Prenatal Δ^9-Tetrahydrocannabinol Exposure in Males Leads to Motivational Disturbances Related to Striatal Epigenetic Dysregulation

Supplement 1

Supplementary Methods and Materials

Δ^9-tetrahydrocannabinol (THC) and vehicle solutions

200 mg/ml THC stock (originally dissolved in 95% ethanol, National Institute on Drug Abuse) was evaporated under 100% nitrogen gas to eliminate the ethanol. The remaining THC was dissolved in 0.9% NaCl with 0.3% Tween 80 to a concentration of 0.75 mg/ml. THC was administered intravenously (i.v). at a dose of 0.15 mg/kg, using an injection volume of 0.20 ml/kg. This dose corresponds to ~16mg of THC, common in cannabis cigarettes, correcting for body weight and route of administration (2). 0.3% Tween 80 dissolved in 0.9% NaCl was used as vehicle (VEH) treatment.

Animals and perinatal THC exposure paradigm

Female dams were intravenously injected with 0.15mg/kg THC or 0.3% Tween 80 vehicle (VEH) from gestational day 5 to postnatal day 2, corresponding to mid-gestation in humans (3). At PND2, male offspring were cross-fostered VEH-treated dams. Four independent cohorts of male rats were used in this study with no more than three pups from each birth litter. Adult male offspring were assigned to different assays with ≤3 pups/litter/group to avoid bias of litter to the biological measures. Rats used in behavioral testing were on a reverse light cycle schedule (7AM-7PM lights off) at the time of weaning. All behavioral testing occurred during the dark phase between 8AM-6PM. Molecular experimental subgroups of animals were euthanized using CO_2, followed by decapitation and brain extraction. NAc was dissected on a -20 °C cold block (Teca AHP-

1200CPV) using a 15-gauge sample corer (Fine Sciences) to punch the tissue. Samples were stored at -80°C until processing. All procedures followed the Icahn School of Medicine Institutional Animal Care and Use Committee approved protocols. Across all behavioral experiments, no changes in overall body weight were observed.

Mice with conditional knockout of Kmt2a in NAc neurons (Camk2a-Cre^{-}, Kmt2a$^{2lox/2lox}$) were bred as previously described (4). Briefly, conditional deletion of Kmt2a was obtained by breeding a Kmt2a$^{2lox/2lox}$ allele into a C57BL/6J background, and during adulthood, stereotactic delivery of adeno-associated virus, serotype 8 (AAV) for expression of a CreGFP fusion protein under control of the *SYNAPSIN1* promoter in NAc resulting in Cre-mediated deletion of Kmt2a in NAc neurons. The coordinates were: anterior-posterior: +1.55, medial-lateral: ±1.54, dorsal-ventral: -4.54. Behavior experiments were performed at least 3-4 weeks after surgery.

Behavioral Testing

Palatable food self-administration: To assess hedonic food reinforcement, animals were placed in operant chambers where lever pressing resulted in the release of a chocolate pellet. Rats were food restricted (10-18g/day) one day before and throughout testing. Self-administration behavior was studied at fixed ratio-1 (FR1; one food pellet in response to one active lever press) to monitor reward-sensitivity, or progressive ratio (PR) reinforcement schedules (5, 6) to assess motivational effort. Animals that never acquired the FR1 task were excluded from the study.

Forced swim test: Behavioral despair paradigm to determine depression-like behavior (7). Rats were placed in cylindrical chambers (30.5 cm x 45.7 cm; MedAssociates) filled with 25°C water. Testing consisted of two phases: a pre-test where animals were placed in the water for 15 min, subsequently dried/warmed and returned to their home cage, followed by the test 24h later, where rats were placed in the same tanks for 5 minutes and movements videotaped. Behavior (active

movement vs. floating) was monitored to determine an immobility score, reflecting time spent floating vs active.

Sucrose Preference: To evaluate hedonic state. A pre-test was conducted where animals were habituated to drink water from two bottles for two days. Subsequently, the animals were exposed to a 1% sucrose solution for two days. Intake of water and sucrose solutions was measured in weight after 24 hours. The positions of the bottles were switched daily, and right/left placement was randomized across animals within each group. Preference for sucrose over water was calculated from the amount of sucrose solution consumed and expressed as a percentage of the total amount of liquid imbibed. Anhedonia-like behavior is operationally defined as a reduction in sucrose intake. No changes in water intake were observed during adulthood.

Kmt2a knockdown in the NAc

Rats were anesthetized (4% isoflurane) and stereotaxically implanted bilaterally with 1-mm guide cannulas (Plastics One, State); anterior-posterior: +1.92 mm, medial-lateral: ±2.8 mm from bregma and dorsal-ventral: -3.9 mm from skull at a 10° angle. One day prior to the behavioral test, animals were lightly sedated, then using an infusion needle that projected 3.2 mm beyond the guide cannula, 1 μl of a 20 pmol solution of small interfering RNA (siRNA, ON-TARGETplus SMART pool, rat *Kmt2a,* L-100995-02-0005), or ON-TARGET nontargeting pool (D-001810-10-05) was injected into the NAc using JetSI transfection reagent (Polyplus Transfection, Illkrich, France) (8-10). The strategy and knock-down efficiency is shown in Fig. S1.

Gene expression measurements using quantitative reverse transcription-PCR (qRT-PCR)

Total RNA was isolated using the RNAqueous-micro kit (Ambion), followed by DNAse treatment (ThermoFisher). RNA purity was confirmed by 280/260 ratios >=1.9 and 260/230 ratios >=1.8 and

gene expression was quantified using the the ΔΔCp method (11), normalized to a TBP as the reference housekeeping gene. The list of chromatin regulators investigated in our initial screen using a universal Locked Nucleic Acid (LNA) customized qRT-PCR approach (Roche), as well as corresponding primer/probe information, can be found in Table S1. Kmt2a expression was also assessed using a commercially available TaqMan assay (Rn01444748_m1, ThermoFisher). Statistical differences between the prenatally THC- and VEH-exposed groups were determined using t-tests at $p<0.05$.

Gene expression measurements using the Nanostring technology

Total RNA was extracted as described above for qRT-PCR and analyzed using the NanoString nCounter technology at the Icahn School of Medicine qPCR Core (12). mRNA was hybridized to a codeset (for complete gene target list see Table S2), followed by scanning of the immobilized mRNA transcripts in the nCounter Digital Analyzer. Statistical differences between the THC-exposed and control groups were determined by standard t-tests at $p<0.05$.

