HHS Public Access

Author manuscript

Science. Author manuscript; available in PMC 2016 September 27.

Published in final edited form as:

Science. 2016 June 24; 352(6293): 1586–1590. doi:10.1126/science.aaf1204.

Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain

Blue B. Lake[1,†], **Rizi Ai**[2,†], **Gwendolyn E. Kaeser**[3,5,†], **Neeraj S. Salathia**[4,†], **Yun C. Yung**[3], **Rui Liu**[1], **Andre Wildberg**[2], **Derek Gao**[1], **Ho-Lim Fung**[1], **Song Chen**[1], **Raakhee Vijayaraghavan**[4], **Julian Wong**[3], **Allison Chen**[3], **Xiaoyan Sheng**[3], **Fiona Kaper**[4], **Richard Shen**[4], **Mostafa Ronaghi**[4], **Jian-Bing Fan**[4,*], **Wei Wang**[2,*], **Jerold Chun**[3,*], and **Kun Zhang**[1,*]

[1]Department of Bioengineering, University of California, San Diego, La Jolla, California, USA

[2]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California, USA

[3]Department of Molecular and Cellular Neuroscience, Dorris Neuroscience Center, The Scripps Research Institute, La Jolla, California, USA

[4]Illumina Inc., San Diego, California, USA

[5]Biomedical Sciences Graduate Program, University of California, San Diego, La Jolla, California, USA

Abstract

The human brain has enormously complex cellular diversity and connectivities fundamental to our neural functions, yet difficulties in interrogating individual neurons has impeded understanding of the underlying transcriptional landscape. We developed a scalable approach to sequence and quantify RNA molecules in isolated neuronal nuclei from post-mortem brain, generating 3,227 sets of single neuron data from six distinct regions of the cerebral cortex. Using an iterative clustering and classification approach, we identified 16 neuronal subtypes that were further annotated on the basis of known markers and cortical cytoarchitecture. These data demonstrate a robust and scalable method for identifying and categorizing single nuclear transcriptomes, revealing shared genes sufficient to distinguish novel and orthologous neuronal subtypes as well as regional identity within the human brain.

Main Text

While significant progress has been achieved in mice (1–3), comprehensive classification of adult human brain neurons on the basis of their single-cell transcriptomes has yet to be realized. Examination of individual neuronal gene expression profiles for functional patterns could provide unbiased insights into subtypes from defined neuroanatomical regions, which are missed by gross anatomical studies that report limited transcriptomic differences across

*Corresponding Authors: Kun Zhang (kzhang@bioeng.ucsd.edu); Jerold Chun (jchun@scripps.edu); Wei Wang (wei-wang@ucsd.edu); Jian-Bing Fan (jianbing_fan@yahoo.com).

†Equally contributed authors.

the neocortex (4–7). Previous analyses of single adult human neurons has been dependent on methods compatible with freshly isolated neurosurgical tissues (8), which can be difficult to obtain, with limited regional sampling and depth. By contrast, post-mortem tissues provide a vastly more accessible source of both normal and diseased brain, wherein challenges to interrogating single neuronal genomes can be overcome using single nuclei (9, 10) combined with RNA sequencing. Here, we report the development of a scalable pipeline from post-mortem brain through nuclear transcriptome analyses that identifies both known and novel neuronal subtypes across the cerebral cortex in humans.

With the goal of defining transcriptomic profiles of single neurons, neuronal nuclear antigen (NeuN) was used (9) to isolate neuronal nuclei (Fig. S1) from the post-mortem brain of a normal, 51-year old female (Fig. 1A). We focused on six classically defined Brodmann Areas (BAs) with well-documented anatomical and electrophysiological properties that were derived from a single cortical hemisphere, since inter-hemispheric and inter-individual transcriptome differences were reported to be minimal (4–7). Isolation of nuclei was used to reduce transcriptomic contamination from other cells or degradation encountered with whole-neuron dissociation or laser caption micro-dissection (Fig. S2). Furthermore, sequencing of RNA from single nuclei on a limited scale has found gene expression values comparable to that of the whole cell (11, 12). Therefore, we developed and implemented a highly-scalable, single nucleus RNA sequencing (SNS) pipeline (13) (Fig. 1A, Fig. S1, Fig. S3–S8) that has broad applicability for post-mortem brains derived from multiple brain banks/repositories (Fig. S4F).

