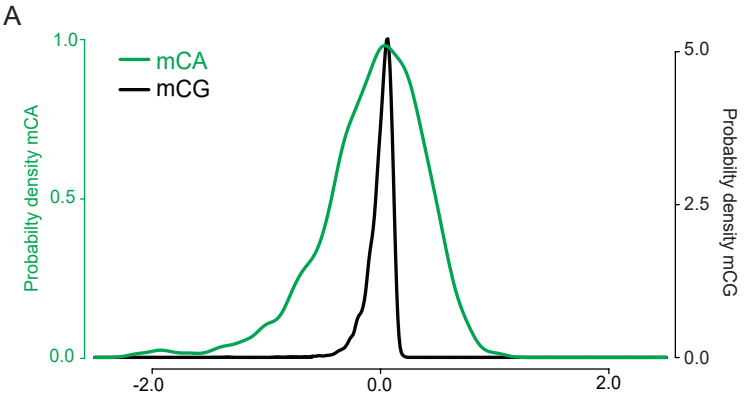
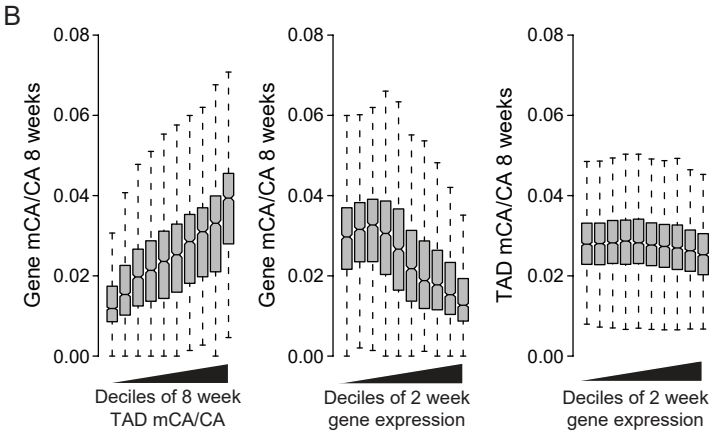
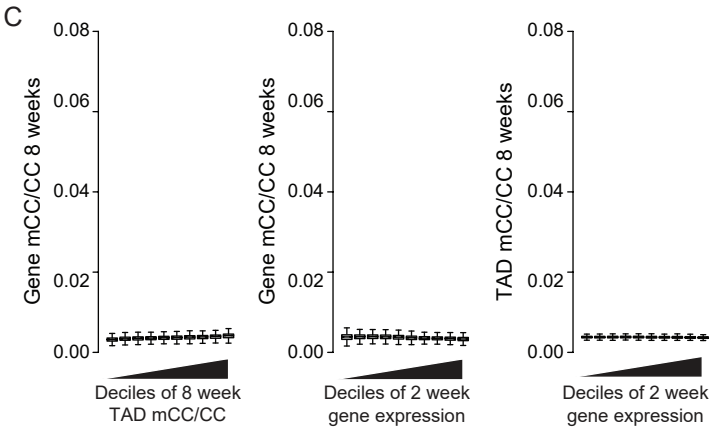
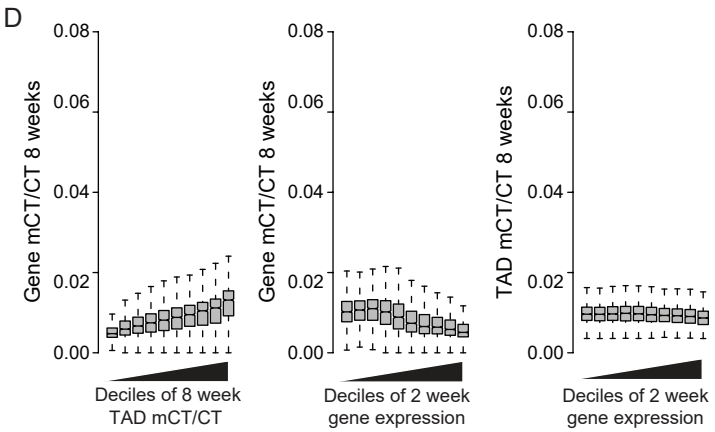
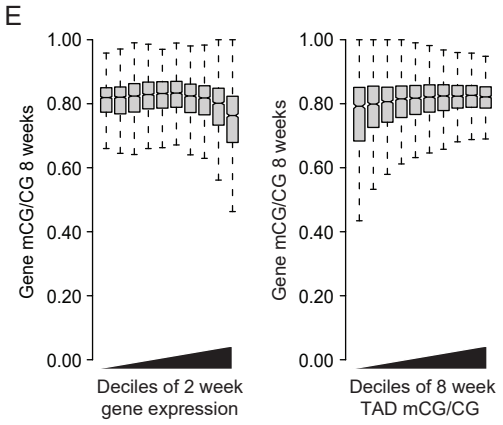
Supplementary Figure 1