Protein isolation and western blot analysis

Tissue punches were homogenized using a Dounce tissue grinder in a modified RIPA buffer (Thermo Scientific) with protease inhibitors (Roche) and phosphatase inhibitors (Thermo Scientific). The homogenate was sonicated and protein was quantified using the BCA assay (Pierce) with a uQuant plate reader (Biotek Instruments) and Gen5 software (Biotek Instruments). For gel electrophoresis, protein aliquots were prepared and samples were heated, centrifuged, migrated into a 4-12% BisTris gel (1.5mm x 16 wells, Life Technologies), and proteins were transferred to a nitrocellulose membrane (Life Science). Total protein was visualized, and the blot was blocked overnight and incubated in anti-Kmt2a primary antibody (1:10000, Bethyl 374A), diluted in 50% 1X PBS and 50% Odyssey blocking buffer, for 1h and 15min at room temperature. The membrane was washed and incubated in IRDye-conjugated secondary antibodies for 1h,

followed by additional washes, then imaged on the Li-Cor Odyssey CLx system. Western blots were analyzed using the ImageJ software to measure integrated density values and then normalized to total protein amount in each sample (13).

RNA-sequencing (RNA-seq)

RNA from rat NAc and neuronal progenitor cell (NPC) samples was extracted using Trizol (Thermo Fisher), followed by DNase treatment (Thermo Fisher), as described by the manufacturer. RNA integrity number (RIN) and concentration were assessed using a Bioanalyzer (Agilent). Sequencing libraries were prepared from 250 ng of total RNA per the manufacturer's protocol. Paired-end sequencing of pooled libraries was performed on an Illumina NovaSeq 6000 S1 flowcell PE50. All libraries were prepared in a single batch and sequenced on a single flow cell. Library preparation and sequencing were carried out at the NYU Langone Health Genome Technology Center.

Data processing and differential gene expression analysis.

Raw sequencing reads were mapped to rn6 (rat) and hg38 (human) genome assemblies using HISAT2. Counts of reads mapping to genes were obtained using the featureCounts software of Subread package against Ensembl v90 annotation. Differentially expressed genes were identified by taking the intersection of transcripts with nominal $p<0.05$ in both outputs generated from DESeq2 (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) (14) and Voom-limma (https://bioconductor.org/packages/release/bioc/html/limma.html) (15). Principal component analysis (PCA) in DESeq2 package was utilized to remove outliers.

Chromatin ImmunoPrecipitation followed by sequencing (ChIP-seq)

ChIP procedure was carried out on bilateral NAc punches as described in (16). Recovered DNA was quantified using a Qubit instrument (ThermoFisher) and 1ng/sample was used for

sequencing library preparation with the TruSeq ChIP Library Preparation Kit (Illumina). Quality of the libraries was assessed on a Bioanalyzer (Agilent). Sequencing was carried out at the NYU Langone Health Genome Technology Center.

Data processing, identification of H3K4me3 peaks and THC effects.

The ChIP-seq data were first checked for quality using the various metrics generated by FastQC (v0.11.2) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Raw sequencing reads were then aligned to the rat Rnor_6.0 genome using default settings of Bowtie (v2.2.0) (17). Only uniquely mapped reads were retained and duplicate reads were removed using the SAMtools package (v0.1.19) (18). Peak-calling was performed using MACS (v2.1.1) (19) with false discovery rate cutoff of 0.05. Differentially bound regions between THC and VEH samples was detected using diffReps (v1.55.6) (20) with p-value cutoff of 0.0001. Annotation of called peaks and differential regions to their genomic features (promoters, gene bodies, intergenic, etc.) was performed using region-analysis (v0.1.2) (https://github.com/shenlab-sinai/region_analysis). All aligned read files were corrected for sequencing depth using the signal extraction method and normalized to the cell-type-specific input to visualize in Integrative Genome Viewer (IGV) browser.

Analysis of H3K4me3 peak distribution relative to genomic features.

Differential H3K4me3 sites were annotated using the R package *ChIPseeker* (21) by whether they were proximal to a gene promoter, gene body, or intergenic. Read count frequencies with reference to distance from transcription start sites were measured for all differential H3K4me3 sites.

Comparison of RNA-seq and ChIP-seq data

To compare the RNA-seq and ChIP-seq datasets, we examined overlapping genes that were both differentially expressed and contained a differential H3K4me3 site. Overlapping genes were

stratified by direction in either dataset (i.e., H3K4me3-up/RNA-up, H3K4me3-down/RNA-up, H3K4me3-up/RNA-down, H3K4me3-down/RNA-down), with counts of genes whose differential H3K4me3 sites were contained in the gene promoter or gene body.

CRISPR-mediated Kmt2a activation (CRISPRa) *in vitro*

Generation of stable selected dCas9-VPR. All CRISPR-based epigenomic editing assays were performed on antibiotic-selected dCas9-VPR (VPR as the tripartite activator, VP64-p65-Rta) human neuronal progenitor cells (NPCs) derived as described in (22) from male patients. For generation of Cas9 stable, selected NPCs, we used a plasmid of lentiCRISPR v2 gifted by Feng Zhang (Addgene plasmid # 52961). DNA sequencing with a U6 primer confirmed identity. Lentiviral production and titration were performed as described previously (Ho et al., 2016). Control S1 and S2 NPCs were spinfected with lentiCRISPR v2 virus as described (22). 48 hours post-transduction, cells were selected by exposure to puromycin at 0.3 μg/mL. Without transduction, all control cells died within around 5 days after the antibiotic addition. The puromycin-selected NPCs were subject to Western blot analysis of Cas9 expression. 30 μg of proteins were electrophoresed in NuPAGE 4-12% Bis-Tris Protein Gels (NP0323PK2, Life Technologies) in 1× MES running buffer, 200 V constant, 35 min. Proteins were transferred onto nitrocellulose membrane (IB23002, Life Technologies) on the iBlot® 2 Dry Blotting System (program P3, 7:00 min). The membranes were incubated with primary antibodies against Cas9 (1:250, monoclonal, clone 7A9, Millipore) and β-Actin (1:10,000, mouse, 1406030, Ambion) overnight at 4°C. Then, membranes were incubated with the IRDye-labeled secondary antibodies for 45 minutes at room temperature, in dark, on a rocker. Fluorescence was visualized using a Li-Cor Odyssey Imaging System.

In vitro transcription and transfection of gRNAs. Guide RNAs (gRNAs) were designed on Benchling (www.benchling.com) using the CRISPR tool. gRNAs were generated via *in vitro*

transcription (IVT) with the GeneArt Precision gRNA Synthesis Kit (Thermo Fisher Scientific) as per manufacturer instructions. Puromycin-selected (1µg/mL in NPC media; Sigma) dCas9-VPR NPCs were seeded at a density of ~450,000 per well on Matrigel-coated (BD Biosciences) 24-well plates. IVT gRNAs (500 ng total RNA/well) and 2 µL EditPro Stem lipofectamine (ThermoFisher, STEM00003) were diluted in 50 µL Opti-MEM (Thermo Fisher Scientific) and added dropwise to each well. Cells were harvested with TRIzol for total RNA extraction 48 hours later. Each data point represents one biological replicate within each condition. CRISPRa results were analyzed on PRISM with a one-way ANOVA across conditions. Cas9 mutagenesis was also performed as described above with the exception of the negative control, which in these experiments consisted of an empty transfection (i.e., lipofectamine + Opti-MEM without any gRNA).

