Science Advances
AAAS

Supplementary Materials for

Multi-omic profiling of the developing human cerebral cortex at the single-cell level

Kaiyi Zhu *et al.*

Corresponding author: Panos Roussos, panagiotis.roussos@mssm.edu

Sci. Adv. **9**, eadg3754 (2023)
DOI: 10.1126/sciadv.adg3754

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Supplementary Text

Protocol for plasmid preparation

Reaction Mix:

- 1 µl guide oligo (100 µM)
- 1 µl reverse complement oligo (100 µM)
- 1 µl 10X T4 DNA Ligase buffer (New England Biolabs, Cat# B0202S)
- 0.5 µl T4 PNK (New England Biolabs, Cat# M0201L)
- 6.5 µl ddH2O
- 10 µl total

Thermocycler:

- 37°C 30 min
- 95°C 5 min
- Ramp down to 25°C at 5°C /min
- 4°C infinite hold

Phospho-annealed oligos were diluted 1:100 before proceeding to the golden gate reaction:

Reaction Mix:

- 12.5 µl Quick Ligation Buffer (2X) (New England Biolabs, Cat# M2200L)
- 0.25 µl BSA (10 mg/ml)
- 1 µl BsmB1 v2 (New England Biolabs, Cat# R0739L) (optimized for golden gate cloning)
- 0.125 µl T7 DNA Ligase (New England Biolabs, Cat# M0318S)
- 1 µl diluted phospho-annealed oligos
- 1 µl of [25ng/ul] LV-EF1a-U6 backbone
- 9.125 µl ddH2O
- 25 µl total

Thermocycler:

- Cycle (30x):
- 37°C 5 min
- 20°C 5 min
- 4°C infinite hold

Ligations were transformed into High Efficiency NEB 10-beta Competent E. coli. (New England Biolabs, Cat# C3019H) according to manufacturer's instructions. Bacteria were plated on LB-Ampicillin agar medium at 37°C overnight. On the following day, colonies were grown in LB-Ampicillin liquid medium at 37°C with shaking. Plasmids were purified with the QIAprep Spin Miniprep kit (Qiagen, Cat# 27106). Positive clones were validated through sanger sequencing by GeneWiz using a universal U6 forward primer. Validated plasmid DNA containing guide RNA inserts were packaged into lentivirus by VectorBuilder with a viral titer of >108 TU/ml in HBSS Buffer.

Design of guides for CRISPRi experimental

- Guide name: NEUROD1 1, sequence: AGTGATAGTCTCATAACCCT, PAM: GGG, sgRNA "cut" Chromosomal position: chr2: 181680471
- Guide name: NEUROD1 2, sequence: TTATGAGACTATCACTGCTC, PAM: AGG, sgRNA "cut" Chromosomal position: chr2: 181680453
- Guide name: NEUROD1 3, sequence: GCAGGAGGCGCGGCGTCCGG, PAM: AGG, sgRNA "cut" Chromosomal position: chr2: 181680484
- Guide name: CUX2 1, sequence: GAGATGCAGCGAGCGCTCCG, PAM: CGG, sgRNA "cut" Chromosomal position: chr12:111033777
- Guide name: CUX2 2, sequence: AGATGCAGCGAGCGCTCCGC, PAM: GGG, sgRNA "cut" Chromosomal position: chr12:111033778
- Guide name: CUX2 3, sequence: AGCGAGCGCTCCGCGGGCCC, PAM: GGG,
- sgRNA "cut" Chromosomal position: chr12:111033784
- Guide name: Scrambled, sequence: GCACTCACATCGCTACATCA

Design of guides for CRISPRi experimental

- Guide name: NEUROD1 1, sequence: AGTGATAGTCTCATAACCCT, PAM: GGG, sgRNA "cut" Chromosomal position: chr2: 181680471
- Guide name: NEUROD1 2, sequence: TTATGAGACTATCACTGCTC, PAM: AGG, sgRNA "cut" Chromosomal position: chr2: 181680453
- Guide name: NEUROD1 3, sequence: GCAGGAGGCGCGGCGTCCGG, PAM: AGG, sgRNA "cut" Chromosomal position: chr2: 181680484
- Guide name: CUX2 1, sequence: GAGATGCAGCGAGCGCTCCG, PAM: CGG, sgRNA "cut" Chromosomal position: chr12:111033777
- Guide name: CUX2 2, sequence: AGATGCAGCGAGCGCTCCGC, PAM: GGG, sgRNA "cut" Chromosomal position: chr12:111033778
- Guide name: CUX2 3, sequence: AGCGAGCGCTCCGCGGGCCC, PAM: GGG,
- sgRNA "cut" Chromosomal position: chr12:111033784
- Guide name: Scrambled, sequence: GCACTCACATCGCTACATCA

