

# Transcriptomic and morphophysiological evidence for a specialized human cortical GABAergic cell type

Eszter Boldog<sup>1,8</sup>, Trygve E. Bakken<sup>2,8</sup>, Rebecca D. Hodge<sup>2,8</sup>, Mark Novotny<sup>3</sup>, Brian D. Aevermann<sup>3</sup>, Judith Baka<sup>1</sup>, Sándor Bordé<sup>1</sup>, Jennie L. Close<sup>2</sup>, Francisco Diez-Fuertes<sup>3</sup>, Song-Lin Ding<sup>2</sup>, Nóra Faragó<sup>1</sup>, Ágnes K. Kocsis<sup>1</sup>, Balázs Kovács<sup>1</sup>, Zoe Maltzer<sup>2</sup>, Jamison M. McCorrison<sup>3</sup>, Jeremy A. Miller<sup>2</sup>, Gábor Molnár<sup>1</sup>, Gáspár Oláh<sup>1</sup>, Attila Ozsvár<sup>1</sup>, Márton Rózsa<sup>1</sup>, Soraya I. Shehata<sup>2</sup>, Kimberly A. Smith<sup>2</sup>, Susan M. Sunkin<sup>2</sup>, Danny N. Tran<sup>3</sup>, Pratap Venepally<sup>3</sup>, Abby Wall<sup>2</sup>, László G. Puskás<sup>4</sup>, Pál Barzó<sup>5</sup>, Frank J. Steemers<sup>6</sup>, Nicholas J. Schork<sup>3</sup>, Richard H. Scheuermann<sup>3,7</sup>, Roger S. Lasken<sup>3</sup>, Ed S. Lein<sup>2\*</sup> and Gábor Tamás<sup>1\*</sup>

**We describe convergent evidence from transcriptomics, morphology, and physiology for a specialized GABAergic neuron subtype in human cortex. Using unbiased single-nucleus RNA sequencing, we identify ten GABAergic interneuron subtypes with combinatorial gene signatures in human cortical layer 1 and characterize a group of human interneurons with anatomical features never described in rodents, having large ‘rosehip’-like axonal boutons and compact arborization. These rosehip cells show an immunohistochemical profile (GAD1<sup>+</sup>CCK<sup>+</sup>, CNR1-SST-CALB2-PVALB<sup>+</sup>) matching a single transcriptomically defined cell type whose specific molecular marker signature is not seen in mouse cortex. Rosehip cells in layer 1 make homotypic gap junctions, predominantly target apical dendritic shafts of layer 3 pyramidal neurons, and inhibit backpropagating pyramidal action potentials in microdomains of the dendritic tuft. These cells are therefore positioned for potent local control of distal dendritic computation in cortical pyramidal neurons.**

Understanding the cellular and circuit organization of the neocortex, the substrate for much of higher cognitive function, has been intensely studied since Ramón y Cajal<sup>1</sup>. Morphophysiological characterization using slice physiology has been the standard for decades<sup>2</sup>, but this approach suffers from undersampling, difficulties in quantitative classification of cell types<sup>3</sup>, and limited scalability to cover neuronal diversity. Single-cell transcriptomics enables unbiased, high-throughput quantitative surveys of molecularly defined cell types<sup>4–6</sup> that can be applied to any species, including human. Initial application to mouse cortex has revealed approximately 50 transcriptomic types, demonstrating both the feasibility of the approach and the complexity of the cortex. There is now great promise in combining these morphoelectric and transcriptomic approaches for the unbiased molecular classification and characterization of these types.

Recent systematic efforts have provided insight into the cellular composition and organization of rodent neocortical circuits, suggesting the presence of several dozen inhibitory and excitatory cell types<sup>3–5,7</sup>. However, conservation of cellular and circuit principles in human cortex is assumed but largely untested to date. Indeed, there is evidence for substantial neuronal differences between rodents and human; for example, distinct membrane<sup>8,9</sup> and synaptic<sup>10–14</sup> properties and dendritic complexity<sup>15–17</sup> of human neurons might

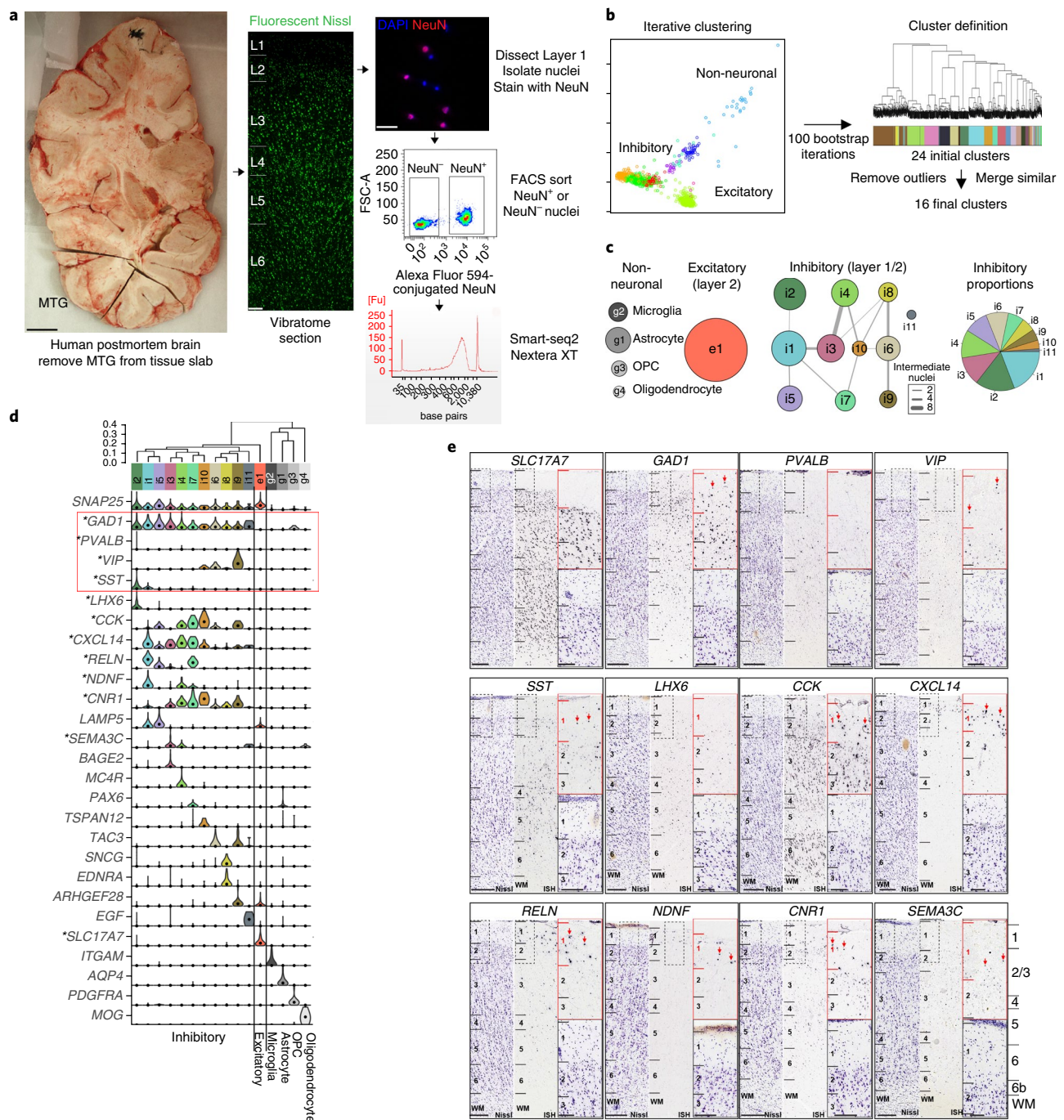
contribute to human-specific signal processing. With the mouse cortex as the dominant model for understanding human cognition, it is essential to establish whether the cellular architecture of the human brain is conserved or whether there are specialized cell types and system properties that cannot be modeled in rodents. Here we combine single-nucleus transcriptomics and slice physiology to study GABAergic neurons in layer 1 of human cortex and provide convergent lines of evidence for identification of a cell type with human-specialized features.

## Results

To allow an unbiased survey of transcriptionally defined cell types in human cortical tissue, we used single-nucleus RNA-sequencing<sup>18,19</sup> to profile large numbers of nuclei from frozen postmortem brain specimens (Fig. 1a). Briefly, this method involved microdissection of regions of interest from fluorescent Nissl-stained vibratome sections of cortex, tissue homogenization to liberate nuclei, NeuN staining and fluorescence-activated cell-sorting isolation, and Smart-seq2-based library preparation<sup>20</sup> (Fig. 1 and Supplementary Fig. 1). We applied this strategy to profile  $n = 769$  NeuN<sup>+</sup> neurons that had passed quality control and  $n = 102$  NeuN<sup>−</sup> non-neuronal cells across two individuals from microdissected layer 1 of the middle temporal gyrus, expected to predominantly contain inhibitory

<sup>1</sup>MTA-SZTE Research Group for Cortical Microcircuits, Department of Anatomy, Physiology and Neuroscience, University of Szeged, Szeged, Hungary.

<sup>2</sup>Allen Institute for Brain Science, Seattle, WA, USA. <sup>3</sup>J. Craig Venter Institute, La Jolla, CA, USA. <sup>4</sup>Laboratory of Functional Genomics, Department of Genetics, Biological Research Center, Hungarian Academy of Sciences, Szeged, Hungary. <sup>5</sup>Department of Neurosurgery, University of Szeged, Szeged, Hungary. <sup>6</sup>Illumina, Inc., San Diego, CA, USA. <sup>7</sup>Department of Pathology, University of California, San Diego, CA, USA. <sup>8</sup>These authors contributed equally to this work: Eszter Boldog, Trygve E. Bakken, Rebecca D. Hodge. \*e-mail: [EdL@alleninstitute.org](mailto:EdL@alleninstitute.org); [gtamas@bio.u-szeged.hu](mailto:gtamas@bio.u-szeged.hu)



**Fig. 1 | Identification of transcriptomic cell types in layer 1 of human temporal cortex.** **a**, Isolation of single nuclei from postmortem adult human cortex for RNA-sequencing. FSC-A, forward-scatter area. Scale bars, 1 cm (left); 20 μm (right). **b**, Left: nuclei were grouped based on similar gene expression profiles using an automated iterative clustering procedure. Clustering was repeated 100 times on random subsets of 80% of nuclei. Right: hierarchical clustering of nuclei that were consistently co-clustered across iterations identified 24 clusters. Sixteen clusters remained after removal of clusters associated with quality control metrics and merging of clusters that lacked at least one binary marker gene. **c**, Four non-neuronal, one excitatory, and 11 inhibitory neuron clusters were identified, although the excitatory cluster and one inhibitory cluster were likely in layer 2 due to incidental capture of superficial layer 2 with layer 1 dissection. For each cluster, the constellation diagram shows the cell-type class (based on canonical marker gene expression), relative frequency (disc area), and discreteness (line thickness proportional to the number of nuclei with ambiguous cluster membership) of clusters. **d**, Clusters arranged by transcriptomic similarity based on hierarchical clustering, with the expression distributions of selective marker genes shown across clusters as violin plots. Expression is on a linear scale and dots indicate median expression. \* marks genes with ISH data shown in **e**. Cluster sample sizes: i2 ( $n=77$ ); i1 ( $n=90$ ); i5 ( $n=47$ ); i3 ( $n=56$ ); i4 ( $n=54$ ); i7 ( $n=31$ ); i10 ( $n=16$ ); i6 ( $n=44$ ); i8 ( $n=27$ ); i9 ( $n=22$ ); i11 ( $n=6$ ); e1 ( $n=299$ ); g2 ( $n=27$ ); g1 ( $n=48$ ); g3 ( $n=18$ ); g4 ( $n=9$ ). **e**, ISH of select marker genes in human temporal cortex at low magnification (left; with near adjacent Nissl stain for cytoarchitectonic laminar identification) and high magnification in layers 1–3 (right). Red arrows, cells expressing genes in layer 1. Note that *LHX6* marks a single cluster (i2) that is not expressed in layer 1 and therefore nuclei in this cluster were likely sampled from upper layer 2. Other clusters are restricted to layer 1 (for example, *NDNF*) or may be distributed across layers 1 and 2. Scale bars, 250 μm (low magnification); 100 μm (high magnification). ISH experiments were conducted on multiple tissue donors: *SLC17A7*, *LHX6*, *CNR1*, *SEMA3C* ( $n=3$ ); *CXCL14* ( $n=5$ ); *GAD1*, *CCK*, *RELN*, *NDNF* ( $n=6$ ); *SST* ( $n=7$ ); *PVALB* ( $n=8$ ); *VIP* ( $n=10$ ). WM, white matter.

neurons. Median gene detection (expression > 0) was 9,937 in neurons and 6,287 in glia. Iterative clustering was used to group nuclei with similar transcriptional profiles, thereby identifying a robust set of transcriptomically defined cell types (Fig. 1b). Based on expression of known marker genes (Supplementary Fig. 2a), clusters corresponded to all major classes of neural cell types that were expected to be captured. These included major non-neuronal cell types (microglia, astrocytes, oligodendrocyte precursor cells (OPCs), and oligodendrocytes) and one excitatory neuron type sampled from upper cortical layer 2 incidentally included in the layer 1 dissection (Fig. 1c). In addition, 11 distinct clusters corresponding to GABAergic neuron subtypes were identified (numbered by relative abundance).

Transcriptomic cell types displayed highly selective gene expression (Fig. 1d and Supplementary Fig. 2a). For example, the pan-neuronal gene synaptosomal-associated protein 25 (SNAP25) clearly differentiated neuronal from non-neuronal types, which were in turn differentiated by highly specific marker genes. Glutamic acid decarboxylase 1 (*GAD1*) clearly delineated the GABAergic neurons. In cortical layers 2–6, most GABAergic neurons have mutually exclusive expression of parvalbumin (*PVALB*), somatostatin (*SST*), or vasoactive intestinal peptide (*VIP*)<sup>21</sup>. In contrast, *Pvalb* and *Sst* are not expressed in mouse layer 1 by in situ hybridization (ISH), while *Vip* labels only sparse cell populations (Supplementary Fig. 2b). Notably, both *SST* and *VIP* (but not *PVALB*) are seen in human middle temporal gyrus (MTG) layer 1 by ISH (Fig. 1e). The layer 1 MTG transcriptomic clusters expressed either *SST* (clusters i1 and i2), *VIP* (clusters i6, i9, and i10), or neither marker, although cluster i2 represented a cell type restricted to layer 2 since it also expresses *LHX6*, which is not found in layer 1 (Fig. 1d,e). Therefore, there appear to be ten inhibitory cell types within layer 1, although it is not clear whether any of these types are completely restricted to layer 1. We compared these layer 1 cell types to eight inhibitory clusters reported by Lake et al.<sup>22</sup> and found increased diversity within several published clusters (clusters In1–4) and decreased diversity of *LHX6*<sup>+</sup> interneuron clusters (clusters In5–8) that are enriched in deeper cortical layers and were not sampled in this study (Supplementary Fig. 3a–c).

These clusters in layer 1 express different combinations of known layer 1 interneuron markers, including cholecystokinin (*CCK*), reelin (*RELN*), neuron-derived neurotrophic factor (*NDNF*), and lysosomal-associated membrane protein family member 5 (*LAMP5*), many of which were confirmed by ISH as being expressed in layer 1 (Fig. 1e). Furthermore, each cluster showed highly selective expression of both known and previously uncharacterized individual marker genes. Notably, given the proximity of layer 1 to the overlying pia, several of these markers appear to be related to interaction with endothelia, including endothelin receptor type A (*EDNRA*) and epidermal growth factor (*EGF*). Furthermore, voltage-gated ion channels and GABA- and glutamate-receptor subunits show diverse expression patterns among interneurons, including highly cell-type-specific expression of *CACNA2D1*, *GABRG1*, *KCNH5*, and *SCN5A* (Supplementary Fig. 4). To summarize, this unbiased transcriptomic approach identified ten GABAergic interneuron subtypes in layer 1 that have distinctive combinatorial and specific gene expression signatures suggestive of distinct morphological and functional properties.

