Deoxycholate, 20% N-lauroylsarcosine), and sonicated with a Bioruptor (Diagenode). Sonicated chromatin was pre-cleared by incubating with Protein A Dynabeads for 2 hours. Pre-cleared chromatin was added to Dynabeads conjugated to specific antibodies and incubating overnight at 4°C, and 1 to 2.5% of chromatin were kept as input material. Antibodies used: DNMT3A (Santa Cruz 20703), H3.1/H3.2 (EMD Millipore ABE154), H3.3 (EMD Millipore 09-838), H2A.Z (EMD Millipore 07-594), H3K4me3 (EMD Millipore 07-473), H3K9me2 (Abcam ab1220), H3K9me3 (Abcam ab8898), H3K27me2 (Cell Signaling 9728), H3K27me3 (EMD Millipore 07-449), H3K36me3 (Abcam ab9050), H3 (Abcam ab1791), MECP2 (Chen et al., 2003), H3K27ac (Abcam ab4729), RNA Pol II (Abcam 817) and RNA Pol II Ser2P (Abcam 5095). Beads were washed twice with Low Salt buffer (0.1% SDS, 1% Triton X-100, 20mM Tris HCl pH8.0, 150mM NaCl, 2mM EDTA), High Salt buffer (0.1% SDS, 1% Triton X-100, 20mM Tris HCl pH8.0, 500mM NaCl, 2mM EDTA) and LiCl buffer (250mM LiCl, 1% NP40, 1mM EDTA, 10mM TrisHCl pH8.0, 1% Sodium Deoxycholate) at 4°C, and washed once with TE buffer at room temperature. Chromatin was eluted off the beads with TE+1%SDS at 65°C, and de-crosslinked by incubating overnight at 65°C. Samples were treated with RNaseA, and then with Proteinase K for 2 hours at 55°C. DNA was extracted by phenol-chloroform extraction and purified with a Minelute PCR cleanup column (QIAGEN). Libraries were constructed by using Ovation Ultralow Library Systems (Nugen) following manufacturer instructions. Libraries were sequenced on an Illumina Hiseq 2500 or Nextseq 500. Reads were mapped to the mm9 genome by allowing up to two mismatches using Bowtie (Langmead et al., 2009). Reads mapping to identical locations were collapsed into one read. For the analysis, reads were normalized to the total number of uniquely mapping reads. To define DNMT3A-enriched regions in the genome, we used SICER (Zang et al., 2009) and compared two-week old DNMT3A-immunoprecipitated samples to input DNA (parameters: W = 1,000; G = 2,000; FDR < 0.05). For building linear models for DNMT3A binding, ChIP-seq densities were calculated within TSS+/−1kb for promoters, 5kb tiles within gene bodies (excluding 3 kilobases downstream of promoters) and 5kb tiles across intergenic regions (at least 10 kb away from annotated genes). To normalize for nucleosome occupancy, histone ChIP-seq data were normalized to histone H3, whereas DNMT3A and H3 were normalized to input DNA. The model was built using the linear regression function lm() in R. To define regions of the genome enriched in histone marks, we used SICER. For broadly enriched histone marks (i.e., H3.1, H3.3, H3K9me2, H3K9me3, H3K27me2, H3K27me3, H3K36me3) or Pol II Ser2P, parameters of W = 500; G = 1,500; FDR < 0.05) was used, and for histone marks present as peaks (i.e., H3K4me3, H3K27ac, H2A.Z) or Pol II,

parameters of W = 500; G = 500; FDR < 0.001 was used. Input DNA was used as controls for H3.1, H3.3, Pol II and Pol II Ser2P, whereas H3 ChIP was used as controls of the other histone marks. For plotting average distributions of ChIP-seq reads across specific genomic elements, DNMT3A, POLII and POLII Ser2P were normalized to input DNA, and histone modifications were normalized to histone H3 (log2 ratio).

Immunoblotting

Cortices were lysed in RIPA buffer, ran on a NuPAGE Novex 4%–12% Bis-Tris gel (ThermoFisher), and transferred to a nitrocellulose membrane, and anti-DNMT3A (1:200, Santa Cruz 20703) and anti-H3 (1:1000, Abcam 1791) antibodies were used for detecting the proteins.

Micrococcal Nuclease Sequencing (MNase-Seq)

Tissue were homogenized in L1 buffer and spun down. Pellets were resuspended in L2 buffer and rotated at room temperature for 10 min. Tubes were spun down and the pellet was resuspended in 100 μl MNase digestion buffer (20mM Tris-Cl, pH 7.5, 10mM NaCl, 2.5mM CaCl2, 0.01 mM PMSF, 1x Roche protease inhibitors) and incubated with 40μg of RNase A (QIAGEN), and incubated at room temperature for 10 min. Treated samples with 1 μl MNase (NEB) for 4min at 37C. The reaction was stopped with addition of 2 μl 0.5M EDTA. Triton X-100 was added to a final concentration of 1%, and samples were rotated at 4°C for 30 min. DNA was subsequently purified using QIAGEN PCR purification kit. Libraries were constructed by using Ovation Ultralow Library Systems (Nugen) following manufacturer instructions. Libraries were sequenced on a Hiseq 2500 or Nextseq 500. Reads were mapped to the mm9 genome by allowing up to two mismatches using Bowtie (Langmead et al., 2009). For the analysis reads were normalized to the total number of uniquely mapping reads. For analyses presented in Figure 1F, significance was assessed by repeating the random grouping and comparing the Spearman's correlation coefficients between the actual data and randomized data.

Whole Genome Bisulfite-Sequencing (WGBS)

WGBS libraries were constructed by using a similar method as previously described (Mo et al., 2015) except that NEXTflex methylated adapters (Bioo Scientific) were used.

Briefly, DNA was sonicated to 200 bp using a Covaris S2. Bisulfite treatment was performed using EZ DNA Methylation-Gold kit (Zymo). Unmethylated lambda DNA (0.1%–0.5%; Promega D1521) was spiked in to assess conversion rates. All datasets generated in this study were confirmed to have conversion error rates of less than 1%. Identical reads were collapsed, and reads were mapped to the mm9 genome using BSmap (Xi and Li, 2009) retaining uniquely mapping reads. Methylation levels were determined by calculating #C/(#C+#T). To calculate methylation within gene bodies, promoter regions, defined as transcription start site (TSS) to TSS+2 kb, were excluded from the analyses. Differentially methylated regions between the cortex and hippocampus were determined by comparing CA methylation levels in 5 kb tiles. To remove tiles with low sequencing coverage and/or with low frequency of CA sequences, tiles that had 100 or greater CA sites that are covered by four reads in both samples were kept. Tiles with absolute mCA difference of 0.02 and Benjamini-Hochberg corrected FDR < 0.01 (Fisher's exact test) were selected. Differentially methylated genes between the adult *Pv* and *Vip* neurons were determined by comparing CA methylation levels in within genes (excluding promoter regions), and genes that had 100 or greater CA sites that are covered by four reads in all samples were kept. Genes with absolute mCA difference of 4% and Benjamini-Hochberg corrected FDR < 0.01 (Fisher's exact test) in all biological replicate comparisons were selected. To identify 100 bp genomic tiles enriched in mCG or mCA, data from biological replicates were combined, and 100 bp tiles with > 20 #C+#T calls for CG (4,466,873) or > 150 #C+#T calls for CA (3,962,712) in all samples (*Pv* 1 wk, 8 wk, and *Vip* 1 wk, 8 wk) were selected. Tiles with greater than 70% mCG, or greater than 10% mCA were compared between the neuronal subtypes. For gene ontology analyses, genes within the top 5 percent in mCA densities within the gene bodies were analyzed using DAVID 6.8 (https://david.ncifcrf.gov/)(Huang da et al., 2009). For the analyses presented in Figures 6E and 7F the subtype defined with lower mCA densities, genes with mCA densities above 5% were excluded from the analyses. (mCA *Pv* > *Vip* genes: n = 345; *Vip* > *Pv* genes: n = 64).