List of materials and reagents used

- Reagent: Accutase; vendor: Innovative Cell Technologies; cat #: AT104
- Reagent: DMEM (Dulbecco’s Modified Eagle Medium); vendor: Gibco; cat #: 11965092
- Reagent: RNAse Free PBS; vendor: Fisher Scientific; cat #: BP24384
- Reagent: Corning Cell Culture Buffer: Dulbecoo's Phosphate-Buffered Salt Solution 1x; vendor: Fisher Scientific; cat #: MT21030CV
- Reagent: Recombinant RNAse Inhibitor; vendor: Takara; cat #: 2313A
- Reagent: BSA; vendor: Miltenyi Biotec; cat #: 130-091-376
- Reagent: Tween-20; vendor: Sigma-Aldrich; cat #: P1379
- Reagent: Countess cell counting chamber slides ; vendor: Invitrogen; cat #: C10283
- Reagent: Trypan Blue; vendor: Gibco; cat #: 15250-061
- Reagent: QuBit RNA HS Assay Kit; vendor: Fisher Scientific; cat #: Q32855
- Reagent: Cell strainer; vendor: Sigma-Aldrich; cat #: BAH13680040
- Reagent: QIAzol Lysis Reagent; vendor: Qiagen; cat #: 79306
- Reagent: Chloroform; vendor: Sigma-Aldrich; cat #: C2432
- Reagent: lentiGuide-Hygro-mTagBFP2; vendor: Addgene; cat #: 99374

- Reagent: 10X T4 DNA Ligase buffer; vendor: New England Biolabs; cat #: B0202S
- Reagent: T4 PNK; vendor: New England Biolabs; cat #: M0201L
- Reagent: Quick Ligation Buffer (2X); vendor: New England Biolabs; cat #: M2200L
- Reagent: BsmB1 v2; vendor: New England Biolabs; cat #: R0739L
- Reagent: T7 DNA Ligase; vendor: New England Biolabs; cat #: M0318S
- Reagent: High Efficiency NEB 10-beta Competent E. coli. ; vendor: New England Biolabs; cat #: C3019H
- Reagent: QIAprep Spin Miniprep kit; vendor: Qiagen; cat #: 27106
- Reagent: Formaldehyde solution (37 wt.% in H2O, contains 10-15% Methanol as stabilizer); vendor: Sigma-Aldrich; cat #: 252549
- Reagent: RNAscope Probe- Hs-NEUROD1-C2; vendor: Advanced Cell Diagnostics; cat #: 437281-C2
- Reagent: RNAscope Probe- Hs-CUX2-C3; vendor: Advanced Cell Diagnostics; cat #: 425581-C3
- Reagent: RNAscope H2O2 and Protease Reagents; vendor: Advanced Cell Diagnostics; cat #: 322381
- Reagent: RNAscope Multiplex Fluorescent Detection Reagents V2; vendor: Advanced Cell Diagnostics; cat #: 323110
- Reagent: RNAscope Wash Buffer Reagents; vendor: Advanced Cell Diagnostics; cat #: 310091
- Reagent: ImmEdge™ Hydrophobic Barrier Pen; vendor: Advanced Cell Diagnostics; cat #: 310018
- Reagent: RNAscope Multiplex TSA Buffer; vendor: Advanced Cell Diagnostics; cat #: 322809
- Reagent: RNAscope Probe Diluent; vendor: Advanced Cell Diagnostics; cat #: 300041
- Reagent: Opal 570 Reagent Pack; vendor: Akoya Biosciences; cat #: FP1488001KT
- Reagent: Opal 690 Reagent Pack; vendor: Akoya Biosciences; cat #: FP1497001KT
- Reagent: Nunc™ Lab-Tek™ II CC2™ Chamber Slide System (8-well); vendor: Thermo Fisher; cat #: 154941
- Reagent: 24 x 50 mm microscope cover glass (thickness No.1); vendor: Fisher Scientific; cat #: 12-545F
- Reagent: Prolong Gold Antifade Reagent ; vendor: Thermo Fisher; cat #: P36934