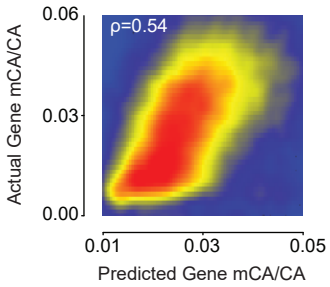




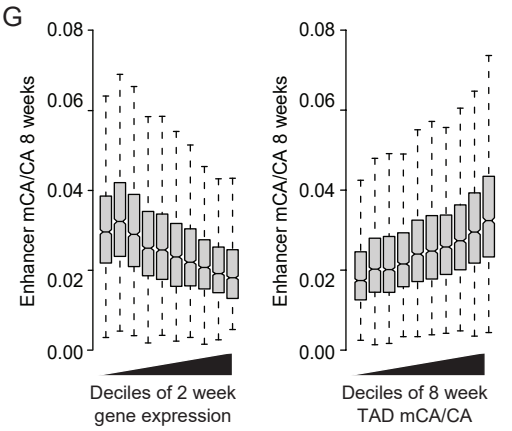


F

$$\text{Gene mCA} = \beta_1(\text{TAD mCA}) + \beta_2(\text{Gene expression}) + \beta_3(\text{Interaction}) + \beta_0$$

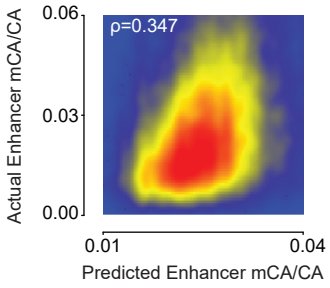


Coefficients:
β_0 = 0.002
β_1 = 0.427
β_2 = -0.306
β_3 = -0.102



H

$$\text{Enhancer mCA} = \beta_1(\text{TAD mCA}) + \beta_2(\text{Gene expression}) + \beta_3(\text{Interaction}) + \beta_0$$



Coefficients:
β_0 = -0.001
β_1 = 0.270
β_2 = -0.199
β_3 = -0.013

Figure S1. Analysis of neuronal mCA and mCG profiles associated with regional mCA and gene expression at genomic regions. Related to Figure 1.

(A) Probability density functions illustrating the relative fluctuations compared to the genomic mean for levels of mCA (green) or mCG (black) in all TADs genome-wide.

(B) Left, boxplots of adult mCA levels in gene bodies plotted for genes sorted into deciles of adult TAD mCA for each gene (TAD mCA is determined by the average level of mCA in the TAD in which the gene resides, see *methods*). Center, boxplot of adult mCA levels in the bodies of genes for genes sorted into deciles of gene expression at 2 weeks of age. Right, boxplot of adult TAD mCA levels for each gene sorted by deciles of gene expression at 2 weeks of age. Comparison of left and center plots indicates that regional TAD mCA set-point and 2-week gene expression in the early postnatal period are each associated with ~2-fold variance in gene body mCA levels genome-wide. Right plot shows little to no association between 2-week expression and regional TAD mCA set-points.

(C) Left, boxplots of adult mCC levels in gene bodies plotted for genes sorted into deciles of adult TAD mCC for each gene. Center, boxplot of adult mCC levels in the bodies of genes for genes sorted into deciles of gene expression at 2 weeks of age. Right, boxplot of adult TAD mCC levels for each gene sorted by deciles of gene expression at 2 weeks of age for each gene.

(D) Left, boxplots of adult mCT levels in gene bodies plotted for genes sorted into deciles of adult TAD mCT for each gene. Center, boxplot of adult mCT levels in the bodies of genes for genes sorted into deciles of gene expression at 2 weeks of age. Right, boxplot of adult TAD mCT levels for each gene sorted by deciles of gene expression at 2 weeks of age. Consistent with low absolute levels of mCC and mCT in neurons, the trends for these methylation sites corroborate mCA, but at a much lower signal magnitude and dynamic range.

(E) Boxplots of mCG levels in gene bodies across deciles of gene expression (left) and TAD mCG (right). mCG levels show much more limited fluctuations associated with these two signals compared to mCA (e.g. compare to Figure 1B).

(F) Linear model predicted gene body mCA versus actual gene body mCA levels, with the associated equation and coefficients for each independent variable shown. Similar coefficient magnitudes supports a roughly equal contribution of these signals to gene mCA levels.

(G) Boxplots of mCA levels at intragenic enhancers across deciles of gene expression (left) and TAD mCA (right).

(H) Linear model predicted enhancer mCA versus actual enhancer mCA levels, with the associated equation and coefficients for each independent variable shown.

Data are wild-type 2-week RNA-seq.[S1] 8-week DNA methylation[S2] and H3K27ac ChIP-seq.[S3]