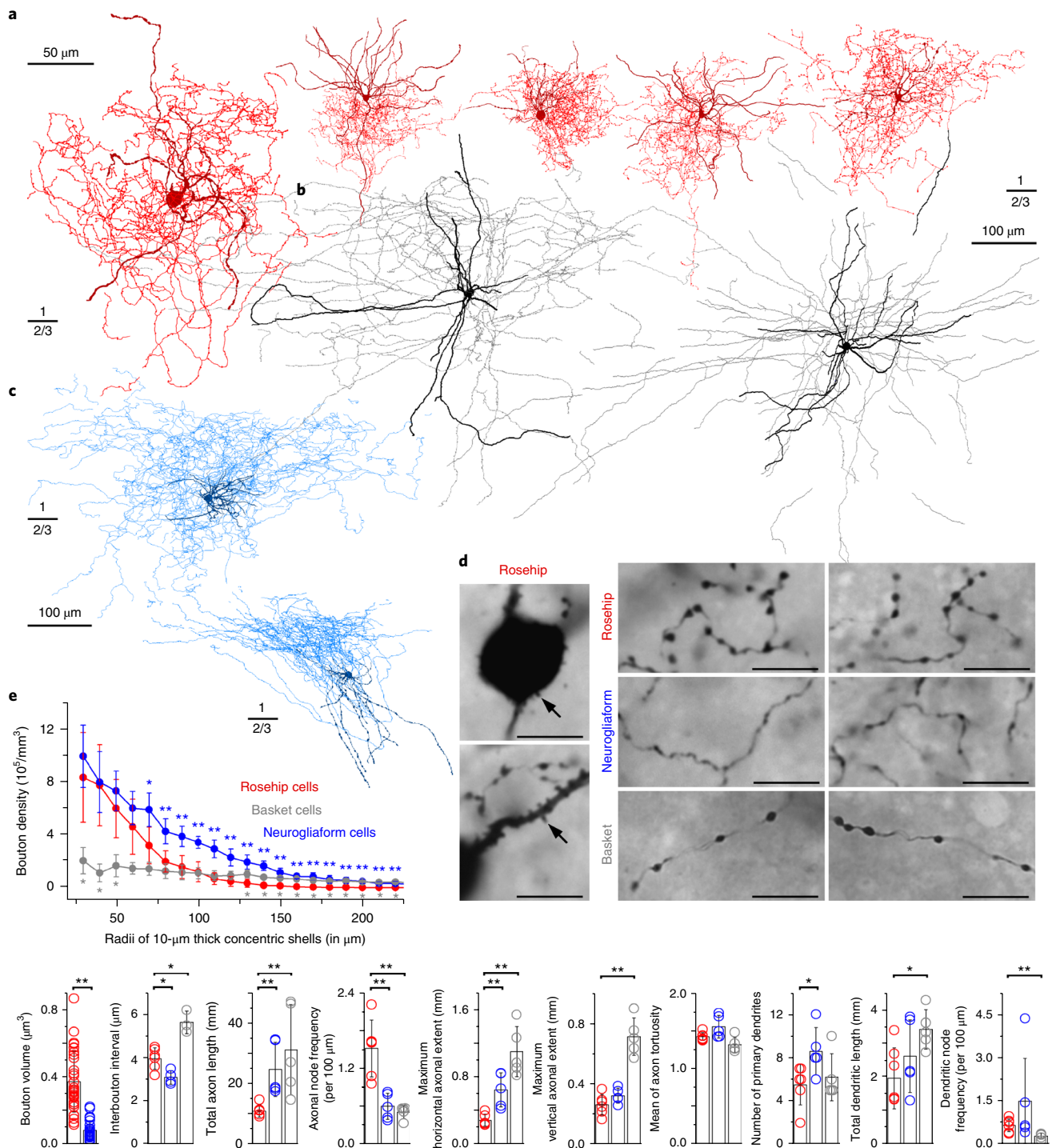
### Rosehip cells: morphological features in layer 1 of the human cerebral cortex

In parallel with the transcriptomic approach, we developed a dataset containing whole-cell-recorded, biocytin-filled interneurons in layer 1 of slices of nonpathological human samples of parietal, frontal, and temporal cortices<sup>10,11,23</sup>. Unbiased recordings of layer 1 cell types yielded a set of interneurons with complete axo-somato-dendritic recovery ( $n = 76$ ). Light-microscopic examination of these

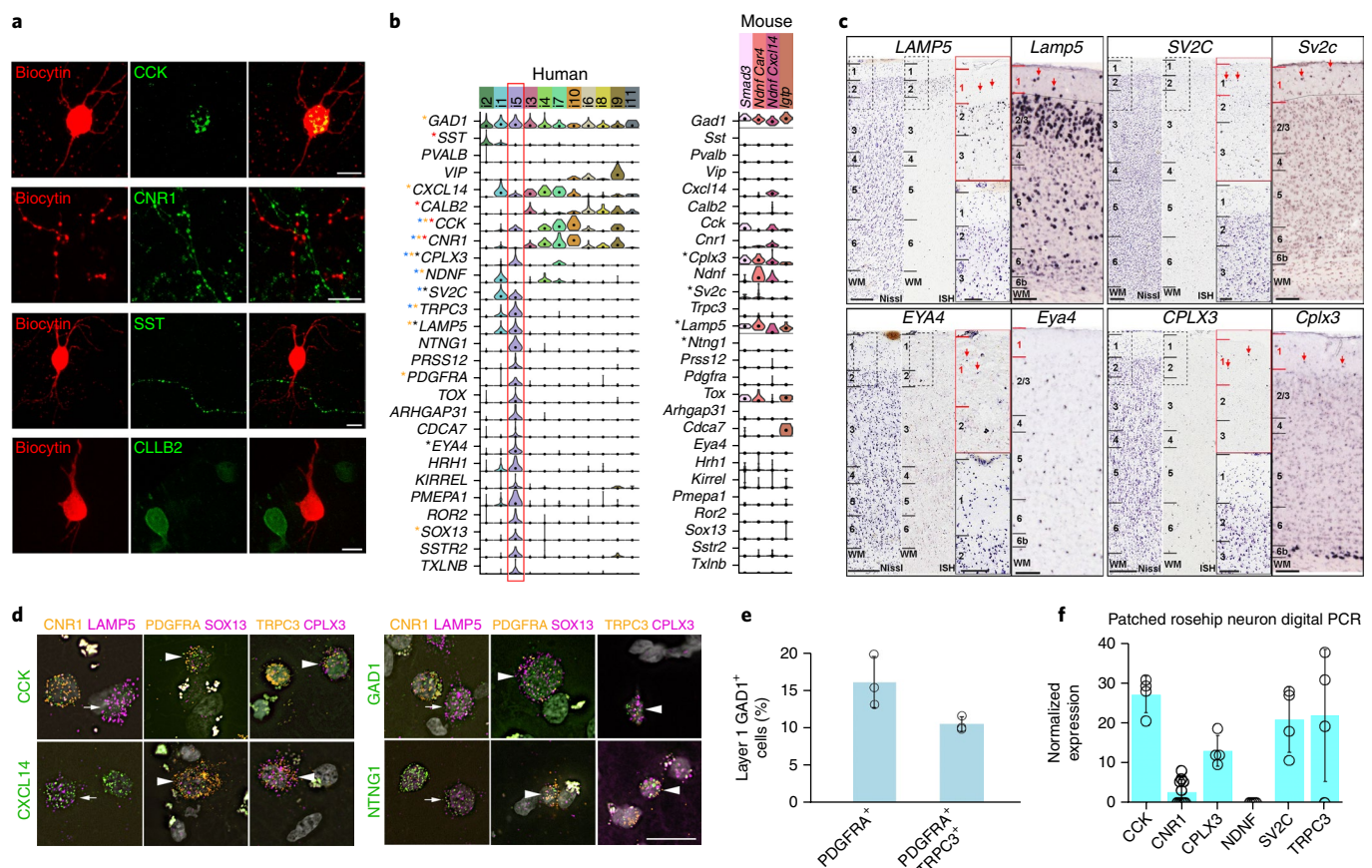
cells identified neurons with previously described morphological features, for example, neurogliaform cells (NGFCs,  $n = 16$ , 21%; Fig. 2c,d)<sup>1,21,24</sup>, as well as a previously undescribed group of interneurons with large, rosehip-shaped axonal boutons forming very compact, bushy arborizations (rosehip cells, RCs,  $n = 10$ , 13%; Fig. 2a,d). To our knowledge, interneurons with the phenotype of RCs as detailed below have not been identified previously in layer 1 of the cerebral cortex. Somata and dendrites of RCs were confined to layer 1, with only distal dendrites occasionally penetrating layer 2. Proximal dendrites and somata of RCs were decorated with stub-like spines. The axon of RCs usually emerged from the basal part of the soma and gave rise to very compact, dense axonal trees predominantly arborizing in layer 1, with tortuous collaterals displaying spindle-shaped boutons with diameters not seen in other types of human layer 1 interneurons in our sample. Targeted recordings increased the number of RCs in our database ( $n = 120$ ), and we quantitatively compared axodendritic parameters of randomly selected and three-dimensionally (3D) reconstructed RCs ( $n = 6$ ) to layer 1 neurogliaform ( $n = 5$ ) and layer 2/3 basket cells (BCs,  $n = 5$ ; Fig. 2b,d)<sup>10,11,24,25</sup>. The number of primary dendrites of RCs ( $5.50 \pm 1.87$ ) was similar to that of BCs ( $6.2 \pm 2.17$ ,  $n = 5$ ) and was significantly fewer compared to NGFCs ( $8.6 \pm 2.19$ ,  $n = 5$ ,  $P < 0.04$ , Mann–Whitney (MW) *U* test). Total dendritic length ( $1.96 \pm 0.90$  mm) and dendritic node frequency per 100  $\mu\text{m}$  ( $0.66 \pm 0.21$ ) of RCs were significantly different from those of BCs ( $3.41 \pm 0.58$  mm,  $P < 0.031$ ;  $0.29 \pm 0.10$ ,  $P < 0.009$ , respectively, MW *U* test) and were similar to those of NGFCs ( $2.62 \pm 1.08$  mm,  $1.50 \pm 1.47$ ). Total length ( $11.13 \pm 1.99$  mm) and maximal horizontal extent of axons ( $287.75 \pm 70.15 \mu\text{m}$ ) of RCs were significantly smaller than those of NGFCs ( $24.74 \pm 8.90$  mm and  $648.68 \pm 202.60 \mu\text{m}$ , respectively;  $P < 0.005$  for both, MW *U* test) and BCs ( $31.16 \pm 14.79$  mm,  $P < 0.009$ ;  $1,102.76 \pm 296.99 \mu\text{m}$ ,  $P < 0.005$ , respectively, MW *U* test). The maximal radial extent of axon of RCs ( $263.42 \pm 69.09 \mu\text{m}$ ) was significantly smaller than that of BCs ( $713.22 \pm 124.87 \mu\text{m}$ ,  $P < 0.005$ , MW *U* test), but not different from that of NGFCs ( $323.18 \pm 49.60 \mu\text{m}$ ). We measured axonal bouton densities of rosehip ( $n = 6$ ), neurogliaform ( $n = 4$ ), and basket ( $n = 3$ ) cells in 10- $\mu\text{m}$ -thick spherical shells of increasing diameter by Sholl analysis corrected with the portion of shells outside the brain slice. The bouton density of rosehip, neurogliaform, and BCs almost monotonously decreased with increasing distances from the soma; however, bouton densities were lower in BCs at 30–50  $\mu\text{m}$  from soma ( $P < 0.04$  for 30–50  $\mu\text{m}$ , MW *U* test), higher in NGFCs 70–220  $\mu\text{m}$  from soma ( $P < 0.02$ , MW *U* test), and higher in BCs 130–220  $\mu\text{m}$  from soma ( $P < 0.03$ ). RCs had longer interbouton intervals compared to NGFCs ( $3.97 \pm 0.49$  and  $3.10 \pm 0.32 \mu\text{m}$ , respectively,  $P < 0.038$ , MW *U* test) and shorter compared to BCs ( $5.63 \pm 0.51 \mu\text{m}$ ,  $P < 0.024$ ), measured as linear distances between neighboring boutons. RC axons branched more frequently, with RCs, NGFCs, and BCs having  $1.52 \pm 0.45$ ,  $0.61 \pm 0.21$ , and  $0.52 \pm 0.10$  nodes along 100  $\mu\text{m}$  length of their axons ( $P < 0.005$  for both, MW *U* test). Axon tortuosity (see Methods) of RCs ( $1.42 \pm 0.05$ ) was similar to that of neurogliaform ( $1.54 \pm 0.15$ ) and BCs ( $1.31 \pm 0.10$ ). Measurements based on serial-section electron microscopy and 3D reconstructions revealed that the volume of RC boutons ( $0.37 \pm 0.18 \mu\text{m}^3$ ,  $n = 31$ ) was approximately four times larger ( $P < 0.001$ ; MW *U* test) than that of NGFC boutons ( $0.08 \pm 0.06 \mu\text{m}^3$ ,  $n = 24$ ; Fig. 2e). The size of active zones in RCs ( $0.11 \pm 0.03 \mu\text{m}^2$ ,  $n = 11$ ) was not correlated to bouton volumes ( $\rho = 0.34$ ,  $P = 0.29$ , Spearman correlation). All fully reconstructed boutons ( $n = 31$ ) formed single synapses targeting dendritic shafts.

To understand the molecular identity of RCs and link them to the transcriptomic clusters, we performed immunohistochemistry on electrophysiologically recorded and anatomically recovered cells for known markers of GABAergic cell types (see Methods for details)<sup>26</sup>. This revealed that RCs were immunopositive for CCK ( $n = 10$ ) but negative for CB1 cannabinoid receptor (CNRI,  $n = 11$ ),





**Fig. 2 | Morphological phenotype of RCs in layer 1 of the human cerebral cortex.** **a**, Anatomical reconstructions of biocytin-filled RCs during whole-cell recordings (somata and dendrites, burgundy; axons, red). **b**, Anatomical reconstructions of layer 2/3 BCs in the human cerebral cortex (somata and dendrites, black; axons, gray). **c**, Anatomical reconstructions of NGFCs in layer 1 of the human cerebral cortex (somata and dendrites, dark blue; axons, light blue). **d**, Left: light micrographs of RCs ( $n=130$ ) showing somata and proximal dendrites with stub-like spines (arrows). Right: axons of RCs arborized densely with large, round boutons (top). Tortuous neurogliaform axons ( $n=16$ ) possess very small boutons (middle). Axons of BCs ( $n=5$ ) form longer segments with less convoluted branches with longer interbouton intervals (bottom). Scale bars, 10  $\mu\text{m}$ . **e**, Quantitative comparison of axonal and dendritic parameters of RC (red,  $n=6$ ), neurogliaform (blue,  $n=5$ ), and basket (gray,  $n=5$ ) cells. Top: bouton densities determined by Sholl analysis in 10- $\mu\text{m}$  thick spherical shells were lower in BCs at 30–50  $\mu\text{m}$  from soma and higher in NGFCs 70–220  $\mu\text{m}$  from soma and in BCs 130–220  $\mu\text{m}$  from soma, compared to those of RCs. Bottom: bouton volume ( $P<0.001$ ) and the number of primary dendrites ( $P<0.04$ ) of RCs were significantly different from those of NGFCs. Maximal vertical extent of axon ( $P<0.005$ ), total dendritic length ( $P<0.031$ ), and dendritic node frequency (per 100  $\mu\text{m}$ ;  $P<0.009$ ) of RCs differed significantly from those of BCs. Axonal tortuosity of RCs was similar to those of the two other cell types, but the frequency of axonal branch points in RCs was 2.5 and 2.95 times that of neurogliaform ( $P<0.005$ ) and BCs ( $P<0.005$ ), respectively. Furthermore, interbouton interval, total axon length, and maximal horizontal extent of the axon were also significantly different (two-sided MW  $U$  test,  $*P\leq 0.05$ ;  $**P\leq 0.01$ ; columns and error bars represent mean and s.d., respectively).