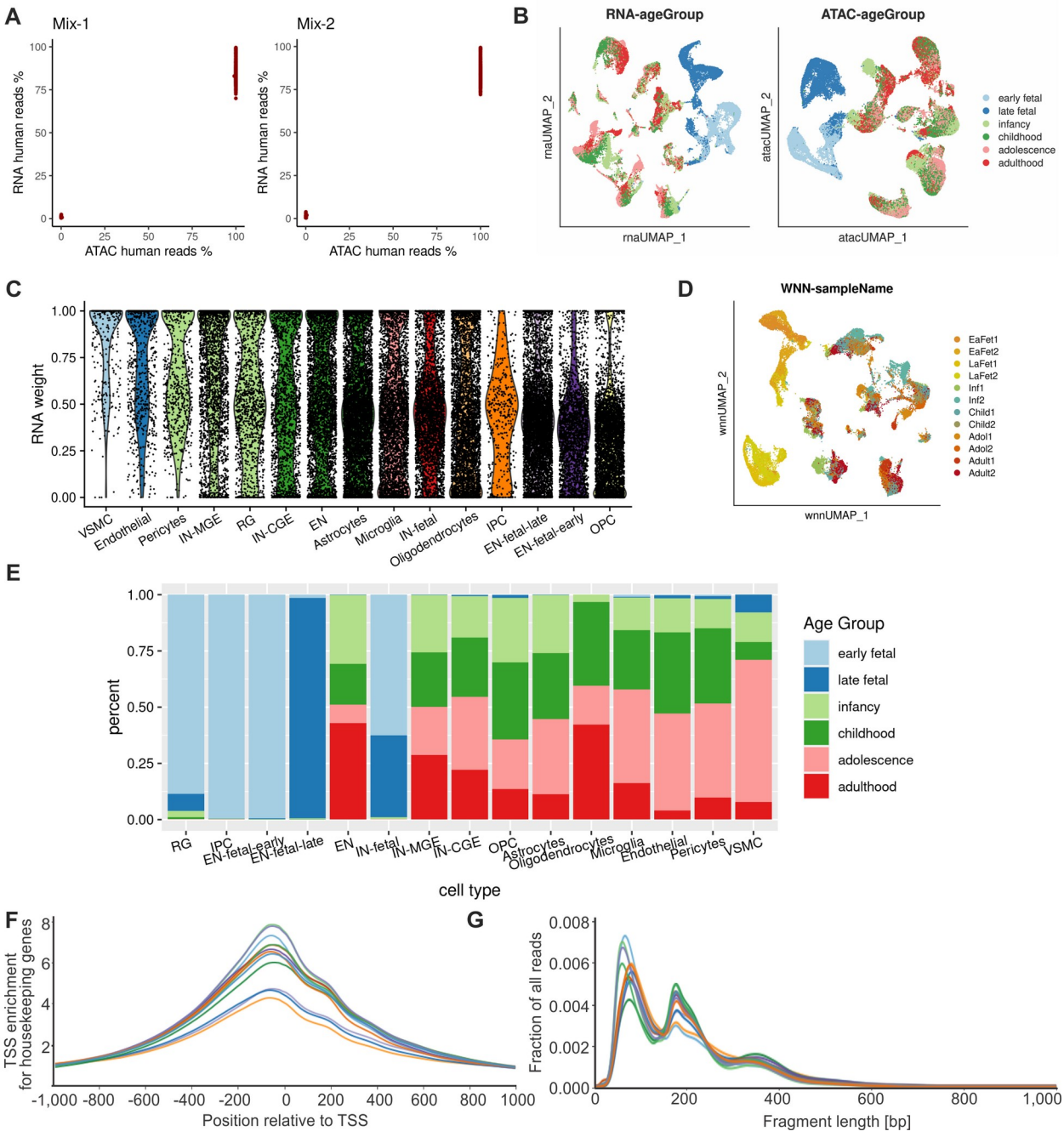

Fig. S1.