Supplementary Figure 2

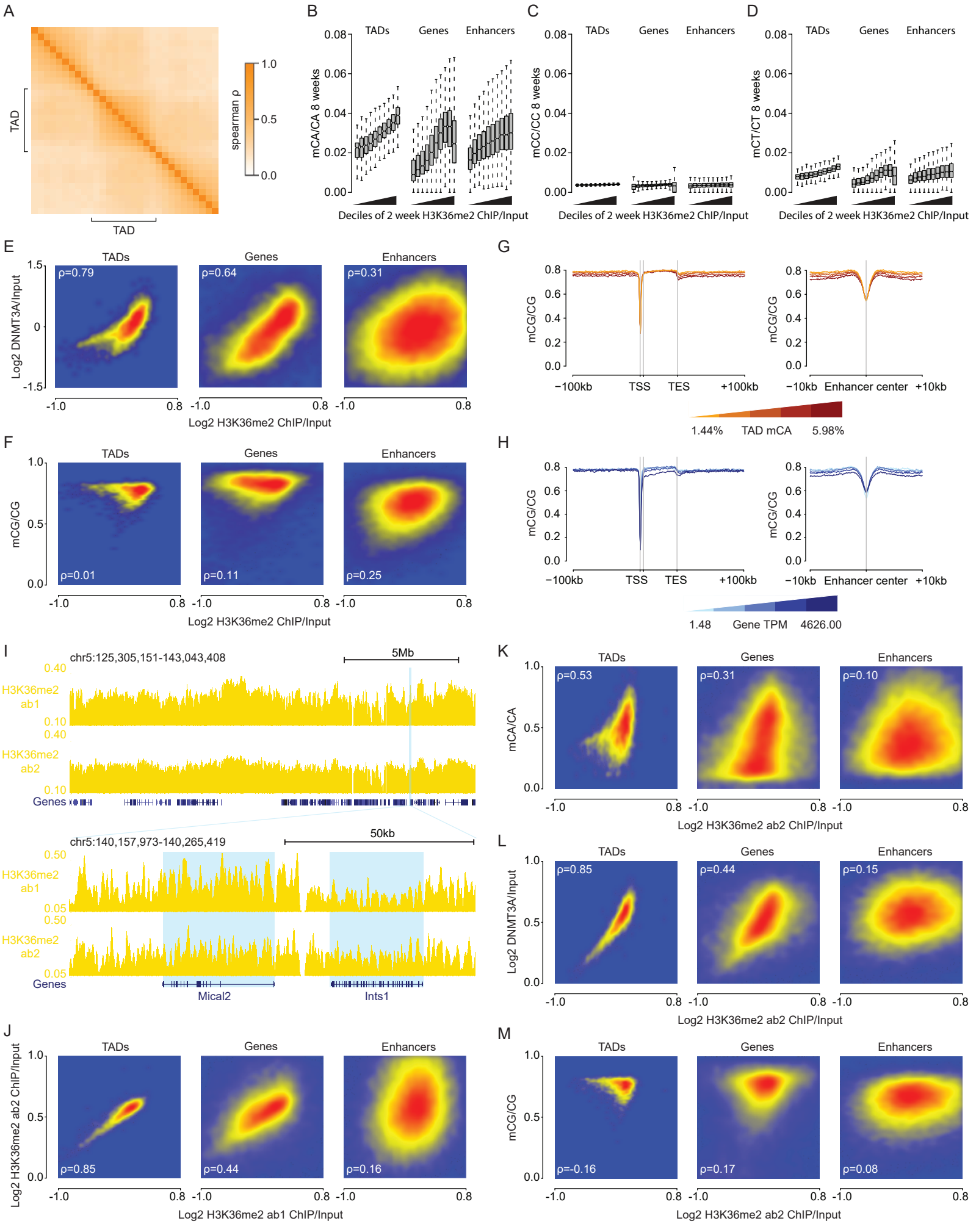


Figure S2. Analysis of H3K36me2 profiles and deposition of neuronal DNA methylation. Related to Figure 2.

(A) Cross-correlation analysis of H3K36me2 in and around all TADs in wildtype cortex. The H3K36me2 ChIP signal was assessed for 10 genomic windows within each TAD as well as for 10 equally sized windows on either side of each TAD. The correlation between the ChIP signal in these windows was assessed for all TADs genome-wide and plotted based on position relative to the TAD (see *methods*). H3K36me2 signal shows a higher correlation within TADs, compared to across TAD boundaries.

(B) Boxplots of adult mCA levels in TADs, gene bodies, and enhancers sorted into deciles of H3K36me2 ChIP signal at 2 weeks of age.

(C) Boxplots of adult mCC levels in TADs, gene bodies, and enhancers sorted into deciles of H3K36me2 ChIP signal at 2 weeks of age.

(D) Boxplots of adult mCT levels in TADs, gene bodies, and enhancers sorted into deciles of H3K36me2 ChIP signal at 2 weeks of age.

(E) Comparison of H3K36me2 and DNMT3A ChIP signal at TADs, genes, and enhancers in wildtype cortex.

(F) Comparison of H3K36me2 ChIP/Input and mCG/CG at TADs, genes, and enhancers in wildtype cortex.

(G) Aggregate mCG levels at genes and enhancers across quintiles of TAD mCA levels.

(H) Aggregate mCG levels at genes and intragenic enhancers across quintiles of gene expression.

(I) Top, broad genome browser view of ChIP-seq signal from wildtype cerebral cortex for two H3K36me2 antibodies, illustrating concordance in megabase-scale fluctuations in H3K36me2 across two independent experiments with ab2 showing reduced dynamic range. Bottom, zoomed in genome browser view of two genes from figure 2 with different levels of expression (highlighted in blue) and similar differential H3K36me2 ChIP signal at each gene across two antibodies.

(J) Comparison of H3K36me2 ChIP signal across two antibodies at TADs, enhancers, and genes. Lower count coverage in analysis of ab2 leads to lower correlations at small scale regulatory regions such as enhancers.

(K) Comparison of H3K36me2 ab2 ChIP signal and mCA/CA at TADs, enhancers, and genes, showing robust genome-wide correlations for these signals that support Figure 2C.

(L) Comparison of H3K36me2 ab2 and DNMT3A ChIP signal at TADs, genes, and enhancers in wildtype cortex.

(M) Comparison of H3K36me2 ab2 ChIP signal and mCG/CG at TADs, genes, and enhancers in wildtype cortex.