Using this pipeline, we processed 86 Fluidigm C1 chips and sequenced 4,488 single nuclei to an average depth of 8.34M reads (Table S1, Fig. S5). Genomic mapping rates revealed a high proportion of reads that corresponded to intronic sequences (Fig. 1A, Fig. S5A). The low percentage of intergenic reads argues against possible genomic contamination. Instead, the intronic reads likely captured an abundance of nascent RNA transcripts present in the nuclei. Intronic reads can be used to predict *de novo* expression (14), as well as whole cell gene transcription levels (15). Additionally, our single nuclei expression data inclusive of intronic reads accurately predicted cellular identity (Fig. S7), thereby providing initial validation for our SNS pipeline.

After quality filtering, including removal of doublets misclassified as single nuclei (13) (Fig. 1A, Fig. S6), we achieved 3,227 data sets across the six cortical regions (Fig. 1A, Table S2). To identify neuronal subtypes, we developed a clustering and classification strategy that was capable of resolving 17 clusters (13) (Fig. S8A) on the basis of differential expression of neuronally annotated marker genes (Tables S3–S4, Fig. S8B). These clusters showed distinct subgroup aggregation (Fig. 1B, Fig. S9A) and unique gene expression profiles associated with neuronal ontologies (Fig. 1C, Fig. S9B, Tables S5–S6). With the exception of a single cluster (NoN, n=44) deriving from one C1 chip having reduced mapping rates, 16 of these clusters were generated independent of detectable batch effects (Table S2, Fig. S10). Differential expression of inhibitory markers associated with GABAergic interneurons (Table S3) distinguished potential inhibitory (In) from excitatory (Ex) neuronal subtypes (Fig. 1B), consistent with mutually exclusive positivity of associated marker genes using a fraction of positive thresholding method (2) (Fig. 2A). As such, our dataset first

differentiated two major classifications within the cerebral cortex: 972 inhibitory neurons that generally encompass interneurons and 2,253 excitatory neurons that generally encompass pyramidal or projection neurons (16). Furthermore, each subgroup within these classifications showed distinct contributions from each brain region (Fig. 2A, Table S7), likely reflecting varied proportions of these neuronal subtypes across BAs, with most variability present in the visual cortex (BA17) that is known to have distinct cytoarchitecture and gene expression profiles (7, 17).

In order to further annotate inhibitory neuron subtypes, we examined expression of known marker genes associated with cortical layers, developmental origin, and interneuron classification (13) (Fig. 2B). On the basis of *in situ* human brain expression data (Fig. S11) (17), our inhibitory neuron subtypes were found to distribute spatially from the pial surface (most superficial boundary) to white matter (deepest boundary) of the neocortex, and could be grouped by the developmental origin of interneurons from subcortical regions of the medial, lateral, or caudal ganglionic eminences (MGE, LGE or CGE) (Fig. 2B) (18, 19). Furthermore, distinct profiles of interneuron classification markers revealed subtypes that parallel those identified from the mouse somatosensory cortex (3) (Fig. 2B–C, Fig. S12A). Cortical regional heterogeneity within subtypes was also observed, as evident by a layer 3 population (In4) that showed a specific absence of *RELN/SST* expression in BA17 (Fig. 2C, Fig. S11B, D). As such, our data distinguished inhibitory neuron subtypes having heterogeneous distributions within the neocortex.

