Science
AAAS

Supplementary Materials for

Brain cell type-specific enhancer-promoter interactome maps and disease risk association

Alexi Nott*, Inge R. Holtman*, Nicole G. Coufal*, Johannes C.M. Schlachetzki, Miao Yu, Rong Hu, Claudia Z. Han, Monique Pena, Jiayang Xiao, Yin Wu, Zahara Keuelen, Martina P. Pasillas, Carolyn O'Connor, Christian K. Nickl, Simon T. Schafer, Zeyang Shen, Robert A. Rissman, James B. Brewer, David Gosselin, David D. Gonda, Michael L. Levy, Michael G. Rosenfeld, Graham McVicker, Fred H. Gage, Bing Ren, Christopher K. Glass†

*These authors contributed equally to this work
†Correspondence to: ckg@ucsd.edu

This PDF file includes:

Other Supplementary Materials for this manuscript include the following:

Table S1 to S8 (Excel)

Materials and Methods

Human tissue collection

Nuclei were isolated from human brain tissue resected for treatment of epilepsy in 10 individuals (Table S1). Brain tissue was obtained with informed consent under a protocol approved by the UC San Diego and Rady Children's Hospital Institutional Review Board (IRB # 171361). For all study participants, informed consent was obtained from the parents and when appropriate assent from the patient. The reported data sets are from sequential samples for which sequencing libraries met technical quality standards. Resected brain tissue was immediately placed on ice, transferred to the laboratory, and either snap frozen and stored at -80 ^{o}C prior to nuclei isolation, or immediately processed for isolation of ex vivo live microglia. Nuclei were collected to generate 4x PU.1, 4x NEUN, 4x OLIG2 and 3x LHX2 H3K27ac ChIP-seq datasets (Table S1). Nuclei were collected to generate 3x PU.1, 4x NEUN, 3x OLIG2 and 3x LHX2 ATAC-seq datasets (Table S1). Nuclei were collected to generate 1x PU.1, 2x NEUN, 3x OLIG2 and 2x LHX2 H3K4me3 ChIP-seq datasets and ex vivo microglia were harvested to generate 1x microglia H3K4me3 ChIP-seq dataset (Table S1). Nuclei were collected to generate 4x NEUN and 3x OLIG2 PLAC-seq datasets and ex vivo microglia were harvested to generate 3x microglia PLAC-seq datasets (Table S1). All datasets represent biological replicates. Experiments were not blinded and no randomization was performed. Sample size was determined by tissue availability and follows ENCODE recommended guidelines. No datasets were excluded for this study; sequencing was performed only when sufficient number of nuclei were obtained.

Nuclei isolation with FANS

Frozen brain tissue (~ 100 mg) was homogenized in 1 ml 1% formaldehyde in Dulbecco's phosphate buffered saline (PBS; Corning) and incubated at room temperature for 10 mins. Fixed tissue was quenched with 0.125 M glycine for 5 mins at room temperature and pelleted at 1,100 xg for 10 mins. Subsequent steps were performed on ice or at 4^{o}C. Homogenates were washed two times with NF1 buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 5 mM $MgCl_2$, 0.1 M sucrose, 0.5 % Triton X-100). Homogenates were incubated on ice in 5 ml NF1 for 30 mins. Nuclei were extracted from homogenates using a 7 ml Wheaton™ Dounce Tissue Grinder (DWK Life Sciences 357542) and passed through a 70-μm strainer. Homogenates were underlaid with a 1.2 M sucrose cushion and centrifuged at 4,000 xg for 30 mins. Pelleted nuclei were washed with NF1 buffer at 1,600 xg for 5 mins. Nuclei were washed with staining buffer (PBS, 1 % bovine serum albumin, 1 mM EDTA) at 1,600 xg for 5 mins. Nuclei were suspended in 0.5 ml staining buffer and incubated overnight at 4^{o}C with the following antibodies: PU.1 Alexa Fluor 647 (1:500; BioLegend, 658004), NEUN Alexa Fluor 488 (1:2,500; Millipore Sigma, MAB377), OLIG2 (1:1,000; Abcam, ab109186) and LHX2 (1:500; Abcam, ab219983). Nuclei were washed with staining buffer. OLIG2 and LHX2 unconjugated antibodies were subsequently stained with goat anti-rabbit Alexa Fluor 647 (1:4,000; ThermoFisher Scientific, A21244) and washed with staining buffer. Nuclei were passed through a 30-μm strainer and stained with 0.5 μg ml^{-1} DAPI. Nuclei for each cell type were sorted with a Beckman Coulter MoFlo® Astrio™ EQ cell sorter and pelleted at 1,600 xg for 15 mins in staining buffer. Nuclei for ChIP-seq and PLAC-seq were stored at -80^{o}C prior to library preparation; nuclei for ATAC-seq were immediately processed.

Ex vivo microglia isolation

Microglia isolation was performed as previously described (*1*).

Chromatin immunoprecipitation-sequencing (ChIP–seq)

Nuclei (500,000 +/- 50,000) were thawed on ice, resuspended in 130 µl LB3 (10 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-Lauroylsarcosine, 1x protease inhibitors) and transferred to microtubes with an AFA Fiber (Covaris, MA). All subsequent steps were performed on ice or at 4°C. Samples were sonicated using a Covaris E220 focused-ultrasonicator (Covaris, MA) for 240 secs (Duty: 5, PIP: 140, Cycles: 200, AMP/Vel/Dwell: 0.0). Samples were adjusted to 250 µl with 1% Triton X-100, centrifuged at 21,000 xg for 10 mins and the pellet was discarded. 1 % of the sample was stored at -20°C for DNA input control. For ChIP, the following were added and rotated at 4°C overnight: 25 µl Protein G Dynabeads and either H3K27ac antibody (2 µl serum; Active Motif, 39135) or H3K4me3 antibody (2 µl; Millipore Sigma, 04-745). Dynabeads were washed 3 times with Wash Buffer 1 (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), three times with Wash Buffer 3 (10 mM Tris-HCl pH, 250 mM LiCl, 1 mM EDTA, 1% Triton X-100, 0.7% Na-Deoxycholate) and three times with TET (10 mM Tris-HCl pH 8, 1 mM EDTA, 0.2% Tween20). Dynabeads were washed once with TE-NaCl (10 mM Tris-HCl pH 8, 1 mM EDTA, 50 mM NaCl) in PCR tubes and resuspended in 25 µl TT (10 mM Tris-HCl pH 8, 0.05% Tween20). Input samples were adjusted to 25 µl with TT and libraries were generated in parallel with ChIP samples. Library NEBNext End Prep and Adaptor Ligation were performed using NEBNext Ultra II DNA Library Prep kit (New England BioLabs) according to manufacturer instructions using barcoded adapters (NextFlex, Bioo Scientific). Libraries were incubated with RNase A and Proteinase K at 55°C for 1 hour and 65°C for overnight. Libraries were PCR amplified for 14 cycles with NEBNext High Fidelity 2X PCR MasterMix (New England BioLabs, NEBM0541). Libraries were size selected for 225 – 350 bp fragments by gel extraction (10% TBE gels, Life Technologies) and were single-end sequenced for 51 cycles on an Illumina HiSeq 4000 (Illumina, San Diego, CA).

Assay for Transposase-Accessible Chromatin-sequencing (ATAC-seq)

We performed ATAC-seq on fixed nuclei samples using a published protocol with slight modifications (*2*). Immediately post-sorting, nuclei (200,000 +/- 20,000) were pelleted at 1,600 xg for 15 mins at 4°C. Pellets were resuspended in lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.5% IGEPAL CA-630) and centrifuged at 1,600 xg for 10 mins. Pellets were resuspended in 50 µl transposition mix (1x Tagmentation buffer (Illumina, 15027866), 2.5 µI Tagmentation DNA Enzyme I (Illumina, 15027865) in 50 µI reaction) and incubated at 37°C for 30 mins. Samples were incubated with 150 µl reverse-crosslinking solution (50 mM Tris-HCl pH 7.4, 200 mM NaCl, 1 mM EDTA, 1% SDS) at 65°C overnight. Samples were incubated with 1 µl Proteinase K (5 ng ml^{-1}) at 55°C for 2 hours. DNA was purified using ChIP DNA and Clean Concentrator kit (Zymo Research; D5205). Libraries were PCR amplified for 8-12 cycles with NEBNext High Fidelity 2X PCR MasterMix (New England BioLabs, NEBM0541) and 1.25 µM Nextera index primer 1 and 1.25 µM Nextera index primer 2-barcode. Libraries were size selected for 155 – 250 bp fragments by gel extraction (10% TBE gels, Life Technologies) and were single-end sequenced for 51 cycles on an Illumina HiSeq 4000 (Illumina, San Diego, CA).

Ribonucleic acid-sequencing (RNA-seq)

PSCs and PSC-derived cells (100,000) were collected in 600 µl Trizol and RNA was purified with Zymo kit (Zymo Research) and eluted in 50 µl water. mRNA was isolated by

incubating with 10 µl oligo(dT) beads in 50 µl 2x DTBB (20 mM Tris-HCl pH 7.5, 1 M LiCl, 2 mM EDTA, 1 % lithium dodecyl sulfate, 0.1 % triton X-100) at 65ºC for 2 mins, followed by an incubation at room temperature for 10 mins, 1 wash with Wash Buffer 1 (10 mM Tris-HCl pH 7.5, 0.15 M LiCl, 1 mM EDTA, 0.1 % lithium dodecyl sulfate, 0.1 % triton X-100) and 1 wash with Wash Buffer 3 (10mM Tris-HCl pH 7.5, 0.15 M NaCl, 1 mM EDTA). mRNA was eluted in 50 µl elution buffer (10mM Tris-HCl pH 7.5, 1mM EDTA) at 80ºC for 2 mins. mRNA isolation was repeated using fresh oligo(dT) beads and was subjected to a final wash with 1x SuperScript III first-strand buffer. mRNA was eluted from oligo(dT) beads and fragmented in 10 µl 2x SuperScript III at 94ºC for 9 mins. Isolated mRNA was used for first-strand cDNA synthesis as follows: 1.5 µg random primers (ThermoFisher Scientific, 48190011), 2 µM oligo d(T) (ThermoFisher Scientific,18418020), 10 U SUPERase-In (AM2696) and 0.8 mM dNTP were added to final volume of 12.5 µl; samples were incubated at 50ºC for 1 min; 0.2 µg Actinomycin D (Millipore Sigma, A1410), 5 mM DTT and 0.01% Tween20 and 100 U SuperScript III (ThermoFisher Scientific, 18080044) were added to final volume of 20 µl; samples were incubated at 25ºC for 10 mins and 50ºC for 50 mins and then on ice. Samples were purified using RNAClean XP beads (Beckman Coulter, A63987) and eluted with 10 µl water. Second-strand synthesis was performed as follows: 1X Blue buffer (Enzymatics, B0110), 0.66 mM PCR mix (Affymetrix, 77330), 0.66 mM dUTP (Affymatrix, 77206), 1 U RNase H (Enzymatics, Y9220L), 10 U DNA polymerase I (Enzymatics, P7050L) were added to a 15 µl final volume; samples were incubated at 16ºC for 2.5 hours. cDNA was purified using 2 µl SpeedBeads (ThermoFisher Scientific, 651520505025) in 28 µl 20% PEG8000/2.5 M NaCl and eluted with 40 µl EB (Zymo Research). Samples were subjected to dsDNA repair using 1x T4 ligase buffer, 0.2 mM dNTP, 0.9 U T4 DNA polymerase (Enzymatics, P7080L), 0.3 U Klenow Fragment (Enzymatics, P7060L), 3 U T4 PNK (Enzymatics, Y9040L) in a final volume of 50 µl and incubated at 20ºC for 30 mins. Samples were purified using 2 µl SpeedBeads in 93 µl 20% PEG8000/2.5 M NaCl and eluted with 15 µl EB. Samples were subjected to dA-tailing using 1x Blue buffer (Enzymatics, B0110), 0.1 mM dATP, 1.5 U Klenow 3'-5' exo- (Enzymatics, P7010-LC-L) in a final volume of 30 µl and incubated at 37ºC for 30 mins. Samples were purified using 2 µl SpeedBeads in 55.8 µl 20% PEG8000/2.5 M NaCl and eluted with 14 µl EB. Samples were subjected to Y-shape adapter ligation using 0.8 µM barcoded adapters (NextFlex, Bioo Scientific), 1x Rapid Ligation buffer (Enzymatics, B1010L), 300 U T4 DNA Ligase HC (Enzymatics, L6030-HC-L) in a final volume of 30 µl and incubated at 21ºC for 15 mins. Samples were purified using 2 µl SpeedBeads in 7.5 µl 20% PEG8000/2.5 M NaCl and eluted with 14 µl EB. Samples were second strand digested with 2 U UDG (Enzymatics, G5010L) at 37ºC for 30 mins. Libraries were PCR amplified for 12-16 cycles with the KAPA HiFi HotStart PCR kit (ThermoFisher Scientific, KK2502). Libraries were size selected for 200 – 475 bp fragments by gel extraction (10% TBE gels, Life Technologies) and were single-end sequenced for 51 cycles on an Illumina HiSeq 4000 (Illumina, San Diego, CA).