Gene network analysis

Network analysis was conducted to examine the commonalities between what is known about the function and molecular relationships of Kmt2a and the gene expression changes observed in our experimental models as the result of either PTE or Kmt2a overexpression. To do this, we used the STRING (‘Search Tool for Retrieval of Interacting Genes/Proteins) database (23) . First, we searched Kmt2a with Homo sapiens as the species, and used STRING to construct a protein interaction network with a confidence threshold of 0.7 for all edges between proteins (nodes), and no more than 50 interactors in the first and second shells. This network was used as a reference for comparisons with networks constructed from our own datasets. Next, we conducted a gene ontology enrichment analysis for differentially expressed genes in our Kmt2a overexpression dataset that had membership in the top 10 most enriched categories of biological processes and components, as well as the PTE-related NAc RNA-seq data, and inputted these genes into STRING to build two data-driven protein-protein interaction networks (one for biological processes, the other for biological components) to compare to the Kmt2a reference network. The

genes in the top 10 GO categories were used because STRING is limited to 2000 input genes, and there were over 4000 DEGs in this dataset. NetworkAnalyzer via Cytoscape (24, 25) was used to quantify degree (connectivity) for each node in these two networks and identified overlapping nodes between these networks and the Kmt2a reference network, along with their respective degrees. Following this, we used the DEGs in the rat ventral striatum RNA-seq dataset, and built a protein-protein interaction network from this gene list, and overlapped this network with the Kmt2a reference network, along with its respective degrees. We used these networks in our subsequent analysis of overlapping proteins between the Kmt2a reference network and our data-driven networks to find their first-shell interaction partners that were enriched for synaptic plasticity-related ontologies.

Differential gene correlation analysis (DGCA)

Differential gene correlation analysis (26) was conducted on the DEGs in the PTE RNA-seq and ChIP-seq and CRISPR RNA-seq datasets. We then identified which differential correlations were common to more than one dataset and compiled the unique genes in these replicated differential correlations.

Comparison of RNA-seq from rat NAc and CRISPR model with publicly available human MDD datasets

Using the Gene Expression Omnibus (27), we searched for publicly available microarray and RNA-seq datasets (GSE87610, GSE102556) from brain tissue in human studies of major depressive disorder. We used GEO2R (28) to conduct differential expression analysis for the microarray datasets, and supplementary data files containing the DEGs from the studies that conducted RNA-seq experiments.

Supplementary Results

Gene ontology analyses of genes showing H3K4me3 enrichment and interrogation in our CRISPR model

We conducted a gene ontology analysis of all genes with H3K4me3 enrichment (Table S16), and identified enriched ontologies (FDR-corrected) including regulation of transcription by RNA polymerase II (0.0396), negative regulation of transcription by RNA polymerase II (p=0.0441), negative regulation of nucleic acid-templated transcription (p=0.0447), and negative regulation of transcription, DNA-templated (p=0.0457). These ontologies included 118 unique genes, 14 of which were also downregulated in our CRISPR model. The overlapping genes are: *Bhlhe22*, *Dusp5*, *Ets1*, *Gata2*, *Gli1*, *Hivep1*, *Mecom*, *Pcgf5*, *Pdgfb*, *Rorb*, *Smad3*, *Tgfb1*, *Tle3*, *Tsc22d1*. *Bhlhb5*, *Mecom*, *Pcgf5*, and *Rorb* have all been shown to play critical roles in cortical development and embryonic stem cell differentiation, and *Smad3* deficiency has been linked to cortical and hippocampal neuronal loss after traumatic brain injury, along with reduced neurogenesis (29-36).

Demographics of public human MDD datasets

The human clinical MDD data shown in Fig. S9 is from a secondary analysis from two datasets (GSE102556, GSE87610).

In the data from GSE102556 (NAc, Subiculum, Insula, Brodmann Areas 11, 25, and 8/9), there were 13 control males (Age: 41.2 ± 11.3), 13 MDD males (46.7 ± 15.7), 9 control females (58.1 ± 19.5), and 13 MDD females (43.7 ± 11.6). Three control males were heavy cigarette smokers. Five MDD males were heavy cigarette smokers, and four were prescribed antidepressants. Four control females were heavy cigarette smokers, two abused alcohol, and two were prescribed antidepressants (for sleep disorders). Six MDD females were heavy cigarette smokers, one was a moderate cigarette smoker, two abused alcohol, and nine were prescribed antidepressants. Detailed clinical information regarding the population is available in Supplementary Table 1 of Labonte et al. (37).

In the data from GSE87610 (L3, L5, L3+5 DLPFC), tetrads were gathered from 10 control males (Age: 22-68) and 8 control females (Ages: 36-65) (72 samples total), 2 MDD males (Ages: 47 and 51) and 1 MDD female (Age: 46) (12 samples), 1 MDD-recurrent male (Age: 29) and 3 MDD-recurrent females (Ages: 26, 28, and 53) (16 samples). Clinical information regarding the subjects is available in Supplementary Table S1 of Arion et al (38).

Supplementary References

1. Tapia MA, Lee JR, Weise VN, Tamasi AM, Will MJ (2019): Sex differences in hedonic and homeostatic aspects of palatable food motivation. *Behav Brain Res*. 359:396-400.
2. Grotenhermen F (2003): Pharmacokinetics and pharmacodynamics of cannabinoids. *Clin Pharmacokinet*. 42:327-360.
3. Bayer SA, Altman J, Russo RJ, Zhang X (1993): Timetables of neurogenesis in the human brain based on experimentally determined patterns in the rat. *Neurotoxicology*. 14:83-144.
4. Jakovcevski M, Ruan H, Shen EY, Dincer A, Javidfar B, Ma Q, et al. (2015): Neuronal Kmt2a/Mll1 histone methyltransferase is essential for prefrontal synaptic plasticity and working memory. *J Neurosci*. 35:5097-5108.
5. Paterson NE, Froestl W, Markou A (2004): The GABAB receptor agonists baclofen and CGP44532 decreased nicotine self-administration in the rat. *Psychopharmacology (Berl)*. 172:179-186.
6. Hodos W, Kalman G (1963): Effects of increment size and reinforcer volume on progressive ratio performance. *J Exp Anal Behav*. 6:387-392.
7. Slattery DA, Cryan JF (2012): Using the rat forced swim test to assess antidepressant-like activity in rodents. *Nat Protoc*. 7:1009-1014.
8. McQuown SC, Barrett RM, Matheos DP, Post RJ, Rogge GA, Alenghat T, et al. (2011): HDAC3 is a critical negative regulator of long-term memory formation. *J Neurosci*. 31:764-774.
9. Aguilar-Valles A, Vaissiere T, Griggs EM, Mikaelsson MA, Takacs IF, Young EJ, et al. (2013): Methamphetamine-Associated Memory Is Regulated by a Writer and an Eraser of Permissive Histone Methylation. *Biol Psychiatry*.
10. Griggs EM, Young EJ, Rumbaugh G, Miller CA (2013): MicroRNA-182 regulates amygdala-dependent memory formation. *J Neurosci*. 33:1734-1740.
11. Livak KJ, Schmittgen TD (2001): Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. 25:402-408.
12. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, et al. (2008): Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol*. 26:317-325.
13. Aldridge GM, Podrebarac DM, Greenough WT, Weiler IJ (2008): The use of total protein stains as loading controls: an alternative to high-abundance single-protein controls in semi-quantitative immunoblotting. *J Neurosci Methods*. 172:250-254.
14. Love MI, Huber W, Anders S (2014): Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 15:550.
15. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. (2015): limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 43:e47.
16. Dinieri J, Wang X, Szutorisz H, Spano S, Kaur J, Casaccia P, et al. (2011): Maternal Cannabis Use Alters Ventral Striatal Dopamine D2 Gene Regulation in the Offspring. *Biological Psychiatry*. 70:763-769.
17. Langmead B, Trapnell C, Pop M, Salzberg SL (2009): Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 10:R25.
18. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. (2009): The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 25:2078-2079.
19. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. (2008): Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 9:R137.
20. Shen L, Shao NY, Liu X, Maze I, Feng J, Nestler EJ (2013): diffReps: detecting differential chromatin modification sites from ChIP-seq data with biological replicates. *PLoS One*. 8:e65598.
21. Yu G, Wang LG, He QY (2015): ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*. 31:2382-2383.