(**A**) The percentage of ATAC-seq or RNA-seq reads aligning to the human genome relative to all reads mapping uniquely to the human or mouse genome in the human-mouse cell line mixtures. We retained cells with total RNA-seq count > 2,000 and < 9,500, total ATAC-seq count > 15,000 and < 200,000, and mitochondrial percentage < 30%. The cell numbers for the two mixed samples are: 465 human cells and 319 mouse cells in Mix-1, 439 human cells and 312 mouse cells in Mix-2. (**B**) UMAP visualizations of single cells defined by RNA-seq and ATAC-seq data, respectively, where cells are annotated in terms of the age groups. (**C**) Single cell RNA modality weights derived from WNN analysis. VSMCs, endothelial cells and pericytes have the

highest RNA weights, consistent with the fact that they can be distinguished only in RNA modality. (**D**) UMAP visualization of single cells based on the WNN-derived graph, where cells are annotated in terms of sample names. (**E**) Bar plots showing the proportion of cells coming from different age groups for each cell type. (**F**) TSS enrichment from ATAC-seq reads. (**G**) Fragment length distribution of ATAC-seq reads.

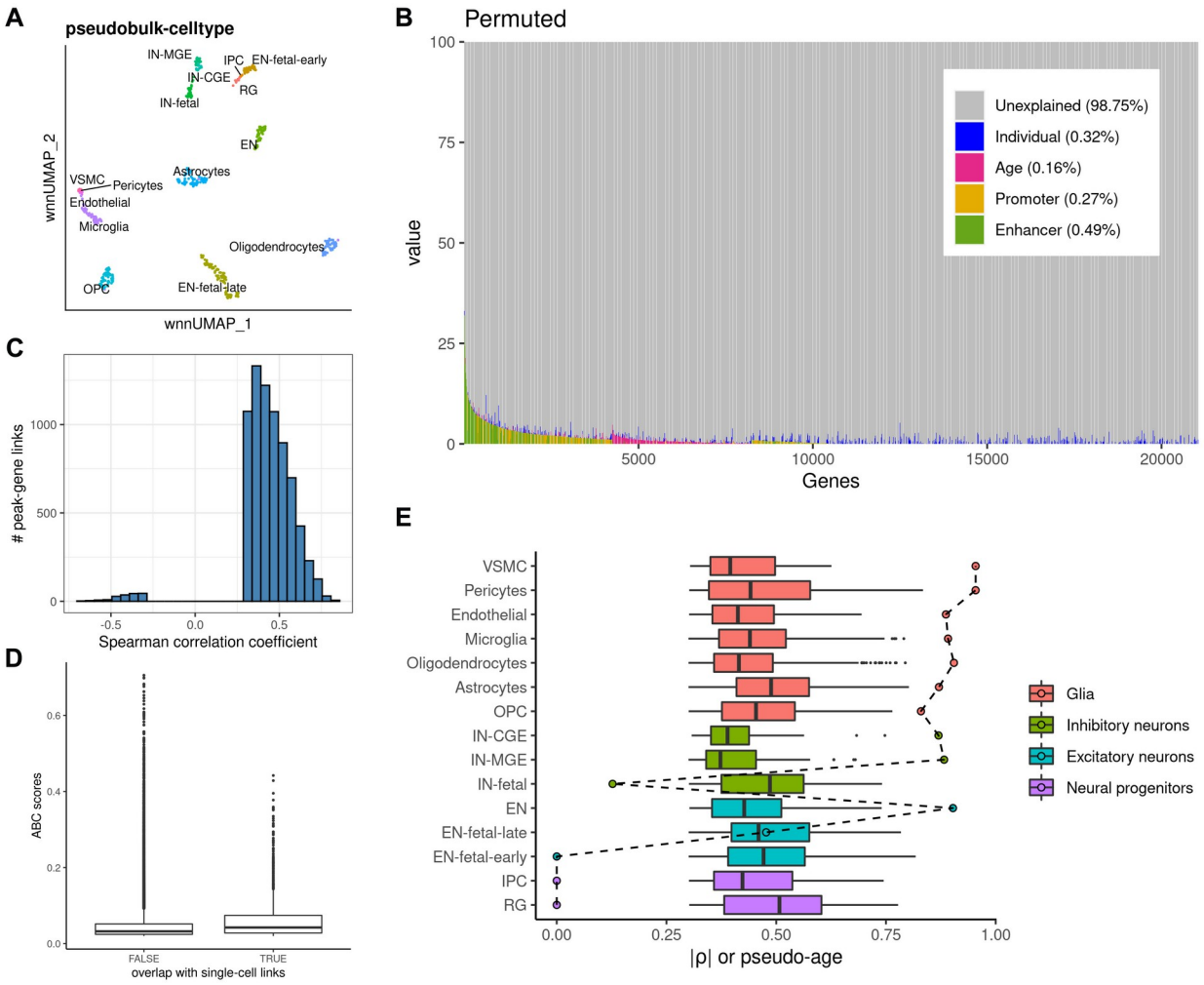

Fig. S2.