Proximity Ligation ChIP-sequencing (PLAC-seq)

PLAC-seq libraries were prepared for ex vivo microglia, NEUN nuclei and OLIG2 nuclei as previously described with minor modifications (*3, 4*). In brief, purified populations were cross-linked for 15 minutes at room temperature with 1% formaldehyde and quenched for 5 mins at room temperature with 0.2 M glycine (Thermo Fisher Scientific). The cross-linked cells/nuclei were centrifuged at 2,500 xg for 5 mins. To isolate nuclei, cross-linked cells were resuspended in 200 µL lysis buffer (10 mM Tris-HCl (pH 8.0), 10 mM NaCl, 0.2% IPEGAL CA-630) and

incubated on ice for 15 minutes. The suspension was then centrifuged at 2,500 xg for 5 mins and the pellet was washed by resuspending in 300 µL lysis buffer and centrifuging at 2,500 xg for 5 mins. The pellet was resuspended in 50 µL 0.5% SDS and incubated for 10 mins at 62°C. 160 µL 1.56% Triton X-100 was added to the suspension and incubated for 15 mins at 37°C. 25 µl of 10X NEBuffer 2 and 100 U MboI were added to digest chromatin for 2 hours at 37°C with rotation (1,000 rpm). Enzymes were inactivated by heating for 20 mins at 62°C. Fragmented ends were biotin labeled by adding 50 µL of a mix containing 0.3 mM biotin-14-dATP, 0.3 mM dCTP, 0.3 mM dTTP, 0.3 mM dGTP, and 0.8 U μl^{-1} Klenow and incubated for 60 mins at 37°C with rotation (900 rpm). Ends were subsequently ligated by adding a 900 µL master mix containing 120 µL 10X T4 DNA ligase buffer (NEB), 100 µL 10% TritionX-100, 6 µL 20 mg ml^{-1} BSA, 10 µL 400 U μl^{-1} T4 DNA Ligase (NEB, high concentration formula) and 664 µL H_2O and incubated for 120 mins at 23°C with 300 rpm slow rotation. Nuclei were pelleted for 5 mins at 4°C with centrifugation at 2,500 xg. For the ChIP, nuclei were resuspended in RIPA Buffer (10 mM Tris (pH 8.0), 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS, 0.1% sodium deoxycholate) with proteinase inhibitors and incubated on ice for 10 mins. Sonication was performed using a Covaris M220 instrument (Power 75W, duty factor 10%, cycle per bust 200, time 10 mins, temperature 7°C) and nuclei were spun for 15 mins at 14,000 rpm at 4°C. 5% of supernatant was taken as input DNA. To the remaining cell lysate was added anti-H3K4me3 antibody-coated Dynabeads M-280 Sheep anti-Rabbit IgG (5 µg antibody per sample, Millipore, 04-745), followed by rotation at 4°C overnight for immunoprecipitation. The sample was placed on a magnetic stand for 1 min and the beads were washed three times with RIPA buffer, two times with high-salt RIPA buffer (10 mM Tris pH 8.0, 300 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS, 0.1% deoxycholate), one time with LiCl buffer (10 mM Tris (pH 8.0), 250 mM LiCl, 1 mM EDTA, 0.5% IGEPAL CA-630, 0.1% sodium deoxycholate) and two times with TE buffer (10 mM Tris (pH 8.0), 0.1 mM EDTA). Washed beads were treated with 10 µg RNase A in extraction buffer (10 mM Tris (pH 8.0), 350 mM NaCl, 0.1 mM EDTA, 1% SDS) for 1 hour at 37°C, and subsequently 20 µg proteinase K was added at 65°C for 2 hours. ChIP DNA was purified with Zymo DNA clean & concentrator-5. For Biotin pull down, 25 µL of 10 mg ml^{-1} Dynabeads My One T1 Streptavidin beads was washed with 400 µl of 1X Tween Wash Buffer (5 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 1 M NaCl, 0.05% Tween) and supernatant removed after separation on a magnet. Beads were resuspended with 2X Binding Buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 2 M NaCl), added to the sample and incubated for 15 mins at room temperature. Beads were subsequently washed twice with 1X Tween Wash Buffer and in between heated on a thermomixer for 2 mins at 55°C with mixing and washed once with 1X NEB T4 DNA ligase buffer. Library prep was prepared using QIAseq Ultralow Input Library Kit (Qiagen, 180492). KAPA qPCR assay was performed to estimate concentration and cycle number for final PCR. Final PCR was directly amplified off the T1 beads according to the qPCR results, and DNA was size selected with 0.5X and 1X SPRI Cleanup and eluted in 1X Tris Buffer and paired-end sequenced.

Human pluripotent stem cell culture

All studies were conducted according to the human stem cell (hESCRO) protocol approved by the Salk Institute for Biological Studies. Human embryonic stem cell (ESC) line H1 (WiCell Research Institute, Madison, WI) (*5*) and induced pluripotent stem cell (iPSC) line EC11, derived from primary human umbilical vein endothelial cells (Lonza, Bioscience) (*6*), were cultured utilizing standard techniques. In brief, cells were cultured in StemMacs iPS-Brew media (Miltenyi Biotech, Auburn, CA) and routinely passaged utilizing Gentle Cell Dissociation Reagent

(STEMCELL Technologies) onto Matrigel-coated (1 mg ml^{-1}) plates. Karyotype was established by standard commercial karyotyping (WiCell Research Institute, Madison, WI).

CRISPR/Cas9 gene editing

CRISPR guides were designed utilizing ChopChop v2 (*7*). *BIN1* enhancer guides: GTAAGTCACTGGCTATGCATAGG and UAGAGUGCCUGCUCAAACGCAGG. ESC/iPSC CRISPR was performed as follows. Guide RNAs (sgRNA) were prepared by dissolving in nuclease-free TE buffer (IDT Technologies) at a final concentration of 100 μM, then complexed with Cas9 (NEB) at a ratio of 2:3 for 10 mins at room temperature. Pluripotent cells were harvested as single cells utilizing Tryple (Gibco). Cells were counted with live/dead discrimination using trypan blue (Sigma), and 150,000 cells were utilized per electroporation. Cells were electroporated utilizing an Amaxa electroporator (Lonza Biosciences) and the Human Stem Cells Nucleofector 2 Kit (Lonza). Each electroporation was plated into 1 well of a 6-well Matrigel-coated plate with 10 μM ROCK inhibitor. After recovery, cells were plated at a low cell concentration as single cells using Tryple for single colony isolation. Once single colonies were obtained, cells were incubated in collagenase IV (Invitrogen) until colonies were free-floating, and individual colonies were plated to a 96-well Matrigel coated plated with 10 μM ROCK inhibitor. Colonies were then screened for the enhancer deletion. Controls were isolated from the same sub-cloning experiment, which did not have an enhancer deletion. CRISPR deletion was confirmed utilizing the following primers outside of the enhancer deletion region. Fwd: CACGGGTGCAGATGGAAT and rev: AGGGCTGGTGTGAAGGTTA, which yielded a 605 bp product if unedited and approximately 242 bp product if the enhancer was deleted. Enhancer deletions were confirmed by isolating the edited product band and cloning into TOPO-TA (Invitrogen) and sequencing the resulting plasmid to confirm deletion. All of the individually isolated deletion lines: ESC (2) and iPSC (1) had the same 363bp deletion:
CCTATGCATAGCCAGTGACTTACGCTGACTTTTCTGGAAACTGGGACTTTAAATAGG
AGTTTTTAAAAAGGAAAGAAATCTCTGTTCTGCTTCTTAAAAACACCCTTTTCCCCTT
TTTACTTTCAGAAGAAGGAACCAGAGAAGACTAGAAACAGGATGGAACGGGAAGG
TGGAGGGAGAGGCGAAGGGGAAGTCTCGATGCTGACTACAAGGCCCCCTCATCCTC
CGGCCTCCGGCTTGCAGGGACTGCCCCTGCCTTGGGAGAGTCCTCACAGGCCACACC
TGCCACTGGGCCAGGCCCAGAGCAGACACTCACGACAGATGATAAGAATGACAGCT
GCCTGCGTTTGAGCAGGCACTCTA.

Pluripotent stem cell differentiation

Microglia were generated as previously described with minor modifications (*8*). Briefly, ESC/iPSCs were plated in iPS-Brew with 10 μM ROCK inhibitor (Stem Cell Technologies) onto Matrigel-coated (1 mg ml^{-1}) 6-well plates (Corning) using ReLeaSR (STEMCELL Technologies). Cells were differentiated to CD43+ hematopoietic progenitors using the StemCell Technologies Hematopoietic Kit (Cat #05310). On day 1, cells were changed to basal media with supplement A (1:200), supplemented with an additional 1 ml/well on day 3, and changed to basal media with supplement B (1:200) on day 3. Cells received an additional 1 ml/well of medium B on days 5, 7, and 10. Nonadherent hematopoietic cells were collected on days 11 and 13. Cells were then replated onto Matrigel-coated plates (1 mg ml^{-1}) at a density of 300,000 cells/well in microglia media. Microglia media consisted of DMEM/F12 (Thermofisher), 2x insulin-transferrin-selenite (Gibco), 2x B27 (Lifetech), 0.5x N2 (Lifetech), 1x Glutamax (Gibco), 2x non-essential amino acids (Gibco), 400 μM monothioglycerol, and 5 μg ml^{-1} insulin (Sigma). Microglia media was

supplemented with 100 ng ml^{-1} IL-34 (Proteintech), 50 ng ml^{-1} TGFβ1 (Proteintech) and 25 ng ml^{-1} M-CSF (Proteintech). Cells were supplemented with microglia media with IL-34, TGFβ1 and M-CSF, with removal of media (1 ml remaining) halfway through differentiation. 25 days after initiation with microglia media, cells were resuspended in microglia media with IL-34, MCSF and TGFβ1 with the addition of CD200 100 ng ml^{-1} (Novoprotein) and CX3CL1 100 ng ml^{-1} (Peprotech). Cells were collected on Day 28 for further experiments.

Astrocyte differentiations were performed as previously described through a glial progenitor cell (GPC) intermediate (*9*) with modifications. Briefly, embryoid bodies were prepared from confluent stem cell cultures by mechanical dissociation with 1 mg ml^{-1} collagenase IV (Invitrogen), plated onto ultra-low attached plates in iPS-Brew media supplemented with 10 μM ROCK inhibitor (Stem Cell Technologies) and incubated overnight with agitation. Cells were then transitioned to astrocyte medium (ScienCell) supplemented with 500 ng ml^{-1} Noggin (R&D) and 10 ng ml^{-1} PDGFAA (Peprotech) for a total of 14 days. Thereafter cells were cultured in astrocyte media with only PDGFAA for a further week. At the conclusion, embryoid bodies were dissociated with papain (Worthington) and the resulting GPCs were expanded on 10 μg ml^{-1} poly-L-ornithine- (Sigma) and 1 μg ml^{-1} laminin-coated (Invitrogen) plates in astrocyte media supplemented with 20 ng ml^{-1} FGF2 (Joint Protein Central) and 20 ng ml^{-1} EGF2 (Humanzyme). Astrocyte differentiation was completed utilizing a serum-free system (*10*). Specifically, GPCs were cultured for 4 weeks in Sato media, which consisted of a 1:1 mixture of DMEM (Thermo Fisher) and Neurobasal media (Thermo Fisher) supplemented with 1x PenStrep (ScienCell), 1 mM of sodium pyruvate (Sigma), 2 mM L-glutamine (Life Sciences), 5 μg ml^{-1} N-acetyl-L-cysteine (Sigma), 5 ng ml^{-1} EGF, 20 ng ml^{-1} CNTF (Proteintech), 20 ng ml^{-1} BMP4-4 (Peprotech), and 4 μg ml^{-1} insulin (Sigma). Sato media was further supplemented with 100 μg ml^{-1} of transferrin (Sigma), 100 μg ml^{-1} bovine serum albumin (Sigma), 16 μg ml^{-1} putrescine (Sigma), 0.2 μM progesterone, and 40 ng ml^{-1} sodium selenite (Sigma). Cultures were maintained on the aforementioned laminin and poly-L-lysine-coated plates throughout.

Neurons were differentiated directly from ESCs/iPSCs as previously described (*11*) with modifications. Briefly, ESC/iPSCs were lentivirally transduced to express Neurogenin 2 with a tetracycline regulatable system pLVX-UbC-rtTA-Ngn2:2A:EGFP (Addgene, #127288), subsequently cultures underwent puromycin selection (0.5 μg ml^{-1}; Gibco) to enrich for transduced cells. For neuronal conversion, cells were transferred to PluriPro Matrix (Cell Guidance Systems) coated plates as a monolayer in PluriPro media (Cell Guidance Systems) with 10 μM ROCK inhibitor and were induced 24 hours after plating with 2 μg ml^{-1} doxycycline (Sigma) for two days. On day 3, cells were transitioned to laminin- and poly-L-lysine-coated plates in neuronal maturation media. Neuronal maturation media consisted of DMEM/F12 1:1 mix with Neurobasal (Gibco both), 1x B27 supplement (Gibco), 1 μg ml^{-1} laminin (Life Technologies), 2 μg ml^{-1} doxycycline (Stem Cell Tech), 0.5 mM dbCAMP (Stem Cell Tech), 20 ng ml^{-1} GDNF (Stem Cell Tech) and 20 ng ml^{-1} BDNF (Stem Cell Tech). On days 4-5, cells were also supplemented with 10 μM Ara C (Sigma) to remove proliferating cells. Neurons were cultured for a total of 14-21 days.

Fluorescence activated cell sorting (FACS)

For neurons, cells were dissociated when mature with Tryple (Gibco) and sorted based on eGFP expression with Zombie Violet (Biolegend) for live/dead discrimination. For microglia, cells were mechanically dislodged from Matrigel plates and treated with FC receptor block (1:20,

Biolegend). Cells were then stained with the following 6 antibodies, all at 1:30 dilution and all from Biolegend: CD64-APC (305014), CX3CR1 PCP-Cy5.5 (341614), CD14-488 (325610), CD11b-PE (301309), HLADR PE-Cy7 (307616), CD45-APC-CY7 (368516) for one hour. Cells were then washed and incubated in Zombie Violet (Biolegend) for live/dead discrimination. Controls consisted of cells incubated with a combination of appropriate isotypes for each antibody (Biolegend). Flow cytometry was performed on a BD InFlux Cytometer (Becton-Dickinson).