22. Ho SM, Hartley BJ, Flaherty E, Rajarajan P, Abdelaal R, Obiorah I, et al. (2017): Evaluating Synthetic Activation and Repression of Neuropsychiatric-Related Genes in hiPSC-Derived NPCs, Neurons, and Astrocytes. *Stem Cell Reports*. 9:615-628.
23. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. (2019): STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 47:D607-D613.
24. Assenov Y, Ramirez F, Schelhorn SE, Lengauer T, Albrecht M (2008): Computing topological parameters of biological networks. *Bioinformatics*. 24:282-284.
25. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. (2003): Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 13:2498-2504.
26. McKenzie AT, Katsyv I, Song WM, Wang M, Zhang B (2016): DGCA: A comprehensive R package for Differential Gene Correlation Analysis. *BMC Syst Biol*. 10:106.
27. Clough E, Barrett T (2016): The Gene Expression Omnibus Database. *Methods Mol Biol*. 1418:93-110.
28. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. (2013): NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res*. 41:D991-995.
29. Wang Y, Symes AJ (2010): Smad3 deficiency reduces neurogenesis in adult mice. *J Mol Neurosci*. 41:383-396.
30. Villapol S, Wang Y, Adams M, Symes AJ (2013): Smad3 deficiency increases cortical and hippocampal neuronal loss following traumatic brain injury. *Exp Neurol*. 250:353-365.
31. Byun H, Lee HL, Liu H, Forrest D, Rudenko A, Kim IJ (2019): Rorbeta regulates selective axon-target innervation in the mammalian midbrain. *Development*. 146.
32. Oishi K, Aramaki M, Nakajima K (2016): Mutually repressive interaction between Brn1/2 and Rorb contributes to the establishment of neocortical layer 2/3 and layer 4. *Proc Natl Acad Sci U S A*. 113:3371-3376.
33. McGrath CL, Glatt SJ, Sklar P, Le-Niculescu H, Kuczenski R, Doyle AE, et al. (2009): Evidence for genetic association of RORB with bipolar disorder. *BMC Psychiatry*. 9:70.
34. Yao M, Zhou X, Zhou J, Gong S, Hu G, Li J, et al. (2018): PCGF5 is required for neural differentiation of embryonic stem cells. *Nat Commun*. 9:1463.
35. van Donkelaar MMJ, Hoogman M, Pappa I, Tiemeier H, Buitelaar JK, Franke B, et al. (2018): Pleiotropic Contribution of MECOM and AVPR1A to Aggression and Subcortical Brain Volumes. *Front Behav Neurosci*. 12:61.
36. Joshi PS, Molyneaux BJ, Feng L, Xie X, Macklis JD, Gan L (2008): Bhlhb5 regulates the postmitotic acquisition of area identities in layers II-V of the developing neocortex. *Neuron*. 60:258-272.
37. Labonte B, Engmann O, Purushothaman I, Menard C, Wang J, Tan C, et al. (2017): Sex-specific transcriptional signatures in human depression. *Nat Med*. 23:1102-1111.
38. Arion D, Huo Z, Enwright JF, Corradi JP, Tseng G, Lewis DA (2017): Transcriptome Alterations in Prefrontal Pyramidal Cells Distinguish Schizophrenia From Bipolar and Major Depressive Disorders. *Biol Psychiatry*. 82:594-600.