Data obtained from wildtype cerebral cortex at 2- and 8-weeks. 2-week, n = 2-3 bioreplicates for H3K36me2 (per antibody) and DNMT3A ChIP-seq. 2-week RNA-seq data obtained from Lister et al., 2013.[S1] 8-week DNA methylation data obtained from Stroud et al., 2017.[S2]

Supplementary Figure 3

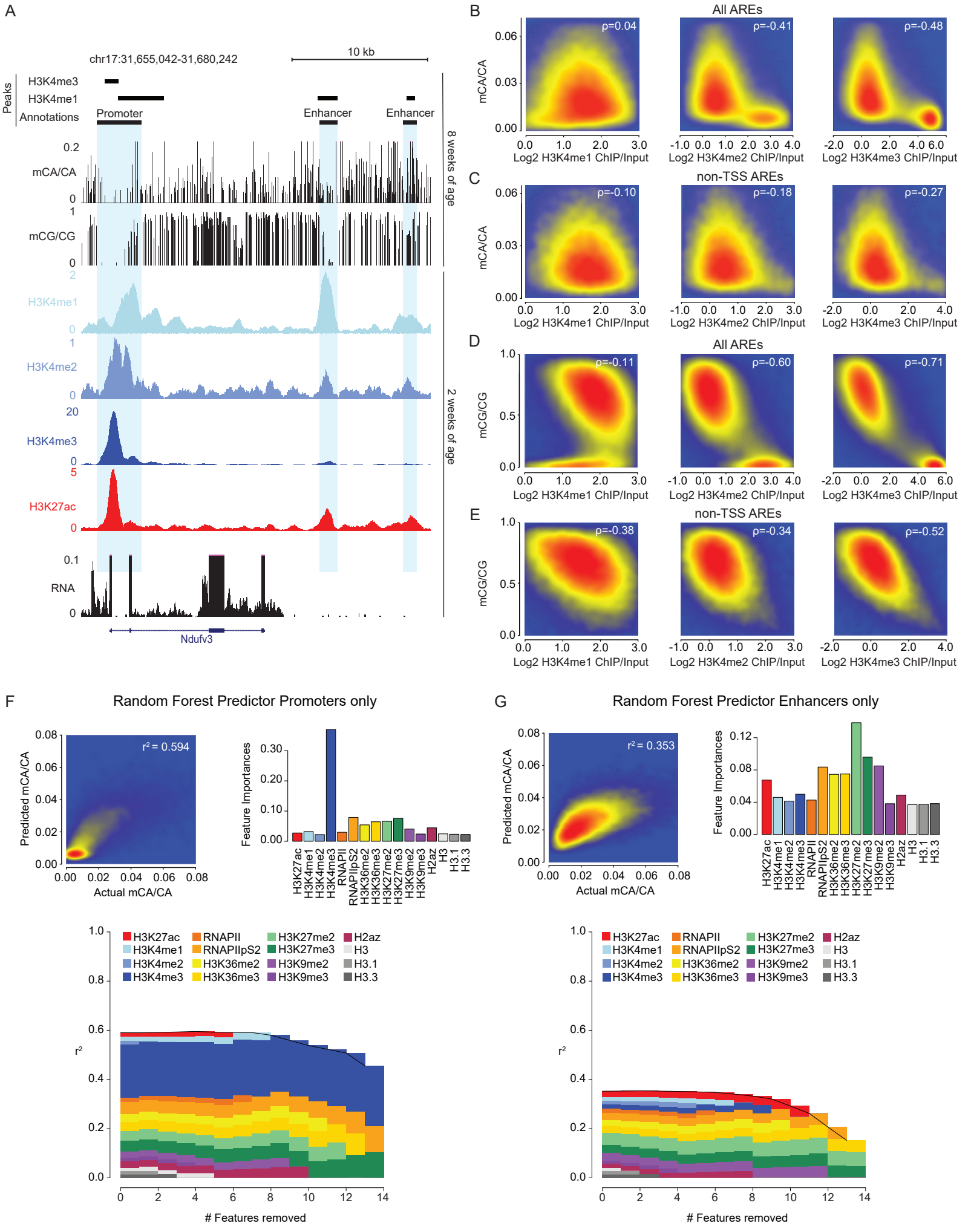


Figure S3. Analysis of mCA prediction by chromatin signals. Related to Figure 3.

(A) Genome browser view illustrating local enrichment of H3K4 methylation at active regulatory elements (AREs) surrounding the *Nudfv3* gene and relative depletion of mCA and mCG at these sites.

(B) Smoothscatter plots showing relative correlation of adult mCA levels with early postnatal H3K4 methylation marks at all active regulatory elements (AREs) across the genome (see *methods*).

(C) Smoothscatter plots showing relative correlation of mCA with H3K4 methylation marks at non-TSS active regulatory elements (AREs) across the genome (see *methods*).

(D) Smoothscatter plots showing relative correlation of mCG with H3K4 methylation marks at all active regulatory elements across the genome (see *methods*).

(E) Smoothscatter plots showing relative correlation of mCG with H3K4 methylation marks at non-TSS active regulatory elements across the genome (see *methods*).

(F) Left, random forest classifier predicted mCA versus actual mCA levels at promoters (+/- 500bp from TSS center, see *methods*). Right, feature importance level of each chromatin signature to the predictive accuracy of mCA levels at promoters by the random forest classifier. Bottom, feature elimination analysis of chromatin signatures in the random forest classifier.

(G) Left, random forest classifier predicted mCA versus actual mCA levels at enhancers (+/- 500bp from enhancer center, see *methods*). Right, feature importance level of each chromatin signature to the predictive accuracy of mCA levels at enhancers by the random forest classifier. Bottom, feature elimination analysis of chromatin signatures in the random forest classifier.