(**A**) WNN-derived UMAP visualization for the 500 pseudobulk aggregate samples, annotated by cell types. (**B**) Variance in gene expression estimated to be explained by chromatin accessibility, as well as other variables, using shuffled data. Genes, in columns, are sorted by decreasing proportion of variance explained by the epigenome (enhancers and promoters), with the mean variance explained by each component shown in parenthesis. (**C**) Histogram showing the Spearman correlation coefficients of the significant peak-gene links. (**D**) Boxplots for comparing the ABC scores of the E-P interactions overlapping peak-gene links, versus those not overlapping. (**E**) Boxplots showing the distribution of the absolute Spearman correlation coefficients of the peak-gene links specific to different cell types; circles connected by a dashed line represent the pseudo-age per cell type; colored in terms of broader cell type classification.

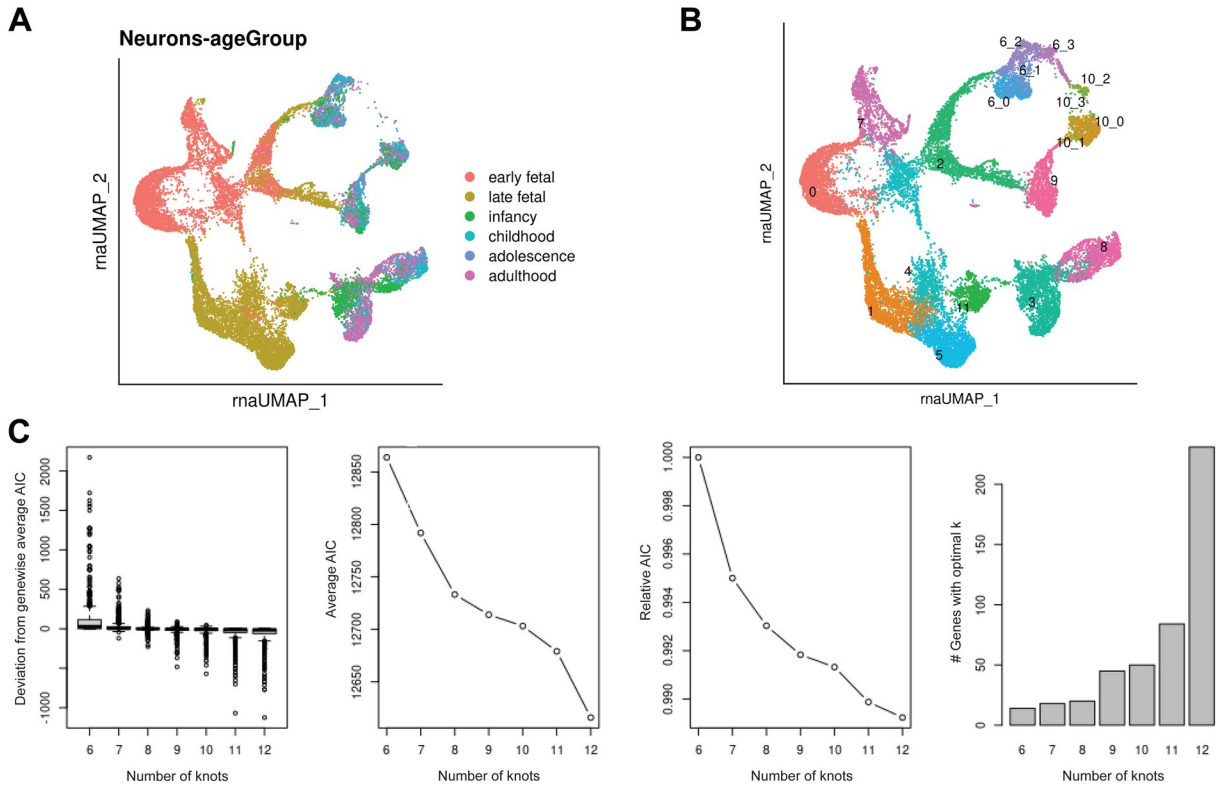

Fig. S3.