Oxidative Bisulfite Sequencing (oxBS-Seq)

OxBS-seq was performed on DNA extracted from 12-week-old cortex. DNA was sonicated with a Covaris S2 to 200 bp and ligated to NEXTflex methylated adapters (Bioo Scientific), then bisulfite conversions and oxidation reactions were performed using the TrueMethyl oxBS-seq kit following the manufacturer instructions (Cambridge Epigenetix). Half the sample was used for BS-seq without oxidation treatment and half the sample was used for oxBS-seq as recommended by the manufacturer.

In Vitro DNA methyltransferase activity assay

100 ng of unmethylated lambda DNA (Promega D1521) was sonicated to 200 bp using a Covaris S2 and was incubated with 1 μg of full-length human DNMT3A (Abcam 170408) along with 5μM biotinylated histone H3 (amino acids 1-20; Epicypher 12-0001) and S-Adenosyl methionine (SAM) (NEB B9003S) in 0.5mg/mL BSA, 25mM Tris-Cl (pH 8). As a control, the same reaction except without the DNMT3A protein was performed in parallel. The reactions were incubated at 37°C for 1 hour. After the incubation, DNA was isolated by phenol-chloroform extraction followed by ethanol precipitation, and bisulfite sequencing libraries were prepared as described above.

RNA Sequencing (RNA-Seq)

Total RNA was extracted from cortices and hippocampi with TRIzol (Invitrogen) and purified with an RNeasy kit (with on-column DNase treatment) (QIAGEN) following manufacturer instructions. Libraries were generated with either the Tru-Seq Strand Specific RNA-Sequencing kit (Illumina) or NEBNext Ultra Directional RNA Library Prep Kits (NEB). Reads were mapped to annotated genes and the genome using Tophat2 (Kim et al., 2013) and quantified with HTseq (Anders et al., 2015). Gene expression levels were quantified by calculating reads per kilobase of transcript per million mapped reads (RPKM). To define KA-induced genes and light-induced genes, gene expression in KA-treated or light-exposed mice were compared to PBS-injected or dark-reared littermate controls, respectively. Because DNMT3A binding levels primarily varied within genes of wild-type expression levels between 1st and 4th deciles, we considered genes with gene expression levels greater than the 5th decile in the treated samples and gene expression levels less than the 1st decile in the control samples. We selected genes that were greater than two-fold in the KA-treated samples compared to control. To define lighted-induced genes, we selected genes that were greater than 1.5-fold in the lighted-exposed samples compared to dark-reared control and a FDR < 0.1 cutoff using DEseq (Anders et al., 2015). To define genes upregulated in *Ezh2* cKO compared to wild-type, we used a two-fold cutoff and a FDR < 0.1 cutoff using DEseq. For generating tracks for the genome browser, we tiled the genome into 20 bp bins, and computed number of reads mapping within the bins, divided this value by total mapping million reads to normalize for sequencing depth, and multiplied by an arbitrary constant.

INTACT Nuclei Isolation

The INTACT method employs a transgenic mouse that expresses in a CRE-dependent manner a SUN1-2xsfGFP-6xMYC fusion protein at the inner nuclear membrane (SUN1 mice). By crossing SUN1 mice with Cre driver lines that promote expression of the SUN1 fusion protein in *Pv* or *Vip* neurons, we obtained progeny that express the SUN1 fusion protein on the nuclear membrane of *Pv* or *Vip* neurons, respectively. SUN1-GFP expressing nuclei were isolated as previously described (Mo et al., 2015) using GFP antibodies (Fisher G10362) coupled with Protein G Dynabeads (Invitrogen 10003D). RNasein Plus RNase Inhibitors (Promega N2611) were added to the buffers to prevent degradation of RNA. One to two mice were pooled for each experiment. Typically ~10 million nuclei were isolated from a cortex from one animal. After the anti-GFP immunoprecipitation, essentially all (~100%) nuclei were GFP positive.

Nuclear RNA-Seq

RNA from SUN1-purified nuclei was extracted using Trizol and prepared as described for the standard RNA-seq (using random primers). For data analysis, duplicated reads were removed, and only uniquely mapping reads were retained. Differential gene expression was determined by comparing one week *Pv* and *Vip* neuron data by DEseq using a FDR cutoff of 0.05 (Anders et al., 2015). Genes transcribed at a high level in *Pv* neurons but not *Vip* neurons were defined as genes with expression levels in *Pv* neurons to be in the 1st quintile of expression among expressed genes (i.e., RPKM > 0) and expression levels in *Vip* neurons to be in the bottom quintile among expressed genes (n = 9). Genes transcribed at a high level in *Vip* neurons but not *Pv* neurons were defined using the same criteria (n = 46). Genes highly expressed in both *Pv* and *Vip* neurons were defined as genes that are within the 1st quintile of expressed genes in both neuron subtypes, and differed in gene expression by less than 10% (FDR > 0.1) (n = 224).

MECP2 ChIP-seq on SUN1 purified nuclei

ChIP-seq on purified nuclei was performed and analyzed as described above using endogenous antibodies against MECP2 (Chen et al., 2003). For gene ontology analyses, genes within the top 10 percent in MECP2 ChIP read densities relative to input DNA within the gene bodies were analyzed using DAVID 6.8 (https://david.ncifcrf.gov/)(Huang da et al., 2009). For the analyses presented in Figures 7A, C and E, significance was assessed by repeating the random grouping of genes and comparing the Spearman's correlation coefficients between the actual data and randomized data.

Single-nuclei RNA sequencing (inDrops)

Adult mice were dark-reared for one week prior to dissecting the visual cortices to reduce variation in transcription caused by light exposure (Mardinly et al., 2016). Nuclei from *Dnmt3a* cKO and *Mecp2 KO* visual cortices along with those from the respective littermate wild-type controls were isolated as previously described (Mo et al., 2015). 0.1% BSA was added to the buffer after nuclei were isolated. Co-encapsulation into microfluidic droplets of nuclei, polyacrylamide gels containing barcoded reverse transcription primers, reverse-transcriptase enzyme, and lysis buffer were performed as previously described to generate barcoded cDNA from single nuclei (Klein et al., 2015; Zilionis et al., 2017). We added 57μL of FD pre-mix (5μL 10X FD buffer; 45μL ddH20; 3μL ExoI (NEB M0293S, 20U/μL); 4μL HinFI (Thermo FD0804) to every 40 μl of sample. Samples were filtered through a 0.45 μm cellulose acetate filter, and incubated at 30 min at 37°C. Samples were purified with 1.2X volume Agencourt AMPure XP beads (Beckman Coulter A63881). Second strand complementary DNA synthesis was performed by adding 10X NEB Second Strand Synthesis Buffer and NEB Second Strand Synthesis Enzyme Mix (New England Biolabs E6111S) and incubated at 16°C for 2.5 hours, then heat inactivated at 65°C for 20min. *In vitro* transcription was performed with NEB T7 High Yield 10X Buffer, dNTPs, NEB T7 High Yield Enzyme Mix (NEB E2040S) and incubating at 37°C for 13 hours. Samples were purified with 1.3X volume AMPure beads. RNA was fragmented with Fragmentation reagent (Ambion AM8740) by incubating at at 70°C for 3 min, stopped with the STOP mix, and purified with the 1.2X volume AMPure beads. Reverse transcription was performed using random hexamers: Fragmented