Data are from wildtype cerebral cortex at 2- and 8-weeks. 2-week, n = 2-3 bioreplicates for H3K36me2, H3K4me1, H3K4me2, and DNMT3A ChIP-seq from this study. 2-week RNA-seq data obtained from Lister et al., 2013.[S1] 2-week ChIP-seq for all other chromatin marks and 8-week DNA methylation data obtained from Stroud et al., 2017.[S2]

Supplementary Figure 4

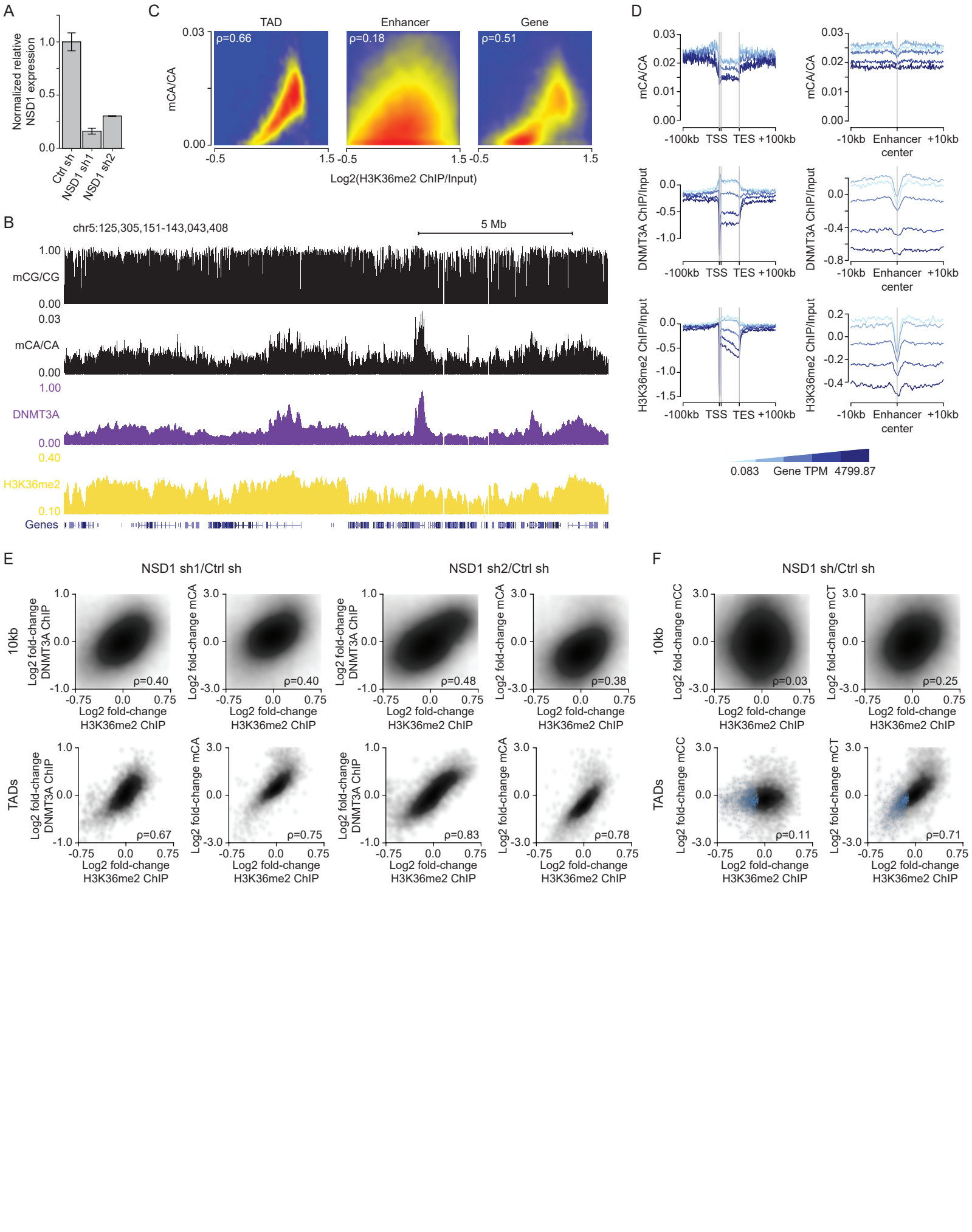


Figure S4. H3K36me2, DNMT3A, and mCA are correlated and concordantly disrupted upon NSD1 loss in primary cortical neurons (PCNs). Related to Figure 4.

(A) qRT-PCR of *Nsd1* over *Actb* normalized to Ctrl shRNA sample for PCNs transduced with one of two independent *Nsd1* shRNAs each (n = 2 per condition).

(B) Genome browser view of DIV 12 ChIP-seq and DIV 18 DNA methylation from shCtrl PCNs. There is strong concordance of fluctuations in H3K36me2, DNMT3A, and mCA across the neuronal genome.

(C) Comparison of H3K36me2 ChIP/Input and mCA/CA at TADs, enhancers, and genes.

(D) Aggregate mCA, DNMT3A/Input, H3K36me2/Input levels at genes and intragenic enhancers across quintiles of PCN gene expression.

(E) Comparison of changes in H3K36me2 and changes in DNMT3A (left) or changes in mCA (right) for each 10kb region (top) and TAD (bottom) across two independent shNSD1 samples.

(F) Comparison of changes in H3K36me2 and changes in mCC (left) or changes in mCT (right) for each 10kb region (top) and TAD (bottom) upon knockdown of NSD1.