Most excitatory cortical projection or pyramidal neurons can be categorized by their layer position established during neocortical development (17) combined with their axonal projections (16) (Fig. 3A). Our excitatory neuron subgroups, which were also in high concordance with subtypes identified in mice (3) (Fig. S12B), expressed known markers associated with a superficial-to-deep cortical distribution (13) (Fig. 3B–D, Fig. S13), with more than one subtype occupying most layers. Our data set was able to resolve cortical region specificity, as seen for the *BHLHE22* positive (Fig. 3C, Fig. S13A,D) layer 4 subtypes Ex2 and Ex3 (Fig. 4A), where Ex2 derived predominantly from rostral regions, BA8 and BA10, and Ex3 from caudal regions, BA17 and BA41/42 (Fig. 2A, Fig. 4B). Consistently, these subgroups showed distinct gene expression (Fig. 4C, Table S8) associated with neuronal electrophysiology and connectivity (Table S9). Furthermore, we were able to resolve intra-subtype heterogeneity, in terms of BA-specific expression patterns, which was observed in all subtypes (Fig. 4B), as for example within the Ex3 subtype between BA17 and BA41/42 regions (Fig. 4B,D; Table S10). As such, regional neurophysiological differences in cortical regions may be attributed to not only variations in the proportions of interneuron and projection neuron subtypes, but also to cell-intrinsic transcriptomic differences amongst single neurons within a subtype. Consistent with this possibility, we found genes having known variability between the visual and temporal cortices from *in situ* hybridization (ISH) studies (17) also had transcriptomic differences that could be attributed to subtypes defined by our data set (13) (Fig. S14A, Table S11). Therefore, our data highlight subtle yet region-defining gene expression signatures amongst specific neuronal subtypes that could not be detected from bulk analyses (Fig. S14B).

To further understand the extent of heterogeneity that may exist within subtypes, we identified genes varying globally (Table S12, Fig. S15A) or expressed differentially within each BA (Table S13, Fig. S15B) for each subgroup. While a subset of *In* and *Ex* subgroup-variable genes were associated with differential expression between brain regions, a large proportion were unique (Fig. S15C). Therefore, the potential exists for not only intra-regional cortical transcriptomic differences, but also further intra-subtype heterogeneity. This might reflect a technical need for increased sampling depth for further subtype resolution, yet may also indicate the potential for even more diversity within subtypes associated with a broader range of individualized neuronal activities. Consistent with these observations, proportions of subgroup variable genes were associated with neuronal subtype classification, post-synaptic function and known regional expression variability (Fig. S15C). These data support further local and regional functional heterogeneity existing amongst defined subtypes.

Our results demonstrate that post-mortem SNS can identify expected and novel neuronal subtypes that provide insight into brain function through distinct profiles of activity defining genes (Fig. S16, Table S14). Furthermore, given that only a very small subset of layer specific markers used in our analyses (*CARTPT, CHRNA7, PDYN, RELN*) were found to have ISH differences between individual donors (17), our subtypes can be expected to be globally representative. Indeed, our subtypes remain highly conserved in mice (3), with differences highlighting evolutionary changes in potential orthologues (Fig. S12). Our data sets reveal shared gene expression signatures that can distinguish subtypes and regional identity, supporting a transcriptional basis for well-known differences in cortical cytoarchitecture. Additional heterogeneity found within single neuronal transcriptomes may further reflect activities of complex neuronal networks that vary with function and time, as well as underlying genomic mosaicism that exists in human cortical neurons (10, 20–23). Our study thus lays the groundwork for high-throughput global human brain transcriptome mapping using nuclei derived from readily available post-mortem tissues for analyses of normal individuals, as assessed here, as well as myriad diseases of brain and mind.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Flow cytometry was performed both at the UCSD Human Embryonic Stem Cell Core and TSRI Flow Cytometry Core. Initial C1 runs were performed at the UCSD Stem Cell Genomics Core. The data tables accompanying this work are provided as Excel files in the supplementary materials. Clustering-and-Classification code used to identify neuronal subtypes and instructions (Readme.txt) for its operation in R are provided as supplementary files. We thank Fluidigm Inc. (M. Ray, R.C. Jones, P. Steinberg) for instrument support and technical advice in adaptation of the C1 protocol for nuclei. Sequencing data has been deposited with dbGaP (accession phs000833.v3.p1), curated by the NIH Single Cell Analysis Program – Transcriptome (SCAP-T) Project (http://www.scap-t.org) and annotated in supplementary material (Table S2). We thank G. Kennedy for help with RNAscope. Funding support was from the NIH Common Fund Single Cell Analysis Program (1U01MH098977). GEK was additionally supported by Neuroplasticity of Aging Training Grant (5T32AG000216-24).