RNA were incubated with 10 μM primers (TCGGCATTCCTGCTGAACCGCTCTTCCGATCTNNNNNN) and 10mM dNTPs for 3 min at 70°C. Then added RNase Inhibitor, 5X PrimeScript Buffer, 0.5 μL PrimeScript Reverse Transcriptase (Takara Clontech 2680A), and incubated at 30°C for 10 min, 42°C for 1 hour, then 70°C for 15 min. The product was purified with 1.2X volume AMPure beads. The product was amplified using 2X KAPA HiFi HotStart ReadyMix (KAPA Biosystems KK2601), 5 μM primer mix (PE1: AATGATACGGC GACCACCGAGATCTACACTCTTTCCCTACACGA; PE2: CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAAC) using the following conditions 98°C 2 min, 2 cycles of (98°C 20sec, 55°C 30sec, 72°C 40sec), 9 to 14 cycles of (98°C 20sec, 65°C 30sec, 72°C 40sec), 72°C 5 min, 4°C hold. The final libraries were purified with 0.7X volume AMPure beads. The libraries were pair-end sequenced for 54 cycles and 21 cycles on a Nextseq 500. We used the published inDrops sequencing read mapping Python pipeline from Klein et al. (Klein et al., 2015)(https://github.com/indrops/indrops) with default parameters (umi quantification - m:10, u:1, d:600, split-ambigs:False, min_non_polyA:15, bowtie [v 1.1.1] - m:200, n:1, l:15, e:80, trimmomatic [v 0.33] - LEADING: ''28,'' SLIDINGWINDOW: ''4:20,'' MINLEN: ''16,'' max_low_complexity_fraction: 0.50) on a custom built transcriptome (made with default parameters of same Python pipeline using the GRCm38.p5 mouse genome assembly with the Ensembl transcript annotation build 85) The data were clustered using Seurat (Satija et al., 2015). Nuclei with at least 500 unique molecular identifiers (UMIs) with at least 200 genes covered were retained for analyses and genes expressed in at least 3 nuclei were kept for analyses. Furthermore, nuclei with greater than 15,000 UMIs or nuclei in which > 10% of UMI came from mitochrondial genes were excluded from the analyses. T-distributed stochastic neighbor embedding (t-SNE) plots were generated using the ''TSNEplot'' function. Gene expression overlays were visualized using the ''FeaturePlot'' function. We used known neuronal and non-neuronal markers ''Slc17a7,'' ''Gad1,'' ''Aldoc,'' ''Olig1,'' ''Cx3cr1,'' ''Cldn5,'' ''Mrc1'' and ''Vtn'' genes, to identify potential doublets. All nuclei expressing more than one of the above markers were excluded from the analyses. Furthermore, nuclei that expressed more that one of either ''Sst,'' ''Pvalb'' and ''Vip'' genes were marked as potential doublets of inhibitory neurons and excluded from the analyses. Average normalized gene expression in a given cell type was quantified using Monocle (Trapnell et al., 2014), and these values were used for all downstream analyses. For the analyses presented in Figure S7D, a $p < 0.05$ cutoff was used to define upregulated genes in either *Dnmt3a* KO and *Mecp2* KO neurons.

QUANTIFICATION AND STATISTICAL AVAILABILITY

Statistical analyses were performed using software noted above, R or in MATLAB.

DATA AND SOFTWARE AVAILABILITY

The accession number for the sequencing data reported in this paper is GEO: GSE103214.

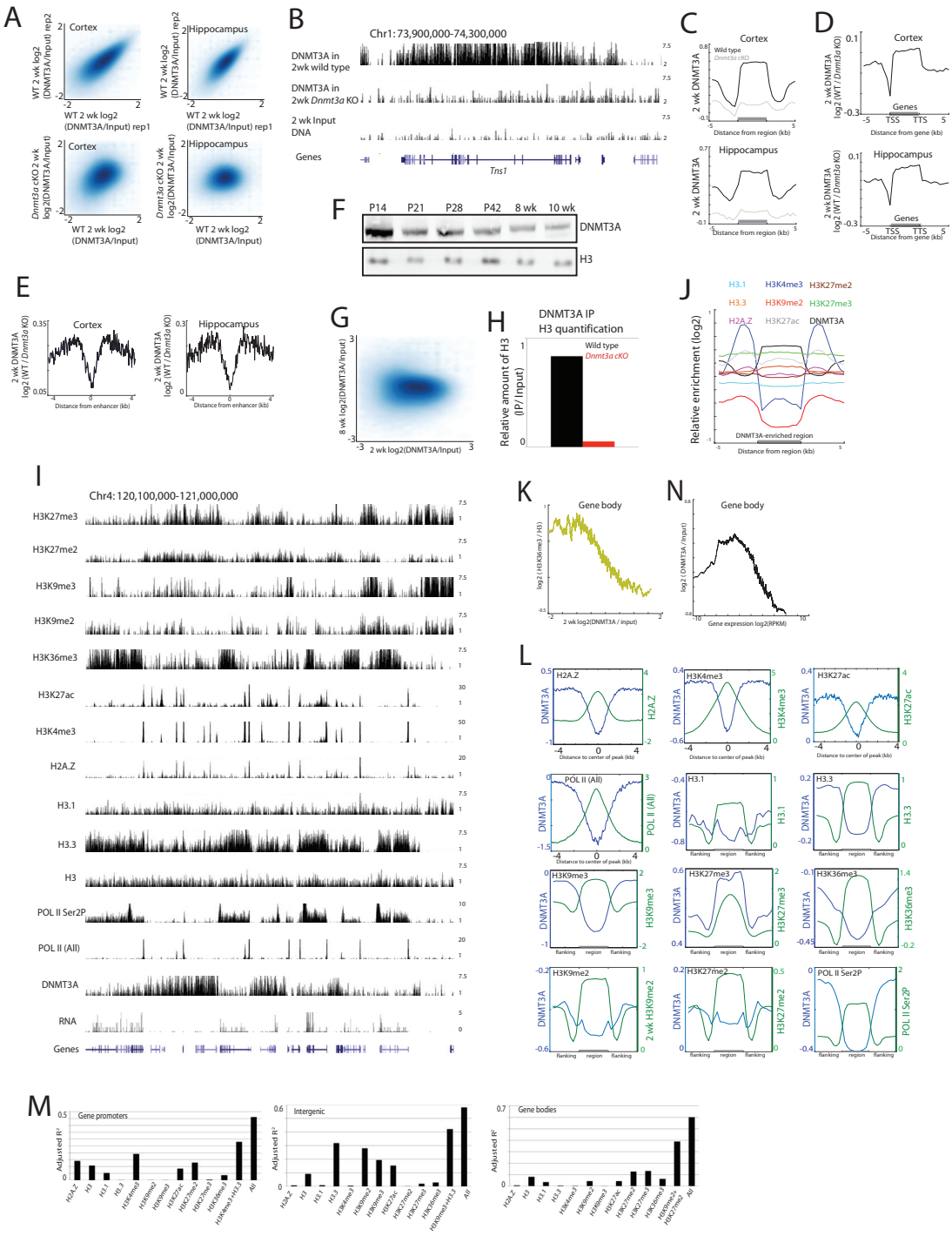


Figure S1. Genomic Profiling of DNMT3A in the Developing Brain, Related to Figure 1

(A) Scatterplots comparing DNMT3A binding between wild-type replicates (top) and between wild-type and *Dnmt3a* cKO (bottom) in the cortex and hippocampus in 5 kilobase tiles across the genome.