(**A**) UMAP visualization of neuronal cell populations on RNA coordinates, colored in terms of different developmental stages. (**B**) UMAP visualization of neuronal cell populations, colored and labeled in the cluster IDs. Cells assigned to clusters '6_3' (n=380) and '10_2' (n=150) were suspected as doublets between IN-MGE and IN-CGE, and were removed from downstream pseudotime analysis. (**C**) Diagnostic plots for selecting the optimal number of knots, k ∈ {6, …, 12}, using the AIC as implemented in the evaluateK function in tradeSeq (*75*). See tradeSeq tutorials for more details.

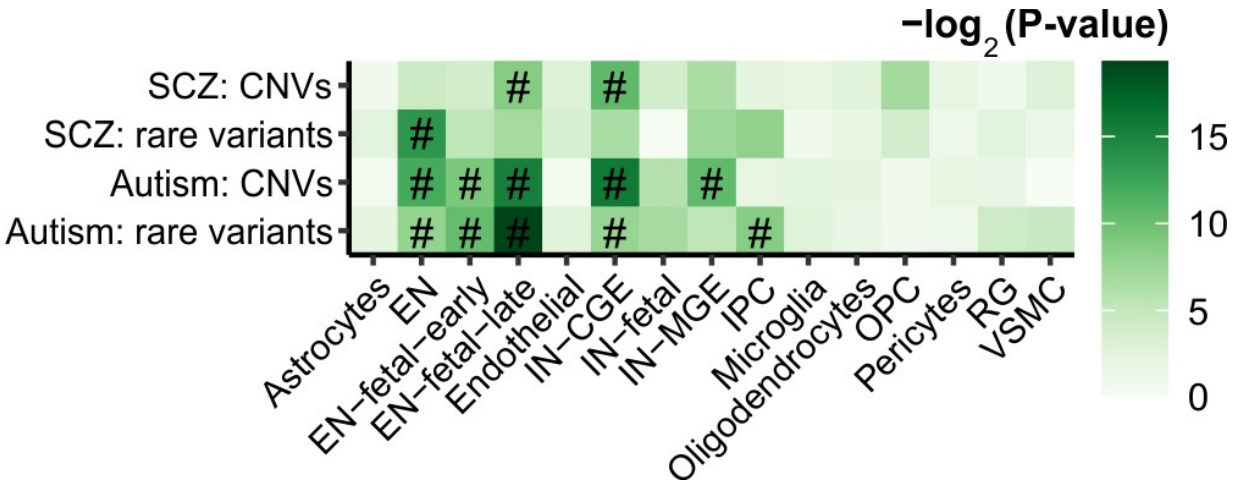

Fig. S4.

Enrichment of brain cell types in genes associated with rare variants and CNVs in autism (*46, 47*) and schizophrenia (*45, 48*). Color intensity is proportional to -$\log_{10}$(*P* value) calculated by Fisher exact test. "#": significant after correction across all tests (Bonferroni-corrected *P* value <0.05).

Table S1. Sample information including brain regions, age, sex, and batches.

Table S2. Differentially expressed genes for each cell type.

Table S3. Differentially accessible peaks for each cell type.

Table S4. List of peak-gene links which were identified with significant associations.

Table S5. List of glia-/neuron-specific super-enhancers overlapped with each DORC.

Table S6. Differentially expressed DORC-regulated genes for each cell type.

Table S7. List of peak-gene links which were identified among neurons with significant associations.

Table S8. GO enrichment analysis results for neuron-specific DORC-regulated genes.

Table S9. TF motif enrichment results for neuronal lineage-specific clusters of genes (km1/2/3/4).

Table S10. Overlap between common risk genetic variants associated with 53 brain and non-brain related traits and cell-specific peaks as calculated by LD-sc method.

Table S11. Overlap between common risk genetic variants associated with 53 brain and non-brain related traits and marker genes (including proximity of those genes) as calculated by MAGMA method.

Table S12. Candidate causal genes for risk variants (GWAS index SNPs and their LD buddies) associated with neuropsychiatric disorders.

Table S13. Overlap between peaksets (“DORC” and “km1-4”) and de novo variants from whole-genome sequencing dataset on ASD cases and controls.

Table S14. Number and relative proportion of doublets per each cell type.