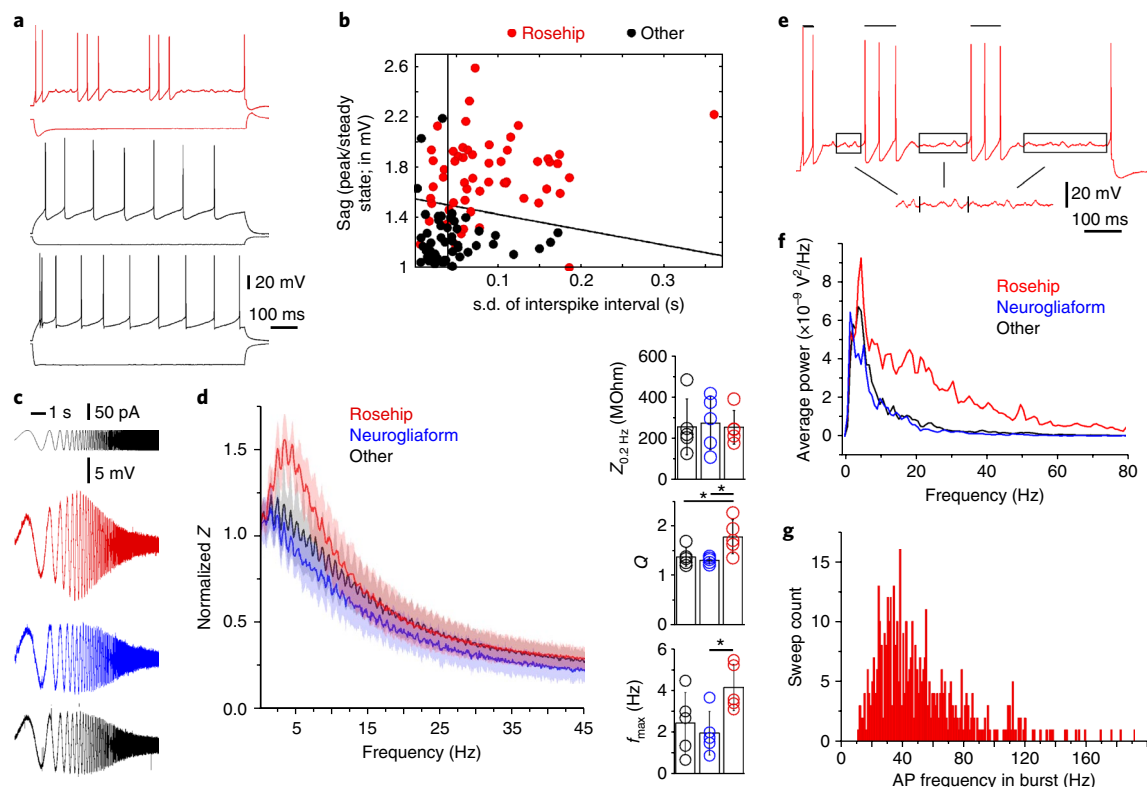
**Fig. 3 | Molecular phenotype of RCs in layer 1 of the human cerebral cortex. a**, Whole-cell recorded and biocytin-filled (red) RCs show CCK immunopositivity (green;  $n=10$ ). All biocytin-labeled RCs (red) tested for CB1 cannabinoid receptors (*CNR1*;  $n=11$ ), somatostatin (*SST*;  $n=9$ ), and calretinin (*CALB2*;  $n=2$ ) were immunonegative, in spite of having labeled cells in the vicinity. Scale bars, 10  $\mu$ m. **b**, Violin plots of gene expression for broad cell type and putative rosehip-specific markers. Expression is on a linear scale and dots indicate median expression. Cluster sample sizes: i2 ( $n=77$ ); i1 ( $n=90$ ); i5 ( $n=47$ ); i3 ( $n=56$ ); i4 ( $n=54$ ); i7 ( $n=31$ ); i10 ( $n=16$ ); i6 ( $n=44$ ); i8 ( $n=27$ ); i9 ( $n=22$ ); i11 ( $n=6$ ); *Smad3* ( $n=12$ ); *Ndnf* + *Car4* ( $n=24$ ); *Ndnf* + *Cxcl14* ( $n=30$ ); *Igtp* ( $n=10$ ). Expression validated for select genes by immunohistochemistry (red stars), colorimetric ISH (black stars), multiplex fluorescent ISH (FISH; orange stars), and single-cell digital PCR (blue) in morphologically identified RCs. **c**, ISH of select marker genes in human temporal cortex (left) and mouse cortex (right). Red arrows highlight cells expressing genes in layer 1. Scale bars, 250  $\mu$ m (low magnification); 100  $\mu$ m (high magnification). ISH experiments were repeated on multiple human donors as follows: *LAMP5* ( $n=2$ ); *EYA4* + *CPLX3* ( $n=3$ ); *SV2C* ( $n=5$ ). For mouse, ISH experiments were repeated on multiple specimens as follows: *Lamp5* + *Sv2c* + *Cplx3* ( $n=2$ ); *Eya4* ( $n=3$ ). **d**, Multiplex FISH validation of rosehip marker co-expression. Arrowheads and arrows show examples of RCs that are triple- and double-positive (i.e., *CNR1*), respectively, for marker genes based on RNA-seq expression data. Scale bar, 25  $\mu$ m. Multiplex FISH experiments were repeated on  $n=2$  tissue donors. **e**, RCs comprise 10–15% of layer 1 interneurons, based on multiplex FISH quantification of 408 *GAD1* cells in 2 subjects. 15% ( $\pm 3$ ) of *GAD1* cells express the rosehip specific marker *PDGFRA*, although a small fraction of these cells may be oligodendrocyte precursor cells (see Supplementary Fig. 5). 10% ( $\pm 1$ ) of *GAD1* cells express *PDGFRA* and a second rosehip marker, *TRPC3*, although some RCs may lack *TRPC3* expression based on RNA-seq. Error bars represent s.d. Cell counts were conducted on  $n=3$  tissue sections from  $n=2$  tissue donors. **f**, Expression of rosehip cluster markers in cytoplasm of whole-cell-recorded RCs, quantified by single-cell digital PCR and reported as a percentage of housekeeping gene (*TBP*) expression in  $n=9$  cells (*CNR1*) or  $n=4$  cells (*CCK*, *CPLX3*, *NDNF*, *SV2C*, *TRPC3*) per gene. Note that *NDNF* expression was not detected in any of the cells tested. Columns and error bars represent mean and s.d.

*SST* ( $n=9$ ), and calretinin (*CALB2*;  $n=2$ ; Fig. 3a). Furthermore, RCs were immunopositive for GABA ( $n=2$ ) and for chicken ovalbumin upstream promoter transcription factor II (*NR2F2*;  $n=2$ ) and negative for parvalbumin ( $n=3$ ), neuronal nitric oxide synthase ( $n=4$ ), neuropeptide Y ( $n=2$ ), calbindin ( $n=2$ ), and choline acetyltransferase ( $n=3$ ; Supplementary Fig. 5a).

Critically, this immunohistochemical profile aligned closely with a single transcriptomic cell type, i5, which was similarly *GAD1*<sup>+</sup>*CCK*<sup>+</sup> but *CNR1*<sup>−</sup>*SST*<sup>−</sup>*CALB2*<sup>−</sup>*PVALB*<sup>−</sup> (Fig. 3b). This putative rosehip transcriptomic type, one of the most distinctive layer 1 GABAergic transcriptomic types, expresses many other genes either highly specifically expressed or co-expressed in only one other layer 1 cell type. Notably, given the rosehip synaptic phenotype, these markers include many genes known to be associated with axon

growth and synaptic structure and function, including synaptic vesicle glycoprotein 2c (*SV2C*), *LAMP5*, transient receptor potential cation channel subfamily C member 3 (*TRPC3*), complexin 3 (*CPLX3*), neurotrophin 3 (*PRSS12*), netrin G1 (*NTNG1*), histamine receptor H1 (*HRH1*), receptor tyrosine kinase-like orphan receptor 2 (*ROR2*), somatostatin receptor 2 (*SSTR2*), and taxilin beta (*TXLNB*).

Since the rosehip anatomical phenotype has not been described in rodents, we asked whether a transcriptomic signature similar to the rosehip transcriptomic type had been observed in a recent large-scale analysis of mouse primary visual cortex using single-cell RNA-seq analysis<sup>4</sup>. We attempted to find homologous cell types between species by correlating the median expression of 212 cell-type-informative genes between all pairs of mouse and human clusters (Supplementary Fig. 3d,e). Expression correlations were quite



**Fig. 4 | Intrinsic electrophysiological properties of RCs.** **a**, Examples of different firing patterns induced by current injections in layer 1 interneurons. Firing pattern of an RC (top), an NGFC (middle), and an unidentified layer 1 interneuron (bottom). **b**, SVM-based wrapper-feature selection of electrophysiological parameters for the identification of RCs. Anatomically identified RCs (red dots) and other types of interneurons with known morphology (black dots) are mapped to the distribution of electrophysiological features ranked as the two best delineators by SVM. Black lines show the best hyperplane separating RCs from other interneuron types. **c,d**, RCs exhibit a distinct impedance profile relative to neurogliaform and other human interneurons in layer 1. **(c)** Individual responses of anatomically identified rosehip (red), neurogliaform (blue), and other (black) interneurons to current injections with an exponential chirp (0.2–200 Hz; top). Traces were normalized to the amplitude of the rosehip response at 200 Hz. **(d)** Left: normalized impedance ( $Z$ ) profiles of distinct groups of interneurons. RCs ( $n=5$ ) had higher impedance in the range of 0.9–12.4 Hz compared to neurogliaform ( $n=5$ ) and other ( $n=5$ ) interneurons. Shaded regions represent s.d. Right: impedances were similar at the lowest frequency ( $Z_{0.2\text{ Hz}}$ ; left), but resonance magnitude ( $Q$ ) calculated as maximal impedance value divided by the impedance at lowest frequency (middle) and frequencies of maximal impedance ( $f_{\text{max}}$ ; right) showed significant differences ( $P < 0.05$ , ANOVA with Bonferroni post hoc correction). **e**, Automated selection of recording periods for assessing subthreshold membrane potential oscillations (boxed segments) and detection of bursts (bars) for measuring intraburst spiking frequency demonstrated on an RC response to near-rheobasic stimulation showing stuttering firing behavior. **f**, Averaged fast Fourier transforms (FFTs) of membrane potential oscillations had higher power between 3.8 and 80 Hz in RCs compared to neurogliaform and other interneurons. **g**, Intraburst frequency of RCs peaked in the gamma range. AP, action potential.

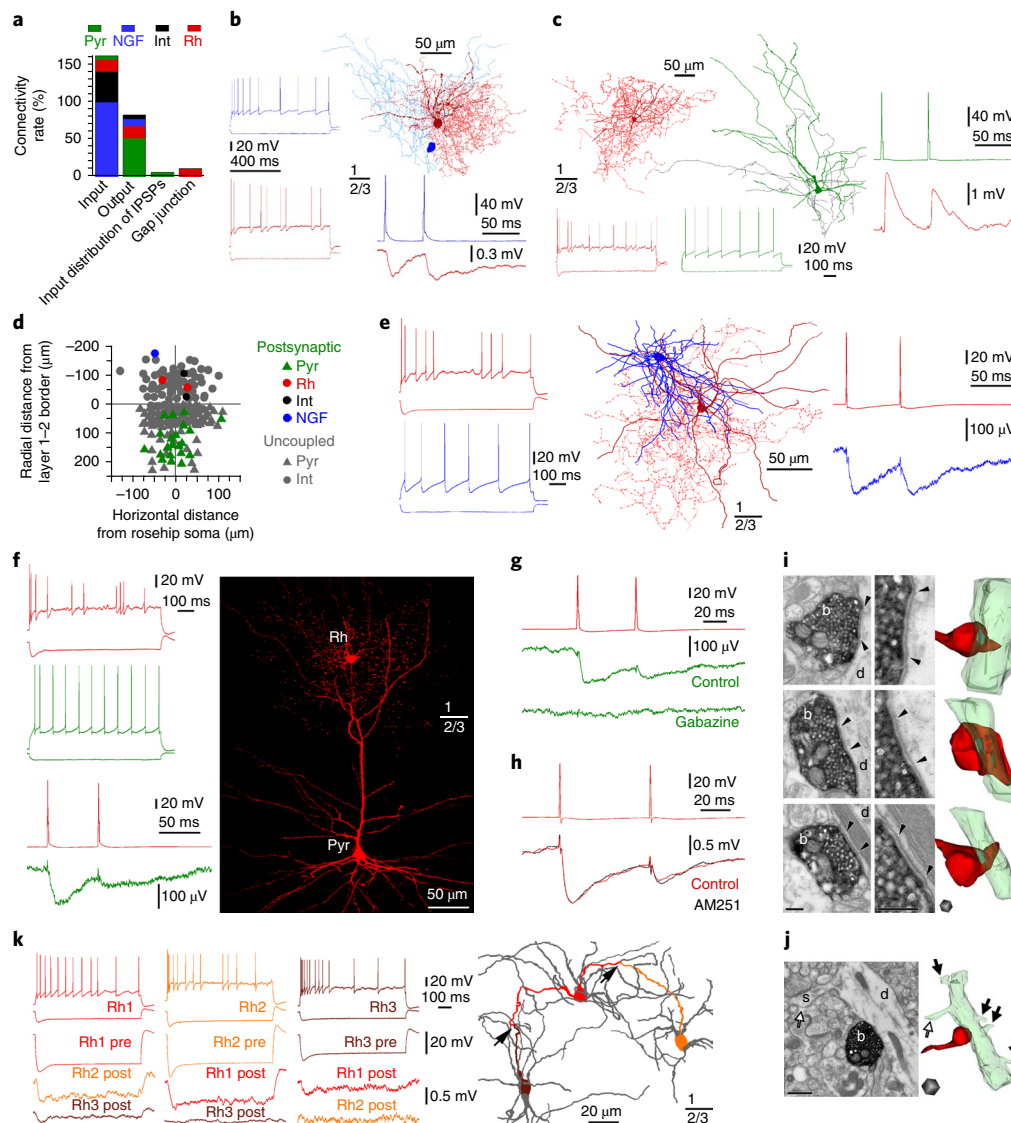
low ( $r < 0.5$ ), and clusters could only be reliably grouped into broad classes of cell types. Human clusters i5, i7, i1, i4, and i3 matched mouse *Smad3*<sup>+</sup> and *Ndnf*<sup>+</sup> clusters. Clusters i6, i8, i9, i11, and i10 matched mouse *Vip*<sup>+</sup> clusters, and cluster i2 matched mouse *Sst*<sup>+</sup> clusters. The rosehip cluster (i5) had a weak ( $r=0.34$ ) reciprocal best match to mouse cluster *Smad3*, although several other human clusters (i7, i1, i4, and i3) matched *Smad3* almost as well ( $r > 0.3$ ).

Many rosehip marker genes are not expressed in *Smad3* or other *Pvalb*<sup>+</sup>*Sst*<sup>+</sup>*Vip*<sup>+</sup> mouse cell types (Fig. 3b) or in the complete set of mouse GABAergic types (Supplementary Fig. 5b). Notably, it is the unique combinatorial expression of many marker genes that defines RCs. For example, expression of *LAMP5*, *SV2C*, *EYA4*, and *CPLX3* is seen by ISH in human layer 1 (Fig. 3c); similarly, as predicted by transcriptomics, three of these four genes are also expressed in mouse layer 1, whereas cells expressing *Eya4* are extremely rare. Many other rosehip-selective genes had no evidence of expression in layer 1 interneurons in mouse, based on single cell transcriptomics (Fig. 3b and Supplementary Fig. 5b).