(B) Genome browser views of DNMT3A binding in wild-type and *Dnmt3a* cKO two-week cortex.

(C) Average DNMT3A distribution across defined DNMT3A-enriched regions in the two week old cortex (n = 22,223) and hippocampus (n = 6,556). DNMT3A signal was normalized to input DNA (log2 ratio). DNMT3A ChIP from wild-type (*Dnmt3a* $^{fl/fl}$) and *Dnmt3a* cKO (*Nestin-cre*; *Dnmt3a* $^{fl/fl}$) cortex and hippocampus was plotted.

(D) Average DNMT3A distribution across all genes in the cortex and hippocampus. DNMT3A signal was normalized to input DNA, and the enrichment of DNMT3A signal in the wild-type (*Dnmt3a* $^{fl/fl}$) compared to *Dnmt3a* cKO controls (*Nestin-cre*; *Dnmt3a* $^{fl/fl}$) (log2 ratio) was plotted.

(legend continued on next page)

(E) Average DNMT3A distribution across putative distal enhancers (Nord et al., 2013) (n = 38,620). The enrichment of DNMT3A signal in the wild-type (*Dnmt3a* $^{fl/fl}$) compared to *Dnmt3a* cKO controls (*Nestin-cre*; *Dnmt3a* $^{fl/fl}$) (log2 ratio) was plotted.

(F) Western blots of DNMT3A on cortical extracts across different stages of postnatal development with histone H3 as a loading control.

(G) Scatterplot comparing DNMT3A binding at 2 weeks and 8 weeks across the genome in 5 kilobase tiles.

(H) Quantification of amount of histone H3 pulled down by immunoprecipitating DNMT3A. The relative amount of H3 in IP compared to input were calculated.

(I) Genome browser view of ChIP-seq data in the two-week cortex.

(J) Average distributions of histone marks over DNMT3A-enriched regions. The log2 ratio of histone marks to H3, and log2 ratio of DNMT3A to input DNA are shown.

(K) Correlation between H3K36me3 and DNMT3A binding within gene bodies. The average H3K36me3 density was binned according to DNMT3A density within genes (200 genes per bin).

(L) Average distributions of DNMT3A over regions of the genome enriched in the indicated histone mark or DNA-binding factor. Number of defined enriched regions (see STAR Methods): H2A.Z (17,020), H3K4me3 (30,072), H3K27ac (58,746), H3.1 (2,631), H3.3 (23,577), H3K9me3 (24,552), H3K27me3 (11,375), H3K36me3 (17,869), H3K9me2 (23,739), H3K27me2 (9,354), RNA POL II (All) (12,330), RNA POL II Ser2P (13,756).

(M) Prediction performance for DNMT3A binding based on various histone modifications. Gene promoters, TSS - 500 bp to TSS + 500 bp. Gene bodies, TSS + 3 kilobases to TTS. Intergenic, regions > 10 kilobases away from annotated genes. Prediction was performed using linear models in 1 kilobase tiles spanning TSS for promoters, 5 kilobase tiles within gene bodies and 5 kilobase tiles across intergenic regions of the genome (see STAR Methods).

(N) Correlation between gene expression and DNMT3A binding within gene bodies. The average DNMT3A density was binned according to gene expression levels (200 genes per bin). Note decrease in DNMT3A occupancy in the extremely lowly expressed genes on the left (e.g., heterochromatic genes) as well as highly expressed genes on the right.

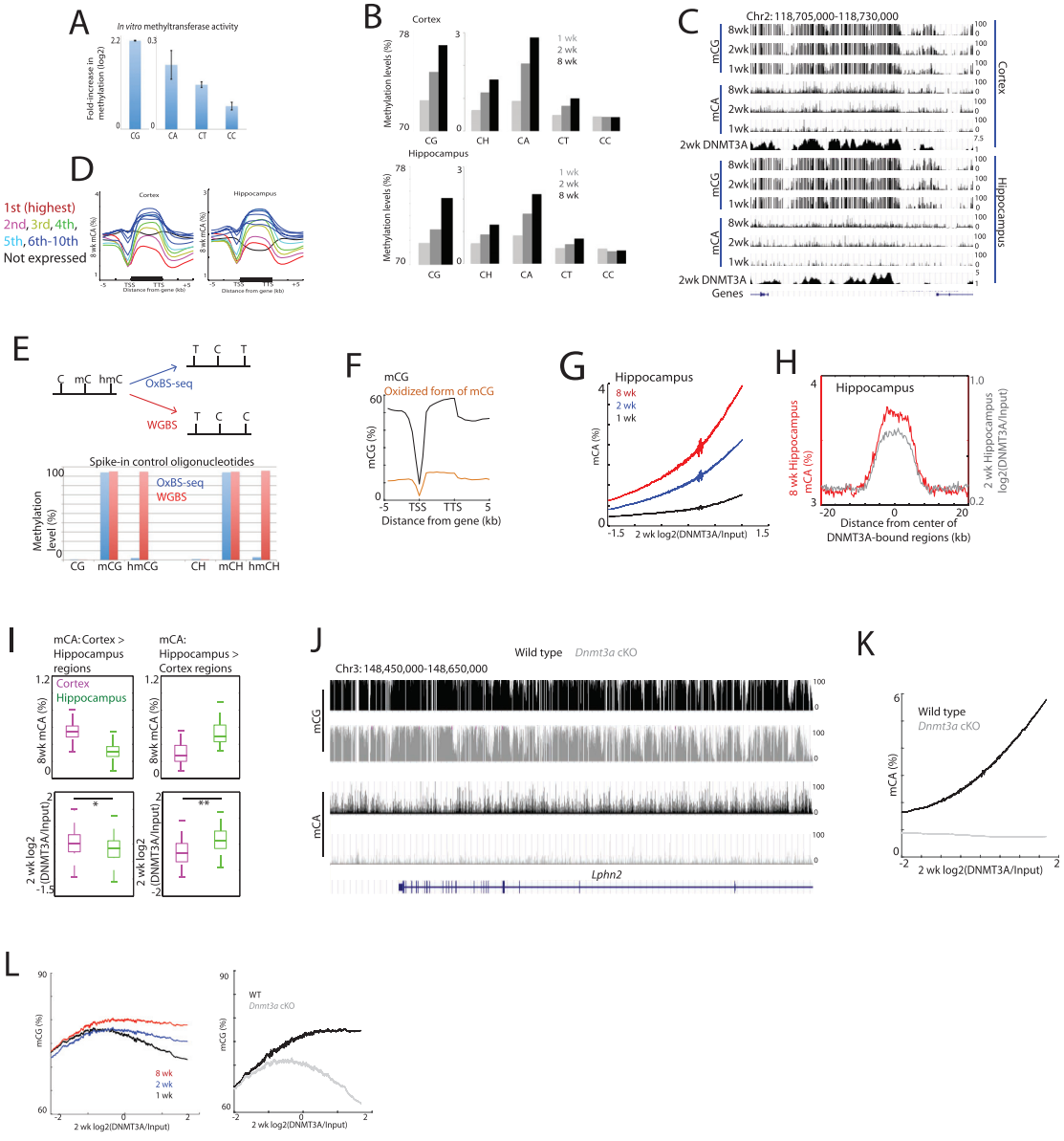


Figure S2. *In Vitro* and *In Vivo* Characterization of DNMT3A Catalytic Activity, Related to Figure 2

(A) Unbiased assessment of the sequence preference of DNMT3A activity. *In vitro* methyltransferase activity of human DNMT3A protein. DNA methylation levels measured by bisulfite sequencing were quantified for each cytosine context along the unmethylated lambda phage genome, and compared with lambda DNA incubated without the DNMT3A protein. Error bars represent s.d. between two technical replicates.