Data are from primary cultured neurons infected with shCtrl or shNSD1 after DIV 1 and collected at DIV 12 and DIV 18 for ChIP-seq and WGBS, respectively. Per time point: n = 2-4 bioreplicates for H3K36me2, DNMT3A ChIP-seq, and DNA methylation. TADs were derived from Hi-C analysis of cortical neurons.[S4]

Supplementary Figure 5

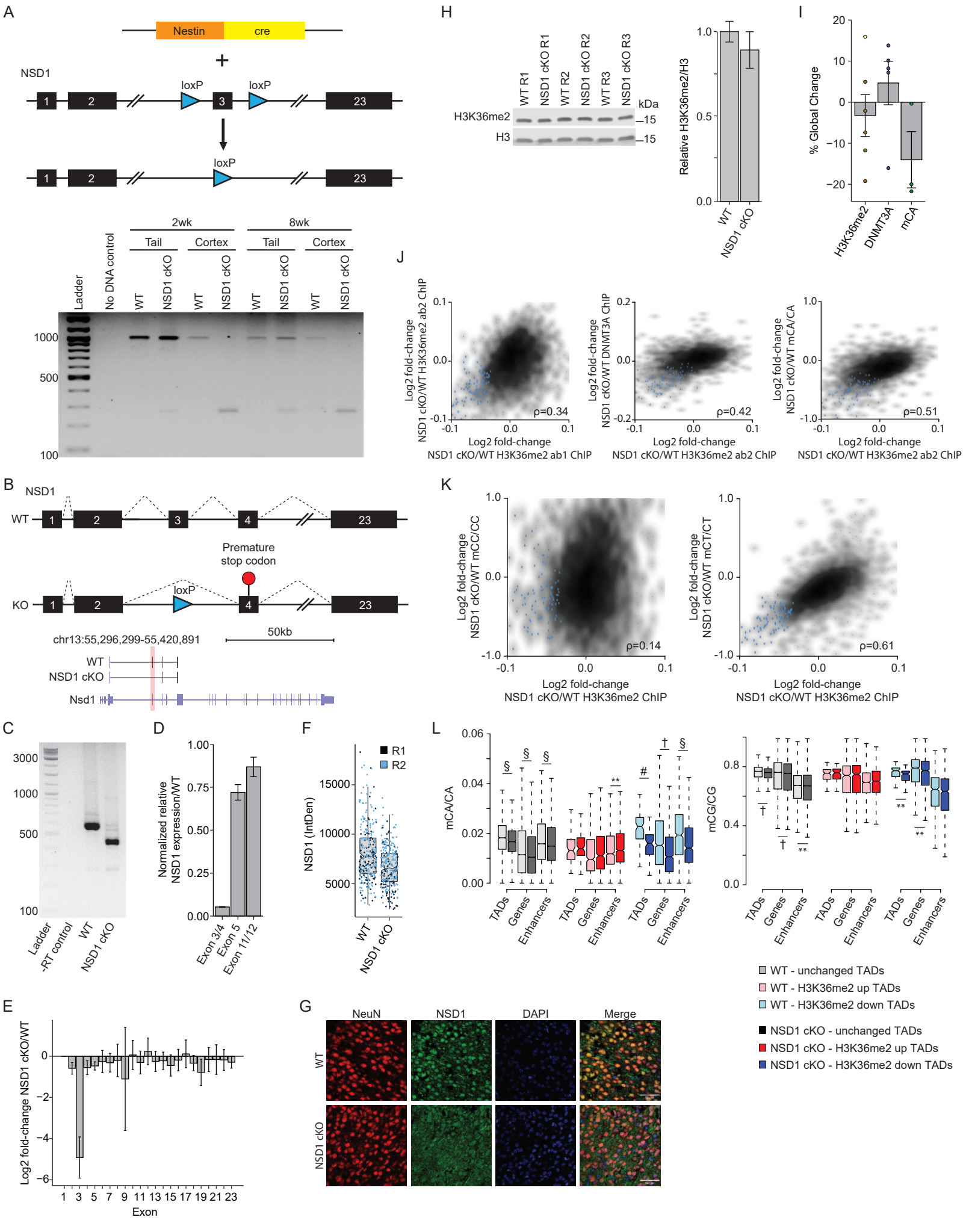


Figure S5. Conditional deletion of *Nsd1* in the brain leads to changes H3K36me2, DNMT3A binding and DNA methylation. Related to Figure 5.

(A) Top, schematic outlining generation of the NSD1 cKO mutant allele. Bottom, genotyping PCR of the wildtype, flx/flx, and mutant alleles from tail or whole cortex at 2- and 8-weeks.

(B) Top, schematic outlining mRNA splicing over excised exon 3 to generate shortened mRNA with premature stop codon and truncated NSD1 protein in the NSD1 cKO. Bottom, genome browser view of Nsd1 gene with sanger sequencing of RT-PCR products from NSD1 cKO and wildtype whole cortex mRNA illustrating successful excision of targeted exon.

(C) PCR of NSD1 cKO and wildtype whole cortex cDNA illustrating successful excision of targeted exon from the major NSD1 isoform. The major product generated corresponds to the exon 2-4 spliced mRNA containing a premature stop codon.

(D) qRT-PCR of *Nsd1* exons (3/4 junction, 5, or 11/12 junction) over *Actb* normalized to wildtype from NSD1 cKO and wildtype whole cortex mRNA (Per condition, n = 6).

(E) Fold-changes in NSD1 mRNA transcript levels per exon in NSD1 cKO versus wildtype as measured by RNA-seq of whole cortex total RNA. Exon number as indicated by UCSC Genes mm9 gene model. Note no value shown for exon 1, as this noncoding exon shows no expression in wild-type or mutant samples.