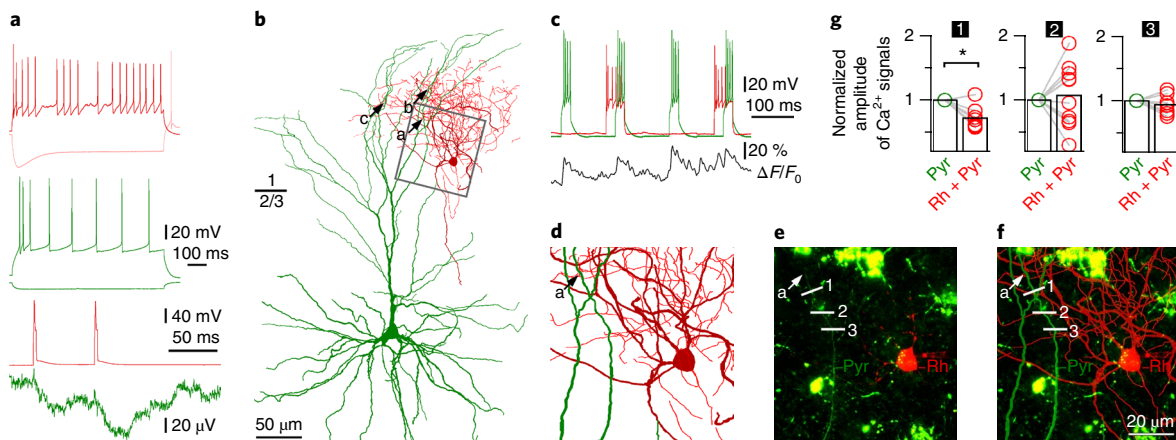
To demonstrate that layer 1 neurons with combinatorial expression patterns predicted by transcriptomics could be found in human

layer 1, and to quantify their proportions, we systematically performed triple-fluorescent ISH on human MTG tissue using discriminating positive and negative gene markers. For all combinations tested we observed cells with the predicted profiles. For example, we observed *CCK*<sup>+</sup>*CNR1*<sup>+</sup>*LAMP5*<sup>+</sup>, *CCK*<sup>+</sup>*PDGFRA*<sup>+</sup>*SOX13*<sup>+</sup>, and *CCK*<sup>+</sup>*TRPC3*<sup>+</sup>*CPLX3*<sup>+</sup> cells, as well as cells in which *CCK* was swapped with other positive rosehip markers (Fig. 3d; additional gene combinations shown in Supplementary Fig. 5c). Quantification of cell proportions using marker expression is complicated by two factors: first, markers for one cell type are often expressed in others; and second, individual markers are often not expressed in every cell in a cluster. We used the combination of *GAD1*, *PDGFRA*, and *TRPC3* to quantify the proportion of RCs among layer 1 GABAergic neurons (Fig. 3e). *PDGFRA* is known to be expressed in OPCs at extremely high levels as well (which is why it appears to only be expressed in OPCs in Fig. 1 but appears high in RCs in Fig. 2 once levels are not normalized across all cell types including OPCs). *PDGFRA*<sup>+</sup> cells represent ~15% of *GAD1*<sup>+</sup> cells, which we therefore considered an upper bound. On the other hand, *TRPC3* is not expressed in all cells in the rosehip cluster. The proportion of *GAD1*<sup>+</sup>





**Fig. 5 | Connections of rosehip cells in the local microcircuit.** **a**, Distribution of local connections mapped in layers 1–3 between RCs (rh, red), pyramidal cells (pyr, green), NGFCs (NGF, blue), and other types of layer 1 interneuron (int, black), based on unbiased targeting of postsynaptic cells. RCs predominantly innervate pyramidal cells, receive monosynaptic EPSPs from layer 2/3 pyramidal cells, and receive monosynaptic IPSPs from neurogliaform and other types of interneurons; IPSPs arriving from RCs were not encountered. In addition, RCs are interconnected by homologous electrical synapses (gap junctions). **b**, Example of an NGFC-to-RC connection. Left: firing patterns of the presynaptic NGFC (blue) and the postsynaptic RC (red). Right: anatomical reconstruction of the recorded NGFC (soma, dark blue; axon, light blue) and RC (soma and dendrites, burgundy; axon, red). Action potentials in the NGFC (blue) elicited slow IPSPs in the RC (red). **c**, Example of a pyramidal cell-to-RC connection. Left: anatomical reconstruction and firing pattern of the presynaptic pyramidal cell (firing, soma and dendrites, green; axon, black) and the postsynaptic RC (firing, soma and dendrites, burgundy; axon, red). Right: action potentials in the pyramidal cell (green) elicited EPSPs in the RC (burgundy). **d**, Spatial distribution of coupled and uncoupled neurons tested as postsynaptic targets of RCs. Note the relative dominance of layer 2/3 pyramidal cells among neurons receiving input from RCs. **e**, The only RC-to-NGFC connection successfully tested for synaptic coupling. Left: firing patterns of the presynaptic RC (burgundy) and the postsynaptic NGFC (blue). Middle: anatomical reconstruction of the RC (soma and dendrites, burgundy; axon, red) and the NGFC (soma and dendrites, blue; axon not shown). Right: action potentials in the RC (red) elicited slow IPSPs in the NGFC (blue). **f**, Example of RC-to-layer 3 pyramidal cell connections ( $n=16$ ). Left: firing patterns of the presynaptic RC (red) and the postsynaptic pyramidal cell (green). Action potentials in the RC (burgundy) elicited IPSPs in the pyramidal cell (green). Right: confocal fluorescence image showing the recorded RC (rh) forming its axonal cloud in the tuft of the apical dendrite of the layer 2/3 pyramidal cell (pyr). **g**, Pharmacological characterization of an RC-to-pyramidal cell connection. Presynaptic spikes in the RC (red) elicited IPSPs in the layer 2/3 pyramidal cell (green), which could be blocked by application of gabazine ( $n=4$ ,  $10\ \mu\text{M}$ ). **h**, Functional test of presynaptic CNR1 expression in RCs shows the absence of modulation by the CNR1 antagonist AM251 ( $n=4$ ). Presynaptic spikes in RC 1 (red, top) elicited IPSPs in RC 2 (red, bottom). Application of AM251 ( $5\ \mu\text{M}$ ) had no effect on IPSPs (black). **i**, Representative electron microscopic images (left) and 3D reconstructions (right,  $n=31$ ) showing axon terminals (b; red) of biocytin-filled RCs ( $n=3$ ) targeting exclusively dendritic shafts (d; green; 100%,  $n=31$ ). Synaptic clefts are indicated between arrowheads. Scale bars, 200 nm. **j**, Representative electron microscopic image (left) and 3D reconstruction (right) of a biocytin-filled RC bouton (b; red) targeting a pyramidal dendritic shaft (d; green), identified based on emerging dendritic spines (s; arrows). Scale bars, 500 nm. **k**, RCs form a network of electrical synapses. Top left: firing patterns of three RCs (red, rh1; orange, rh2; burgundy, rh3). Bottom left: hyperpolarization of RC rh1 was reciprocally transmitted to RCs rh2 and rh3, confirming electrical coupling. Right: route of the hyperpolarizing signals through putative dendrodendritic gap junctions (arrows) between RCs rh1, rh2, and rh3 shown by corresponding colors in the dendritic network of the three cells (gray). Pre, before hyperpolarization; post, after hyperpolarization.



**Fig. 6 | Human rosehip interneurons perform segment-specific regulation of action potential backpropagation to apical dendritic tufts of pyramidal cells.** **a**, Top: firing patterns of a presynaptic RC (burgundy) and a postsynaptic pyramidal cell (green). Bottom: action potentials in the RC (burgundy) elicited IPSPs in the pyramidal cell (green). **b**, Anatomical reconstruction of the RC (soma and dendrites, burgundy; axon, red) and the layer 2/3 pyramidal cell (soma and dendrites, green; axon not shown). Presynaptic axonal boutons of the RC formed close appositions (a, b, and c) with three separate branches on the tuft of the pyramidal apical dendrite. **c**, Repetitive burst firing was triggered to initiate backpropagating  $\text{Ca}^{2+}$  signals in the pyramidal cell (green) while the output of the RC (red) was switched on and off, timed prior to and during every second pyramidal burst. Simultaneously,  $\text{Ca}^{2+}$  dynamics of the pyramidal apical dendritic tuft was measured at several locations, and signals detected at location no. 1 in **e** and **f** are shown in black. **d**, Boxed area in **b** shows the dendritic branch of the apical tuft of the pyramidal cell (green) with a putative synaptic contact (a) arriving from the RC. **e**, Confocal z-stack image of the same area shown in **d** taken during paired whole-cell recordings. The soma of the RC (rh, red), the dendrite of the pyramidal cell (pyr, green), the putative synaptic contact (a), arriving from the RC to the pyramidal cell, and sites of line scans performed across the dendrite (1, 2, and 3) are indicated. Cytoplasmic lipofuscin autofluorescence characteristic to human tissue is seen as green patches. The experiment was repeated independently with similar results in  $n=4$  cell pairs. **f**, Superimposition of the anatomical reconstruction in **d** and the confocal image in **e**. **g**, Normalized amplitudes of  $\text{Ca}^{2+}$  signals during pyramidal cell firing with and without coactivation of the RC detected at the three sites of line scans (1, 2, and 3) on the pyramidal dendrite. Rosehip input simultaneous with the backpropagating pyramidal action potentials was significant ( $P=0.02$ ) in suppressing  $\text{Ca}^{2+}$  signals only at site 1, which was closest ( $8\mu\text{m}$ ) to the putative synapse between the two cells; no effect ( $n=10$  trials,  $P=1.0$  and  $P=0.27$ , respectively, two-sided Wilcoxon test) of the RC was detected at sites 2 and 3, located at distances of 21 and  $28\mu\text{m}$ , respectively, from the putative synaptic contact.

cells that are  $\text{PDGFRA}^+\text{TRPC3}^+$  was  $\sim 10\%$ , which we therefore considered a lower bound. The triple-positive cells for this combination were sparsely distributed across layer 1, although not restricted to this layer (Supplementary Fig. 5d). To determine whether cells with the transcriptional signature of RCs could be found in cortical regions outside of MTG, we conducted triple-fluorescent ISH on tissue sections from Brodmann Area (BA) 9 (frontal cortex) and BA40 (parietal cortex) using several combinatorial gene panels. We found that  $\text{GAD1}^+$  cells expressing rosehip marker genes with low or absent expression of  $\text{CNR1}$  were present in layer 1 of both BA9 and BA40 (Supplementary Fig. 6), suggesting that this cell type is found in the cortical areas sampled for morphoelectric profiling. Furthermore, Lake et al.<sup>22</sup> identified cluster In4 (the best match to rosehip cluster i5) in all six cortical areas sampled (frontal, temporal, and visual cortex).

Finally, to more concretely link morphologically and transcriptionally defined RCs, we performed digital PCR for additional marker genes on cellular content extracted from individual rosehip neurons. As predicted by the transcriptome data, RCs were positive for *CCK*, *CPLX3*, *SV2C*, and *TRPC3*, and low (*CNR1*) or absent (*NDNF*) for genes not expressed by cells in that cluster (Fig. 3f). Together, these data strongly link the anatomically defined rosehip phenotype with a highly distinctive transcriptomic cell type signature that is found in human but apparently not in mouse layer 1.

### Intrinsic electrophysiological properties of rosehip cells

Anatomically identified RCs responded to long (800ms) suprathreshold current injections with stuttering or irregular spiking firing patterns<sup>2</sup> when activated from resting membrane potential ( $-61.34 \pm 5.8\text{mV}$ ; Fig. 4a). Analysis of silent and suprathreshold

periods during rheobasic firing of RCs indicated that membrane oscillations and firing of RCs were tuned to beta and gamma frequencies (Fig. 4e–g). The power of averaged fast Fourier transforms of subthreshold membrane potential oscillations<sup>27</sup> was higher between 3.8 and 80 Hz in RCs compared to neurogliaform and other interneurons (Fig. 4f), and intraburst frequency of stuttering firing also peaked in the beta–gamma range (Fig. 4g). The s.d. of interspike intervals was higher in RCs ( $87 \pm 64\text{ms}$ ,  $n=55$ ) compared to neurogliaform ( $41 \pm 34\text{ms}$ ,  $n=16$ ,  $P<0.001$ , Wilcoxon test) or unclassified interneurons ( $47 \pm 41\text{ms}$ ,  $n=36$ ,  $P<0.001$ , Wilcoxon test), indicating alternating silent and active periods during rheobasic stimulation. As described previously<sup>23</sup>, human interneurons recorded in layer 1 had a characteristic lag when responding to hyperpolarizing current pulses. However, the amplitude of the lag measured in anatomically classified RCs ( $1.73 \pm 0.30$ ,  $n=55$ ) exceeded that of interneurons morphologically identified as NGFCs ( $1.19 \pm 0.12$ ,  $n=16$ ,  $P<0.001$ , Wilcoxon test) or unclassified interneurons ( $1.29 \pm 0.28$ ,  $n=36$ ,  $P<0.001$ , Wilcoxon test). Input resistances of RCs ( $139.6 \pm 54.1\text{M}\Omega$ ) were similar to those of NGFCs ( $160.1 \pm 55.9\text{M}\Omega$ ) and lower compared to other interneurons ( $216.3 \pm 84.4\text{M}\Omega$ ,  $P<0.001$ , Wilcoxon test); however, time constants of RCs ( $7.3 \pm 3.7\text{ms}$ ) were similar to those of neurogliaform cells ( $8.9 \pm 2.4\text{ms}$ ,  $P<0.001$ ) and faster compared to other cells ( $11.1 \pm 12.5\text{ms}$ ,  $P<0.001$ ). Anatomically identified RCs showed distinct impedance profiles relative to other layer 1 interneurons in response to current injections, with an exponential chirp (0.2–200 Hz; Fig. 4c,d). Impedance at the lowest frequency ( $Z_{0.2\text{Hz}}$ ) was similar in layer 1 interneurons (rosehip,  $258 \pm 81\text{M}\Omega$ ; neurogliaform,  $279 \pm 128\text{M}\Omega$ ; unclassified,  $261 \pm 133\text{M}\Omega$ ; Lilliefors test followed by one-way ANOVA with Bonferroni correction). The resonance magnitude ( $Q$ ; see Methods)



of RCs ( $1.77 \pm 0.34$ ) was significantly higher compared to those of NGFCs ( $1.31 \pm 0.07$ ;  $P < 0.021$ , Lilliefors test followed by one way ANOVA with Bonferroni correction) and unclassified interneurons ( $1.37 \pm 0.19$ ;  $P < 0.049$ ). In addition, frequencies of maximal impedance ( $f_{\max}$ ) in RCs ( $4.17 \pm 1.1$  Hz) were significantly higher than in NGFCs ( $1.98 \pm 1.04$  Hz;  $P < 0.045$ ) but the difference was not significant compared to unclassified interneurons ( $2.47 \pm 1.47$  Hz,  $P < 0.142$ ). We did not find substantial differences between NGFCs and unclassified interneurons in impedance parameters. Support vector machine (SVM)-based wrapper-feature selection of electrophysiological parameters ranked the amplitude of the lag and the s.d. of interspike intervals as the two best delineators out of  $n = 200$  measured electrophysiological parameters for separating anatomically identified rosehip, neurogliaform, and unclassified interneurons in layer 1 (Fig. 4b). Indeed, the best hyperplane separating RCs from other interneuron types according to SVM analysis had a false positive rate of 0% for identifying RCs ( $n = 37$ ) in the total population of anatomically recovered layer 1 interneurons ( $n = 107$ ). Thus, we included cells defined by the hyperplanes of SVM analysis, referred to as 'SVM-identified RCs', when anatomical recovery was lacking in some experiments, as indicated below.

### Function of rosehip cells in local microcircuits

To assess functional connectivity of RCs in the local microcircuit, we established recordings from RCs and then searched for potential pre- and postsynaptic partners without any cell-type preference in an area of brain slices within a horizontal and vertical radius of  $\sim 100 \mu\text{m}$  and  $\sim 200 \mu\text{m}$ , respectively (Fig. 5a–f). Monosynaptic input connections ( $n = 226$ ) were tested on anatomically ( $n = 43$ ) and SVM-identified ( $n = 24$ ) RCs. Presynaptic layer 1 interneurons were connected to RCs with an overall coupling ratio (CR) of 45%. GABAergic cells evoking inhibitory postsynaptic potentials (IPSPs) on RCs included layer 1 NGFCs ( $n = 10$ , CR 100%), RCs ( $n = 2$ , CR 17%), and unclassified interneurons ( $n = 14$ , CR 40%); however, none of the tested interneurons ( $n = 9$ ) with somata in layer 2 (defined as  $< 70 \mu\text{m}$  below the layer 1–2 border) were connected to RCs. Fast components of IPSPs arriving to RCs evoked by different presynaptic interneurons had similar amplitudes ( $0.982 \pm 0.705$  mV,  $0.915 \pm 0.594$  mV, and  $1.504 \pm 1.308$  mV, respectively) and showed paired-pulse depression with paired-pulse ratios of  $0.42 \pm 0.48$ ,  $0.27 \pm 0.04$ , and  $0.71 \pm 0.26$ , respectively. RCs received local excitatory inputs from layer 2/3 pyramidal cells sporadically ( $n = 8$ , CR 5%), with monosynaptic excitatory postsynaptic potential (EPSP) amplitudes of  $3.357 \pm 1.458$  mV and paired-pulse ratios of  $0.68 \pm 0.12$ . Very large unitary EPSPs, described as driving human basket and axo-axonic cells to suprathreshold postsynaptic responses<sup>10,11,14</sup>, were not encountered on RCs. Thus, local inputs to RCs appear to be predominantly GABAergic, with the caveat that some axon collaterals of pyramidal cells were cut during the slicing procedure (Fig. 5c), leading to a potential underrepresentation of pyramidal cell-triggered EPSPs.