References

1. Macosko EZ, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015; 161:1202–1214. [PubMed: 26000488]
2. Usoskin D, et al. Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nat Neurosci. 2015; 18:145–153. [PubMed: 25420068]
3. Zeisel A, et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science. 2015; 347:1138–1142. [PubMed: 25700174]
4. Hawrylycz MJ, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. Nature. 2012; 489:391–399. [PubMed: 22996553]
5. Kang HJ, et al. Spatio-temporal transcriptome of the human brain. Nature. 2011; 478:483–489. [PubMed: 22031440]
6. Johnson MB, et al. Functional and evolutionary insights into human brain development through global transcriptome analysis. Neuron. 2009; 62:494–509. [PubMed: 19477152]
7. Hawrylycz M, et al. Canonical genetic signatures of the adult human brain. Nat Neurosci. 2015; 18:1832–1844. [PubMed: 26571460]
8. Darmanis S, et al. A survey of human brain transcriptome diversity at the single cell level. Proc Natl Acad Sci U S A. 2015; 112:7285–7290. [PubMed: 26060301]
9. Bushman DM, et al. Genomic mosaicism with increased amyloid precursor protein (APP) gene copy number in single neurons from sporadic Alzheimer's disease brains. Elife. 2015; 4
10. Gole J, et al. Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. Nat Biotechnol. 2013; 31:1126–1132. [PubMed: 24213699]
11. Grindberg RV, et al. RNA-sequencing from single nuclei. Proc Natl Acad Sci U S A. 2013; 110:19802–19807. [PubMed: 24248345]
12. Krishnaswami SR, et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. Nat Protoc. 2016; 11:499–524. [PubMed: 26890679]
13. See supplementary text, materials and methods available at *Science* Online.
14. Graf A, et al. Fine mapping of genome activation in bovine embryos by RNA sequencing. Proc Natl Acad Sci U S A. 2014; 111:4139–4144. [PubMed: 24591639]
15. Gaidatzis D, Burger L, Florescu M, Stadler MB. Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation. Nat Biotechnol. 2015; 33:722–729. [PubMed: 26098447]
16. Greig LC, Woodworth MB, Galazo MJ, Padmanabhan H, Macklis JD. Molecular logic of neocortical projection neuron specification, development and diversity. Nat Rev Neurosci. 2013; 14:755–769. [PubMed: 24105342]
17. Zeng H, et al. Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. Cell. 2012; 149:483–496. [PubMed: 22500809]
18. Hansen DV, et al. Non-epithelial stem cells and cortical interneuron production in the human ganglionic eminences. Nat Neurosci. 2013; 16:1576–1587. [PubMed: 24097039]
19. Ma T, et al. Subcortical origins of human and monkey neocortical interneurons. Nat Neurosci. 2013; 16:1588–1597. [PubMed: 24097041]
20. Lodato MA, et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. Science. 2015; 350:94–98. [PubMed: 26430121]
21. McConnell MJ, et al. Mosaic copy number variation in human neurons. Science. 2013; 342:632–637. [PubMed: 24179226]
22. Rehen SK, et al. Constitutional aneuploidy in the normal human brain. J Neurosci. 2005; 25:2176–2180. [PubMed: 15745943]
23. Westra JW, et al. Neuronal DNA content variation (DCV) with regional and individual differences in the human brain. J Comp Neurol. 2010; 518:3981–4000. [PubMed: 20737596]
24. Spalding KL, Bhardwaj RD, Buchholz BA, Druid H, Frisen J. Retrospective birth dating of cells in humans. Cell. 2005; 122:133–143. [PubMed: 16009139]
25. Edelstein AD, et al. Advanced methods of microscope control using muManager software. J Biol Methods. 2014; 1