(F) Quantification of NSD1 fluorescent signal from NeuN+ NSD1 cKO and wildtype cortical nuclei at 2-weeks (Per condition, n = 2).

(G) Representative images of immunohistochemistry in NSD1 cKO and wildtype coronal brain sections at 2-weeks. Scale bar = 50μm. NSD1 antibody shows substantial non-specific non-nuclear background but displays loss of nuclear signal, supporting disruption of NSD1 in these mice.

(H) Left, representative immunoblot from NSD1 cKO and wildtype cortical whole cell lysate at 2-weeks (lane 1 ladder not shown). Right, quantification of relative immunoblot signal of H3K36me2 over total Histone H3 normalized to wildtype (Per condition, n = 3).

(I) Mean percent global change in H3K36me2, DNMT3A, and mCA observed in the NSD1 cKO using ChIP-Rx normalization.[S5] Each point denotes a batch of bioreplicates were processed, with two colors (dark orange = ab1, yellow = ab2) indicating different antibodies used for the H3K36me2 group.

(J) Smoothscatter showing changes in H3K36me2 across two H3K36me2 ChIP antibodies (left) as well as changes in H3K36me2 ab2 and DNMT3A (center) or mCA (right) for each TAD. TADs with significantly reduced H3K36me2 by edgeR (FDR<0.1) analysis of H3K36me2 ab1 ChIP are highlighted in blue.

(K) Smoothscatter showing changes in H3K36me2 and mCC (left) or mCT (right) for each TAD. TADs with significantly reduced H3K36me2 by edgeR (FDR<0.1) are highlighted in blue.

(L) DNA methylation levels at TADs with significantly changed H3K36me2 (edgeR, FDR <0.1) and kilobase-scale genomic elements that reside within these TADs in NSD1 cKO and wildtype cerebral cortex.

Data are from NSD1 cKO and control mouse whole cortex at two or eight weeks. Per time point: 2-weeks, n = 5 bioreplicates for H3K36me2 ab1 (n = 2 for H3K36me2 ab2), H3K36me3, DNMT3A ChIP-Rx. 8-weeks, n = 3 bioreplicates for DNA methylation, n = 6 bioreplicates for total RNA-seq. TADs were derived from Hi-C analysis of cerebral cortex.[S3,S6] *$p < 0.05$, **$p < 0.01$, #$p < 10^{-5}$, †$p < 10^{-10}$, §$p < 10^{-15}$.