In turn, monosynaptic output connections triggered by anatomically ( $n = 49$ ) and SVM-identified ( $n = 13$ ) RCs rarely innervated postsynaptic interneurons (overall CR 8%). Even though an NGFC ( $n = 1$ , CR 10%), RCs ( $n = 2$ , CR 17%), unclassified layer 1 interneurons ( $n = 2$ , CR 5%), and superficial layer 2 pyramidal cells ( $n = 5$ , CR 5%) were targeted when testing a total number of  $n = 197$  connections, RC outputs were predominantly directed toward layer 3 pyramidal cells ( $n = 16$ , CR 46%) having somata  $> 70 \mu\text{m}$  below the layer 1–2 border. IPSPs elicited by RCs were mediated by GABA<sub>A</sub> receptors, based on experiments showing blockade of IPSPs by application of the GABA<sub>A</sub>-receptor antagonist gabazine ( $n = 4$ ,  $10 \mu\text{M}$ ; Fig. 5g). Amplitudes of RC-triggered IPSPs arriving to interneurons ( $0.428 \pm 0.370$  mV) were larger compared to those targeting layer 3 pyramidal cells ( $0.087 \pm 0.059$  mV,  $P < 0.05$ , MW  $U$  test), in agreement with dendritic filtering of distally elicited IPSPs

during signal propagation along the apical dendrite to the somatically placed electrode. The results above indicate that RCs in layer 1 might preferentially target pyramidal cells sending terminal branches of their apical dendrites to layer 1. Indeed, when randomly sampling the output formed by RCs ( $n = 6$ ) using serial electron microscopic sections, we found that axon terminals ( $n = 64$ ) exclusively targeted dendritic shafts (Fig. 5i). Moreover, further ultrastructural analysis of postsynaptic dendrites ( $n = 46$ ) revealed dendritic spines and sparse innervation by symmetrical synapses on the shaft, suggesting that these dendrites belonged to pyramidal cells ( $n = 41$ , 86%; Fig. 5j). The remaining  $n = 5$  (11%) dendrites had no spines and received asymmetric synapses on the shaft; they were likely formed by interneurons.

Previous studies on rodent cortical interneurons containing CCK show functional presynaptic expression of the CB1 cannabinoid receptor<sup>28</sup>; however, application of the CB1-receptor antagonist AM251 was ineffective in modulating RC evoked IPSPs ( $n = 4$ ; Fig. 5h), supporting our results from single-cell digital PCR, immunohistochemistry, and ISH data (Fig. 3). Earlier reports on human microcircuits identified single-cell-triggered polysynaptic network events<sup>10,11,14</sup>. We found that RCs were involved in single-cell-activated ensembles detected through disynaptic IPSPs triggered by layer 2 ( $n = 1$ ) and layer 3 ( $n = 2$ ) pyramidal cells and through polysynaptic EPSPs triggered by an axo-axonic cells (data not shown). In addition to mono- and polysynaptic chemical synaptic communication, human interneurons are also involved in gap-junctional signaling<sup>23</sup>. RCs also formed homologous electrical synapses ( $n = 5$ , CR 57%) between each other (Fig. 5k) and established convergent heterologous electrical synapses ( $n = 2$ , CR 11%) with an unclassified layer 1 interneuron. When applying hyperpolarizing current steps in the first neuron to elicit response in the second neuron, coupling coefficients for gap junctions ( $0.05 \pm 0.05$ ) were found to be similar to those observed previously between human and rodent interneurons<sup>23</sup>.

Preferential placement of output synapses on distal dendritic shafts of pyramidal cells reaching layer 1 suggest that RCs might specialize in the control of dendritic signal processing. RCs established  $2.6 \pm 1.5$  (range: 1–4) close appositions on dendrites of layer 3 pyramidal cells at distances of  $290 \pm 98 \mu\text{m}$  (range: 94–455  $\mu\text{m}$ ) from the somata of postsynaptic pyramidal cells ( $n = 5$ ). We found a correlation between the rise times of IPSPs arriving to the postsynaptic pyramidal cells ( $n = 5$ ,  $7.3 \pm 2.4$  ms, range: 3.8–10.1 ms) and the distances of close axodendritic appositions from the somata ( $\rho = 0.90$ ,  $P = 0.037$ , Spearman correlation). In dual recordings of synaptically connected RCs to pyramidal cell pairs ( $n = 6$ ), we loaded RCs with Alexa Fluor 594 to label presynaptic axons and filled the postsynaptic pyramidal cells with Oregon Green BAPTA 1 to structurally map the course of dendrites and to measure dendritic  $\text{Ca}^{2+}$  dynamics (Fig. 6). The amplitude of the IPSPs triggered by the first action potential of RCs and evoked on distal dendrites of the postsynaptic pyramidal was  $35.6 \pm 24.7 \mu\text{V}$  at the soma (Fig. 6a,b). Backpropagation of action potentials to dendrites of human neurons has been shown in previous studies<sup>13,29</sup>, and we confirmed these results by detecting dendritic  $\text{Ca}^{2+}$  responses following somatically elicited burst firing (100-ms current injections, four spikes per burst) in layer 3 pyramidal cells. Changes in  $\Delta F/F$  ( $17.2 \pm 7.3\%$ ) in distal branches of the apical dendrites in layer 1 were consistently detected at multiple ( $17 \pm 8$ ) locations on the postsynaptic neurons, confirming action potential backpropagation into distal apical dendritic branches of human pyramidal cells (Fig. 6c). We chose regions of interest on Oregon Green BAPTA 1-filled branches of the postsynaptic apical dendrites overlapping with the Alexa Fluor 594-labeled axonal arborization of presynaptic RCs and triggered somatically evoked bursts in the pyramidal cells alone, for control, and together with bursts in the RC, in an alternating fashion (Fig. 6b–f). Inputs from RCs simultaneous with backpropagating

action potentials were effective in suppressing the amplitude of  $\text{Ca}^{2+}$  signals relative to control ( $n=6$ ,  $12.8 \pm 4.6\%$  vs.  $18.8 \pm 5.7\%$   $\Delta F/F$ ,  $P < 0.02$ , Wilcoxon test; Fig. 6c) in one or two locations heuristically line-scanned on dendrites of postsynaptic cells (Fig. 6g). The anatomical arrangement of presynaptic axons and imaged segments of postsynaptic dendrites was recovered in  $n=4$  pairs. Rosehip inputs simultaneous with backpropagating pyramidal cell action potentials were effective in suppressing  $\text{Ca}^{2+}$  signals only at sites that were neighboring ( $8 \pm 5 \mu\text{m}$ ) the putative synapses between the two cells. No effect of RCs was detected at dendritic sites one step further in distance ( $21 \pm 14 \mu\text{m}$ ; Fig. 6d–f). This suggests that RCs specialize in providing tightly compartmentalized control of dendritic  $\text{Ca}^{2+}$  electrogenesis of human pyramidal cells, thereby enforcing inhibitory microdomains in dendritic computation.

## Discussion

Understanding the cellular makeup of the cortex and understanding its conservation across species represents twin challenges difficult to address in human tissue. Historically, forming a representative overview of cell-type diversity in a particular brain region has been achieved based on molecular marker expression cross-referenced to axonal and dendritic morphology<sup>3,4,21,30</sup>. Conserved patterns of molecular and morphophysiological features for particular cell classes have been reported<sup>2,23,24</sup>, but interspecies variation<sup>26,31–34</sup> and cell types potentially characteristic to several species have also been described<sup>35–37</sup>. Recent studies have overcome some of the difficulties associated with the scarcity of human tissue of sufficient quality<sup>9–14,22,24,25,31,38</sup>, propelling understanding of human circuits. Here we demonstrate the strength of a modern version of this approach that can be applied to human postmortem and neurosurgical tissues. Single-nucleus transcriptomics provides the scale for an unbiased survey of molecular expression, while human slice physiology characterizes the functional properties of those types. Together these approaches provide convergent evidence for robust cell-type identities and concomitant evidence for species conservation or specialization.

The targeted application of single-nucleus sequencing reported here has demonstrated a substantially higher degree of GABAergic neuron complexity in just one layer of the human cerebral cortex (ten types) than what has previously been described in all of the cortex (eight types<sup>22</sup>). This difference is likely due to a combination of improved sequencing technique and increased sampling in a targeted anatomical domain enriched in GABAergic neurons. This diversity also appears to be higher than described for layer 1 in mouse<sup>4</sup>, although by covering all layers the study by Tasic, et al., likely underrepresented layer 1. Indeed, a recent characterization of rat cortex<sup>3</sup> described six morphological and 17 morphoelectric types in layer 1, so our results are consistent with the neuronal diversity described using other methods. RCs represent a type with a highly distinctive transcriptomic signature; a highly distinctive morphological, physiological, and connectional phenotype; and a strong correspondence between these properties. In this respect, it appears similar to other highly specialized and distinctive cortical cell types, such as chandelier cells<sup>39</sup>. To our knowledge, a similar anatomical cell type has not been described in rodent. While we cannot prove the negative, given the extent of cellular studies of rodent cortex, such cells would have to be either extremely rare or experimentally difficult to study to have escaped detection to date. Similarly, the rosehip molecular marker signature appears highly distinctive from any published data from rodent. Although the transcriptomic comparison is between human temporal cortex and mouse visual cortex, regional differences seem unlikely to account for this difference, as we found the anatomically defined rosehip type in multiple human regions. A complete comparison of all cortical cell types and assessment of relative similarities between cell types should be possible in the future as more

comprehensive transcriptome data become available and linked to other cellular phenotypes in multiple species. Our study is based on a relatively limited number of multimodally characterized cells due to the scarcity of high-quality human samples, and further systematic analyses of human cell types in well-defined cytoarchitectonic areas using increased sample sizes are needed to substantiate further interpretations.

It is widely accepted<sup>28</sup> that CCK<sup>+</sup> cells in the rodent show selectively high expression of cannabinoid receptors and are involved in perisomatic inhibition. The bouton morphology and/or the compact axonal field of RCs resembles that of cell types described in deeper layers of the cat cortex that innervate relatively proximal dendrites (dendrite targeting and clutch cells)<sup>40,41</sup>. In contrast, RCs are CCK<sup>+</sup> but cannabinoid receptor-negative, and appear to selectively target distal dendrites of pyramidal neurons. Moreover, when assessing layer 1 canonical inhibitory pathways with high throughput electrophysiology capable of sampling all cell types in layer 1 in rodent, Lee et al.<sup>42</sup> found two interneuron types and two canonical pathways involving feedforward interneuron-to-interneuron connections. Thus, the monosynaptic pyramidal cell-preferring pathway initiated by RCs does not appear to have a homolog in the rodent layer 1 circuit. Furthermore, focal intralayer inhibition restricted by the compact axonal arbor of RCs to distal dendrites of a column of pyramidal cells is also missing from the rodent; rather, mouse feedforward inhibitory connections are vertically spread to all somatodendritic domains<sup>42</sup>.

Addition of new human cell types, or specialization of existing types through major modification of cellular features, would be expected to alter circuit function<sup>3,43,44</sup> and therefore cannot be studied in rodents. Dissimilarities of RCs and other dendrite-targeting interneurons cannot be fully understood without further experiments testing these differences directly. RCs may be of particular importance in compartmental control of backpropagating action potentials and their pairing with incoming excitatory inputs. The uniquely small membrane capacitance ( $C_m$ ) found in human pyramidal cells<sup>8</sup> promotes backpropagation of action potentials and increases excitability in human dendrites<sup>13,29</sup> relative to rodent dendrites having larger  $C_m$ . Action potentials backpropagate to distal dendrites of human pyramidal cells and can be attenuated by RC activation. Thus, RCs may provide the supplementary inhibitory control required to balance the potentially higher excitability in human dendrites<sup>8</sup> and modulate interactions between long-range excitatory connections arriving to layer 1 and the backpropagating action potentials suggested to participate in interhemispheric modulation<sup>45</sup>. The sharp resonance in the theta-range detected in individual RCs and its potential spread through gap junctions to a rosehip network could phase-selectively interact with long-range inputs similarly to mechanisms suggested (for example) in oscillation dependent memory consolidation<sup>30,46</sup>. The function of neuron types specific to the human circuit could be important in understanding pathological alterations of network functions. For example, several highly selective markers for RCs have been implicated as risk factors for neuropsychiatric disease, including netrin G1 (NTNG1) for Rett syndrome<sup>47</sup> and neurotrypsin (PRSS12) for intellectual disability<sup>48</sup>. A better understanding of human cellular and circuit organization may help counteract the current lack of success in translating promising rodent results to effective treatment against human neuropsychiatric disorders<sup>49,50</sup>.

## Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41593-018-0205-2>.

Received: 3 November 2017; Accepted: 14 June 2018;

Published online: 27 August 2018

## References

- DeFelipe, J. & Jones, E. G. *Cajal on the Cerebral Cortex*. (Oxford University Press, Oxford, UK, 1988).
- Ascoli, G. A. et al. Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nat. Rev. Neurosci.* **9**, 557–568 (2008).
- Markram, H. et al. Reconstruction and simulation of neocortical microcircuitry. *Cell* **163**, 456–492 (2015).
- Tasic, B. et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016).
- Zeisel, A. et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
- Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
- Jiang, X. et al. Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* **350**, aac9462 (2015).
- Eyal, G. et al. Unique membrane properties and enhanced signal processing in human neocortical neurons. *eLife* **5**, e16553 (2016).
- Wang, B. et al. A subtype of inhibitory interneuron with intrinsic persistent activity in human and monkey neocortex. *Cell Rep.* **10**, 1450–1458 (2015).
- Molnár, G. et al. Complex events initiated by individual spikes in the human cerebral cortex. *PLoS Biol.* **6**, e222 (2008).
- Molnár, G. et al. Human pyramidal to interneuron synapses are mediated by multi-vesicular release and multiple docked vesicles. *eLife* **5**, e18167 (2016).
- Testa-Silva, G. et al. Human synapses show a wide temporal window for spike-timing-dependent plasticity. *Front. Synaptic Neurosci.* **2**, 12 (2010).
- Verhoog, M. B. et al. Mechanisms underlying the rules for associative plasticity at adult human neocortical synapses. *J. Neurosci.* **33**, 17197–17208 (2013).
- Szegedi, V. et al. Plasticity in single axon glutamatergic connection interneurons regulates complex events in the human neocortex to GABAergic. *PLoS Biol.* **14**, e2000237 (2016).
- Elston, G. N. Pyramidal cells of the frontal lobe: all the more spinous to think with. *J. Neurosci.* **20**, RC95 (2000).
- Morales, J. et al. Random positions of dendritic spines in human cerebral cortex. *J. Neurosci.* **34**, 10078–10084 (2014).
- Mohan, H. et al. Dendritic and axonal architecture of individual pyramidal neurons across layers of adult human neocortex. *Cereb. Cortex* **25**, 4839–4853 (2015).
- Grindberg, R. V. et al. RNA-sequencing from single nuclei. *Proc. Natl. Acad. Sci. USA* **110**, 19802–19807 (2013).
- Lacar, B. et al. Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat. Commun.* **7**, 11022 (2016).
- Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
- Kawaguchi, Y. & Kubota, Y. GABAergic cell subtypes and their synaptic connections in rat frontal cortex. *Cereb. Cortex* **7**, 476–486 (1997).
- Lake, B. B. et al. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586–1590 (2016).
- Oláh, S. et al. Output of neurogliaform cells to various neuron types in the human and rat cerebral cortex. *Front. Neural Circuits* **1**, 4 (2007).
- Kisvárdy, Z. F. et al. Synapses, axonal and dendritic patterns of GABA-immunoreactive neurons in human cerebral cortex. *Brain* **113**, 793–812 (1990).
- Oláh, S. et al. Regulation of cortical microcircuits by unitary GABA-mediated volume transmission. *Nature* **461**, 1278–1281 (2009).
- Varga, C., Tamas, G., Barzo, P., Oláh, S. & Somogyi, P. Molecular and electrophysiological characterization of GABAergic interneurons expressing the transcription factor COUP-TFII in the adult human temporal cortex. *Cereb. Cortex* **25**, 4430–4449 (2015).
- Zemankovics, R., Káli, S., Paulsen, O., Freund, T. F. & Hájos, N. Differences in subthreshold resonance of hippocampal pyramidal cells and interneurons: the role of h-current and passive membrane characteristics. *J. Physiol. (Lond.)* **588**, 2109–2132 (2010).
- Katona, I. & Freund, T. F. Multiple functions of endocannabinoid signaling in the brain. *Annu. Rev. Neurosci.* **35**, 529–558 (2012).
- Kerekes, B. P. et al. Combined two-photon imaging, electrophysiological, and anatomical investigation of the human neocortex in vitro. *Neurophotonics* **1**, 011013 (2014).
- Freund, T. F. & Buzsáki, G. Interneurons of the hippocampus. *Hippocampus* **6**, 347–470 (1996).
- Hawrylycz, M. J. et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
- Lein, E. S. et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
- Xu, Q., Cobos, I., De La Cruz, E., Rubenstein, J. L. & Anderson, S. A. Origins of cortical interneuron subtypes. *J. Neurosci.* **24**, 2612–2622 (2004).
- DeFelipe, J. Types of neurons, synaptic connections and chemical characteristics of cells immunoreactive for calbindin-D28K, parvalbumin and calretinin in the neocortex. *J. Chem. Neuroanat.* **14**, 1–19 (1997).
- von Economo, C. Eine neue Art Spezialzellen des Lobus cinguli und Lobus insulae. *Zschr. ges. Neurol. Psychiatry* **100**, 706–712 (1926).
- Nimchinsky, E. A., Vogt, B. A., Morrison, J. H. & Hof, P. R. Spindle neurons of the human anterior cingulate cortex. *J. Comp. Neurol.* **355**, 27–37 (1995).
- Gabbott, P. L. A. Subpial Fan Cell™ - a class of calretinin neuron in layer 1 of adult monkey prefrontal cortex. *Front. Neuroanat.* **10**, 28 (2016).
- Miller, J. A. et al. Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014).
- Somogyi, P. A specific 'axo-axonal' interneuron in the visual cortex of the rat. *Brain Res.* **136**, 345–350 (1977).
- Tamás, G., Buhl, E. H. & Somogyi, P. Fast IPSPs elicited via multiple synaptic release sites by different types of GABAergic neuron in the cat visual cortex. *J. Physiol. (Lond.)* **500**, 715–738 (1997).
- Kisvárdy, Z. F., Martin, K. A. C., Whitteridge, D. & Somogyi, P. Synaptic connections of intracellularly filled clutch cells: a type of small basket cell in the visual cortex of the cat. *J. Comp. Neurol.* **241**, 111–137 (1985).
- Lee, A. J. et al. Canonical organization of layer 1 neuron-led cortical inhibitory and disinhibitory interneuronal circuits. *Cereb. Cortex* **25**, 2114–2126 (2015).
- Tremblay, R., Lee, S. & Rudy, B. GABAergic interneurons in the neocortex: from cellular properties to circuits. *Neuron* **91**, 260–292 (2016).
- Gjorgjieva, J., Drion, G. & Marder, E. Computational implications of biophysical diversity and multiple timescales in neurons and synapses for circuit performance. *Curr. Opin. Neurobiol.* **37**, 44–52 (2016).
- Palmer, L. M. et al. The cellular basis of GABA(B)-mediated interhemispheric inhibition. *Science* **335**, 989–993 (2012).
- Klausberger, T. & Somogyi, P. Neuronal diversity and temporal dynamics: the unity of hippocampal circuit operations. *Science* **321**, 53–57 (2008).
- Borg, I. et al. Disruption of Netrin G1 by a balanced chromosome translocation in a girl with Rett syndrome. *Eur. J. Hum. Genet.* **13**, 921–927 (2005).
- Molinari, F. et al. Truncating neurotrophin mutation in autosomal recessive nonsyndromic mental retardation. *Science* **298**, 1779–1781 (2002).
- Mak, I. W., Evaniew, N. & Ghert, M. Lost in translation: animal models and clinical trials in cancer treatment. *Am. J. Transl. Res.* **6**, 114–118 (2014).
- Cavanaugh, S. E., Pippin, J. J. & Barnard, N. D. Animal models of Alzheimer disease: historical pitfalls and a path forward. *ALTEX* **31**, 279–302 (2014).