Immunocytochemistry

Cell cultures were fixed with 4% paraformaldehyde solution for 15 mins at room temperature. Antigen blocking and cell permeabilization were performed using 10% horse serum and 0.1% Triton X-100 in PBS for 1 hour at room temperature. Primary antibodies were incubated in 10% horse serum overnight at 4°C, and secondary antibodies (1:250, Jackson Laboratories) were incubated in the same solution for 1 hour at room temperature. The cells were counterstained with DAPI for nuclei detection. The following primary antibodies were used: rabbit anti-S100β (1:100; Abcam, ab52642), chicken anti-GFAP monoclonal (1:1000; Abcam, ab5541), mouse βIII Tubulin (1:500, Biolegend, 801202), chicken Map2 a+b (1:1000, Abcam, ab5392), goat Iba-1 (1:1000, Novus, NB100-1028), rabbit PU.1 (1:200, Cell Signaling, 2266), mouse TRA 1-60 (1:200, EMD Millipore, MAB4360), mouse TRA 1-81 (1:200, EMD Millipore, MAB4381), goat Nanog (1:200, R&D Systems, AF1997), and rabbit SOX2 (1:500, Cell Signaling, 27482).

Western blotting

Western blotting was performed using standard procedures. The following antibodies were utilized: mouse GAPDH (1:10,000; Fitzgerald, 10R-G109A) and rabbit monoclonal BIN1 (1:5,000; Abcam, ab182562). All secondary antibodies were purchased from Jackson ImmunoResearch, mouse or rabbit HRP (1:5,000).

Imaging

Imaging data were captured and processed using a confocal microscope (Zeiss LSM780). Image processing was performed using Zen software (Zeiss). Western image quantification was performed using ImageJ (NIH). Flow cytometry images were processed using FloJo Software (FlowJo).

Data analysis

Data preprocessing

FASTQ-files for ATAC-seq and ChIP-seq were processed with the official ENCODE ATAC-seq (https://github.com/ENCODE-DCC/atac-seq-pipeline) and ChIP-seq (https://github.com/ENCODE-DCC/chip-seq-pipeline2) pipelines, respectively. Both pipelines were installed with a local anaconda environment. The Build_genome_data scripts were run to build the local hg19 reference genome and other dependencies. These pipelines execute a number of programs to process and align sequencing reads and to generate quality control statistics. In brief, FASTQ-files were trimmed with cutadapt (Version 1.9.1) and aligned with Bowtie2 (*12*) to hg19. SAMtools (*13*) (Version 1.2), MarkDuplicates (Picard Version 1.126), and bedtools (*14*) (Version 2.26) were used to filter and clean data post alignment. MACS2 (Version 2.1.1) was used for peak calling (*15*). The ATAC-seq pipeline also runs IDR (*16*) to detect peaks with high

reproducibility. The optimal peak sets that came out of these pipelines were used for downstream analysis.

RNA-seq data processing

FASTQ-files were mapped to the UCSC genome build hg19 with STAR (Version 2.5.3) using default parameters (*17*), and converted into HOMER tag directories (*18*). The HOMER function "analyzeRepeats" was used to quantify raw reads with the parameters "-raw -count exons -condenseGenes" and normalized reads, as either counts per million (CPM) values using "-cpm -count exons -condenseGenes", or as transcripts per million (TPM) values using "-tpm -count exons -condenseGenes." The base-2 logarithm of the TPM/CPM values was taken after adding a pseudocount of 1 TPM/CPM to each gene. For the iPSC experiments, differential analysis was done with the HOMER function "getDiffExpression" with the parameters "-repeats -DESeq2 -AvsA" with an FDR < 0.05 and $\log_2$ fold change > 1.

Enhancer and promoter calling

For each cell type, active promoters were identified by H3K4me3 peaks that overlap H3K27ac within 2,000 bp to a nearest transcription start site (TSS). Active enhancers were identified by H3K27ac peaks that were outside of H3K4me3 peaks.

LDSC regression analysis

To determine whether specific brain cell type annotations were enriched for heritability of specific neurological disorders, psychiatric disorders and neurobehavioral traits, stratified LD score regression was applied (*19, 20*). We obtained GWAS summary statistics for neurological disorders [AD (*21, 22*), amyotrophic lateral sclerosis (ALS) (*23*), multiple sclerosis (MS) (*24*), Parkinson's disease (PD) (*25*), and epilepsy (*26*)], psychiatric disorders [autism (*27*), major depressive disorder (MDD) (*28*), bipolar disorder (BD) (*29*), schizophrenia (SCZ) (*30*), and attention deficit hyperactivity disorder (ADHD) (*31*)] and neurobehavior traits [intelligence (*32*), cognitive function (*33*), risk behavior (*34*), neuroticism (*35*) and insomnia (*36*)]. Table S3 contains an overview of all the GWAS datasets that were used. The cell type-specific enhancer and promoter peak sets were tested for enrichment of heritability while controlling for the full baseline model. The corresponding P-values were FDR multiple testing corrected for the number of GWAS studies and number of cell type-specific annotations.

Quantification of ChIP-seq and ATAC-seq datasets

HOMER (*18*) (Version v4.9.1) was used to make tag directories of H3K27ac and H3K4me3 ChIP-seq, and ATAC-seq datasets belonging to each individual sample, for all combined samples per cell type, and for published ex vivo microglia datasets (*1*). Quantification of H3K27ac and H3K4me3 ChIP-seq and ATAC-seq signal at peaks merged for all cell types was performed using the HOMER annotatePeaks function. The HOMER annotatePeaks function was used to quantify H3K27ac reads around TSSs (+/- 2000 bp) for differential gene activation analysis.

Differential analyses

To identify gene promoters that showed increased activation in a given cell type, we performed a differential analysis on the H3K27ac at TSSs with Limma Voom (*37, 38*) (Version 3.34.9). A linear model was set up with four contrasts: each cell type versus the average of the

other cell types. Genes that were at least 4-fold induced with an FDR corrected P-value below 0.01 were combined and used to generate cell type-specific gene signatures. For genes with multiple TSSs, the TSS quantification with the highest number of reads was selected as a proxy for gene expression. These cell type-specific gene activation signatures were subsequently used for Metascape enrichment analyses (*39*).

Correlations across assays

RNA-seq FASTQ-files from different cell types of the brain were downloaded from SRA and processed as described above (*10*). Hippocampal, fetal, bulk whole-cortex, endothelial, and tumor-derived tissues were excluded. Spearman correlations between $\log_2$(CPM) of RNA-seq with H3K27ac and H3K4me3 ChIP-seq and ATAC-seq around TSSs were calculated for all replicate combinations. Correlations were classified as 'cell type correlations' when RNA-seq from one cell type was compared with either H3K27ac ChIP-seq, H3K4me3 ChIP-seq or ATAC-seq from the same cell type. Correlations were classified as 'aspecific correlations' when RNA-seq from one cell type was compared with either H3K27ac ChIP-seq, H3K4me3 ChIP-seq or ATAC-seq from a different cell type. The significance of the difference of these distributions of correlations was calculated with Kruskall-Wallis between group test.

Cell type purity analyses

Gene signatures for individual brain cell types were collected from PsychENCODE (*40*). For the 'inhibitory neurons' and 'excitatory neurons', all the gene signature sets corresponding to excitatory (Ex1-8) and inhibitory (In1-8) neurons were combined. The distributions of H3K27ac $\log_2$(CPM) values at TSSs for each pooled tag directory were compared between cell types and within cell types using Kruskall-Wallis between group test. The Dunn post-hoc test was applied to calculate the significance of each pairwise combination with a Bonferroni multiple testing correction. The distributions were visualized as a violin plot, with asterisk on the bottom that correspond to the Kruskall-Wallis between group p-values, and lines and asterisk between individual significant pairwise differences at the top. *** = $p < 1e\text{-}8$; ** = $p < 5e\text{-}5$, * $p < 5e\text{-}2$. Gene sets for microglia, monocytes, CNS-associated macrophages and dendritic cells were collected from a recent manuscript (*41*), and the distributions of H3K27ac signal at the promoters were compared to PU.1 nuclei as described above. Mouse genes were converted to human orthologues.

Motif analysis

To find transcription factors that were associated with each cell type, we performed a de novo motif analysis on accessible chromatin peaks within 500 bp of enhancers. Motif enrichments were performed with HOMER function findMotifsGenome. The output report links de novo motifs to known motifs and gives a motif matching score. Known motifs that matched to the de novo motif with a score higher than 0.7 were linked to transcription factors based on the human transcription factor catalog (*42*). The human transcription factor catalog was used to determine transcription factors that were associated with disease based upon proximal GWAS signal (*42*). N.C., a board-certified pediatrician and pediatric intensivist, manually curated this list for conditions that are related to the brain.

Gene Ontology (GO) Enrichment analysis

Metascape (http://metascape.org), with input species set to Homo sapiens, was used to identify GO terms and pathways that were overrepresented in particular gene sets (*39*) such as cell type-specific H3K27ac ChIP-seq at TSSs and cell type-specific PLAC-seq-linked genes. The resulting FDR corrected p-values were depicted as -log10(q) bar plots.

Fine mapping

Fine mapping was performed with PAINTOR Version 3 (*43*) and CAVIAR Version 2.2 (*44*) using AD GWAS variants from Kunkle Stage 1. Loci were selected for fine mapping that contained genome-wide significant SNPs (P < 5E-8), apart from the APOE and HLA loci which were excluded. Loci genomic coordinates were defined by the position of the outermost significant SNP within each locus. To account for LD structure at loci that were smaller than 20,000 bp, these small loci were extended to a minimum size of 20,000 bp. Fine mapping was performed using all the available genotyped SNPs within these selected loci regardless of p-values. Linkage Disequilibrium (LD) values between all SNPs in these loci (including non-significant SNPs) were calculated with PAINTOR function CalcLD_1KG_VCF, with the 1000G European reference cohort. The microglia H3K27ac peak set was used to guide fine mapping with PAINTOR, with an enumerate value of 1. CAVIAR was run on the same LD data with max causal-variants value of 1. The number of variants in the 95% credible set were 378 and 261 for CAVIAR and PAINTOR, respectively. All 261 SNPs in the 95% credible set from PAINTOR were also identified by CAVIAR and were used for downstream analysis.

Identification of genes PLAC-seq connected to AD-risk variants

Putative AD-risk genes were identified according to whether they had active promoters that overlapped with AD risk variants and/or had active promoters PLAC-seq connected to AD risk variants. Three separate sets of AD-risk variants were used for this analysis: (1) fine-mapped 95% credible set variants from stage 1 Kunkle (see 'Fine mapping' section), (2) variants from Kunkle Stage 2 meta-analysis with a genome-wide significance P < 5e-8 (*21*), and (3) variants from Jansen meta-analysis with a genome-wide significance P < 5e-8 (*22*). The analysis was performed in two steps. First, it was determined whether each variant intersected with an active promoter in each cell type, which were then termed 'promoters with AD-risk SNPs'. Second, each variant was extended by 2,500 bp in both directions to correspond with the PLAC-seq bin size of 5,000 bp and then intersected with PLAC-seq bins that had significant interactions in each cell type. If a variant-region overlapped one or more PLAC-seq bin, it was determined whether an active promoter was distally linked in the corresponding cell type, which were then termed 'Promoters PLAC-linked to AD-risk SNPs'.

STRING analysis

The Cytoscape (Version 3.7.2) plugin for the STRING app (Version 1.5.0) was used to generate protein-protein interaction networks (PPI) for each cell type using putative AD-risk genes identified as described in 'Identification of genes PLAC-seq connected to AD-risk variants'. The networks were overlaid with a color according to the number of times that the protein was connected to risk variants and was assigned a diamond shape if it was a fine mapped credible set SNP. Edges were colored to indicate different evidence codes used to connect proteins: purple for experimentally derived, turquoise for curated databases, black for co-expression, and grey for remaining evidence. The network was visualized using the 'Compound Spring Embedder' clustering algorithm. STRING enrichment analysis of the network was performed and the top 25

most significant GO Process terms across enrichment analyses for PPI networks for the three cell types were depicted as a heatmap.

PLAC-seq processing

FASTQ-files were trimmed with TrimGalore (version 0.4.5), which is a wrapper around Cutadapt (version 1.1.5). PLAC-seq data was processed with MAPS (*45*) (downloaded from github on October 4th, 2018). In short, MAPS aligned the trimmed FASTQ-files with BWA to a hg19 reference genome for the R1 and the R2 reads separately. Aligned reads were paired and valid pairs were used for downstream application. MAPS normalizes chromatin contact frequencies to detect significant chromatin interactions anchored at genomic regions at the merged H3K4me3 peaks (for Olig2, NeuN and PU.1) to identify long-range chromatin interactions at 5,000 bp resolution. Significant interactions were identified with FDR corrected p-value cutoff of 0.01. MAPS was run on all samples individually and on a merge of the samples that belong to the same cell type. The significant interactions for all samples belonging to each cell type were merged in a general interaction set, and for each individual sample these interactions were quantified. The quantified matrix of all significant interactions for all cell types was used as input for Limma differential analysis. A linear model was fit, with three pairwise contrasts. Interactions that were significantly increased in both comparisons of each cell type vs other cell types (FDR < 0.05, and $\log_2$ FC>2) were taken as cell type-specific increased interactions. Genes with active promoters and 2 PLAC-seq loops or more were used for Metascape analysis (*39*).

To classify the functional elements belonging to the loops and calculate the proportions, the interactions were classified in a hierarchical fashion for genomic annotations belonging to the same cell type. First, it was determined if both bins contained an active promoter (subsequently classified as a promoter-promoter (P-P) interaction). If not, it was determined if the loops contained an active promoter in one bin and an enhancer in the other bin (promoter-enhancer (P-E) interaction), or if the loop links a promoter to an ATAC-seq peak (promoter-ATAC (P-ATAC) interaction), or if the promoter was not linked to any peaks (promoter to none (P-none)). If there were no active promoters within either bin of the interactions, we determined if there was a TSS-distal H3K4me3 peak in either bin (classified as distal H3K4me3 bound interaction). Interactions without classifications were assigned 'unclear'.