Supplementary Figure 6

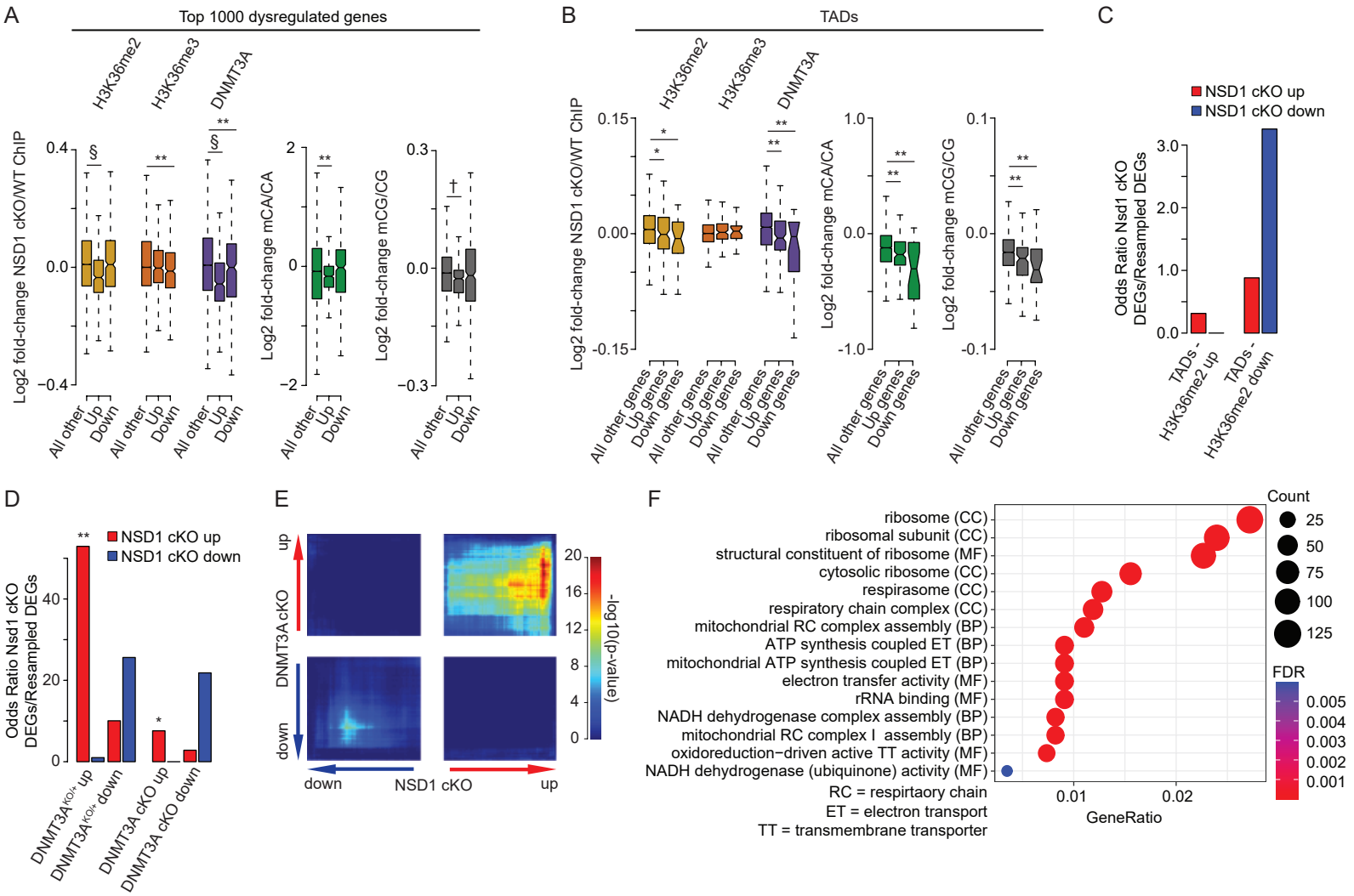


Figure S6. Analysis of altered gene expression and associated epigenetic changes upon conditional deletion of *Nsd1* in the brain. Related to Figure 6.

(A) Fold-changes of H3K36me2, H3K36me3, DNMT3A, and DNA methylation at the top 500 upregulated and top 500 downregulated genes identified (DESeq2, ranked by FDR) by RNA-seq in the NSD1 cKO. (Wilcoxon test).

(B) Fold-changes of H3K36me2, H3K36me3, DNMT3A, and DNA methylation at TADs that overlap significantly dysregulated genes in the NSD1 cKO.

(C) Odds ratio of the overlap of NSD1 cKO significantly dysregulated genes with TADs showing significantly changed H3K36me2. (Fisher's exact test, observed versus background estimated by resampling).

(D) Odds ratio of the overlap of NSD1 cKO significantly dysregulated genes with $DNMT3A^{KO/+}$ and DNMT3A Baf53-cKO significantly dysregulated genes.[S3,S7] (Fisher's exact test, observed versus background estimated by resampling).

(E) RRHO[S8,S9] of all gene expression changes in the NSD1 cKO versus the DNMT3A Baf53-cKO in which DNMT3A is deleted from all postmitotic neurons and mCA is blocked from accumulating genome-wide.[S3]

(F) Top five Gene Ontology terms from "Biological Process", "Molecular Function", and "Cellular Components" from RRHO-determined "dd" (downregulated-downregulated) genes from NSD1 cKO and $DNMT3A^{KO/+}$ using clusterProfiler.[S10,S11]

Data are from NSD1 cKO and control mouse whole cortex at two or eight weeks. Per time point: 2-weeks, n = 5 bioreplicates for H3K36me2, H3K36me3, DNMT3A ChIP-Rx. 8-weeks, n = 3 bioreplicates for DNA methylation, n = 6 bioreplicates for total RNA-seq. TADs were derived from Hi-C analysis of cerebral cortex.[S3,S6] *$p < 0.05$, **$p < 0.01$, #$p < 10^{-5}$, †$p < 10^{-10}$, §$p < 10^{-15}$.

Supplemental References

S1. 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. 10.1126/science.1237905.

S2. Stroud, H., Su, S.C., Hrvatin, S., Greben, A.W., Renthal, W., Boxer, L.D., Nagy, M.A., Hochbaum, D.R., Kinde, B., Gabel, H.W., et al. (2017). Early-Life Gene Expression in Neurons Modulates Lasting Epigenetic States. Cell *171*, 1151-1164.e16. 10.1016/j.cell.2017.09.047.

S3. Clemens, A.W., Wu, D.Y., Moore, J.R., Christian, D.L., Zhao, G., and Gabel, H.W. (2019). MeCP2 Represses Enhancers through Chromosome Topology-Associated DNA Methylation. Mol Cell *77*, 279-293.e8. 10.1016/j.molcel.2019.10.033.

S4. Bonev, B., Cohen, N.M., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.-P., Tanay, A., et al. (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell *171*, 557-572.e24. 10.1016/j.cell.2017.09.043.

S5. Orlando, D.A., Chen, M.W., Brown, V.E., Solanki, S., Choi, Y.J., Olson, E.R., Fritz, C.C., Bradner, J.E., and Guenther, M.G. (2014). Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome. Cell Reports *9*, 1163–1170. 10.1016/j.celrep.2014.10.018.

S6. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380. 10.1038/nature11082.

S7. Christian, D.L., Wu, D.Y., Martin, J.R., Moore, J.R., Liu, Y.R., Clemens, A.W., Nettles, S.A., Kirkland, N.M., Papouin, T., Hill, C.A., et al. (2020). DNMT3A Haploinsufficiency Results in Behavioral Deficits and Global Epigenomic Dysregulation Shared across Neurodevelopmental Disorders. Cell Reports *33*, 108416. 10.1016/j.celrep.2020.108416.

S8. Cahill, K.M., Huo, Z., Tseng, G.C., Logan, R.W., and Seney, M.L. (2018). Improved identification of concordant and discordant gene expression signatures using an updated rank-rank hypergeometric overlap approach. Sci Rep-uk *8*, 9588. 10.1038/s41598-018-27903-2.

S9. Plaisier, S.B., Taschereau, R., Wong, J.A., and Graeber, T.G. (2010). Rank–rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures. Nucleic Acids Res *38*, e169–e169. 10.1093/nar/gkq636.

S10. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation *2*, 100141. 10.1016/j.xinn.2021.100141.

S11. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. Omics J Integr Biology *16*, 284–287. 10.1089/omi.2011.0118.