## Acknowledgements

The authors thank the Allen Institute for Brain Science founders, Paul G. Allen and Jody Allen, for their vision, encouragement, and support. The authors thank L. Christiansen and F. Zhang from Illumina, Inc. for their assistance with RNA sequencing. This work was supported by the ERC Interimpact project (G.T.), the National Institute of Mental Health (USA) RFA MH 17 210 (E.S.L., G.T.), the Hungarian Academy of Sciences (G.T.), the National Research, Development and Innovation Office of Hungary (GINOP-2.3.2-15-2016-00018, VKSZ-14-1-2015-0155), and by the National Brain Research Program, Hungary (G.T.).

## Author contributions

Conceptualization, E.S.L., R.S.L., G.T. Methodology, R.D.H., M.N., J.L.C., P.B., L.G.P., G.T. Validation, J.L.C., S.-L.D., G.M., G.T. Formal analysis, T.E.B., B.D.A., J.M.M., J.A.M., P.V., M.R., S.B., R.H.S., E.B., J.B., G.O., G.M. Investigation, R.D.H., M.N., J.L.C., F.D.-F., S.I.S., K.A.S., A.W., D.N.T., Z.M., E.B., J.B., Á.K.K., N.F., B.K., M.R., G.M., A.O., G.O., G.T. Resources, E.S.L., F.J.S., N.J.S., R.H.S., R.S.L., G.T. Data curation, T.E.B., B.D.A., J.M.M., J.A.M., P.V., S.B. Writing – original draft, T.E.B., R.D.H., J.L.C., J.A.M., E.S.L., E.B., G.O., G.T. Writing – review and editing, T.E.B., R.D.H., J.A.M., R.H.S., R.S.L., E.S.L., G.T. Visualization, T.E.B., R.D.H., J.L.C., S.-L.D., J.A.M., E.B., G.M., G.O. Supervision, E.S.L., F.J.S., N.J.S., R.H.S., R.S.L., G.T. Project administration, E.S.L., S.M.S., G.T. Funding acquisition, E.S.L., F.J.S., R.S.L., G.T.

## Competing interests

The authors declare no competing interests.

## Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41593-018-0205-2>.

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints).

Correspondence and requests for materials should be addressed to E.S.L. or G.T.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



## Methods

**Postmortem human brain specimens.** After obtaining permission from decedent next-of-kin, postmortem adult human brain tissue was collected by the San Diego Medical Examiner's office and provided to the Allen Institute for Brain Science. All tissue collection was performed in accordance with the provisions of the Uniform Anatomical Gift Act described in Health and Safety Code §§ 7150, et seq., and other applicable state and federal laws and regulations. The Western Institutional Review Board reviewed tissue collection processes and determined that they did not constitute human subjects research requiring IRB review. The tissue specimens used in this study were prescreened for known neuropsychiatric and neuropathological history and underwent routine serological testing and toxicology screening. Specimens were further screened for RNA quality and had an RNA integrity number (RIN)  $\geq 7$ . The specimens used for RNA-sequencing in this study were from two individual control Caucasian male donors, aged 50 and 54 years. Postmortem interval was 24 h for both specimens. For multiplex fluorescent *in situ* hybridization (FISH) on frontal and parietal brain regions, postmortem tissue was obtained from two different male Caucasian donors, aged 60 (24 h PMI) and 66 (22 h PMI) years.

**Tissue processing for nuclei isolation.** Whole postmortem brain specimens were bisected through the midline, and individual hemispheres were embedded in alginate for slabbing. Coronal brain slabs were cut at 0.5- to 1-cm intervals through each hemisphere, and the slabs were frozen in a bath of dry ice and isopentane and stored at  $-80^{\circ}\text{C}$ . For RNA-sequencing experiments, middle temporal gyrus (MTG) was identified on slabs of interest and removed for further sectioning. MTG tissue was then thawed in a buffer containing PBS supplemented with 10 mM DL-Dithiothreitol (DTT, Sigma Aldrich), mounted on a vibratome (Leica), and sectioned at 500  $\mu\text{m}$  in the coronal plane. Sections were transferred to a fluorescent Nissl staining solution (NeuroTrace 500/525, ThermoFisher Scientific) prepared in PBS with 10 mM DTT and 0.5% RNasin Plus RNase inhibitor (Promega). After staining for 5 min, sections were visualized on a fluorescence dissecting microscope (Leica), and layer 1 was microdissected using a needle-blade microknife (Fine Science Tools).

**Nuclei isolation and FACS.** Microdissected sections of layer 1 from MTG were transferred into nuclei isolation medium containing 10 mM Tris, pH 8.0 (Ambion), 250 mM sucrose, 25 mM KCl (Ambion), 5 mM  $\text{MgCl}_2$  (Ambion), 0.1% Triton-X 100 (Sigma Aldrich), 1% RNasin Plus, 1  $\times$  protease inhibitor (Promega), and 0.1 mM DTT and placed into a 1-mL dounce homogenizer (Wheaton). Tissue was homogenized to liberate nuclei using 10 strokes of the loose dounce pestle followed by 10 strokes of the tight pestle. Homogenate was strained through a 30- $\mu\text{m}$  cell strainer (Miltenyi Biotec) and centrifuged at 900 g for 10 min to pellet nuclei. Nuclei were then resuspended in staining buffer containing 1  $\times$  PBS (Ambion), 0.8% nuclease-free BSA (Omni-Pur, EMD Millipore), and 0.5% RNasin Plus. Mouse monoclonal anti-NeuN antibody (EMD Millipore, MAB377) was applied to nuclei preparations at a concentration of 1:1,000, and samples were incubated for 30 min at  $4^{\circ}\text{C}$ . Control samples were incubated with mouse IgG1-k isotype control (BD Pharmingen). Samples were then centrifuged for 5 min at 500 g to pellet nuclei, and pellets were resuspended in staining buffer as described above. Nuclei samples were incubated with secondary antibody (goat anti-mouse IgG, Alexa Fluor 594, ThermoFisher Scientific) for 30 min at  $4^{\circ}\text{C}$ , centrifuged for 5 min at 500 g, and resuspended in staining buffer. DAPI (4',6'-diamidino-2-phenylindole, ThermoFisher Scientific) was applied to nuclei samples at a concentration of 1  $\mu\text{g}/\text{mL}$ .

Single-nucleus sorting was carried out on a BD FACSria Fusion instrument (BD Biosciences) using a 130  $\mu\text{m}$  nozzle. Nuclei were first gated on DAPI and then passed through doublet-discrimination gates before being gated on the NeuN (Alexa Fluor 594) signal. Approximately 10% of nuclei were intentionally sorted as NeuN<sup>+</sup> to allow for the collection of non-neuronal nuclei. Single nuclei were sorted into 96-well PCR plates (ThermoFisher Scientific) containing 2  $\mu\text{L}$  of lysis buffer (0.2% Triton-X 100, 0.2% NP-40 (Sigma Aldrich), 1 U/ $\mu\text{L}$  RNaseOut (ThermoFisher Scientific), PCR-grade water (Ambion), and ERCC spike-in synthetic RNAs (Ambion)). The 96-well plates containing sorted nuclei were then snap frozen and stored at  $-80^{\circ}\text{C}$ . Positive controls (10 nuclei pools and/or 10 pg and 1 pg total RNA) were included on every 96-well plate of sorted nuclei.

**cDNA and sequencing library preparation.** cDNA libraries from single nuclei were prepared using Smart-seq2<sup>20</sup> with minor modifications. Briefly, Protoscript II (New England Biolabs) was used for reverse transcription, the final dilution of ERCCs in the reverse transcription reaction was 1:55 million, and the template switching oligonucleotide was 5'-biotinylated. Additionally, the number of PCR cycles used for cDNA amplification was increased to 21 to compensate for the lower RNA content in single nuclei. cDNA yield was quantified using PicoGreen (ThermoFisher Scientific) and a subset of single-nuclei libraries was screened for quality on a Bioanalyzer (High Sensitivity DNA Chip, Agilent Technologies). cDNA library quality was further assessed using qPCR for a housekeeping gene (*ACTB*) and an ERCC spike-in control RNA (ERCC-00009)<sup>51</sup>.

Sequencing libraries were prepared using a Nextera XT (Illumina) with minor modifications. Briefly, the input amount of cDNA was 250 pg per

reaction, reactions were carried out at 1/4 the volume recommended by the manufacturer, and the tagmentation step was extended to 10 min. Sequencing library concentration was determined using PicoGreen, and 53–57 samples were pooled per sequencing lane. Pooled libraries were purified using Ampure XP beads and eluted to a concentration of 5 nM. Following purification, the pooled library size, using a Bioanalyzer and Kapa Library QC, was used to determine nM concentrations. Final library pools were then diluted to 3 nM final concentration. Pooled samples were sequenced on a HiSeq 4000 instrument (Illumina) using 150-base paired-end reads at a mean untrimmed read depth of ~19 million reads/sample and a mean trimmed read depth of ~16 million reads/sample.

**RNA-seq processing.** The RNA-seq data obtained from single nuclei was processed and analyzed according to the procedure described in detail previously<sup>51</sup>. Briefly, following the demultiplexing of the barcoded reads generated on the Illumina HiSeq platform, the amplification (cDNA and PCR) and sequencing primers (Illumina) and the low-quality bases were removed using Trimmomatic 0.35 software<sup>52</sup>. The trimmed reads were mapped to the human reference genome, version GRCh38 (Ensembl), guided by the version 21 annotations obtained from the GENCODE repository. RSEM 1.2.31<sup>53</sup>, TOPHAT 2.1.1<sup>54</sup>, and CUFFLINKS 2.2.1<sup>54</sup> were used to quantify transcript expression at the transcriptome (exon) and the whole genome (exon plus intron) level, respectively. Software packages fastQC 0.10.1 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), FASTX 0.0.14 ([http://hannonlab.cshl.edu/fastx\\_toolkit/download.html](http://hannonlab.cshl.edu/fastx_toolkit/download.html)), RSeQC 2.6.1<sup>55</sup>, and RNA-seq-QC 1.1.8<sup>56</sup> were used to generate various sequence and alignment quality metrics used for classifying sample quality. A novel pipeline (SCavenger, J.M.M., unpublished) was created to automate execution across statistical analysis tools, integrate preformatted laboratory and clustering metrics, and calculate new statistics specific to biases identified in the single-nuclei lab and sequence preparation protocol. The normalized expression counts (FPKM/TPM) generated at both gene and isoform level by RSEM and TOPHAT-CUFFLINKS analyses and the raw counts generated from the RSEM/TOPHAT alignment (BAM) files by the HTSeq-count program<sup>57</sup> were used for differential expression analysis.

**RNA-seq quality control.** To remove data from low-quality nuclei samples before downstream analysis, we implemented a random forest machine-learning classification approach as described in detail in Aeversmann et al.<sup>58</sup>. The overall workflow for sample quality classification and filtering was to (i) establish a training set using a representative subset of samples, (ii) collect a series of 108 quality control metrics (for example, percent unique reads, percent reads surviving trimming, transcript isoform counts) spanning both the laboratory and data analysis workflows as model features, (iii) use these training data and quality control metrics to build a classification model using the random forest method, and (iv) apply the model to the entire dataset for quality classification and data filtering.

A training set of 196 samples, including 169 single-nuclei samples, was selected, and a set of high confidence pass/fail calls for individual samples was determined based on the qualitative assessment of data produced by fastQC, which includes quality Phred scores, GC content,  $K_{\text{mer}}$  distributions, and sequence over-representation information. Pass samples (152 samples, including single-nuclei and purified bulk-RNA positive controls) were identified as having high average quality per read across the entire length of the sequenced fragment and a unimodal average GC content around 40%, reflecting the GC content of the expressed human transcriptome. In contrast, two types of Fail samples were identified. One type of Fail samples (29 samples) exhibited a substantial number of reads with low mean Phred quality and average Phred quality scores that fell off down the length of the sequence read. A second type of Fail samples (15 samples) showed a second peak in the GC content distribution with a mean around 48% GC; this peak appears to have been generated from ERCC reads, which are derived from bacterial genome sequences.

The quality control metrics for these training data were then used as features to construct a random forest model to distinguish these three quality classes (Pass, Fail-Phred, and Fail-ERCC) comprised of 100,000 decision trees generated by standard bagging methods as implemented in KNIME v3.1.2. Using this random forest classification model, all 196 samples in the training set were classified correctly with high confidence scores. To test the classification accuracy of the resulting random forest model, we used an independent test set of 185 single-nuclei samples classified using the same fastQC evaluation criteria applied to the training data, with 135 determined to be Pass samples, 29 determined to be Fails, and 21 determined to be Marginals. Application of the random forest model to these test Pass and Fail samples resulted in only 8 misclassifications (4.9%), for a classification accuracy of 95%. The random forest model was then applied to the remaining data and final classifications were determined. A Pass confidence cutoff of 0.6 or greater was used to select single-nuclei data for downstream analysis. Using this random forest model applied to the entire layer 1 dataset, including contaminating layer 2 excitatory and inhibitory nuclei, 79% of 1,154 single-nuclei samples passed quality control. For these Pass samples, the average number of reads after trimming was 16,383,881 ( $\pm 19,810,661$ ), the number of ERCC transcripts detected was 43.76 ( $\pm 3.77$ ), and the number of genes detected at FPKM  $> 1$  was 6,337 ( $\pm 1,659$ ), giving an average coverage of 879 reads per human gene detected.