26. Bardou P, Mariette J, Escudie F, Djemiel C, Klopp C. jvenn: an interactive Venn diagram viewer. BMC Bioinformatics. 2014; 15:293. [PubMed: 25176396]
27. Wang F, et al. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. J Mol Diagn. 2012; 14:22–29. [PubMed: 22166544]
28. Ramskold D, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol. 2012; 30:777–782. [PubMed: 22820318]
29. Wu AR, et al. Quantitative assessment of single-cell RNA-sequencing methods. Nat Methods. 2014; 11:41–46. [PubMed: 24141493]
30. Zhou W, Abruzzese RV, Polejaeva I, Davis S, Ji W. Amplification of nanogram amounts of total RNA by the SMART-based PCR method for high-density oligonucleotide microarrays. Clin Chem. 2005; 51:2354–2356. [PubMed: 16306094]
31. Baker SC, et al. The External RNA Controls Consortium: a progress report. Nat Methods. 2005; 2:731–734. [PubMed: 16179916]
32. Fan J, et al. Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. Nat Methods. 2016; 13:241–244. [PubMed: 26780092]
33. Zhang Y, et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J Neurosci. 2014; 34:11929–11947. [PubMed: 25186741]
34. Vogt D, et al. Lhx6 directly regulates Arx and CXCR7 to determine cortical interneuron fate and laminar position. Neuron. 2014; 82:350–364. [PubMed: 24742460]
35. McKinsey GL, et al. Dlx1&2-dependent expression of Zfhx1b (Sip1, Zeb2) regulates the fate switch between cortical and striatal interneurons. Neuron. 2013; 77:83–98. [PubMed: 23312518]
36. Schaerlinger B, Hickel P, Etienne N, Guesnier L, Maroteaux L. Agonist actions of dihydroergotamine at 5-HT2B and 5-HT2C receptors and their possible relevance to antimigraine efficacy. Br J Pharmacol. 2003; 140:277–284. [PubMed: 12970106]
37. Ross SE, et al. Bhlhb5 and Prdm8 form a repressor complex involved in neuronal circuit assembly. Neuron. 2012; 73:292–303. [PubMed: 22284184]
38. Bayes A, et al. Characterization of the proteome, diseases and evolution of the human postsynaptic density. Nat Neurosci. 2011; 14:19–21. [PubMed: 21170055]

Summary

Single-nucleus RNA sequencing of neurons from the adult human cerebral cortex revealed transcriptomic signatures sufficient to identify neuronal subtypes and neuroanatomical areas while also revealing transcriptomic heterogeneity.

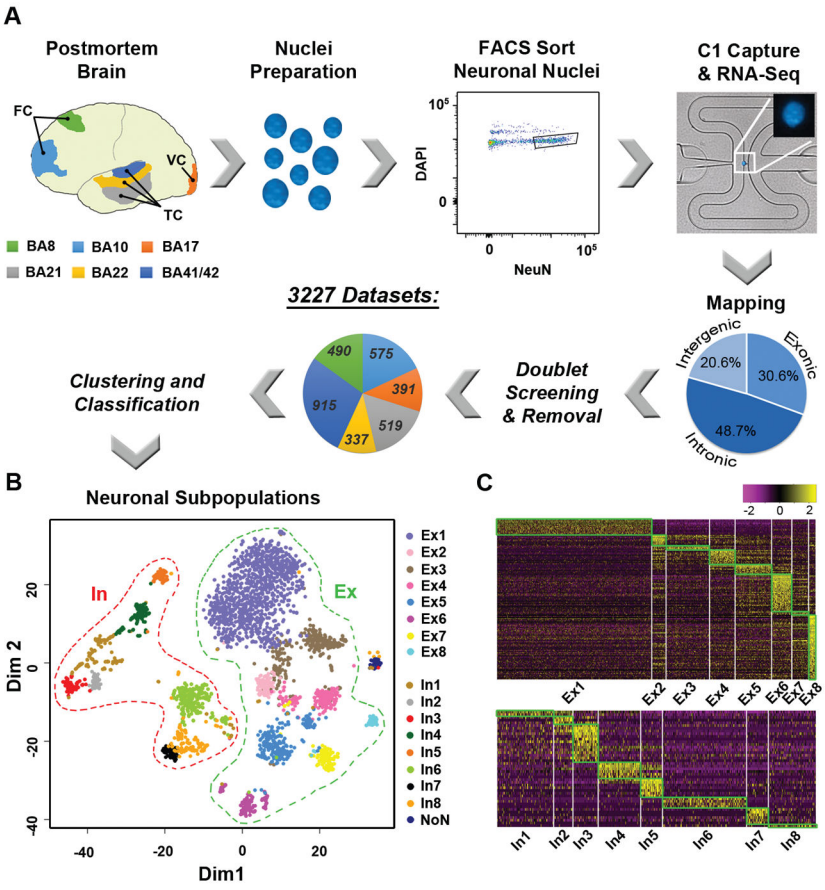

Fig. 1.