(B) Genome-wide DNA methylation levels in one-week, two-week, and eight-week cortex and hippocampus.

(C) Genome browser views of DNMT3A binding and DNA methylation at different developmental stages in the cortex and hippocampus.

(D) Average mCA patterns across genes of different expression levels in the wild-type cortex and hippocampus.

(E) Methylation calls for spike-in controls of known methylation patterns were calculated. Scheme illustrating the oxBS-seq, adapted from (Song et al., 2012) also shown.

(F) Average distribution of mCG and oxidized forms of mCG over genes, determined by OxBS-seq.

(G) Correlation between two week DNMT3A binding and mCA levels across the genome at different stages of development of the hippocampus.

(H) Distribution of mCA in eight-week hippocampus over DNMT3A-enriched regions in the 2-week hippocampus (n = 6,565).

(I) Boxplots of eight-week mCA and two-week DNMT3A in cortex and hippocampus within genomic tiles defined to be differentially CA methylated (see STAR Methods). mCA cortex > hippocampus: n = 13,438; hippocampus > cortex: n = 58. *p = 2.3e-278, **p < 1.7e-5, Wilcoxon rank-sum test.

(J) Genome browser views of DNA methylation in wild-type and *Dnmt3a* cKO cortex (10 weeks).

(K) Distribution of mCA in wild-type and *Dnmt3a* cKO cortex binned by DNMT3A densities across the genome in 5 kilobase tiles (5,000 tiles per bin).

(L) Distribution of mCG in wild-type cortex (left) and 10-week wild-type and *Dnmt3a* cKO cortex (right) binned by DNMT3A densities across the genome in 5 kilobase tiles (5,000 tiles per bin).

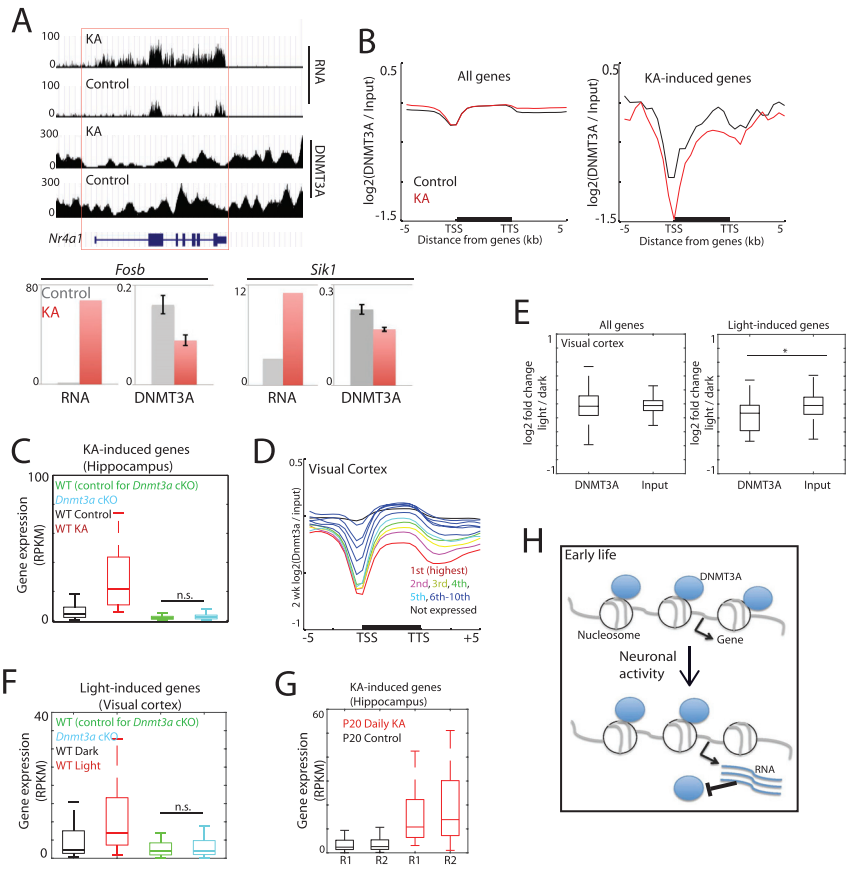


Figure S3. Causal Relationship between DNMT3A Binding and Gene Expression, Related to Figure 3

(A) Genome browser view for DNMT3A binding and gene expression with KA treatment in two-week hippocampus on the activity-dependent *Nr4a1* locus (top). Quantification of the average densities of DNMT3A ChIP-seq reads within the *Fosb* and *Sikl* gene bodies (bottom). Error bars represent s.d. for two biological replicates.

(B) Average distribution of 2-week hippocampal DNMT3A ChIP-seq reads across all genes and KA-induced genes (n = 51).

(C) Gene expression levels of genes defined as KA-induced. RNA-seq data for two-week *Dnmt3a* cKO and littermate wild-type controls are also shown.

(D) Average distribution of DNMT3A over genes of different expression levels in the wild-type visual cortex.

(E) Effect of light exposure on DNMT3A binding over all genes and light-induced genes (n = 33). DNMT3A ChIP-seq and input DNA read densities of two-week visual cortex within gene bodies were calculated, and log2 ratios between light exposed (light) and control (dark) samples are shown. *p = 0.01, Wilcoxon rank-sum test.

(F) Boxplots of gene expression levels of genes defined as light-induced. RNA-seq data for two-week *Dnmt3a* cKO and littermate wild-type controls are also shown.

(G) KA-induced genes are still inducible after ten days of daily KA administration. Boxplots of gene expression levels of genes defined as KA-induced for daily KA-treated and control hippocampus at the final day of injections (P20). The RNA was collected an hour after injections. Data for two biological replicates are shown.

(H) Model of activity-dependent gene induction disrupting DNMT3A binding during early life.