Supplementary Figures

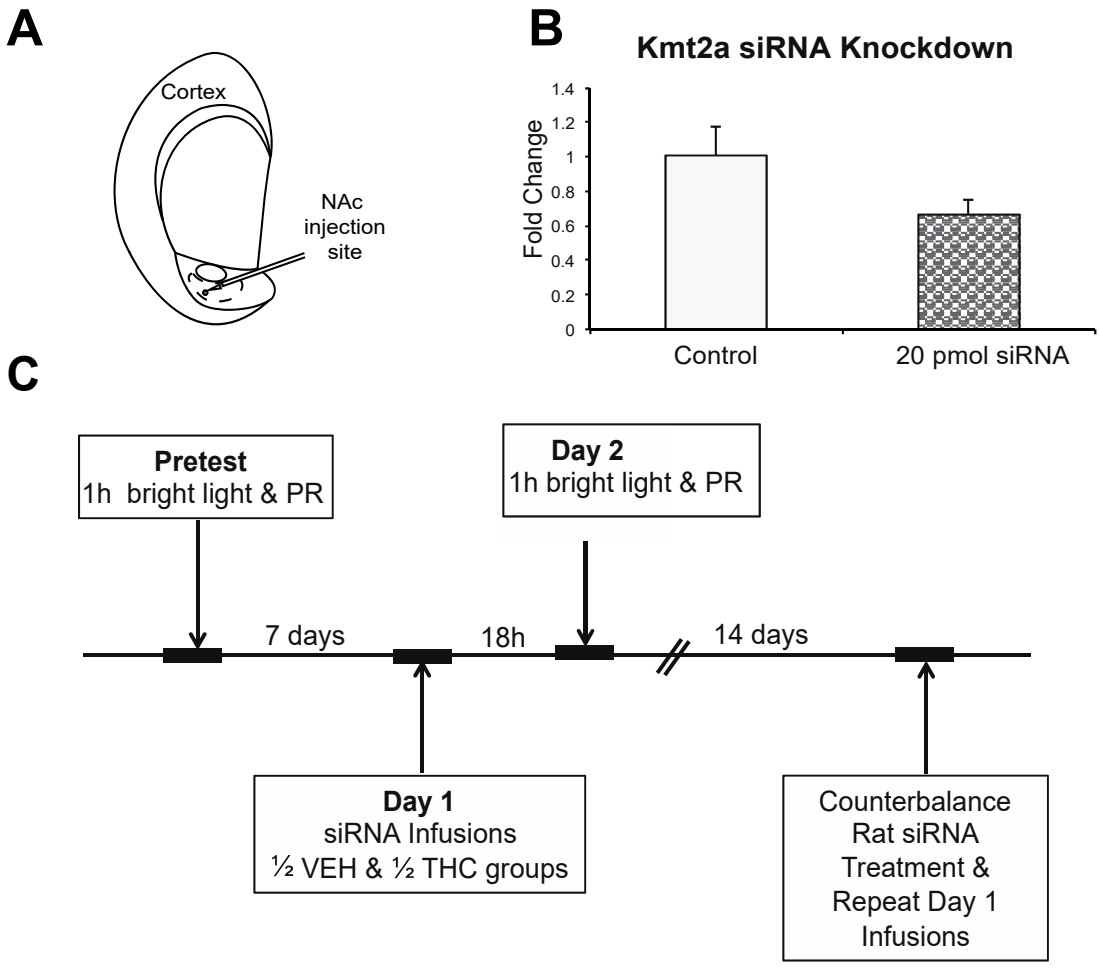

Figure S1. siRNA-mediated knockdown of Kmt2a in the NAc. (A) Schematic of the infusion into the NAc. (B) qRT-PCR analysis of Kmt2a mRNA expression following the injection of 20 pmol siRNA. The manipulation achieves an effect size similar to prenatal THC. (C) siRNA treatment and counterbalanced testing design. Adult rats with prenatal THC or VEH exposure were first reintroduced to the PR food SA task, followed by a 7-day period of only handling. Half of the VEH and THC rats were infused with either Kmt2a or scrambled siRNA, then 14 days later injected with the other siRNA and retested, to ensure effects were due to the specific siRNA manipulation and not simply individual differences.

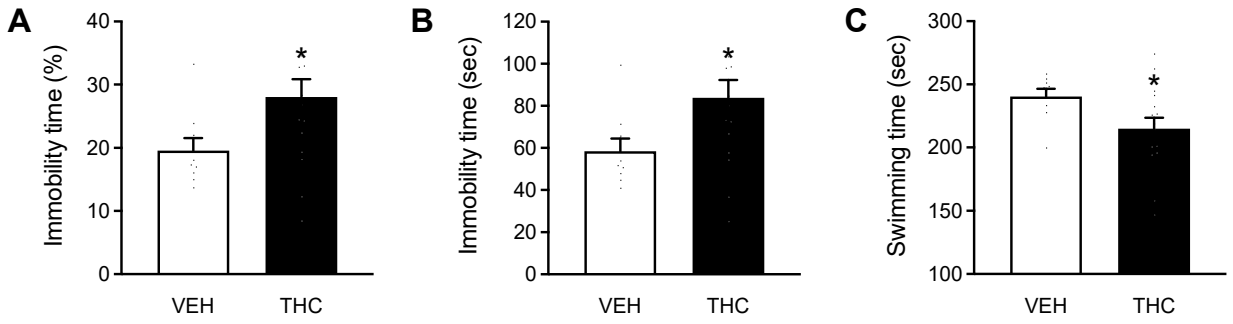

Figure S2. PTE significantly increases immobility time in the forced swim test. (A) Immobility time percentage is increased ($F_{(1,22)} = 4.14$, p=0.047) along with (B) immobility time ($F_{(1,22)} = 4.14$, p=0.047). (C) Similar to the primary cohort, swimming time is decreased ($F_{(1,22)} = 4.14$, p=0.047) in the forced swim test in a separate cohort of male rats.

Figure S3. PTE affects Kmt2a expression in NAc of male rats but not females. A two-way ANOVA revealed a trending significant interaction between the effects of PTE and sex ($F_{(1, 1)} = 4.05$, p = 0.0548). Simple main effects analysis showed significant effects of both PTE (p = 0.031) and sex (p = 0.007). Post-hoc comparisons with Bonferroni correction showed statistically significant effects of PTE in males (p = 0.043), and of sex in the THC groups (p = 0.007). Groups consisted of 6-8 rats.

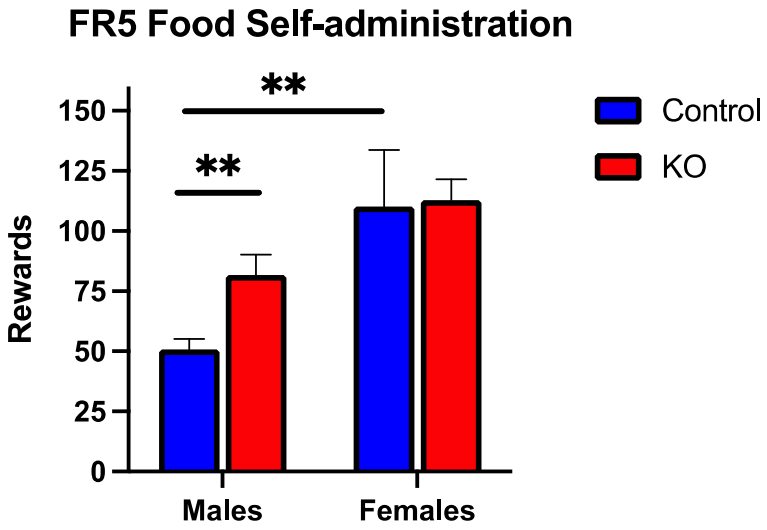

Figure S4. Kmt2a NAc knockout dysregulates food motivation in a sex-specific manner. Fixed ratio 5 food self-administration is increased in male knockout mice ($t_{(11)} = 4.39$, p=0.001), but not females ($t_{(6)} = 0.143$, p=0.89). Additionally, there were baseline differences between the male and female control groups ($t_{(9)} = 3.38$, p=0.008), in line with published sex differences (1). Groups consisted of 4-7 mice.