**Gene expression calculation.** For each nucleus, expression levels were estimated based on the scaled coverage across each gene. Specifically, bam files were read into *R*<sup>59</sup> using the readGAlignmentPairs function in the GenomicAlignments library, and genomic coverage was calculated using the coverage function in GenomicRanges<sup>60</sup>. All genes in GENCODE human genome GRCh38, version 21 (Ensembl 77; 09-29-2014) were included, with gene bounds defined as the start and end locations of each unique gene specified in the gtf file (<https://www.encodegenes.org/releases/21.html>). Total counts for each gene (including reads from both introns and exons) were estimated by dividing total coverage by twice the read length (150 bp, paired end). Expression levels were normalized across nuclei by calculating counts per million (CPM) using the cpm function in edgeR<sup>61</sup>.

**Clustering nuclei.** Nuclei that passed quality control were grouped into transcriptomic cell types based on an iterative clustering procedure. For each gene,  $\log_2(\text{CPM} + 1)$  expression was centered and scaled across nuclei. Gene expression dropout was more likely to occur in nuclei with lower-quality cDNA libraries and for genes with lower average expression in nuclei isolated from the same cell type. Expression noise models were estimated for each nucleus based on the eight most-similar nuclei using the knn.error.models function of the scde *R* package as described in ref. <sup>62</sup>. These noise models were used to select significantly variable genes (adjusted variance > 1.25) and to estimate a zero-weight matrix that represented the likelihood of dropouts based on average gene expression levels. Dimensionality reduction was performed with principal components analysis (PCA) on variable genes, and the covariance matrix was adjusted by the zero-weight matrix to account for gene dropouts. Principal components (PCs) were retained for which more variance was explained than the broken-stick null distribution or PCs based on permuted data. If more than two PCs were retained, dimensionality was further reduced to two dimensions using *t*-distributed stochastic neighbor embedding (t-SNE)<sup>63</sup> with a perplexity parameter of 80.

After dimensionality reduction, nuclei were clustered using a conservative procedure that attempted to split them into the fewest number of clusters possible. Nearest-neighbor distances between all nuclei were calculated and sorted, and segmented linear regression (using the 'segmented' *R* package) was applied to estimate the distribution breakpoint to help define the distance scale for density clustering. Next, density clustering (dbSCAN *R* package<sup>64</sup>) was applied to nuclei, and the number of clusters calculated for a range of ten nearest-neighbor distances (parameter epsilon), starting from the maximum distance between nuclei to the distance breakpoint identified in the last step. If only one cluster was found using all values of epsilon, the above procedure was repeated using perplexity parameters of 50, 30, and 20 for t-SNE and stopping when more than one cluster was detected. Finally, if no cluster splitting was possible using t-SNE, then a final density clustering was applied to the first two significant PCs. If more than one cluster was identified, then the statistical significance of each cluster pair was evaluated with the *R* package sigclust<sup>65</sup>, which compares the distribution of nuclei to the null hypothesis that nuclei are drawn from a single multivariate Gaussian. Iterative clustering was used to split nuclei into subclusters until the occurrence of one of four stop criteria: (i) more than nuclei in a cluster (because it cannot be split due a minimum cluster size of 3), (ii) no significantly variable genes, (iii) no significantly variable PCs, or (iv) no significant subclusters.

To assess the robustness of clusters, the iterative clustering procedure described above was repeated 100 times for random subsamples of 80% of nuclei. A co-clustering matrix was generated that represented the proportion of clustering iterations in which each pair of nuclei was assigned to the same cluster. Average-linkage hierarchical clustering was applied to this matrix, followed by dynamic branch cutting (*R* package WGCNA) with cut heights ranging from 0.01 to 0.99 in steps of 0.01. A cut height resulting in 25 clusters was selected to balance cohesion (average within cluster co-clustering) and discreteness (average between cluster co-clustering) across clusters. Finally, gene markers were identified for all cluster pairs, and clusters were merged if they lacked binary markers (gene expressed in > 50% nuclei in first cluster and < 10% in second cluster) with average CPM > 1 (see also "Marker gene selection" below).

**Cluster visualization.** The relationships between cell-type clusters were represented as a constellation diagram in which the area of each disc is proportional to the number of nuclei in each cluster and the width of the lines connecting clusters is proportional to the number of 'intermediate nuclei' between these clusters, as described below and in ref. <sup>66</sup>. To define core and intermediate nuclei, we used a nearest-centroid classifier, which assigns a nucleus to the cluster whose median is most highly correlated based on expression of the 1,200 best marker genes, as described below. We performed five-fold cross-validation 100 times: in each round, the nuclei were randomly partitioned into five equally sized groups, and the nuclei in each group were classified by a nearest centroid classifier trained using the remaining nuclei. Nuclei classified to the same cluster fewer than 90 times or classified to a cluster different from the originally assigned cluster were defined as core, while the others were designated as intermediate. In total, 443/470 (94.3%) of nuclei were defined as core.

Next, clusters were arranged by transcriptomic similarity based on hierarchical clustering. First, the average expression level of each gene was calculated for each cluster. Genes were then sorted based on variance and the top 2,000 genes

were used to calculate a correlation-based distance matrix,  $D_{xy} = 1 - (\text{cor}(x,y))/2$ , between each cluster average. A cluster tree was generated by performing hierarchical clustering on this distance matrix (using hclust with default parameters), and then reordered to show inhibitory clusters first, followed by excitatory clusters and glia, with larger clusters first, while respecting the tree structure. Note that this measure of cluster similarity is complementary to the co-clustering similarity described above. For example, two clusters with high transcriptomic similarity but few distinct marker genes may have low co-clustering similarity.

**Marker gene selection.** Initial sets of marker genes for each pair of clusters were selected by assessing the significance of differential expression using the limma<sup>67</sup> *R* package, and then filtering these sets of significant genes to include only those expressed in more than 50% of nuclei in the 'on' cluster and fewer than 20% of nuclei in the 'off' cluster. Potential marker genes for individual clusters were chosen by ranking the significance of pairwise marker genes, summing the ranks across all possible pairs for a given cluster, and sorting the resulting gene list in ascending order by summed rank. The final set of marker genes was selected by comparing the gene expression distribution for the top-ranked marker genes for each cluster using the visualization described below.

**Scoring marker genes based on cluster specificity.** Many genes were expressed in the majority of nuclei in a subset of clusters. A marker score (beta) was defined for all genes to measure how binary expression was among clusters, independently of the number of clusters labeled. First, the proportion ( $x_i$ ) of samples in each cluster that expressed a gene above background level (CPM > 1) was calculated. Scores were then defined as the squared differences in proportions normalized by the sum of absolute differences plus a small constant ( $\epsilon$ ) to avoid division by zero. Scores ranged from 0 to 1, and a perfectly binary marker had a score equal to 1.

$$\beta = \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - x_j)^2}{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j| + \epsilon}$$

**Matching clusters based on marker gene expression.** Human MTG layer 1 clusters were compared to published cell types from human cortex<sup>22</sup> and mouse primary visual cortex<sup>68</sup>. The proportion of nuclei or cells expressing each gene with CPM > 1 was calculated for all clusters. Genes selected were cluster-specific (beta score > 0.3) in this study and in the published human and mouse studies. Weighted correlations were calculated between all pairs of clusters across these genes and weighted again by beta scores to increase the influence of more informative genes. Heatmaps were generated to visualize all cluster correlations, and pairs of clusters that were reciprocal best matches were labeled. Finally, scatter plots were generated to compare the expression detection of marker genes in these labeled cluster pairs.

**Gene expression visualization.** Gene expression (CPM) was visualized using heat maps and violin plots, which both show genes as rows and nuclei as columns, sorted by cluster. Heat maps display each nucleus as a short vertical bar, color-coded by expression level (blue = low; red = high), and clusters were ordered as described above. The distributions of marker gene expression across nuclei in each cluster were represented as violin plots, which are density plots turned 90 degrees and reflected on the *y* axis. Black dots indicate the median gene expression in nuclei of a given cluster; dots above  $y=0$  indicate that a gene is expressed in more than half of the nuclei in that cluster.

**Colorimetric in situ hybridization.** In situ hybridization data for human temporal cortex and mouse cortex was from the Allen Mouse Brain Atlas<sup>32</sup> and a comparable study in human temporal cortex<sup>68</sup>. All data is publicly accessible through <https://www.brain-map.org>. Data was generated using a semiautomated technology platform, as described<sup>32</sup>, with modifications for working with postmortem human tissues, as described in ref. <sup>68</sup>. Digoxigenin-labeled riboprobes were generated for each human gene such that they would have > 50% overlap with the orthologous mouse gene in the Allen Mouse Brain Atlas<sup>32</sup>. Mouse ISH data shown is from the region most closely corresponding to human temporal cortex, corresponding to the medial portion of TeA in the Watson & Paxinos Atlas<sup>69</sup>.

**Multiplex fluorescent in situ hybridization (FISH).** Tissue specimens used for multiplex FISH came from either neurosurgical resections (all MTG tissue) or postmortem brain specimens (frontal and parietal regions). Tissue procurement from donors undergoing surgery was performed at hospitals, fully outside of the supervision of the Allen Institute. Tissue was provided to researchers under the supervision and authority of the Internal Review Board (IRB) of each participating hospital. All surgical tissue donors met with a hospital-appointed surgical case coordinator to review the option of tissue donation and voluntarily signed an IRB-approved informed consent form. Tissue donors for these experiments ranged in age from 28 to 37 years old. Tissue from surgical resections was transported in chilled, oxygenated ACSF and then mounted for slice preparation on a Compresstome VF-200 or VF-300 vibrating microtome (Precisionary Instruments) to be sliced perpendicular to pial surface. Slices (350  $\mu\text{m}$ ) were embedded in OCT

(optimal cutting temperature medium), rapidly frozen, and subsectioned at 20  $\mu\text{m}$  on a Leica cryostat.

For experiments using postmortem tissue, coronal brain slabs containing Brodmann Area 9 (rostradorsal portion of dorsolateral prefrontal cortex) and Brodmann Area 40 (rostral division of posteroventral parietal cortex) were identified, and regions of interest were removed from frozen slabs and subdivided into small blocks. Blocks were embedded in OCT and sectioned at 16  $\mu\text{m}$  using a Leica cryostat.

An RNAscope multiplex fluorescent kit was used according to the manufacturer's instructions for fresh frozen tissue sections (ACD Bio), with the exception that fixation at 4 °C with 4% PFA was performed for 60 min on 16- to 20- $\mu\text{m}$  human brain sections, and the protease treatment step was shortened to 15 min. Probes used to identify specific cell types in layer 1 were designed antisense to the following human genes: CCK (hs-539041, NM\_000729.4), CNR1 (hs-591521, NM\_001160226.1), CPLX3 (hs-487681-C3, NM\_001030005.2), GAD1 (hs-404031 and hs-404031-C3, NM\_000817.2), LAMP5 (hs487691-C3, NM\_012261.3), SV2C (hs448361-C3, NM\_014979.3), PRSS12 (hs-493931-C3, NM\_003619.3), SOX13 (hs-493941-C3, NM\_005686.2), TRPC3 (hs-427641-C2, NM\_001130698.1), NTNG1 (hs-446101, NM\_00113226.1), CXCL14 (hs-425291, NM\_004887.4), PDGFRA (hs-604481-C2, NM\_006206.4), SOX9 (hs-404221-C2, NM\_000346.3). Positive controls (POLR2A, UBC, and PPIB) were used on each tissue sample to ensure RNA quality (ACD Bio, 320861). Following hybridization and amplification, FISH sections were imaged using a 40 $\times$  oil-immersion lens on a Nikon TiE fluorescent microscope. RNA spots in each channel were quantified manually using the ImageJ 1.51 cell-counting plug-in. To count the percentage of RCs in layer 1, GAD1<sup>+</sup> cells were first identified, followed by the PDGFRA<sup>+</sup> cells within that population, followed by the TRPC3<sup>+</sup> cells in that population. These counts were used to calculate the percentage of the GAD1<sup>+</sup> cells expressing PDGFRA and TRPC3. A total of 408 GAD1<sup>+</sup> cells were identified from two individuals for this quantification.

**Electrophysiological recordings.** All procedures were performed according to the Declaration of Helsinki with the approval of the University of Szeged Ethical Committee. We used neocortical tissue surgically removed from patients ( $n=42$ ,  $n=22$  female and  $n=20$  male, aged  $49 \pm 18$  years) over a course of 5 years as part of the treatment protocol for aneurysms ( $n=9$ ) and brain tumors ( $n=33$ ). Patients with a history of epilepsy were excluded from this study. Anesthesia was induced with intravenous midazolam and fentanyl (0.03 mg/kg, 1–2 mg/kg, respectively). A bolus dose of propofol (1–2 mg/kg) was administered intravenously. To facilitate endotracheal intubation, the patient received 0.5 mg/kg rocuronium. After 120 s, the trachea was intubated and the patient was ventilated with a mixture of  $\text{O}_2$  –  $\text{N}_2\text{O}$  at a ratio of 1:2. Anesthesia was maintained with sevoflurane at monitored anesthesia care (MAC) volume of 1.2–1.5. Tissue blocks were removed from prefrontal ( $n=16$ ), temporal ( $n=6$ ), and parietal ( $n=10$ ) areas. Blocks of tissue were immersed in ice-cold solution containing (in mM) 130 NaCl, 3.5 KCl, 1  $\text{NaH}_2\text{PO}_4$ , 24  $\text{NaHCO}_3$ , 1  $\text{CaCl}_2$ , 3  $\text{MgSO}_4$ , 10 D (+)-glucose, saturated with 95%  $\text{O}_2$  and 5%  $\text{CO}_2$  in the operating theatre. Slices were cut perpendicular to cortical layers at a thickness of 350  $\mu\text{m}$  with a vibrating blade microtome (Microm HM 650 V) and were incubated at room temperature for 1 h in the same solution. The solution used during recordings differed only in that it contained 2 mM  $\text{CaCl}_2$  and 1.5 mM  $\text{MgSO}_4$ . Somatic whole-cell recordings were obtained at approximately 16 °C from up to four concomitantly recorded cells visualized by infrared differential interference contrast videomicroscopy at depths of 60–130  $\mu\text{m}$  from the surface of the slice. Signals were filtered at 8 kHz, digitized at 16 kHz, and acquired with Patchmaster software. Micropipettes (5–7 M $\Omega$ ) were filled with a low [Cl]<sub>i</sub> solution containing (in mM) 126 potassium-gluconate, 4 KCl, 4 ATP-Mg, 0.3 GTP- $\text{Na}_2$ , 10 HEPES, 10 phosphocreatine, and 8 biocytin (pH 7.20; 300 mOsm). Presynaptic cells were stimulated with brief (2–10 ms) suprathreshold pulses delivered at > 7-s intervals, to minimize intertrial variability. For pharmacological experiments, 10  $\mu\text{M}$  gabazine and 5  $\mu\text{M}$  1-(2,4-dichlorophenyl)-5-(4-iodophenyl)-4-methyl-N-1-piperidinyl-1H-pyrazole-3-carboxamide (AM251) were applied (Sigma-Aldrich). Membrane properties of human neurons did not show significant changes for up to 20 h after slicing, but recordings included in the analysis were arbitrarily terminated 15 h after slice preparation. Data were analyzed with Fitmaster (HEKA) and Origin 7.5 (OriginLab) Data are given as mean  $\pm$  s.d.. The Mann–Whitney  $U$  test was used to compare datasets; differences were considered significant if  $P < 0.05$ . Data collection and analysis were not performed blind to the conditions of the experiments.

**Firing classification analysis.** First, a set of  $n=200$  electrophysiological features was calculated for each cell identified based on light-microscopic investigation of the axonal arbor. Then a wrapper-feature selection method using Support Vector Machine (SVM) was used on the cells (RCs:  $n=55$ ; non-RCs:  $n=52$ ) to find the best feature set separating the group of RCs from the group of other cells. The best feature set consisted of two features: the maximal s.d. of interspike intervals (ISI SD) and the amplitude of lag in response to hyperpolarization (–100 pA). Sweeps with < 5 spikes were discarded for the calculation of ISI s.d. The lag value was calculated as the ratio of the voltages at the onset of the hyperpolarizing step to those during steady state.