Single nucleus RNA sequencing (SNS) identified 16 neuronal subtypes over 6 neocortical regions. **A.** Overview of SNS pipeline. Post-mortem tissue from Brodmann Areas (BA) 8, 10, 17, 21, 22, and 41/42 were dissociated to single nuclei for NeuN+ and DAPI+ sorting and capture on C1 chips. Resultant libraries were sequenced, mapped to the reference genome (pie chart showing averaged proportions) and screened for doublet removal before clustering and classification. BA proportions are shown. FC = Frontal Cortex; TC = Temporal Cortex; VC = Visual Cortex. **B.** Neuronal subtypes (excitatory (Ex) and inhibitory (In)) shown by multidimensional plotting using 10-fold or greater differentially expressed genes (Table S3); NoN (no nomenclature), low expression outlier cluster. **C.** Heatmap showing unique marker gene expression (Table S5).

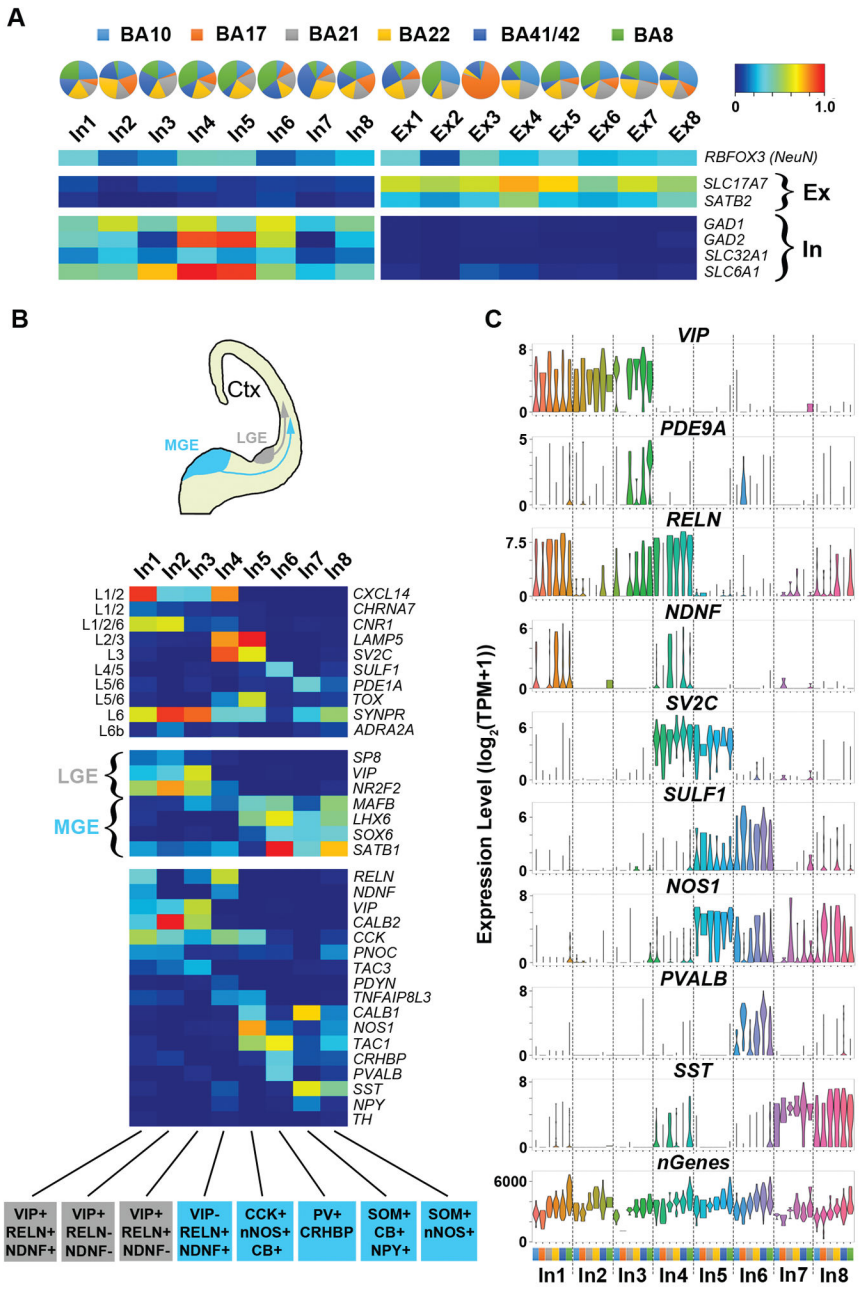

Fig. 2.