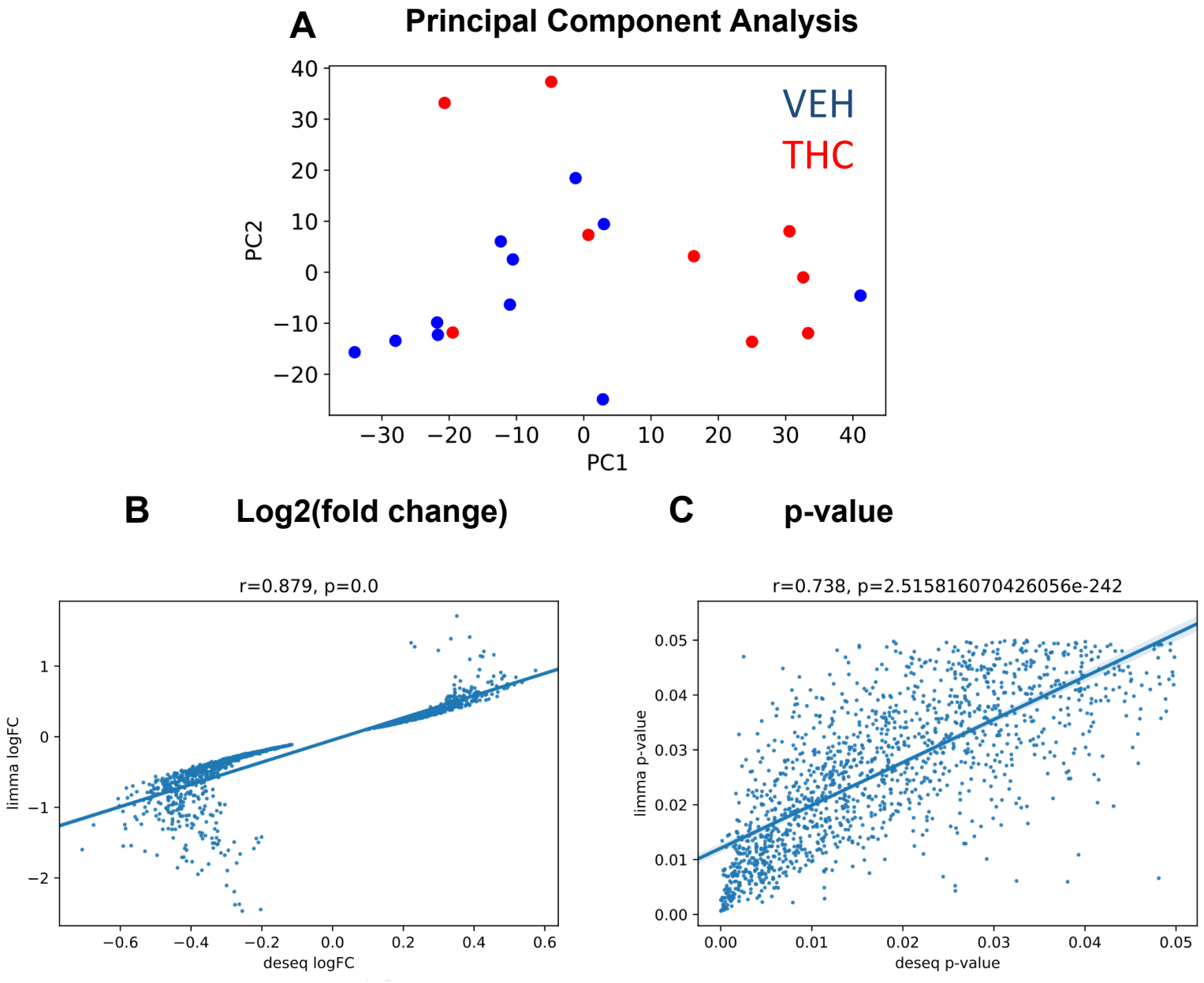

Figure S5. Several factors considered in the PTE RNA-seq data analysis. (A) Principal Component Analysis of RNA-seq reads. Each dot corresponds to a prenatally VEH- or THC-exposed sample. N=9-11/group. Pearson correlation analysis indicated high level of agreement in (B) $\log2_{(\text{Fold change})}$ and (C) p-values. To increase rigorousness, we used differentially expressed genes (DEGs) shared by both analysis at $p<0.05$ in all subsequent data mining and comparisons.

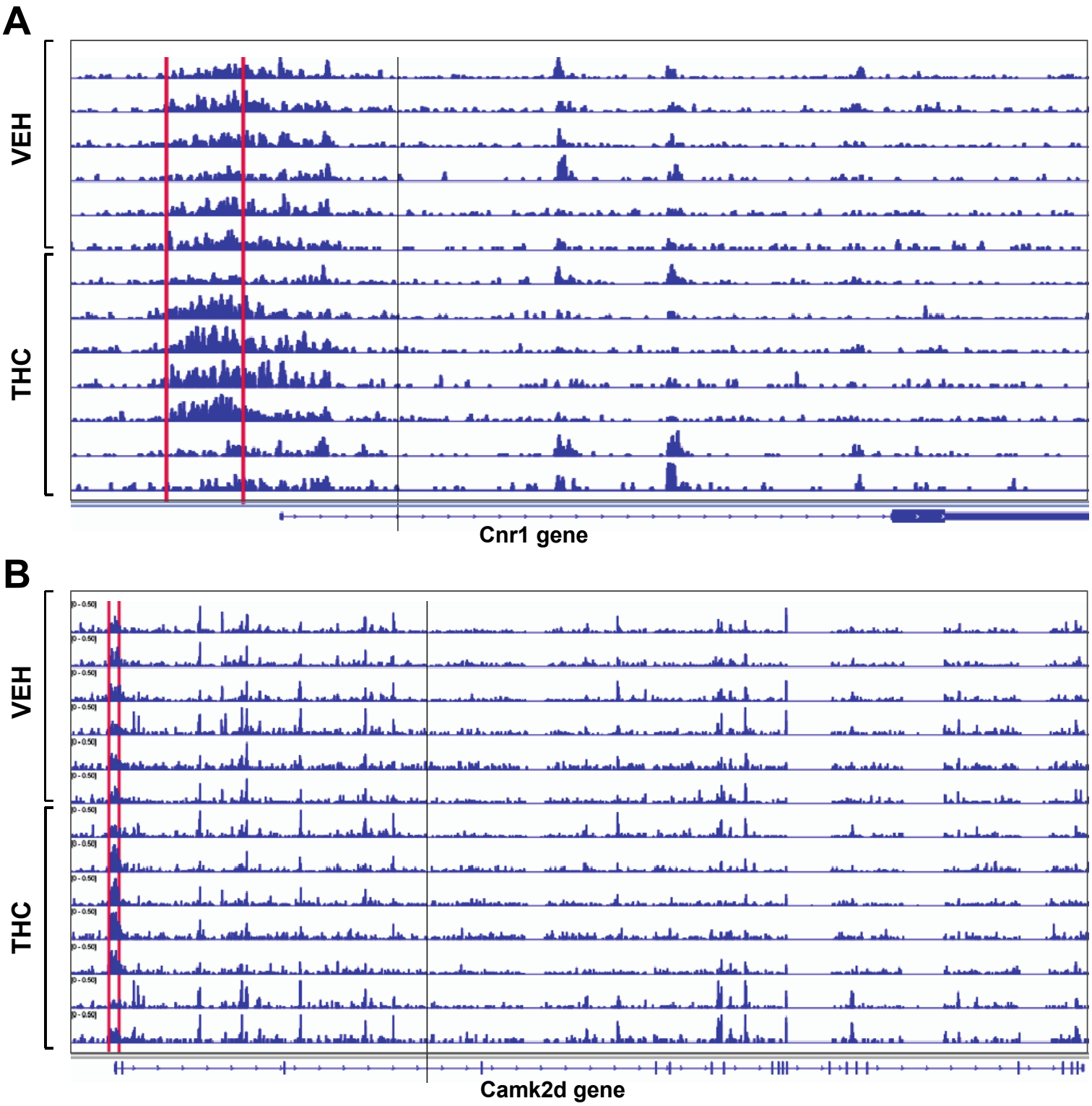

Figure S6. ChIP-seq peak pattern at the entire (A) Cnr1 and (B) Camk2d gene loci. Individual rows correspond to samples within the VEH and THC groups. Red bars span the promoter regions significantly affected by prenatal THC (also shown in Fig. 4G). Gene structures are displayed below the peak distributions in each panel.