UCSC genome browser

For ATAC-seq and ChIP-seq signals, the HOMER function makeMultiWigHub was used to generate UCSC track hubs from HOMER tag directories. For PLAC-seq interactions, significant MAPS interactions were converted into bigbed with UCSCtool bedToBigBed (*46*) with type = bed5+13 for the merged samples belonging to each cell type. UCSC genome browser (*47*) was used to generate genome browser shots of loci of interest.

Super-enhancers

Super-enhancers were identified using HOMER's findPeaks command on the pools of H3K27ac reads with pooled input reads as background with the parameters "-style super -L 1". As a means to orthogonally validate the super-enhancer and enhancer maps, active genes were classified according to whether they were linked to super-enhancers, regular enhancers or not linked to enhancers. The distribution of H3K27ac at the promoters were compared as described

above with Kruskall Wallis between group test and Dunn post-hoc Bonferroni correction. The data was visualized as a violin plot.

Heatmaps

Heatmaps were made in R with pheatmap (Version 1.0.8) (https://CRAN.R-project.org/package=pheatmap).

Circos plot

Circos is a circular visualization tool for genomic data that allows comparison of localized regions of interest that are separated by large regions of uninterest. The Circos plot was generated with BioCircos (*48*) (0.3.3) R Package.

Chow Ruskey plots

Chow-Ruskey plots are area proportional (or scaled) Venn diagrams using polygons to represent the intersection of more than 4 datasets. Chow-Ruskey plots were made with R package Vennerable (https://github.com/js229/Vennerable Version 3.1.0.9).

Additional visualizations

Violin plots, bar plots and donut plots (pie charts) were made with Seaborn (Version 0.8.1) and matplotlib (*49*) (Version 2.2.2). Violin plots illustrate the probability density (or distribution) of the data, as well as the median (white circle), the interquartile range (box), and the one standard deviation above and below the mean (whiskers).

Python and R environments

The Jupyter notebook was used for all analyses from Anaconda (Conda version 4.6.9) with python (version 3.6.5) and R (version 3.4.3). Statistical analyses were performed with Scipy (version 1.1.0) and Numpy (version 1.1.5) with Pandas libraries (version 0.23.4). Post-hoc tests were calculated with scikit-posthocs library (Version 0.6.1).

Supplementary Figure Legends

Fig S1. Cell type-specific isolation of nuclei populations from the brain. (A) Nuclei isolation strategy for generating cell type-specific datasets of the human brain. (B) Microglia, neuronal and oligodendrocyte populations were defined as $PU.1^+$, $NEUN^+$ and $OLIG2^+$ nuclei, respectively. The astrocyte population was defined as $NEUN^{neg}$ $LHX2^+$ nuclei. (C) Heatmap of Pearson's correlation of H3K27ac ChIP-seq log_2(CPM) in ex vivo microglia and nuclei of microglia, neurons, astrocytes and oligodendrocytes. (D) Heatmap of Pearson's correlation of ATAC-seq log_2(CPM) in ex vivo microglia and nuclei of microglia, neurons, astrocytes and oligodendrocytes. (E) Heatmap of Pearson's correlation of H3K4me3 ChIP-seq log_2(CPM) in ex vivo microglia and in nuclei of microglia, neurons, astrocytes and oligodendrocytes. (F) Violin plots depicting the distribution of Pearson's correlations of RNA-seq log_2(CPM) values with ATAC, H3K27ac and H3K4me3 log_2(CPM) values at gene promoters between replicates of the same cell types (cell type correlations) and between replicates of different cell types (aspecific correlations). (G) Violin plots (left) of brain nuclei populations showing log_2(CPM) values of H3K27ac signal at the promoters of cell type signature genes for microglia, excitatory and inhibitory neurons, astrocytes and oligodendrocytes as defined by PsychENCODE (*40*). Violin plot (middle) of oligodendrocyte nuclei showing log_2(CPM) values of H3K27ac signal at the promoters of cell type signature genes for OPCs and oligodendrocytes. Violin plot (right) of neuronal nuclei showing log_2(CPM) values of H3K27ac signal at the promoters of cell type signature genes for subclasses of excitatory and inhibitory neurons. Kruskal-Wallis-between group test and Dunn post-hoc test; bottom asterisks show Kruskall-Wallis between group p-values; upper asterisks show pairwise significant p-values; *** = p < 1e-8; ** = p < 5e-5, * p < 5e-2. (H) Violin plot of PU.1 nuclei showing log_2(CPM) values of H3K27ac signal at the promoters of genes elevated in CNS-associated macrophages (CAM), dendritic cells, microglia and monocytes (*41*). Kruskal-Wallis-between group test and Dunn post-hoc test; bottom asterisks show Kruskall-Wallis between group p-values; upper asterisks show pairwise significant p-values; * p < 5e-2. PCC, Pearson's correlation coefficient.

Fig. S2. Cell type-specific isolation of nuclei populations from the brain. (A) Heatmap of z-score normalized log_2(CPM) values of promoter H3K27ac signal for genes with cell type-enriched promoter H3K27ac versus other brain cell types (FDR < 0.01 and logFC > 2). (B) Heatmap of z-score normalized log_2(CPM) of ATAC-seq and H3K4me3 signal at gene promoters identified and shown in the same order as the H3K27ac heatmap in (A). (C) Heatmap of z-score normalized log_2(CPM) values of cell type-specific RNA-seq (*10*) at genes corresponding to promoters identified and shown in the same order as the H3K27ac heatmap in (A). (D) Metascape analyses of genes with differentially active promoters as identified in (A) depicted as $-log_{10}$(q)-values.

Fig. S3. Enrichment of de novo motifs at brain cell type enhancers. Significant de novo motifs identified by HOMER in regions of open chromatin at active enhancers in nuclei of microglia, neurons, astrocytes and oligodendrocytes. For each de novo motif is displayed the $-log_{10}$(P) values and the percentage of peaks containing the motif.

Fig. S4. Differential promoter H3K27 acetylation of human transcription factor genes in cell types of the brain. Heatmap of z-score normalized log_2(CPM) values of H3K27ac at promoters of the top 148 transcription factors with cell type differential H3K27ac signal (FDR < 0.01 and

logFC > 3) in microglia, neurons, astrocytes and oligodendrocytes using a compilation of 1639 human transcription factors (*42*).

Fig. S5. Motif enrichment at active genomic regulatory regions of brain cell types and corresponding candidate transcription factors. Tables show de novo motifs enriched at ATAC-seq peaks associated with H3K27AC-marked enhancers for the indicated brain cell types. Z-score normalized promoter H3K27ac tag counts, ex vivo microglia RNA expression log_2(transcripts per million (TPM)) (15) and brain cell type RNA expression log_2(TPM) (21) for transcription factors predicted to bind the corresponding motifs are shown as heat maps to the right. Transcription factors with an asterisk are associated with GWAS variants identified for brain disorders (Table S4). Motif enrichment at active genomic regulatory regions of brain cell types and corresponding candidate transcription factors. Tables show de novo motifs enriched at ATAC-seq peaks associated with H3K27AC-marked enhancers for the indicated brain cell types. Z-score normalized promoter H3K27ac tag counts, ex vivo microglia RNA expression log_2(transcripts per million (TPM)) (15) and brain cell type RNA expression log_2(TPM) (21) for transcription factors predicted to bind the corresponding motifs are shown as heat maps to the right. Transcription factors with an asterisk are associated with GWAS variants identified for brain disorders (Table S4).

Fig. S6. Chromatin interactions in cell types of the brain. (A) Donut plots (pie charts) represent the fraction of total PLAC-seq interactions located between promoters ('P') and the following gene regulatory regions: promoters ('P'), enhancers ('E'), open chromatin outside of enhancers (ATAC-seq) or regions absent of active enhancers or open chromatin ('none'). PLAC-seq interactions located at H3K4me3 regions not associated with promoters were classified as distal H3K4me3. (B) Pearson's correlation heatmap of PLAC-seq log_2(CPM) values for microglia, neuron and oligodendrocyte replicates; PCC, Pearson's correlation coefficient. (C) Scatterplots of H3K4me3 log_2(CPM) values (y-axis) and PLAC-seq interaction counts (x-axis) for microglia (top), neurons (middle) and oligodendrocytes (bottom). Each data point represents a gene, the mean effect size is depicted in the color gradient of each data point and the cumulative log10(q-values) is depicted as the size of each data point. (D) Heatmap showing z-score normalized log_2(counts per million (CPM)) values of PLAC interactions upregulated in microglia, neurons and oligodendrocytes.

Fig. S7. Chromatin interactions in cell types of the brain. (A) Table of super-enhancers linked to cell type-specific genes by PLAC-seq interactions. Asterisks denotes genes with PLAC-seq interactions to super-enhancers that colocalize with disease-risk variants identified by GWAS shown in Fig. 1F (Table S6); SE, super-enhancer. (B) Violin plot showing log_2(CPM) values of promoter-H3K27ac signal from microglia (left), neurons (middle) and oligodendrocytes (right) at active promoters that are 'not linked to enhancers', 'linked to enhancers' and 'linked to SEs' by PLAC-seq. Kruskal-Wallis-between group test and Dunn post-hoc test; bottom asterisks show Kruskall-Wallis between group p-values; upper asterisks show pairwise significant p-values; *** p < 1e-11. SE, super-enhancer.

Fig. S8. Table of GWAS-identified regions and extended PLAC-seq linked genes assigned to AD risk loci. (A-C) Table of genes that have either (1) promoters with AD-risk variants, (2) promoters PLAC-linked to AD-risk variants or (3) promoters with AD-risk variants and PLAC-linked to AD-risk variants. The cell type(s) with active promoters for each gene is indicated for microglia (M), neurons (N) and oligodendrocytes (O). (A) Data analysis using fine mapped Kunkle

GWAS variants. (B) Data analysis using Kunkle GWAS stage 2 meta-analysis variants with p-value < 5e-8 (*21*). (C) Data analysis using Jansen GWAS meta-analysis variants with p-value < 5e-8 (29).

Fig. S9. Chromatin interactions at AD loci for cell types of the brain. (A) Chow-Ruskey plot of genes that are GWAS-assigned and PLAC-seq linked to Kunkle AD-risk variants (p < 5e-8) in microglia, neurons and oligodendrocytes (*21*). (B) Chow-Ruskey plot of genes that are GWAS-assigned and PLAC-seq linked to Jansen AD-risk variants (p < 5e-8) in microglia, neurons and oligodendrocytes (*22*). (C, D, E) STRING Protein-protein interaction (PPI) network of microglia (C), neuron (D) and oligodendrocyte (E) active genes associated with AD-risk. The node color corresponds to the number of GWAS gene-sets that the gene is linked to AD variant (yellow = 3, dark yellow = 2, grey = 1). The node shapes correspond to whether the gene is linked to a credible set variant (oval = not linked to credible set variant, diamond = linked to credible set variant). The edge color corresponds to different evidence codes: purple for experimentally derived, turquoise for curated databases, black for co-expression, and grey for remaining evidence codes. (F) Heatmap of STRING gene ontology analyses of AD-risk genes identified in microglia, neurons and oligodendrocytes showing -$\log_{10}$(q) values (red).

Fig. S10. Deletion of a microglia enhancer harboring a lead AD-risk variant upstream of the *BIN1* locus. (A) Schematic of microglia-specific BIN1 enhancer deletion in PSCs harboring the lead rs6733839 AD-risk variant. Control and enhancer deletion PSCs were differentiated into astrocytes, microglia and neurons. (B) PCR amplification of the *BIN1* enhancer region in control and *BIN1*$^{enh\_del}$ PSCs delineating the 363bp enhancer deletion. (C) UCSC genome browser visualization of an enhancer at the *BIN1* locus that overlaps the rs6733839 AD-risk variant shows ATAC-seq and H3K27ac ChIP-seq signal restricted to microglia. A 363 bp CRISPR/Cas9-deleted region is shown as a black bar (highlighted yellow). (D) All control and *BIN1*$^{enh\_del}$ PSCs retained a normal karyotype.

Fig. S11. Control and *BIN1*$^{enh\_del}$ PSCs and derived microglia, neurons and astrocytes express typical cell lineage markers. Representative immunohistochemistry images of PSCs (A), microglia (B), neurons (C) and astrocytes (D) in control and *BIN1*$^{enh\_del}$ lines stained for the indicated cell lineage markers.

Fig. S12. Deletion of a microglia enhancer harboring a lead AD-risk variant affects *BIN1* expression in PSC-derived microglia. (A) Control and *BIN1*$^{enh\_del}$ neurons were sorted for expression of Neurogenin2 with 2A peptide-linked eGFP (B) Control and *BIN1*$^{enh\_del}$ microglia were sorted for expression of CD11B^{high}, CD45^{intermdetiate}, CX3CR1^{+}, CD64^{+}, CD14^{+}, and HLADR^{+}. (C) Principal component analysis of RNA-seq $\log_2$(TPM) for control and *BIN1*$^{enh\_del}$ PSCs, microglia, neurons and astrocytes. (D) Heatmap of RNA-seq $\log_2$(TPM) for control and *BIN1*$^{enh\_del}$ PSCs, microglia, neurons and astrocytes. (E) UCSC browser visualization of RNA-seq tag counts at the *BIN1* locus for control and *BIN1*$^{enh\_del}$ PSCs, microglia, neurons and astrocytes. (F) Western blot of BIN1 and GAPDH in control and *BIN1*$^{enh\_del}$ PSCs (top), and derived astrocytes (middle) and hematopoietic progenitors (bottom).

Supplementary Tables (tables published separately online in Excel format)

Table S1. Patient information and assays. Table summarizing patient information including age, sex, ethnicity, brain region and pathology diagnosis. Table summarizes isolation method for either ex vivo cells or nuclei, cell type of origin and assays performed.

Table S2. Differential promoter H3K27ac analysis. Table with results of differential H3K27ac around TSS. Columns represent log_2FC of each cell type versus the other cell types and FDR multiple testing corrected p-values. The Boolean 'is_TF' column represents whether a gene is annotated as a transcription factor.