SNS reveals distinct interneuron subtypes. **A.** Pie charts display relative proportions of subtypes amongst BAs, and fraction of positive (FOP) heatmaps for inhibitory (In) and excitatory (Ex) marker genes. **B.** Diagram of subpallial origins of interneurons from either the lateral or medial ganglionic eminence (LGE, MGE) with FOP heatmaps (see **A** for scale) for marker genes associated with cortical layer (L) (upper panel), subpallial origin (middle panel) and interneuron classification (bottom panel). Potential interneuron subtypes are indicated below. SOM, *somatostatin or SST*; NPY, *neuropeptide Y*; CB, *calbindin-D-28k or CALB1*; VIP, *vasoactive intestinal peptide*; RELN, *reelin*; nNOS, *neuronal nitric oxide*

synthase or NOS1; PV, *parvalbumin or PVALB*; CCK, *cholecystokinin;* NDNF, *neuron-derived neurotrophic factor;* CRHBP*, corticotropin releasing hormone binding protein*. **C.** Violin plots showing select marker gene expression values by BA (colors indicated in **A**) for each inhibitory neuron subtype. *nGenes,* total number of genes identified.

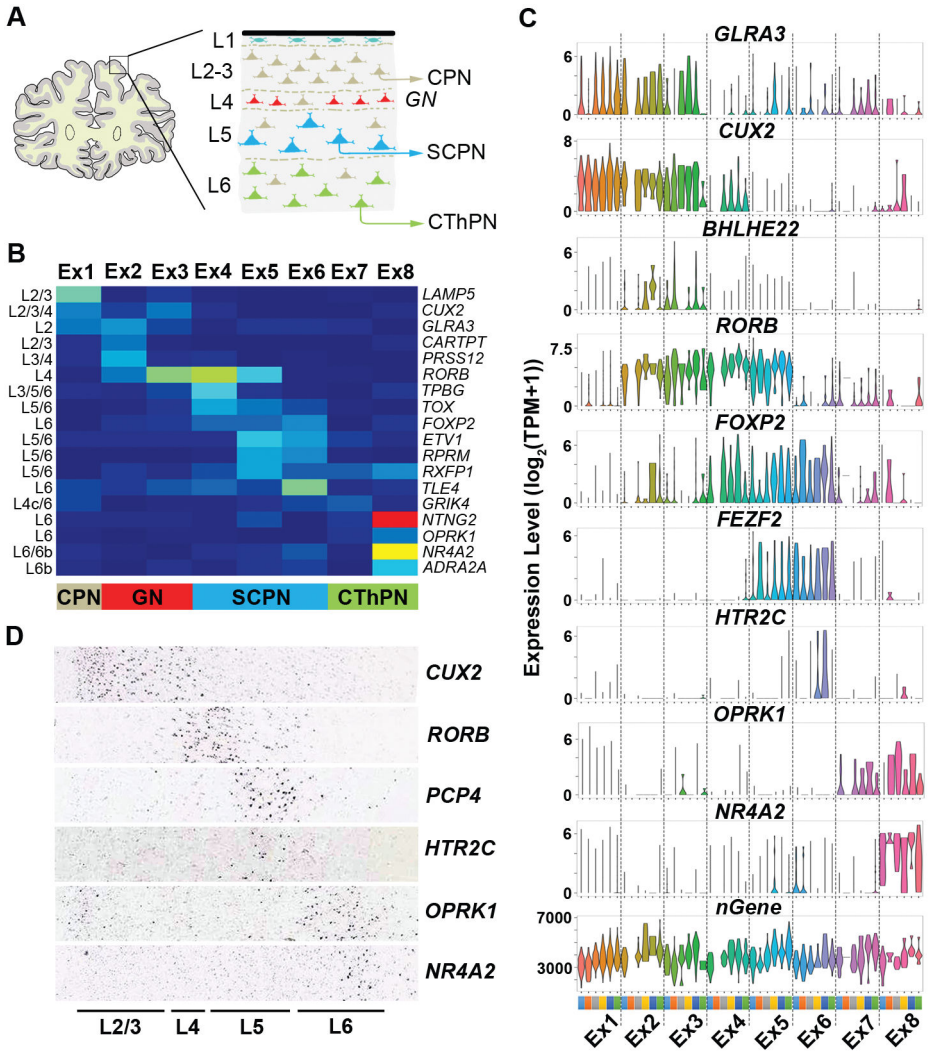

Fig. 3.