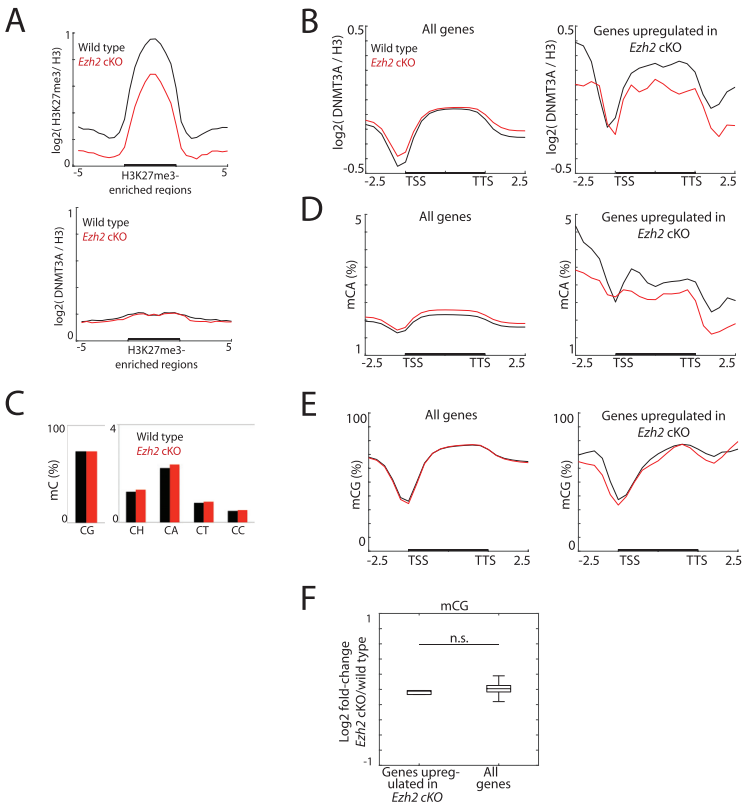


Figure S4. Effect of a Genetic Mutation in the *Ezh2* Gene on DNMT3A and mCA Deposition, Related to Figure 4

(A) Top, average H3K27me3 distribution in *Ezh2* cKO cortex compared to wild-type across defined H3K27me3-enriched regions in wild-type (n = 11,375). Bottom, average DNMT3A distribution across defined H3K27me3-enriched regions in wild-type.

(B) Average distribution of DNMT3A ChIP-seq reads across all genes and genes upregulated in *Ezh2* cKO (n = 11).

(C) Genome-wide DNA methylation levels in *Ezh2* cKO and wild-type.

(D) Average distribution of mCA across all genes and genes upregulated in *Ezh2* cKO.

(E) Average distribution of mCG across all genes and genes upregulated in *Ezh2* cKO.

(F) Boxplots of the difference between mCG in *Ezh2* cKO and wild-type control cortices within all genes and genes upregulated in *Ezh2* cKO.

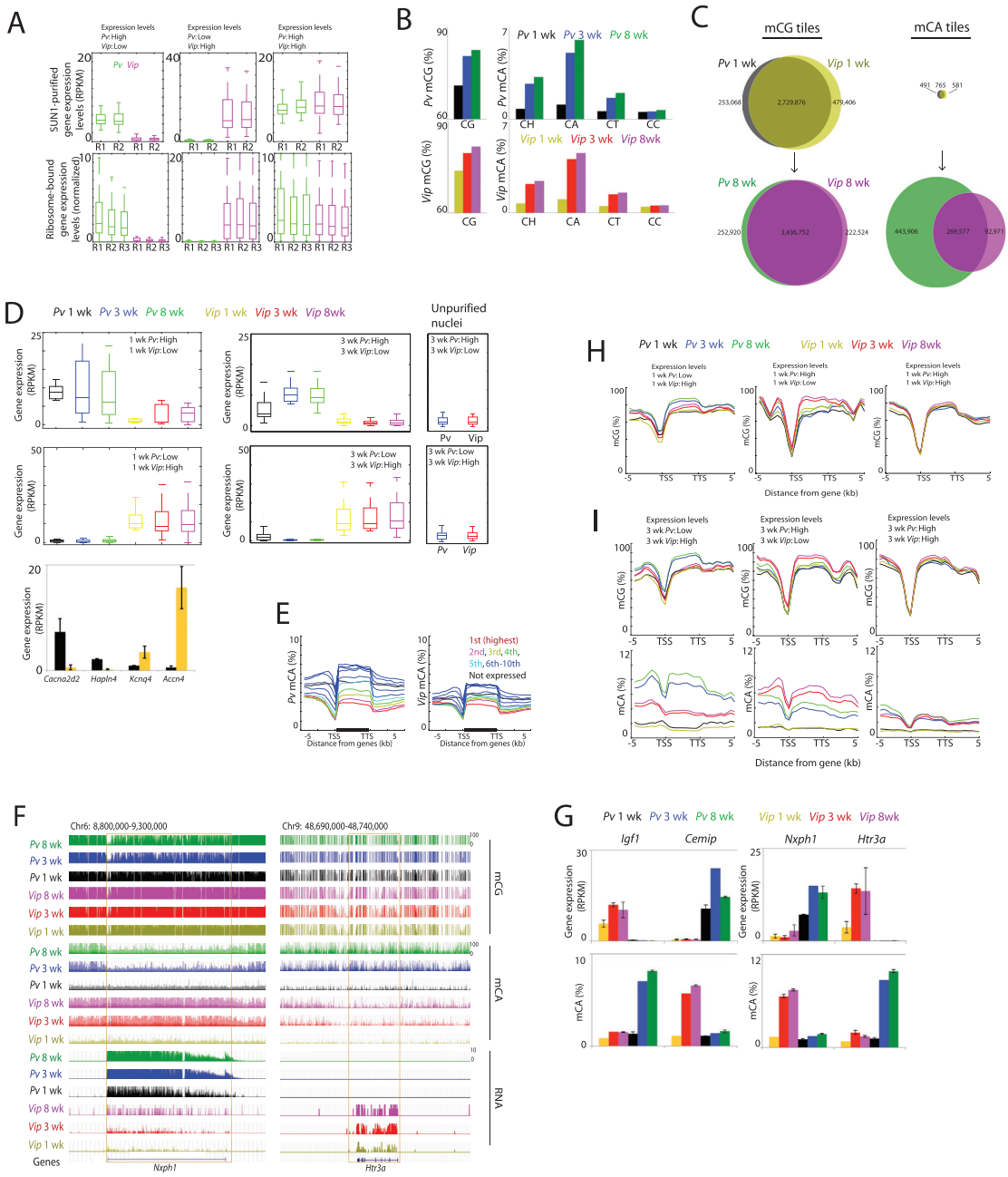


Figure S5. DNA Methylome and Transcriptome Data in Two Neuron Subtypes at Three Developmental Time Points, Related to Figure 5

(A) Comparison of INTACT-purified nuclear RNA-seq with ribosome-bound mRNA-seq approach (Mardinly et al., 2016). Genes were classified by the INTACT-purified RNA-seq dataset (*Pv* high, *Vip* low genes: n = 39; *Vip* high, *Pv* low genes: n = 72; *Pv* high, *Vip* high genes: n = 105), and gene expression levels were plotted. Data for two biological replicates for the INTACT-purified nuclear RNA-seq and three biological replicates for the ribosome-bound mRNA-seq is shown. Genome-wide correlation between the different datasets: *Pv*: $\rho = 0.73$, $p < 1e\text{-}4$; *Vip* $\rho = 0.67$, $p < 1e\text{-}4$.

(B) Genome-wide DNA methylation levels in *Pv* and *Vip* neurons.

(C) Overlaps of genomic 100 bp tiles with high levels of mCG (right) or mCA (left) between *Pv* and *Vip* neurons. Tiles were selected if they were greater than 70% for mCG, and greater than 10% for mCA (see STAR Methods). Gene ontology (GO) analyses of genes with high mCA densities within the gene bodies (top 5% of all genes) for both *Pv* and *Vip* neurons suggested that, in addition to annotated neuronal functions such as axon guidance, regulation of neuronal development, genes involved in various biological processes such as transcriptional regulation, response to stimuli, and cellular differentiation were also over-represented, consistent with the finding that mCA coats the neuronal genome.