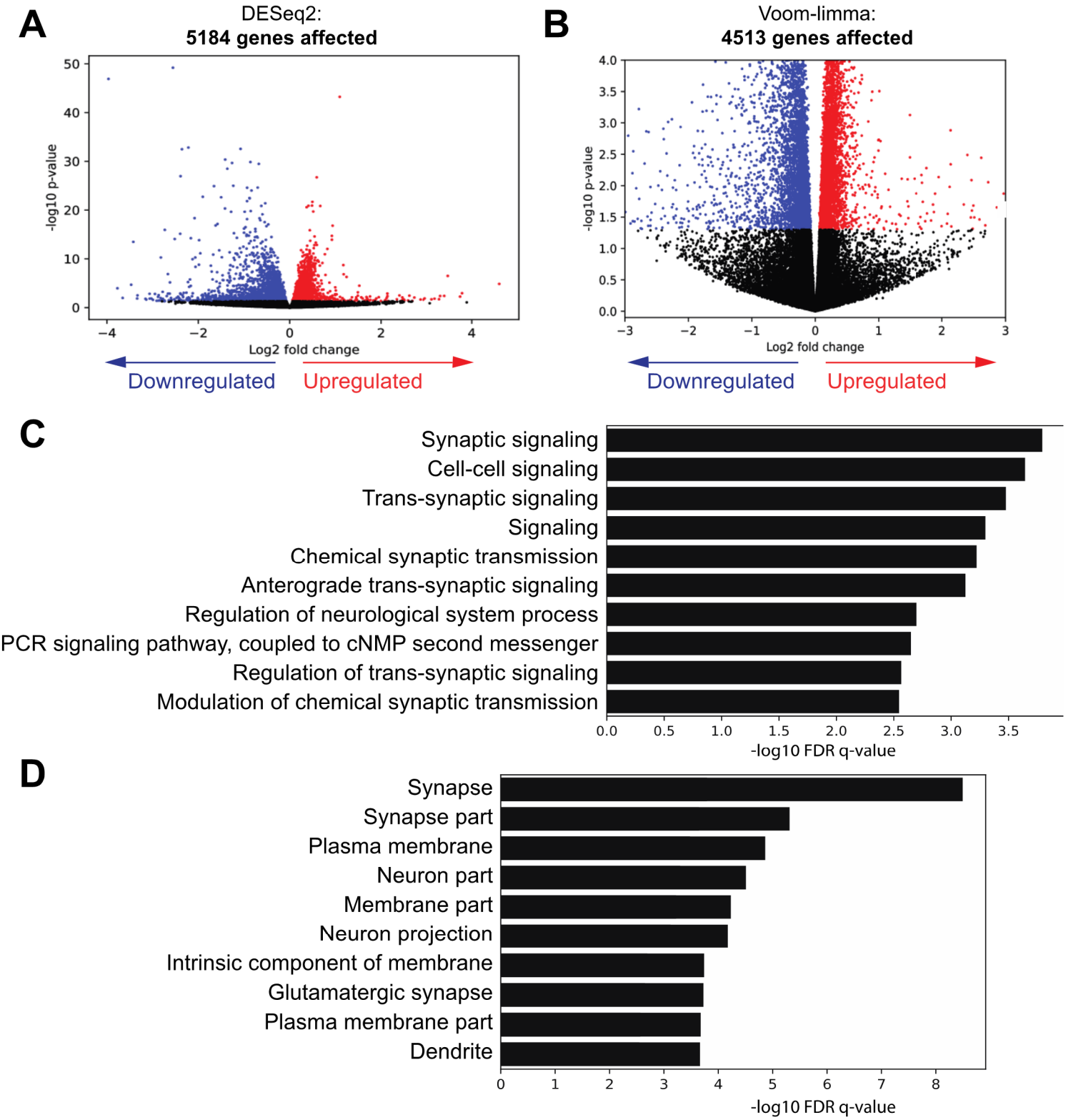

Figure S7. Effects of KMT2A overexpression on the cellular transcriptome and biological pathways. Distribution of DEGs obtained using the DESeq2 (A) and Voom-limma (B) analysis tools. Each dot corresponds to a single gene; colors indicate significant up- (red) or down- (blue) regulation. Gene ontology analysis using Gorilla on DEGs shared in the prenatal THC and CRISPR experiments, graphs show the top 10 “process” (C) and “component” (D) GO categories.