Table S3. Summary of GWAS studies. Table summarizing GWAS studies used for LDSC regression analysis in this study. We obtained GWAS summary statistics for neurological disorders [AD (*21, 22*), amyotrophic lateral sclerosis (ALS) (*23*), multiple sclerosis (MS) (*24*), Parkinson's disease (PD) (*25*), and epilepsy (*26*)], psychiatric disorders [autism (*27*), major depressive disorder (MDD) (*28*), bipolar disorder (BD) (*29*), schizophrenia (SCZ) (*30*), and attention deficit hyperactivity disorder (ADHD) (*31*)] and neurobehavior traits [intelligence (*32*), cognitive function (*33*), risk behavior (*34*), neuroticism (*35*) and insomnia (*36*)].

Table S4. Cell type-specific transcription factors and disease association. Table summarizing a subset of cell type-specific transcription factors that were paired with enhancer motifs of the corresponding cell type. Table includes information on diseases associated with transcription factors based upon proximal GWAS signal (*42*), which has been subset for conditions that are related to the brain.

Table S5. PLAC-seq promoter interactome map. Promoter centric interactome map that links promoters active in one or multiple cell types to distal regulatory regions that contain one or multiple active promoters and or active enhancers.

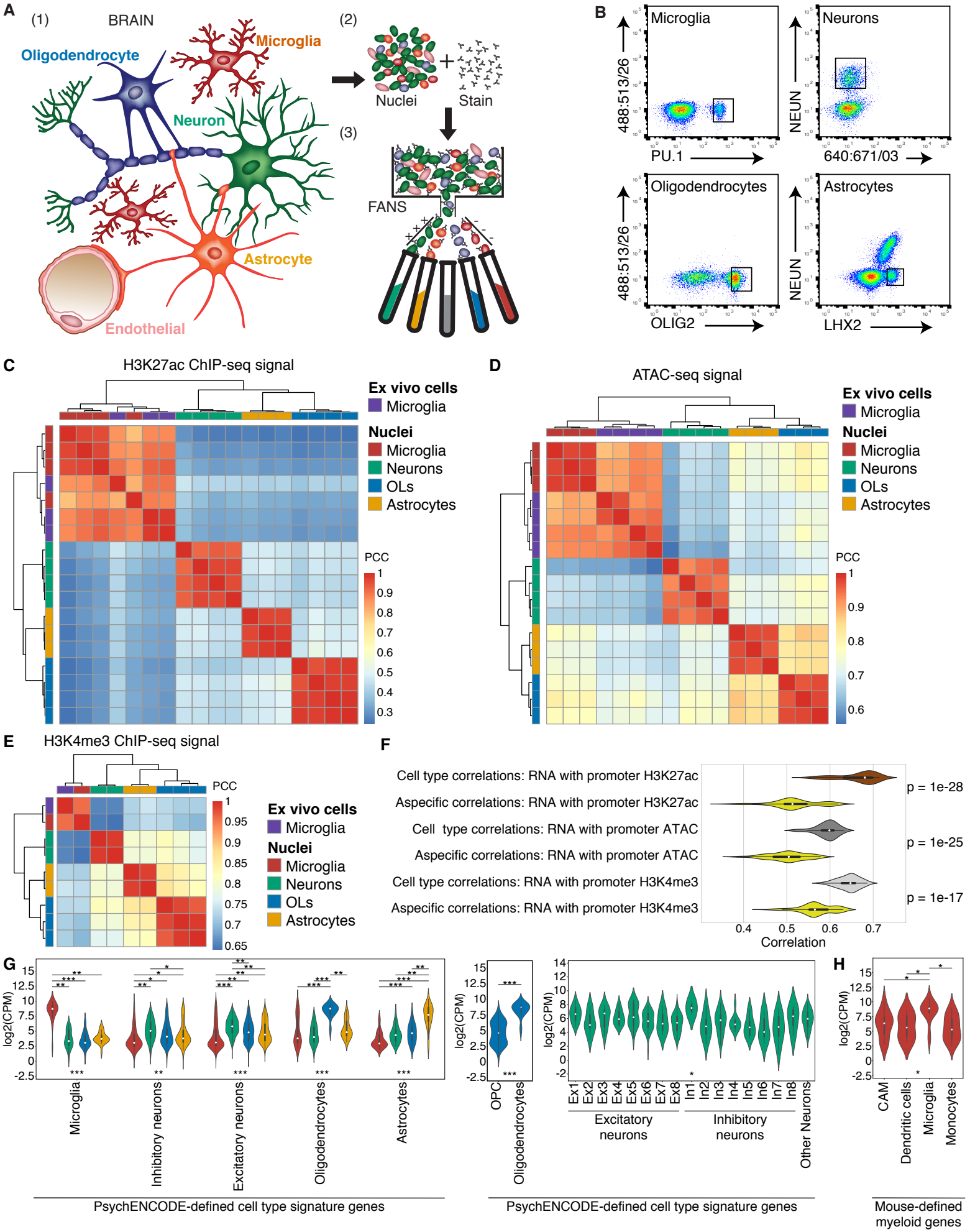
Table S6. PLAC-seq interactions at super enhancers. Table summarizing PLAC-seq promoter interactions with super enhancers for each cell type, including information of GWAS variants localized to PLAC-seq linked super-enhancers. The Boolean columns indicate whether the super-enhancer overlaps with GWAS variants for neurodegenerative diseases (AD, amyotrophic lateral sclerosis, epilepsy and autism), psychiatric conditions (major depressive disorder, schizophrenia), and neurobehavioral traits (intelligence, cognitive function, risk behavior, neuroticism and insomnia). The numerical columns indicate the number of GWAS variants that overlap within a super-enhancer for each condition.

Table S7. Active gene promoters PLAC-seq linked to AD variants. Table of the analyses linking AD GWAS variants to active genes for the Kunkle credible set SNPs, the Kunkle meta-analysis SNPs and the Jansen meta-analysis SNPs. For each analysis we have provided both a summary and detailed results page. The summary page contains the following information for each locus: start and end position of the locus, GWAS-assigned gene, active gene promoters that overlap AD-risk SNPs and active gene promoters that are PLAC-seq linked to AD-risk SNPs. The detailed results page contains the following additional information: variant IDs, fine mapped posterior

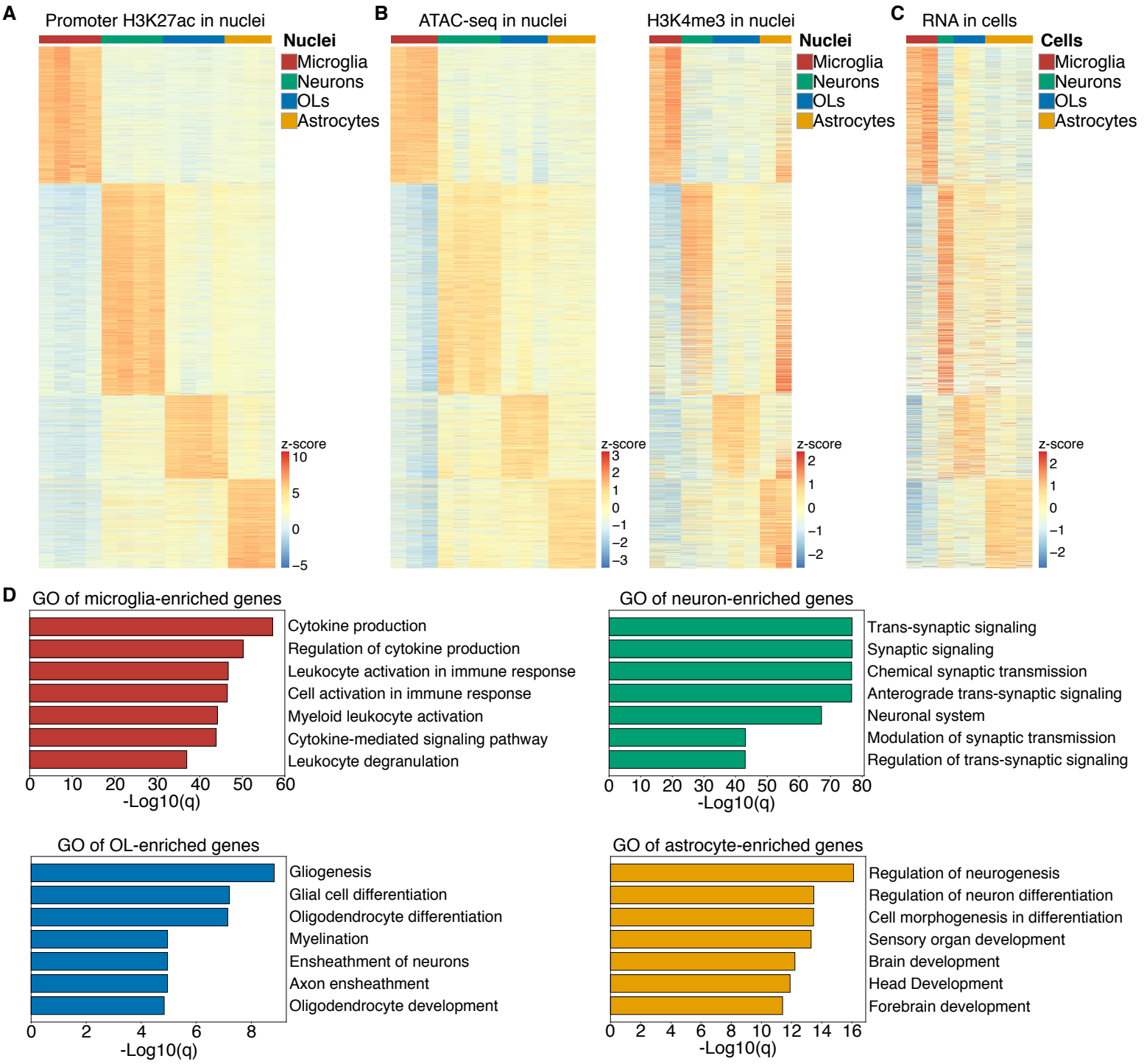
probabilities, GWAS summary stats, overlap of variants with active enhancers and promoters, genes associated with active promoters, and PLAC-seq linked genes.

Table S8. Differential gene expression of *BIN1* enhancer deletion lines. Table provides DESeq2 output as rlog transformed data of RNA-seq data for PSCs and PSC-derived microglia, neurons and astrocytes for each gene. Table includes $\log_2$ fold change values, p-values and adjusted p-values for each gene between control and *BIN1* enhancer deletion groups for PSCs and PSC-derived microglia, neurons and astrocytes.