(D) Expression levels of genes that are differentially expressed between *Pv* and *Vip* neurons during the early postnatal development. As a control for potential genetic differences between the two mouse strains that may potentially lead to differential gene expression, we performed RNA-seq on unpurified nuclei (i.e., before immunoprecipitation of the nuclei) from each mouse strain and observed no significant differences in gene transcription between these two mouse strains

(legend continued on next page)

(right), confirming that the cell-type-specific differences in gene expression we observe are not driven by genetic differences between the two strains of mice. Examples of neuronal subtype-specific genes at one week are also shown (below). Error bars represent s.d. for two biological replicates for *Pv* neurons and three biological replicates for *Vip* neurons.

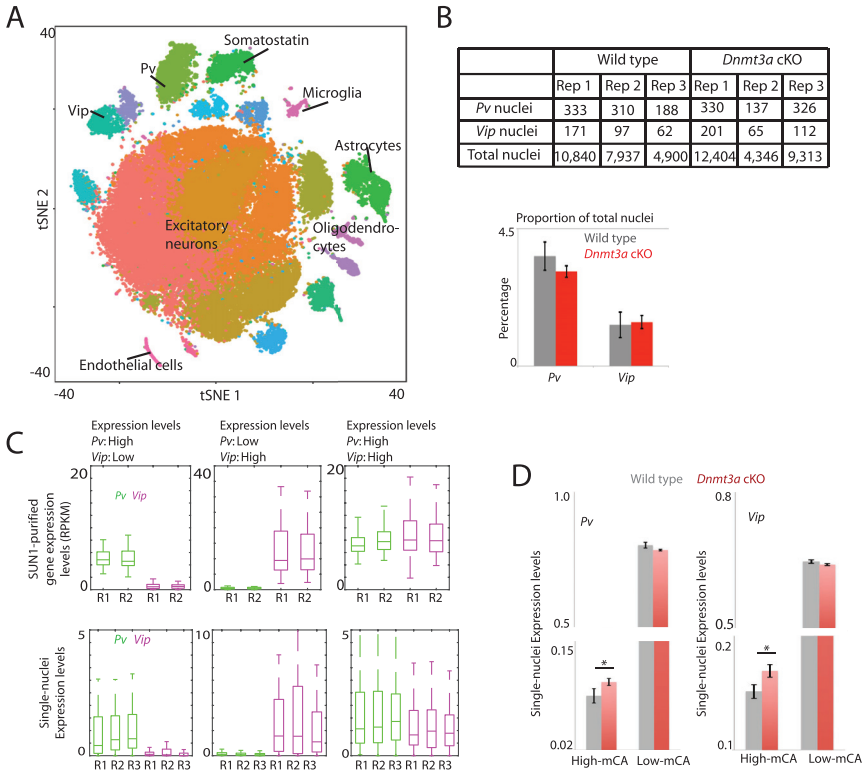
(E) Average distribution of mCA in *Pv* and *Vip* neurons over genes of different expression levels at 3 weeks.

(F) Genome browser views of DNA methylation and transcription data in cortical *Pv* and *Vip* neurons at indicated developmental time points.

(G) Quantifications of intragenic mCA and gene expression levels of the genes shown in the browser tracks in Figures 5C and S5F. Error bars represent s.d. for two to three biological replicates.

(H) Average distribution of mCG over genes differentially expressed in *Pv* and *Vip* neurons at one week.

(I) Average distribution of mCG and mCA over genes differentially expressed in *Pv* and *Vip* neurons at three weeks.



	Wild type			*Dnmt3a* cKO		
	Rep 1	Rep 2	Rep 3	Rep 1	Rep 2	Rep 3
Pv nuclei	333	310	188	330	137	326
Vip nuclei	171	97	62	201	65	112
Total nuclei	10,840	7,937	4,900	12,404	4,346	9,313

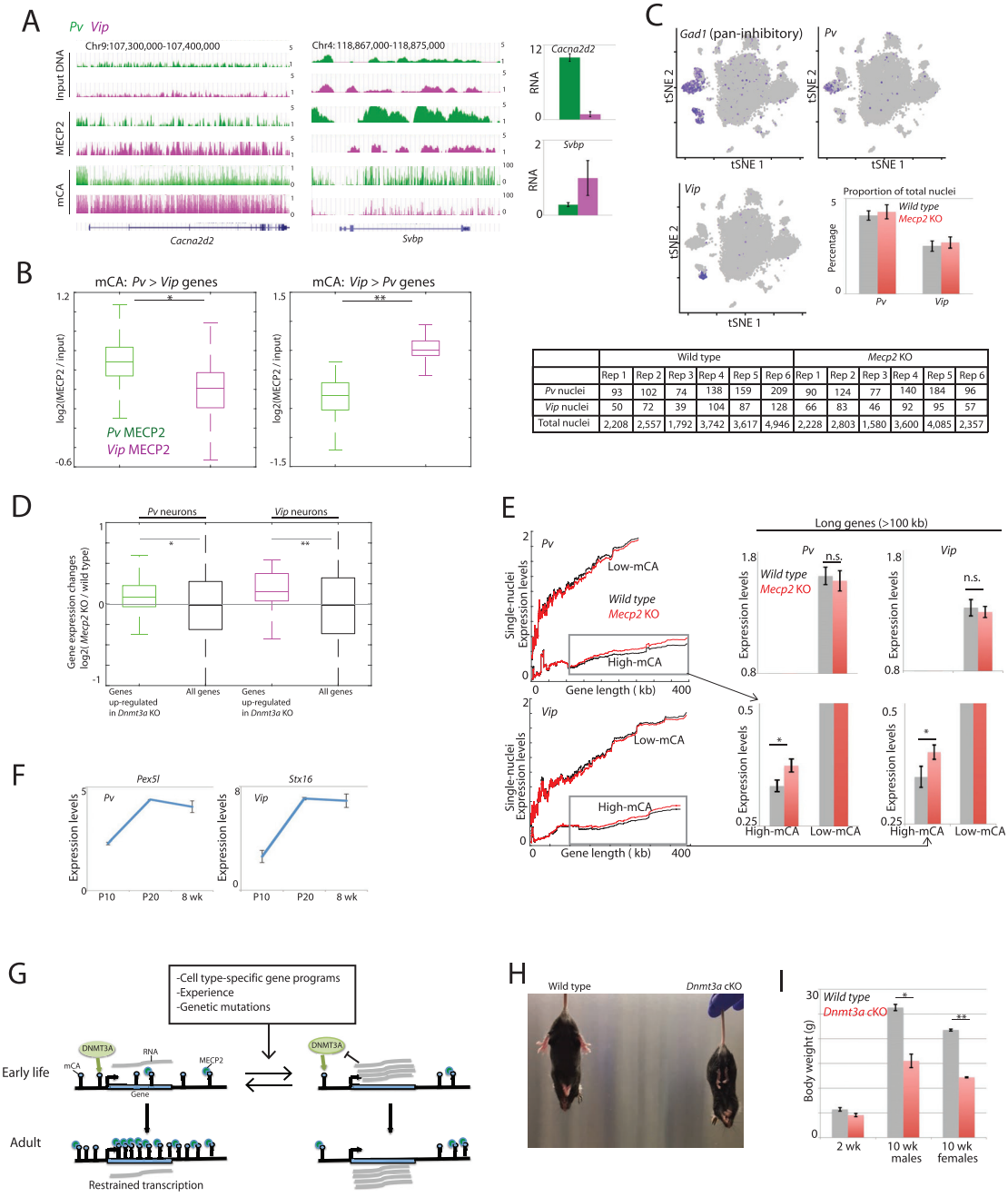
Figure S6. High-Throughput Single-Nuclei Profiling of Neurons Deficient in DNMT3A, Related to Figure 6

(A) Seurat t-SNE output of distinct cell populations.