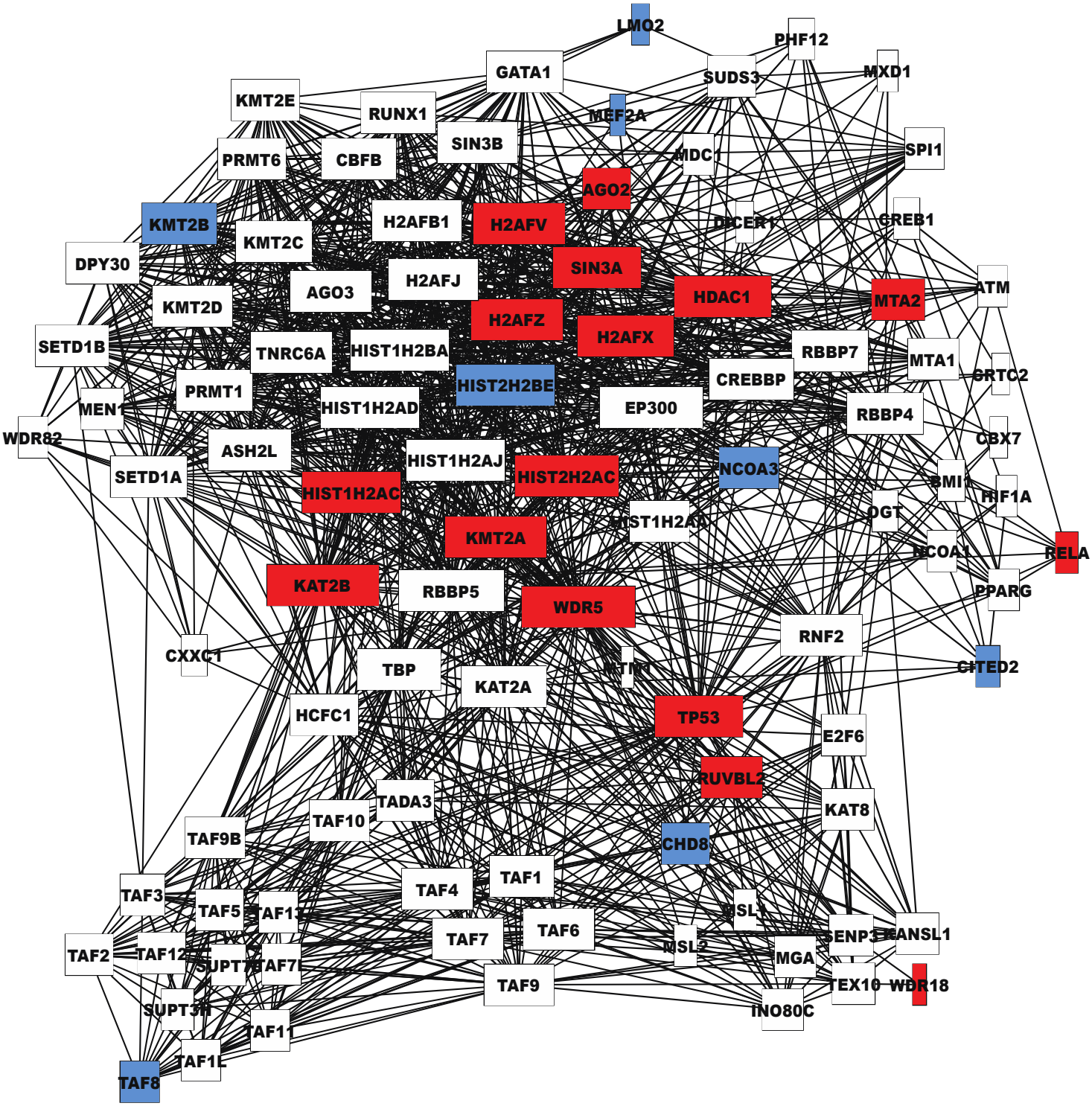

Figure S8. Kmt2a reference network from the STRING database. Rectangle size corresponds to the number of connections for each node (degree). Colors indicate genes dysregulated either in our Kmt2a CRISPR or prenatal THC studies. Red: upregulation, blue: downregulation.

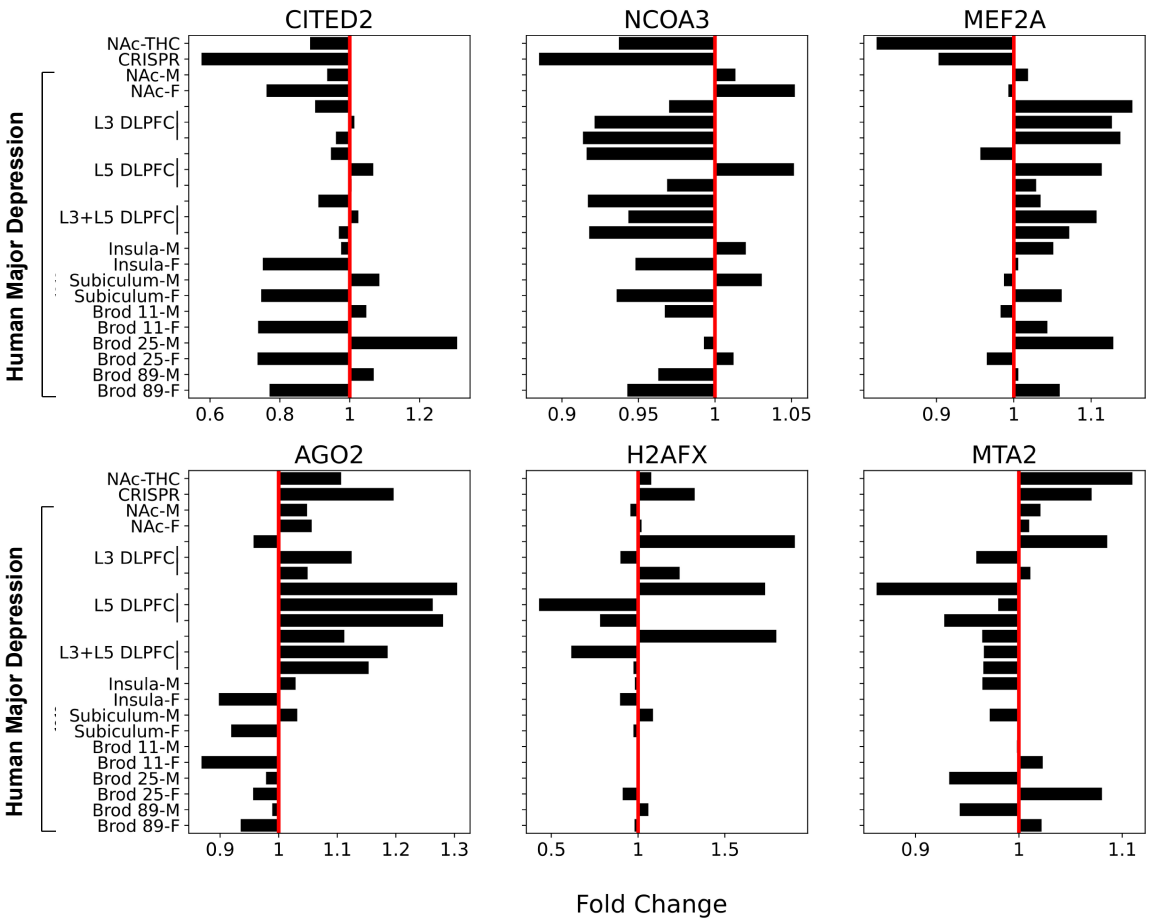

Figure S9. Expression of genes with known Kmt2a interactions in human MDD datasets. mRNA expression changes of six potential transcriptional repressors identified in the Kmt2a reference network. The first two datasets (NAc-THC, CRISPR) refer to our PTE and Kmt2a upregulation networks. RNA-seq or microarray data from post-mortem MDD patients in multiple brain regions were compared to our NAc and CRISPR-based Kmt2a overexpression data. Datasets labeled “L3 DLPFC,” “L5 DLPFC,” and “L3+L5 DLPFC” are from layer 3 and/or 5 of dorsolateral prefrontal cortex, and each comprise three comparisons: control vs. MDD, control vs. MDD-recurrent (i.e., 2+ lifetime episodes of MDD), and control vs. MDD and MDD-recurrent combined. NAc=Nucleus accumbens. Brod=Brodmann area. All sex-specific data (ending in “-M” or “-F”) come from GSE102556: 13 control males (Age: 41.2 ± 11.3), 13 MDD males (46.7 ± 15.7), 9 control females (58.1 ± 19.5), and 13 MDD females (43.7 ± 11.6). The data from L3 and/or L5 DLPFC come from GSE87610, with sexes combined: Tetrads were gathered from 10 control males and 8 control females, 2 MDD males and 1 MDD female, 1 MDD-recurrent male and 3 MDD-recurrent females.