Supplementary Figure 1

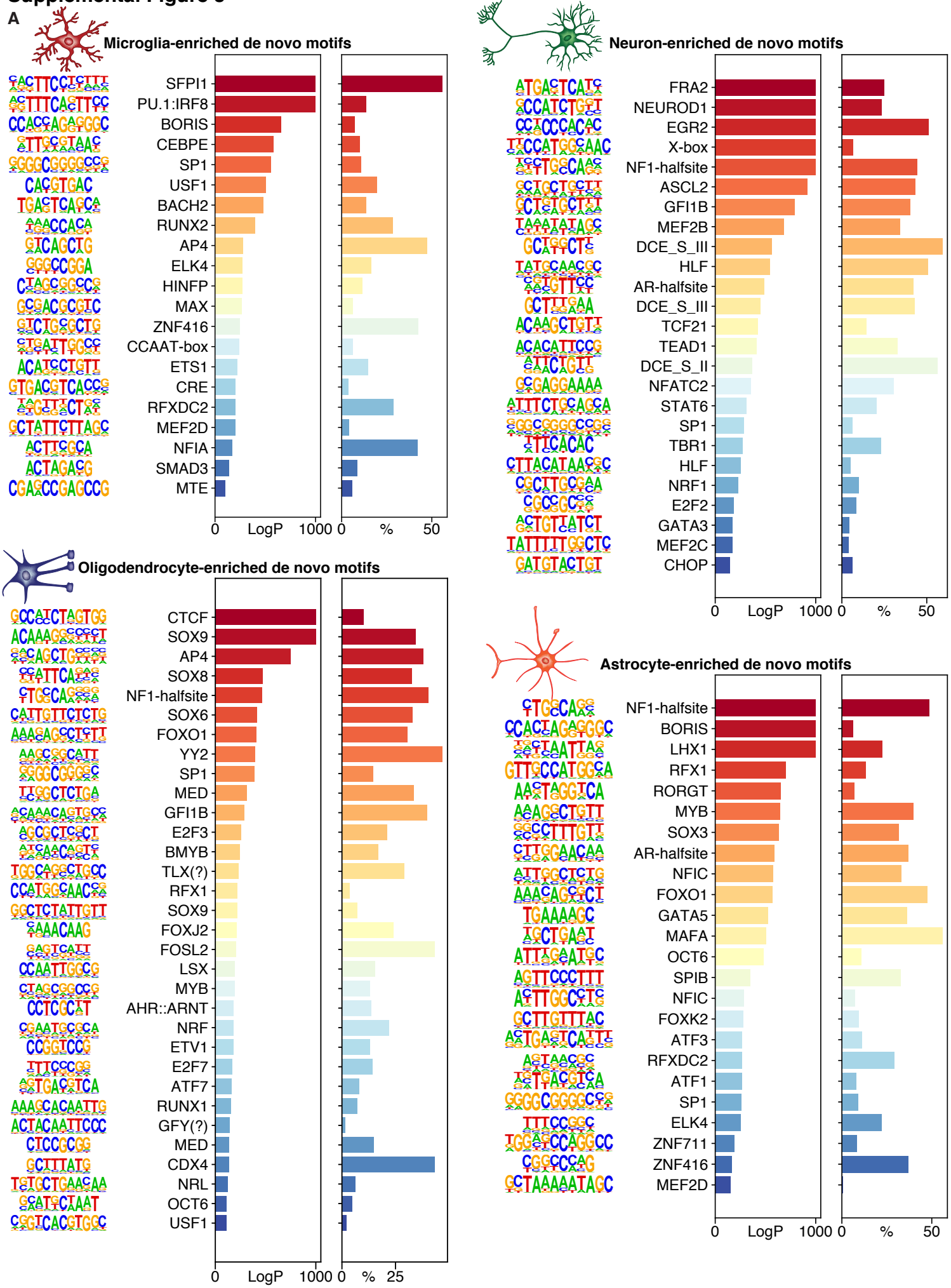


Supplementary Figure 2



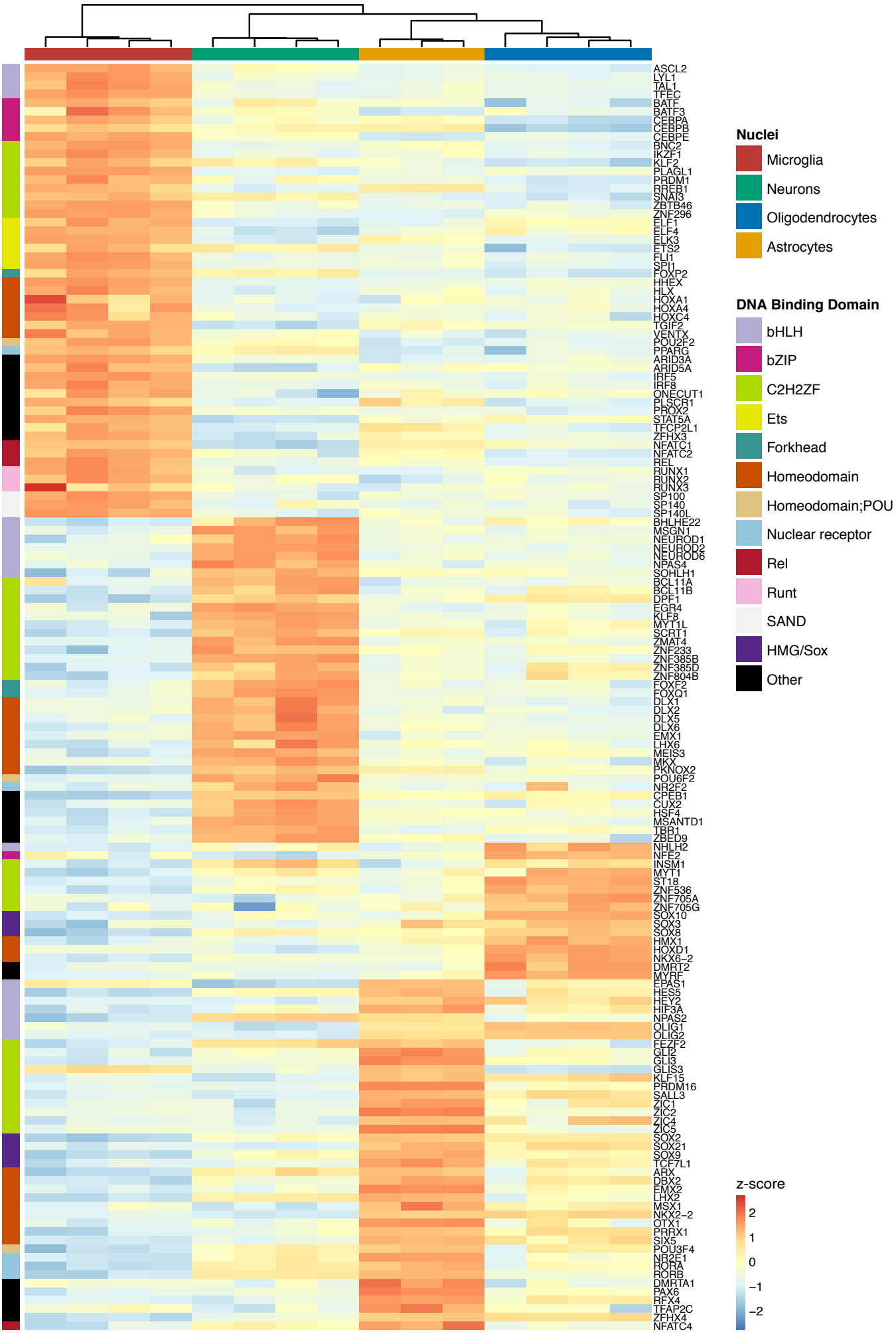
Supplemental Figure 3

A

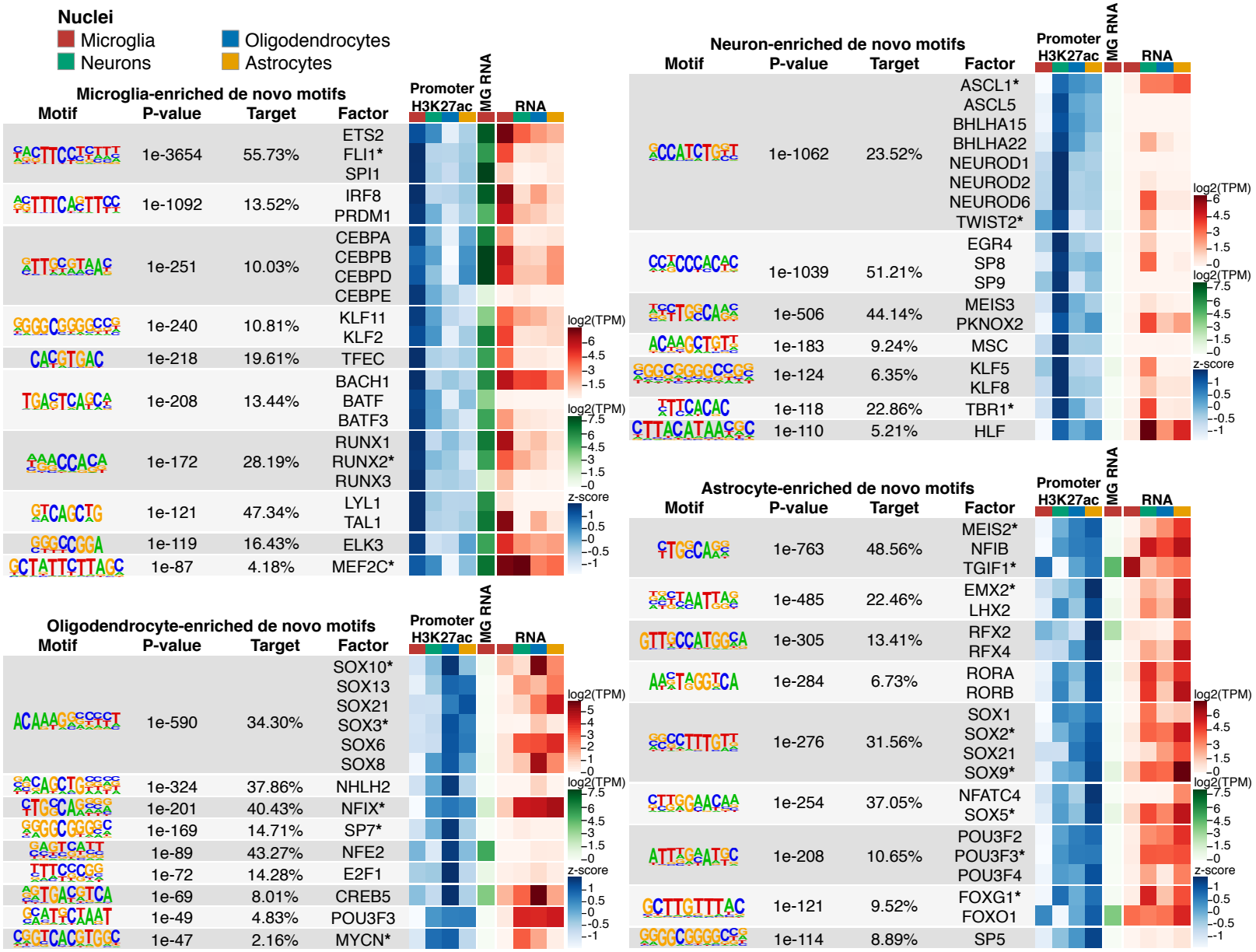


Supplemental Figure 4

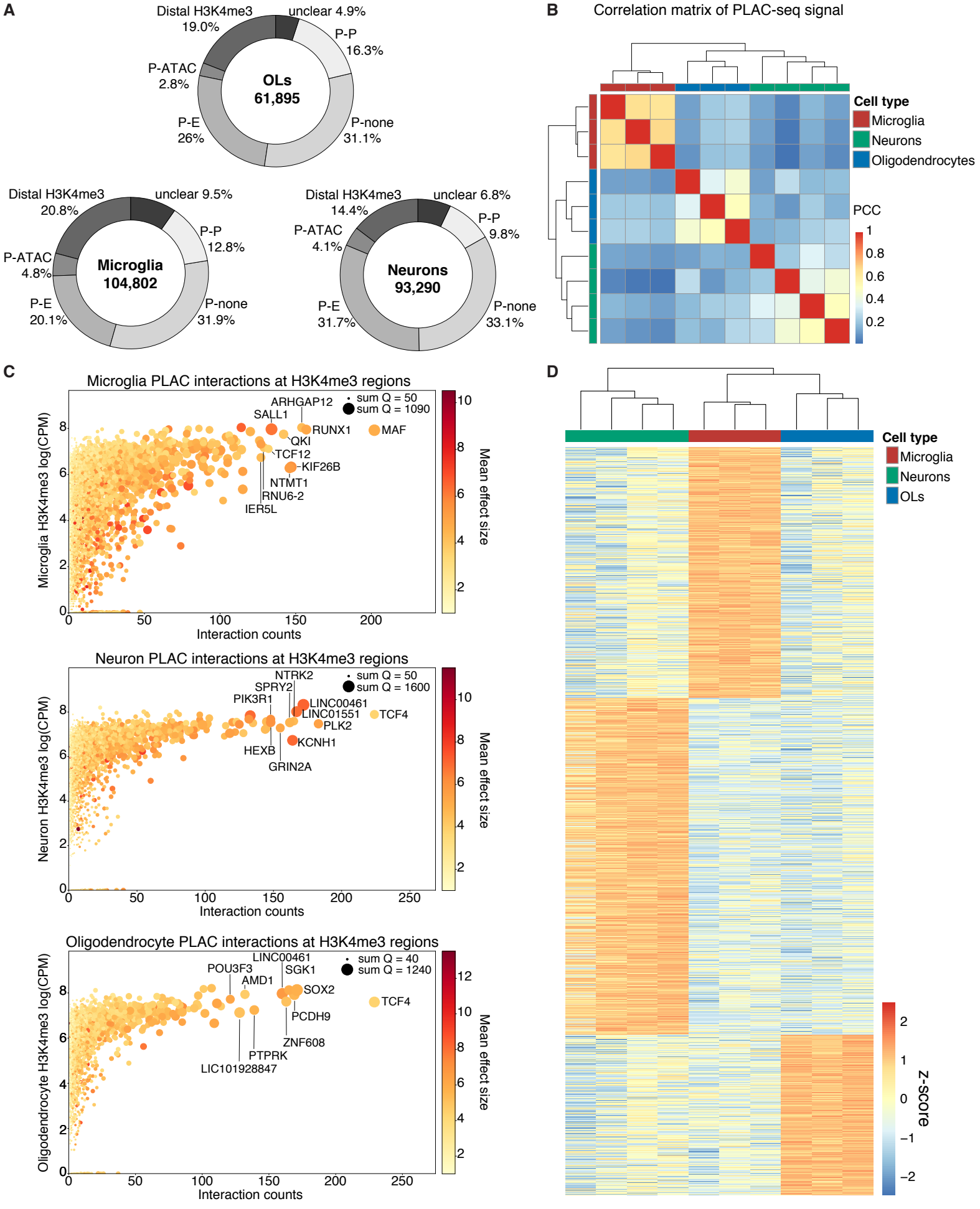
A



Supplemental Figure 5



Supplementary Figure 6

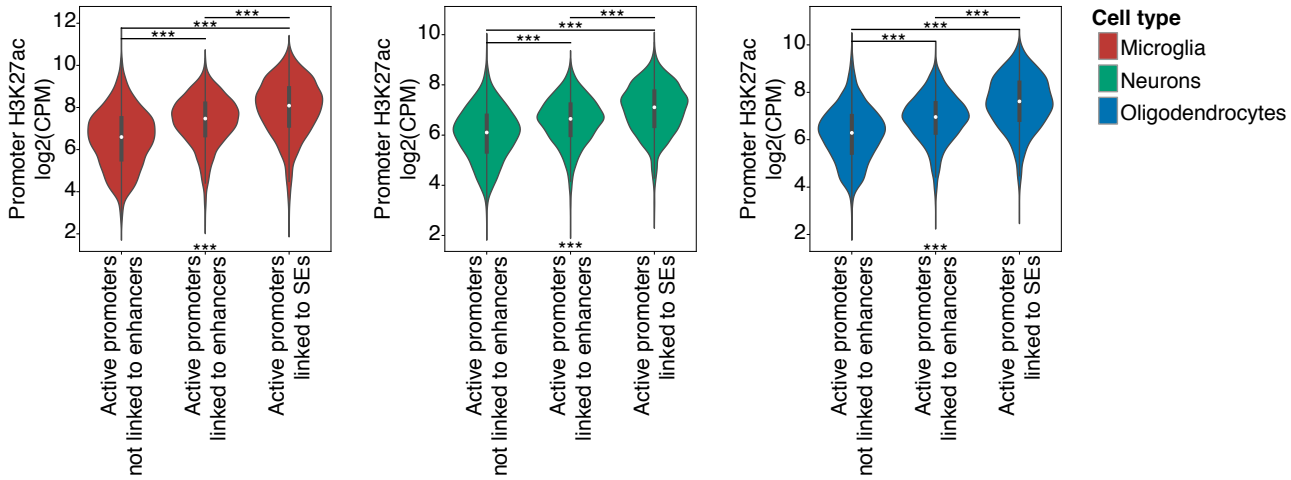


Supplementary Figure 7

A

Source	SEs	Percent of SE with PLAC links	Percent of enhancers with PLAC links	Example of cell type-specific genes PLAC-linked to SEs
Microglia	631	88%	59%	**TFs**:*CEBPA,EGR2,ETS2,FLI1,FOS,IRF8,MEF2C*,RUNX1,RUNX2,SPI1,STAT6** **Receptors**:*CD14*,CSF1R,CX3CR1,ITGAM*,ITGAX*,P2RY12* **Others**:*AIF1*,ALOX5,APOE*,B2M,CASS4*,IL6,INPP5D*,PICALM*,TGFB1,TNF**
Neurons	1170	81%	34%	**TFs**:*ASCL5,DLX1,EMX1*,MYT1L,NEUROD2,NR4A1,SATB2*,TBR1** **Receptors**:*CHRNB2,EPHB6,GRIN2A*,GRIN2B,GABRA1,GABBR2** **Others**:*ARC,CACNA1A,CACNA1B,CAMK2B,KCNB1*,NGEF*,NRXN3,SLC4A10*,SYT1**
Oligodendrocytes	1153	81%	41%	**TFs**:*NKX2-2,NFIX*,OLIG1,OLIG2,POU3F3,SOX8,SOX10* **Membrane proteins**:*FGFR2,GPR37*,GPR37L1*,GRID1,P2RX7* **Others**:*ABCA8,ELMO1,MOB3B,MOBP,MOG*,SEMA3B**