(B) Number of post-filtered nuclei analyzed (top). Percentage of identified *Pv* and *Vip*-expressing nuclei among all nuclei (bottom). Error bars represent s.d. between three biological replicates. DNMT3A disruption did not result in strong changes in the proportions of *Pv* and *Vip*-expressing neurons in the *Dnmt3a* cKO cortex compared to wild-type, suggesting that DNMT3A likely does not play a major role in specification of neuronal subtypes, consistent with our finding that mCA is deposited after specification. Having similar numbers of *Pv* and *Vip*-expressing neurons enabled us to carefully compare gene expression profiles between the two subtypes in the *Dnmt3a* cKO and wild-type mice.

(C) Comparison between single-nuclei RNA-seq and INTACT-purified nuclear RNA-seq data. Genes were classified by the INTACT-purified RNA-seq dataset (*Pv* high, *Vip* low genes: n = 34; *Vip* high, *Pv* low genes: n = 56; *Pv* high, *Vip* high genes: n = 102), and gene expression levels determined by single-nuclei RNA-seq were plotted. Genome-wide correlation between the different datasets: Pv data: r = 0.79, p < 1e-4; Vip data: r = 0.75, p < 1e-4.

(D) Association between mCA densities within gene bodies and gene expression levels in wild-type and *Dnmt3a* KO *Pv* and *Vip* neurons (average single-nuclei RNA-seq data). Genes were separated based on the mCA densities within gene bodies: High-mCA: > 9.55% (*Pv*) (n = 1,716) or > 6.67% (*Vip*) (n = 1,716); Low-mCA: < 5% (*Pv:* n = 9,586 and *Vip:* n = 12,878). p < 0.05, two-tailed t test. Error bars represent s.d. between three biological replicates.



	Wild type						*Mecp2* KO					
	Rep 1	Rep 2	Rep 3	Rep 4	Rep 5	Rep 6	Rep 1	Rep 2	Rep 3	Rep 4	Rep 5	Rep 6
Pv nuclei	93	102	74	138	159	209	90	124	77	140	184	96
Vip nuclei	50	72	39	104	87	128	66	83	46	92	95	57
Total nuclei	2,208	2,557	1,792	3,742	3,617	4,946	2,228	2,803	1,580	3,600	4,085	2,357

Figure S7. Cell-Type-Specific MECP2 Binding, and Single-Nuclei RNA-Seq in *Mecp2* KO Brains, Related to Figure 7

(A) Genome browser views of MECP2 and mCA data in 8-week cortical *Pv* and *Vip* neurons. Gene expression level determined by INTACT-purified RNA-seq is shown on right. Error bars represent s.d. between two biological replicates.

(B) Boxplots of MECP2 densities within genes with differential CA methylation levels between *Pv* and *Vip* neurons (see STAR Methods). (mCA *Pv* > *Vip* genes: n = 872; *Vip* > *Pv* genes: n = 84). *p = 1.6e-100, **p = 4.4e-27 Wilcoxon rank-sum test. Consistent with the fact that mCA marks a wide variety of genes involved in range of biological processes, GO analyses on genes enriched for MECP2 binding within the gene bodies for both *Pv* and *Vip* neurons suggested that MECP2 binding is not restricted to genes involved in neuronal functions, but also binds across genes involved in other biological processes, such as development, cell proliferation, and transcriptional regulation (see STAR Methods).

(C) Seurat t-SNE plots of all nuclei in samples. Nuclei expressing indicated marker genes are depicted in purple. Number of *Pv* and *Vip*-expressing nuclei analyzed (bottom), and percentage of identified *Pv* and *Vip*-expressing nuclei among all nuclei (right). Error bars represent s.d. between six biological replicates.

(D) Overlap between genes misregulated in *Dnmt3a* and *Mecp2* KO *Pv* and *Vip* neurons using a threshold approach. Genes with mCA levels greater than 5% mCA within gene bodies in wild-type that became significantly upregulated in *Dnmt3a* KO ($p < 0.05$ cutoff) were defined to be a small subset of genes that are putatively directly repressed by DNMT3A (*Pv*: n = 93; *Vip*: n = 25). The expression levels of this small subset of DNMT3A-target genes in *Mecp2* KO and littermate wild-type

(legend continued on next page)

neurons were calculated. The expression levels of all genes were also plotted. *p = 0.001, **p = 0.022, Wilcoxon rank-sum test. The genes that are upregulated in *Dnmt3a* or *Mecp2* KO neurons are partially overlapping between the two KO strains compared to wild-type mice but they can differ between *Pv* and *Vip* neurons, and they include genes that are involved in specific neuronal processes that presumably display some neuronal subtype-specificity with regard to their function. For example, one gene that is selectively misregulated in *Pv*, *Sez6* modulates synaptic connectivity and is associated with febrile seizures in humans (Gunnersen et al., 2007; Yu et al., 2007). A gene, *Epha5* that is selectively upregulated in *Vip* neurons encodes a protein that regulates axon guidance and mediates aggressive behavior in mice possibly via its selective expression in *Vip* neurons (Mamiya et al., 2008; Wilkinson, 2001).

(E) Correlation between gene length and gene expression levels in wild-type and *Mecp2* KO neurons (average single-nuclei RNA-seq data) (500 genes per bin). Long genes (> 100 kb) were separated based on the density of mCA within the gene bodies. High mCA: > 5% (*Pv*: n = 1,294; *Vip*: n = 948); Low mCA: < 5% (*Pv*: n = 859; *Vip*: n = 1,205). *p < 0.001, two-tailed t test. Error bars represent s.d. between six biological replicates.

(F) Gene expression levels of *Pex5l* (highly CA-methylated in *Pv* neurons) and *Stx16* (highly CA-methylated in *Vip* neurons) at three postnatal time points. Error bars represent s.d. between two to three biological replicates. The reverse correlation between level of gene expression and extent of mCA and MECP2 binding holds true for the majority of genes but not all genes. Intriguingly, we find that a small subset of high-mCA genes is highly expressed. Inspection of the highly expressed, high-mCA genes indicated that a subset of these genes including *Pex5l* and *Stx16* are moderately expressed during early life and subsequently become highly expressed in adults. The moderate expression of these genes in early life may facilitate methylation across their gene bodies, and these genes may then be upregulated by other transcriptional control mechanisms. On the other hand, a small subset of genes that are highly expressed during early life nevertheless become highly CA-methylated, suggesting that there may be unknown gene-specific DNMT3A-targeting mechanisms that remain to be identified.

(G) Model of how early-life transcription modulates lasting mCA and MECP2 patterns across genes. Once mCA is deposited it is bound by MECP2 to fine-tune the expression of genes.

(H) Hind limb clasping in adult *Dnmt3a* cKO mice.

(I) Developmental weight loss in *Dnmt3a* cKO mice. Young (two weeks) and adult (ten weeks) *Dnmt3a* cKO (*Nestin-cre*; *Dnmt3a*$^{fl/fl}$) were compared to littermate wild-type mice (*Dnmt3a* $^{fl/fl}$). Error bars represent s.d. between two to three biological replicates. *p = 3.2e-5, **p = 3.2e-4, two-tailed t test.