Excitatory neuronal subtypes show distinct spatial organization. **A.** Schematic of the prefrontal cortex showing projection neuron layers (L) and expected axonal projection destinations (layer 4 granule neurons typically receive outside inputs for distribution of signals locally). **B.** FOP heatmap (see Fig. 2A for scale) for layer specific marker genes showing expected cortical layer identity (L2–L6b) and excitatory neuron sub-classification. CPN = cortical projection neuron; GN = granule neuron; SCPN = subcortical projection neuron; CThPN = corticothalamic projection neuron. **C.** Violin plots showing selected marker gene expression values by Ex subtype and BA represented by colors (see Fig. 2A). *nGenes* = total number of genes identified. **D.** RNA ISH showing layer-specific expression of selected markers in the temporal cortex (Allen Human Brain Atlas, Table S11).

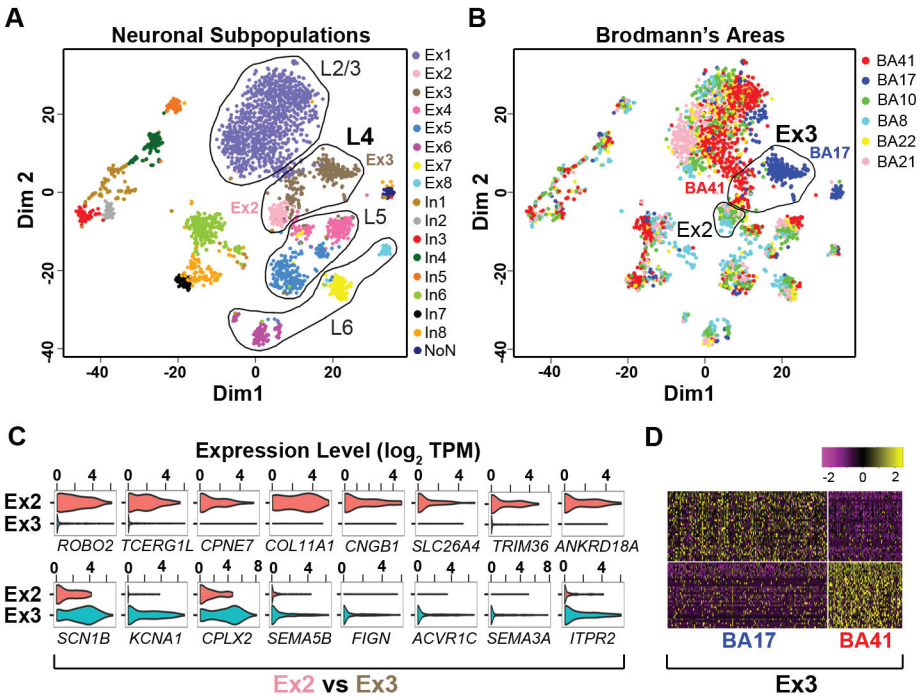

Fig. 4.

Neuronal subtypes reveal heterogeneity amongst BAs. **A.** Multidimensional plot showing projection neuron subtypes distributed according to their predicted cortical layer (L) identity. Layer 4 Ex2 and Ex3 subtypes are indicated. **B.** Clusters shown in (**A**) colored by BA and with BA41/42 and BA17 subpopulations of Ex3 indicated. **C.** Violin plots showing differentially expressed genes between Ex2 and Ex3 subtypes (Table S8). **D.** Heatmap showing genes differentially expressed between BA17 and BA41/42 within the Ex3 subtype (Table S10).