B



Supplementary Figure 8

A Summary table for Kunkle fine mapped credible set variants

Locus	Chr	Gene	Start	End	Promoters with AD-risk SNPs	Promoters PLAC-linked to AD-risk SNPs	Promoters with AD-risk SNPs and PLAC-linked to AD-risk SNPs
1	1	CR1	207679455	207850485		PFKFB2(O),YOD1(O)	
2	2	BIN1	127826632	127895415		BIN1(M,O)	
3	6	TREM2	40706556	41128693		TREM2(M),LRFN2 (N)	
4	6	CD2AP	47391648	47594722	CD2AP(M/O)		CD2AP(N)
5	7	EPHA1	143094496	143114496		LOC100507507(O),ZYX (O)	
6	8	PTK2B	27195263	27226756			
7	8	CLU	27451550	27471550	CLU(O/N)	PTK2B(M),TRIM35(M),CCDC25(O/N), ESCO2(O/N),SCARA3(O)	
8	10	ECHDC3	11711114	11731114		USP6NL(M)	
9	11	SPI1	47277094	47915038	MYBPC3(O/N)	ACP2(M),DDB2(M),MADD(M),NR1H3(M), ARFGAP2(O),C1QTNF4(O),CELF1(O), KBTBD4(O/N),NDUFS3(O/N),PACSIN3(O)	SLC39A13(M/O/N),SPI1(M), PSMC3(M/O/N)
10	11	PICALM	85654274	85869737		PICALM(M/O)	
11	11	SORL1	121425562	121445562		SC5D(M/N),SORL1(M/N)	
12	11	MS4A2	59831716	60103112		MS4A7(M)	
13	14	SLC24A4	92929076	92949076		ATXN3(M),CPSF2(M),NDUFB1(M),TRIP11(M)	
14	19	ABCA7	1039509	1063306	ABCA7(O/N)	ARHGAP45(M),ATP5F1D(M),STK11(M)	ABCA7(M)

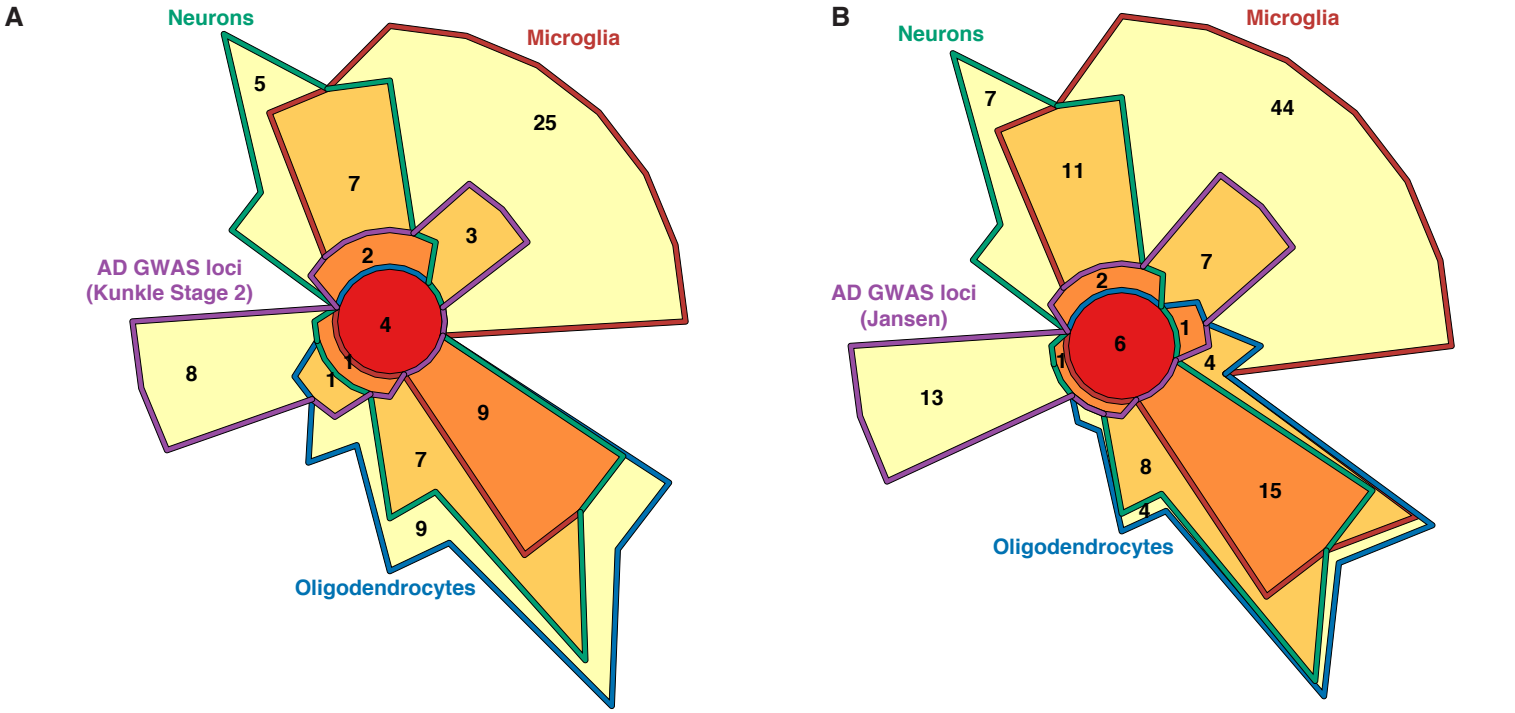
B Summary table for Kunkle significant SNPs identified by stage 2 meta analysis

Locus	Chr	Gene	Start	End	Promoters with AD-risk SNPs	Promoters PLAC-linked to AD-risk SNPs	Promoters with AD-risk SNPs and PLAC-linked to AD-risk SNPs
1	1	CR1	207653395	207806730		PFKFB2(N),YOD1(N)	
2	2	BIN1	127793239	127895573	BIN1(N)		BIN1(M,O)
3	2	INPP5D	233981912	233981912		ATG16L1(M),INPP5D(M)	
4	6	HLA -DRB1	32224388	32684387		HLA-DQA1(M),HLA-DQB1(M),HLA-DRA(M),HLA-DRB1(M), HLA-DRB5(M),PSMB8(M),PSMB9(M/N),TAP1(M/N)	
5	6	CD2AP	47379843	47633098	CD2AP(M/O)		CD2AP(N)
6	7	NYAP1	99971313	100091795		AGFG2(O),C7orf61(O),FBXO24(O), MEPCE(M),LRCH4(O),NYAP1(O), TSC22D4(O),ZCWPW1(M)	PILRA(M)
7	7	EPHA1	143099107	143131854		LOC100507507(O),ZYX(O)	
8	8	PTK2B	27195121	27337045		PTK2B(M)	CHRNA2(N),CLU(N)
9	8	CLU	27430506	27467686	CLU(O)	CCDC25(O/N),EPHX2(N),ESCO2(O/N), PTK2B(N),TRIM35(M/N),SCARA3(O)	
10	10	ECHDC3	11720308	11720308		USP6NL(M)	
11	11	SPI1	47232038	47910823	MTCH2(M/N),MYBPC3(O/N), NUP160(M/O/N),SLC39A13(O)	ACP2(M),ARFGAP2(O/N),CELF1(M/O/N), C1QTNF4(O/N),DDB2(M/N),FNBP4(O/N),KBTBD4(M/O/N), MADD(M/N),NDUFS3(M/O/N),NR1H3(M),PACSIN3(O)	MTCH2(O),PSMC3(M/O/N), SLC39A13(M/N),SPI1(M)
12	11	PICALM	85652826	85871214	PICALM(N)	CCDC83(M),CREBZF(M),EED(M/O/N),HIKESHI(M/N)	PICALM(M/O)
13	11	SORL1	121435587	121435587		SC5D(M/N),SORL1(M/N)	
14	11	MS4A2	59831716	60103385		MS4A7(M)	
15	14	FERMT2	53390015	53400629			
16	14	SLC24A4	92926952	92934269		ATXN3(M),CPSF2(M),NDUFB1(M),TRIP11(M)	
17	17	ACE	61503610	61538148		DCAF7(N),MAP3K3(O/N),TACO1(M/O/N)	
18	19	ABCA7	1028149	1075979	ABCA7(O/N)	ARHGAP45(M),ATP5F1D(M),GPX4(M), MIDN(M),POLR2E(M),STK11(M)	ABCA7(M)
19	20	CASS4	54997568	54997568			

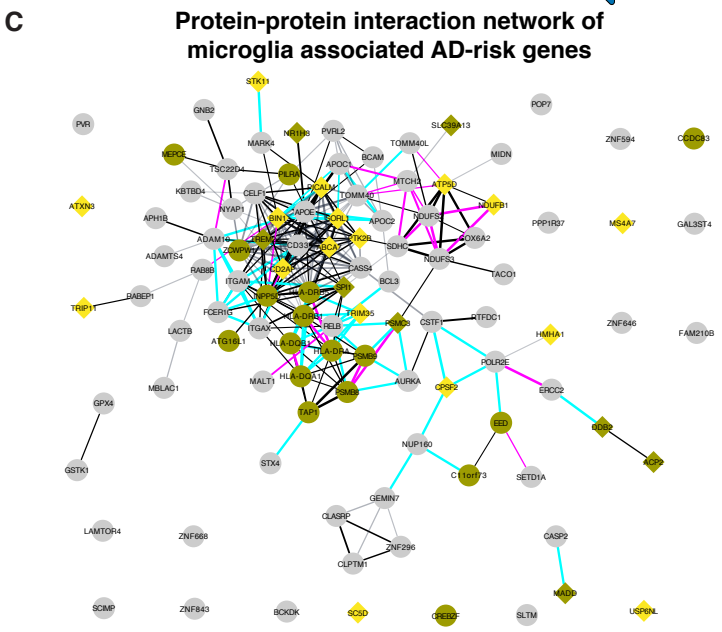
C Summary table for Jansen meta analysis significant SNPs

Locus	Chr	Gene	Start	End	Promoters with AD-risk SNPs	Promoters PLAC-linked to AD-risk SNPs	Promoters with AD-risk SNPs and PLAC-linked to AD-risk SNPs
1	1	ADAMTS4	161155392	161156033		ADAMTS4(M/N),FCER1G(M),NDUFS2(M/N), SDHC(M),TOMM40L(M)	TOMM40(M)
2	1	CR1	207679307	207850539		PFKFB2(N/O),YOD1(N/O)	
3	2	BIN1	127826533	127895487	BIN1(N)		BIN1(M/O)
4	2	INPPD5	233981912	234082384		ATG16L1(M),INPP5D(M),NGEF(N)	
5	3	HESX1	57226150	57226150			
6	4	CLNK	11022799	11040406			
7	6	HLA-DRB1	32340070	32682429		HLA-DQB1(M),HLA-DRA(M),HLA-DRB1(M), HLA-DRB5(M),PSMB8(M),PSMB9(M/N),TAP1(M/N)	HLA-DQA1(M)
8	6	TREM2	40904030	41129252	TREM2(M)		
9	6	CD2AP	47427281	47590476	CD2AP(M/O)		CD2AP(N)
10	7	ZCWPW1	99774122	100175473		AGFG2(O),C7orf61(O),FBXO24(O),GAL3ST4(M),GNB2(M), LRCH4(O),LAMTOR4(M),MBLAC1(M),MEPCE(M), NYAP1(M/N/O),POP7(M),TSC22D4(M/N/O),ZCWPW1(M)	PILRA(M)
11	7	EPHA1	143083661	143124782		CASP2(M),GSTK1(M),LOC100507507(N/O),ZYX(N/O)	
12	7	CNTNAP2	145793482	146047392			
13	8	CLU/PTK2B	27195121	27494366		CCDC25(N/O),ESCO2(N/O),PTK2B(M/N), SCARA3(N/O),TRIM35(M/N)	CHRNA2(N),CLU(N/O)
14	10	ECHDC3	11717397	11720308		USP6NL(M)	
15	11	SORL1	121435587	121469981		SC5D(M/N/O),SORL1(M/N/O)	
16	11	MS4A6A	59826677	60103385		MS4A7(M)	
17	11	PICALM	85652251	85869737	PICALM(N)	CCDC83(M),CREBZF(M),EED(M/N/O),HIKESHI(M/N)	PICALM(M/O)
18	14	SLC24A4	92931045	92938855		ATXN3(M),CPSF2(M),NDUFB1(M),TRIP11(M)	
19	15	ADAM10	58972995	59174539	ADAM10(M/N/O)	SLTM(M/O)	
20	15	APH1B	63569902	63569902	APH1B(M/N/O)	LACTB(M/O),RAB8B(M/O),RPS27L(N/O)	
21	16	KAT8	31132250	31133449		ARMC5(N),BCKDK(M),COX6A2(M/N),ITGAM(M), ITGAX(M),SETD1A(M),STX4(M),ZNF646(M), ZNF668(M),ZNF843(M/N)	
22	17	SCIMP	4999207	5259765	RABEP1(M/N/O),SCIMP(M)	NUP88(N),RPAIN(N),ZNF594(M)	
23	17	ABI3	47450775	56398006			
24	18	ALPK2	56189459	56189459		LOC101927322(M),MALT1(M)	
25	19	CD33	51727962	51736383	CD33(M)		
26	19	ABCA7	1039323	1053524		ABCA7(M),ARHGAP45(M),ATP5F1D(M),STK11(M)	
27	19	APOE	45020859	45734751	APOC1(N),BCAM(O),BCL3(O), CLASRP(N),CLPTM1(N),GEMIN7(O), PVR(M/N/O),RELB(N/O),TOMM40(N/O), ZNF296(N/O)	BLOC1S3(N),ERCC2(M),MARK4(M/N), PPP1R37(M/N),TRAPPC6A(N)	APOC1(M/O),APOC2(M),APOE(M/O), BCAM(M),BCL3(M/N),CLASRP(M/O), CLPTM1(M/O),GEMIN7(M/N),NECTIN2(M/N/O), RELB(M),ZNF296(M)
28	19	AC074212.3	45830947	46241841			
29	20	CASS4	54982351	55040537		AURKA(M/N),CSTF1(M/N),FAM210B(M),RTF2(M)	CASS4(M)

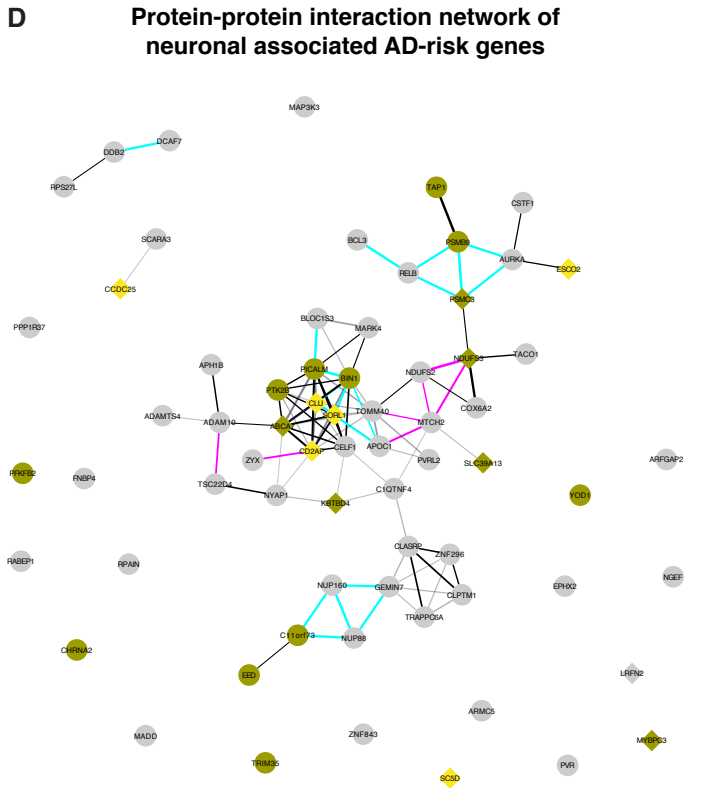
Supplementary Figure 9



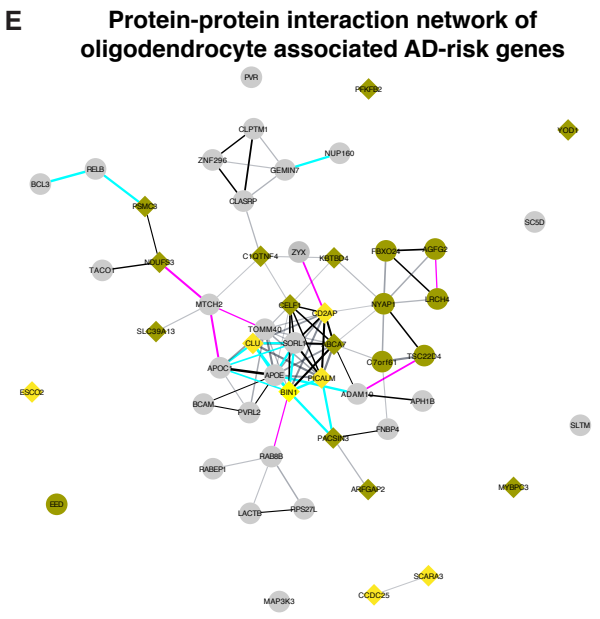
C



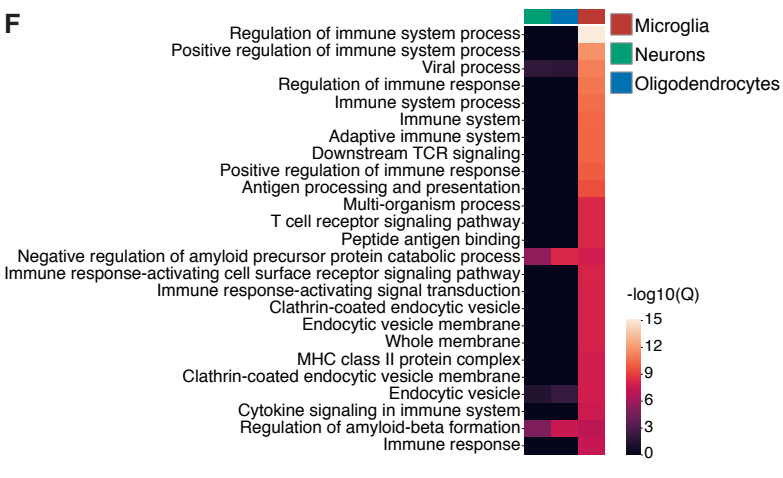
D



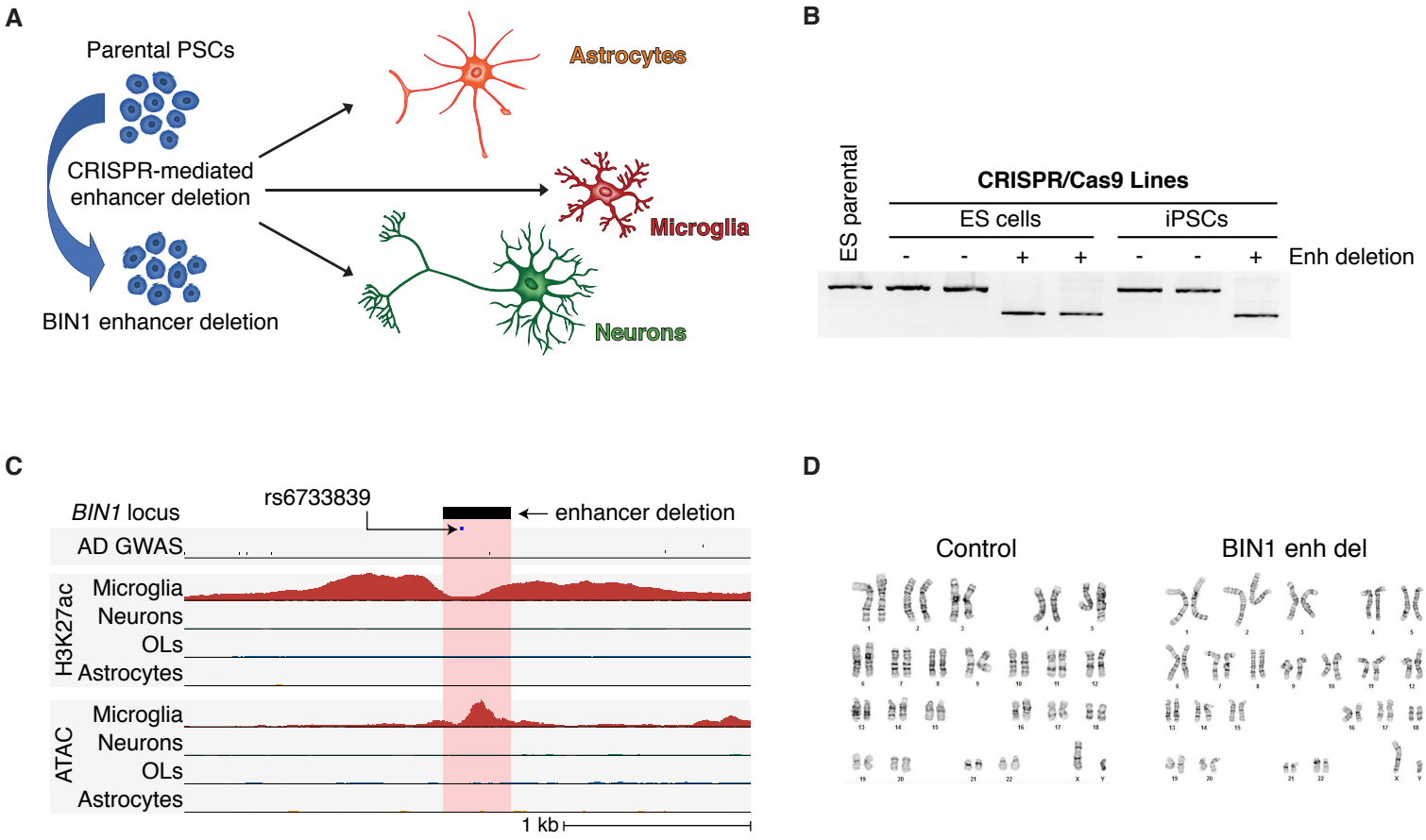
E



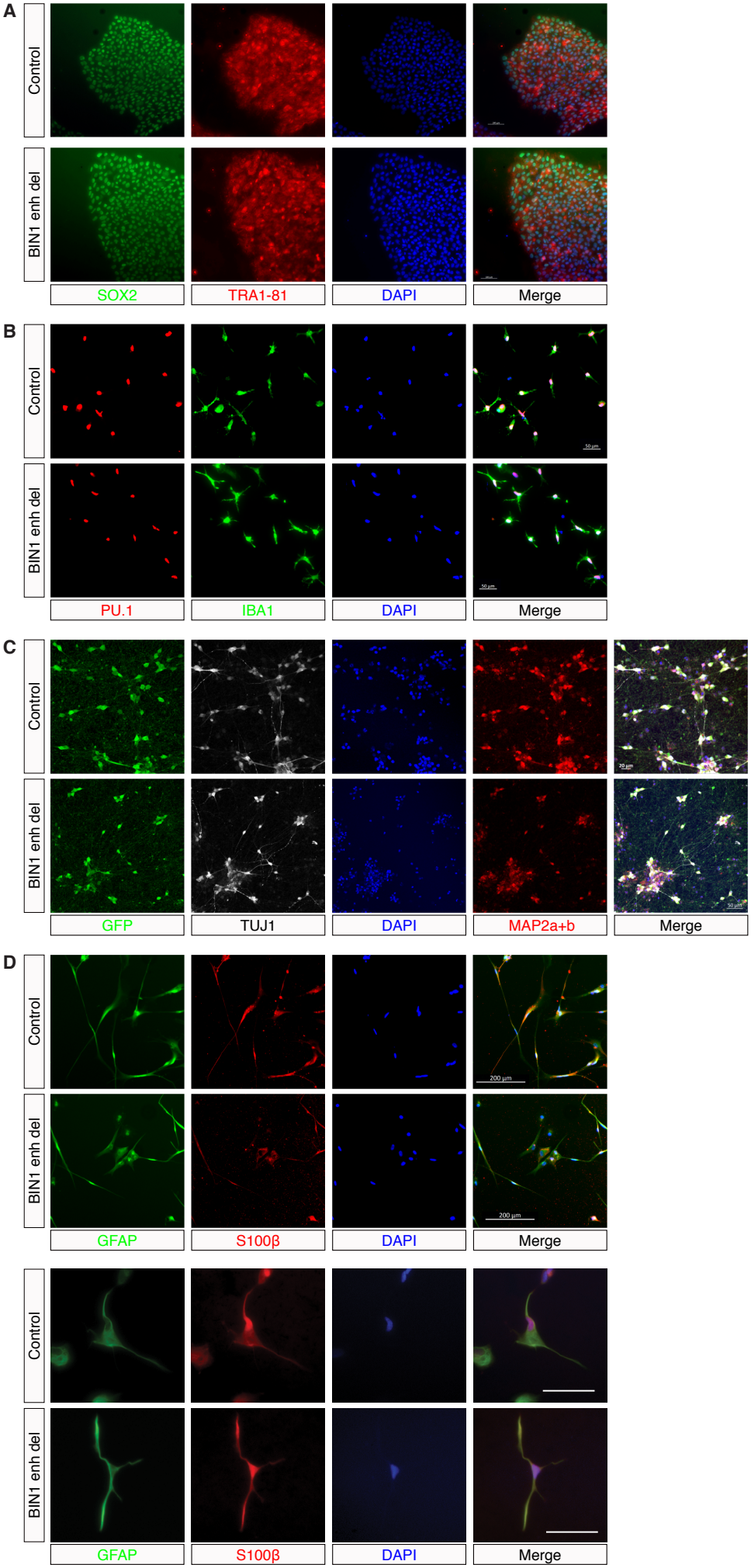
F



Supplementary Figure 10



Supplementary Figure 11



Supplementary Figure 12