**Measurement of impedance profile.** The impedance profile was determined by sinusoidal current injections using a standard exponential chirp pattern (0.2–200 Hz, 10 s duration) generated with Patchmaster (HEKA). Measurements (7–10 traces per cell) were made at resting membrane potential, and the peak-to-peak amplitude of the command current waveform was tuned between 40 and 100 pA to test subthreshold voltage responses. The impedance profile ( $Z$ ) was determined for each trace by calculating the fast Fourier transform (FFT) of the voltage response and dividing by the FFT component of the corresponding command current, and then the impedance profiles were normalized to the value at 200 Hz. After anatomical identification of the recorded cells, the dataset was pooled on the basis of three defined cell types, and then the averaged impedance was plotted against input frequency. For statistical comparison of the impedance profiles, four parameters were considered: impedance at lowest frequency ( $Z_{0.2\text{Hz}}$ ); resonance magnitude ( $Q$ , the impedance magnitude at the resonance peak, i.e., the maximal impedance value divided by the impedance magnitude at the lowest input frequency of 0.2 Hz); and the frequency at maximum impedance ( $f_{\text{max}}$ ).

**Two-photon calcium imaging.** Structural labeling of RCs was based on 40  $\mu\text{M}$  Alexa Fluor 594 (Molecular Probes). We also applied 100  $\mu\text{M}$  of Oregon Green 488 BAPTA-1 (Molecular Probes) to measure intracellular  $\text{Ca}^{2+}$  dynamics of pyramidal cell dendrites in the intracellular solution (see above). Imaging with multiphoton excitation was performed using a modified Zeiss LSM7 MP (Oberkochen, Germany) two-photon laser scanning system and a FemtoRose 100 TUN (R&D Ultrafast Lasers, Hungary) titanium–sapphire laser with Finesse4 pumping laser (Laser Quantum, UK) providing 100-fs pulses at 80 MHz at a wavelength of 820 nm. Fluorescence images were acquired through a 40 $\times$  water-immersion objective (0.8 NA; Olympus, Japan).

**Single-cell reverse-transcription and digital PCR.** At the end of electrophysiological recordings, the intracellular content was aspirated into the recording pipettes by application of a gentle negative pressure while maintaining the tight seal. Pipettes were then delicately removed to allow outside-out patch formation, and the content of the pipettes (~1.5  $\mu\text{L}$ ) was expelled into a low-adsorption test tube (Axygen) containing 0.5  $\mu\text{L}$  SingleCellProtect (Avidin Ltd., Szeged, Hungary) solution to prevent nucleic acid degradation and to be compatible with the direct reverse-transcription reaction. Samples were snap-frozen in liquid nitrogen and stored or immediately used for reverse transcription. Reverse transcription of individual cells was carried out in two steps. The first step was performed for 5 min at 65 °C in a total reaction volume of 7.5  $\mu\text{L}$  containing the cell collected in 4  $\mu\text{L}$  SingleCellProtect (Avidin Ltd., Cat.No. SCP-250), 0.45  $\mu\text{L}$  TaqMan Assays (Thermo Fisher), 0.45  $\mu\text{L}$  10 mM dNTPs (Thermo Fisher, Cat.No. 10297018), 1.5  $\mu\text{L}$  5 $\times$  first-strand buffer, 0.45  $\mu\text{L}$  0.1-mol/L DTT, 0.45  $\mu\text{L}$  RNase inhibitor (Thermo Fisher, Cat.No. N8080119), and 100 U of reverse transcriptase (Superscript III, Thermo Fisher, Cat.No. 18080055). The second step of the reaction was carried out at 55 °C for 1 h and then the reaction was stopped by heating at 75 °C for 15 min. The reverse transcription reaction mix was stored at –20 °C until PCR amplification.

For digital PCR analysis, the reverse transcription reaction mixture (7.5  $\mu\text{L}$ ) was divided into two parts: 6  $\mu\text{L}$  was used for amplification of the gene of interest and 1.5  $\mu\text{L}$  cDNA was used for amplifying the housekeeping gene GAPDH. Template cDNA was supplemented with nuclease-free water to a final volume of 8  $\mu\text{L}$ . We then mixed 2  $\mu\text{L}$  TaqMan Assays (Thermo Fisher), 10  $\mu\text{L}$  OpenArray Digital PCR Master Mix (Thermo Fisher, Cat.No. 4458095), and nuclease-free water (3  $\mu\text{L}$ ) to obtain a total volume of 20  $\mu\text{L}$ , and the mixture was evenly distributed over 4 subarrays (256 nanocapillary holes) of an OpenArray plate using the OpenArray autoloader. Processing of the OpenArray slide, cycling in the OpenArray NT cycler, and data analysis were done as previously described<sup>70</sup>. For our dPCR protocol amplification, reactions with CT values less than 23 or greater than 33 were considered primer dimers or background signals, respectively, and excluded from the dataset.

**Histology and reconstruction.** Following electrophysiological recordings, slices were immersed in a fixative containing 4% paraformaldehyde (for immunohistochemistry) or 4% paraformaldehyde, 15% (v/v) saturated picric acid, and 1.25% glutaraldehyde (for reconstructions) in 0.1 M phosphate buffer (PB; pH = 7.4) at 4 °C for at least 12 h. After several washes with 0.1 M PB, slices were frozen in liquid nitrogen then thawed in 0.1 M PB, embedded in 10% gelatin, and further sectioned into 60- $\mu\text{m}$  slices. Sections were incubated in a solution of conjugated avidin–biotin horseradish peroxidase (ABC; 1:100; Vector Labs) in Tris-buffered saline (TBS, pH = 7.4) at 4 °C overnight. The enzyme reaction was revealed by 3′3′-diaminobenzidine tetra-hydrochloride (0.05%) as chromogen and 0.01%  $\text{H}_2\text{O}_2$  as oxidant. Sections were postfixed with 1%  $\text{OsO}_4$  in 0.1 M PB. After several washes in distilled water, sections were stained in 1% uranyl acetate and dehydrated in an ascending series of ethanol. Sections were infiltrated with epoxy resin (Durcupan) overnight and embedded on glass slides. Three-dimensional light-microscopic reconstructions were carried out using a Neurolucida system (MicroBrightField) with a 100 $\times$  objective. Reconstructed neurons were quantitatively analyzed with NeuroExplorer software (MicroBrightField). Axon tortuosity was measured as the average ratio of the actual axonal path and linear distance between neighboring nodes.



**Immunohistochemistry of biocytin-labeled cells.** The recorded cells were first visualized with incubation in Cy3-conjugated streptavidin (Jackson ImmunoResearch) for 2 h, diluted at 1:400 in TBS. After examination by epifluorescence microscopy, the sections containing the soma of the labeled neurons were incubated in 20% normal horse serum in TBS to block nonspecific antibody-binding sites. Free-floating sections containing the soma were incubated in primary antibodies dissolved in TBS containing 0.05% Na<sub>2</sub>S<sub>2</sub>O<sub>8</sub> for 72 h at room temperature (22 °C). The following primary antibodies were used: rabbit anti-pro-cholecystokinin (1:2,000, polyclonal, lot TL1, gift from A. Varro, Liverpool University); mouse anti-CNR1 (1:4,000, IMG-CB1R-mAb001, clone IMG-3C2, lot CJ03ImmunoGenes); rabbit anti-GABA (1:1,000, A2052, polyclonal, lot 056M4834V, Sigma-Aldrich); mouse anti-NR2F2 (1:700, Ab41859, clone H7147, lot GR15505-8, Abcam); mouse anti-PV (1:1,500, S235, clone 235, lot 10-11 F, Swant); rabbit anti-nNOS (1:1,000, 160870, polyclonal, lot 189901, Cayman Chemical); rabbit anti-NPY (1:300, T-4069, polyclonal, A00721, Peninsula Laboratories); rat anti-somatostatin (1:50, MAB354, clone YC7, lot 2294201, Merck Millipore); rabbit anti-calbindin (1:2,000, CB-38a, polyclonal, lot 9.03, Swant); goat anti-calretinin (1:700, CG1, polyclonal, lot 10.1, Swant); and goat anti-acetyltransferase (1:100, AB144P, polyclonal, 0608037072, Merck Millipore). After several washes in TBS, the immunoreactions were visualized with A488- or Cy5-conjugated secondary antibodies (1:500; 711-545-152, lot 119191; 715-485-150, lot 89132; 712-485-150, lot 88670; 705-485-003, lot 89133; 715-175-150, lot 105884; 705-175-147, lot 114786; Jackson ImmunoResearch). The sections were mounted on slides in Vectashield (Vector Laboratories). Images were taken by confocal laser scanning microscope (LSM 880, Zeiss) using a 40× oil-immersion objective (1.4 NA). After photography, the sections were demounted, washed in 0.1 M PB, and biocytin was visualized with the avidin-biotinylated horseradish peroxidase method described above. For quantification of positive and negative immunoreactions, we measured the mean fluorescence intensity of the immunostaining in the soma or in the axon terminals of the biocytin-filled RCs and the fluorescence intensity of the background using the thresholding tool by ImageJ 1.48. Fluorescence intensity measurements were corrected with the background in every image. Finally, we classified fluorescence intensity values < 2 AU as negative immunoreactions. The fluorescence intensity of cells measured positive and negative with this method were significantly different ( $P < 0.01$ , MW  $U$  test).

**Electron microscopy.** Axonal boutons of biocytin-filled RCs ( $n = 6$ ) and NGFCs ( $n = 2$ ), identified based on distinctive electrophysiological properties and light microscopic investigation of the axonal arbor, were re-embedded and resectioned at 70-nm thickness. Digital images of serial EM sections were taken at magnifications ranging from 8,000× to 50,000× with a JEOL JEM-1400Plus electron microscope equipped with an 8-megapixel CCD camera (JEOL Ruby). Axon terminals were reconstructed in 3D and their volumes were measured using the Reconstruct software (<http://synapses.clm.utexas.edu/>;  $n = 31$  RC boutons;  $n = 24$  NGFC boutons). The areas of active zones of RCs were measured at perpendicularly cut synapses, where the rigid apposition of the pre- and postsynaptic membranes was visible ( $n = 11$  active zones).

**Statistics.** No sample-size calculation was performed; all data were subject to statistical tests to decide whether parametric or nonparametric tests should be applied. Two-tailed tests were used throughout. Single nuclei were isolated from postmortem brains of two donors. This allowed us to collect nuclei from high-quality specimens that met stringent quality control metrics while also confirming that transcriptomic clusters were consistent between donors and not driven by technical artifacts. No data with successful quality control were excluded from the analyses. Data from low-quality nuclei samples were removed before downstream analysis, using a random forest machine-learning classification approach as described<sup>58</sup>. No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications<sup>10,11,14</sup>. All human specimens were controls (nonpathological) and were therefore allocated into the same experimental group. Randomization was not used. Human specimens were

de-identified and assigned a unique numerical code. Researchers had access to basic information about donors (age, sex, ethnicity) as well as the unique numerical code assigned to each donor; data collection and analysis were not performed blind to the conditions of the experiments.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data and code availability.** Custom R code and data used to generate transcriptomics related figures can be downloaded from [https://github.com/AllenInstitute/L1\\_rosehip](https://github.com/AllenInstitute/L1_rosehip). The data that support the findings of this study are available from the corresponding author upon reasonable request.

## References

- Krishnaswami, S. R. et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat. Protoc.* **11**, 499–524 (2016).
- Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
- Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
- Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
- Wang, L., Wang, S. & Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185 (2012).
- DeLuca, D. S. et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* **28**, 1530–1532 (2012).
- Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
- Aevermann, B. et al. Production of a preliminary quality control pipeline for single nuclei RNA-seq and its application in the analysis of cell type diversity of post-mortem human brain neocortex. *Pac. Symp. Biocomput.* **22**, 564–575 (2017).
- R Development Core Team. *R: a Language and Environment for Statistical Computing*. (R Found. Stat. Comput., Vienna, Austria, 2016).
- Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLOS Comput. Biol.* **9**, e1003118 (2013).
- Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
- Fan, J. et al. Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. *Nat. Methods* **13**, 241–244 (2016).
- Van Der Maaten, L. J. P. & Hinton, G. E. Visualizing high-dimensional data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
- Ester, M., Kriegel, H. P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD* **2**, 226–231 (1996).
- Liu, Y., Hayes, D. N., Nobel, A. & Marron, J. S. Statistical significance of clustering for high-dimension, low-sample size data. *J. Am. Stat. Assoc.* **103**, 1281–1293 (2008).
- Tasic, B. et al. Shared and distinct transcriptomic cell types across neocortical areas. Preprint at *bioRxiv* <https://doi.org/10.1101/229542> (2017).
- Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
- Zeng, H. et al. Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. *Cell* **149**, 483–496 (2012).
- Paxinos, G. & Franklin, K. *The Mouse Brain in Stereotaxic Coordinates*. (Academic Press, Cambridge, MA, USA, 2012).
- Faragó, N. et al. Digital PCR to determine the number of transcripts from single neurons after patch-clamp recording. *Biotechniques* **54**, 327–336 (2013).

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- ☐ ☒ The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- ☐ ☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☐ ☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☐ ☒ Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated
- ☐ ☒ Clearly defined error bars  
*State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on [statistics for biologists](#) may be useful.

### Software and code

Policy information about [availability of computer code](#)

Data collection	BD Diva software v8.0 (Flow Cytometry data). Patchmaster 2x90 and 2x73 versions (electrophysiology)
Data analysis	Trimmomatic 0.35, RSEM 1.2.31, TOPHAT 2.1.1, CUFFLINKS 2.2.1, fastQC 0.10.1, FASTX 0.0.14, RSeQC 2.6.1, RNA-seq-QC 1.1.8 (RNA-Seq processing). Custom R code written for clustering single nucleus RNA-seq data and marker gene analysis and using open source R packages. Available from <a href="https://github.com/AllenInstitute/L1_rosehip">https://github.com/AllenInstitute/L1_rosehip</a> . Image J 1.48 and 1.51 (immunohistochemistry and cell counting). Fitmaster 2x73 (HEKA) and Origin 7.5 (OriginLab) for electrophysiological data analysis.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Custom R code written for clustering single nucleus RNA-seq data and marker gene analysis available from [https://github.com/AllenInstitute/L1\\_rosehip](https://github.com/AllenInstitute/L1_rosehip). The datasets generated and analysed in figures 1-6 are available from the corresponding author upon request.

## Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculation was performed, all data were subject to statistical tests to decide whether parametric or nonparametric tests should be applied. Single nuclei were isolated from post-mortem brains of 2 donors. This allowed us to collect nuclei from high quality specimens that met stringent quality control metrics while also confirming that transcriptomic clusters were consistent between donors and not driven by technical artifacts.
Data exclusions	No data with successful quality control were excluded from the analyses. Data from low quality nuclei samples was removed prior to downstream analysis, using a Random Forest machine learning classification approach as described in detail in Aevermann et al., 2016 and in the Methods section of the manuscript.
Replication	Experimental protocols were reliably reproduced and the reproducibility of findings from each protocol can be judged from the figures and corresponding statistics. In particular, flow cytometry data were reproducible across human tissue specimens and across different nuclei isolations from individual tissue donors.
Randomization	All human specimens were controls (nonpathological) and were therefore allocated into the same experimental group. Randomization was not used.
Blinding	Human specimens were de-identified and assigned a unique numerical code. Researchers had access to basic information about donors (age, sex, ethnicity) as well as the unique numerical code assigned to each donor.

## Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	n/a
Research sample	n/a
Sampling strategy	n/a
Data collection	n/a
Timing	n/a
Data exclusions	n/a
Non-participation	n/a
Randomization	n/a



# Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	n/a
Research sample	n/a
Sampling strategy	n/a
Data collection	n/a
Timing and spatial scale	n/a
Data exclusions	n/a
Reproducibility	n/a
Randomization	n/a
Blinding	n/a
Did the study involve field work?	<input type="checkbox"/> Yes <input type="checkbox"/> No

## Field work, collection and transport

Field conditions	n/a
Location	n/a
Access and import/export	n/a
Disturbance	n/a

# Reporting for specific materials, systems and methods

## Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants

## Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Unique biological materials

Policy information about [availability of materials](#)

Obtaining unique materials	n/a
----------------------------	-----

## Antibodies

Antibodies used	<p>Primary antibody Supplier name Catalog number Clone number Lot number Dilution</p> <p>Rabbit-anti-pro-cholecystokinin gift from Andrea Varro, Liverpool University N.A. polyclonal N.A. 1:2000</p> <p>Mouse-anti-CNR1 ImmunoGenes IMG-CB1R-mAb001 IMG-3C2 CJ03 1:4000</p> <p>Rabbit-anti-GABA Sigma-Aldrich A2052 polyclonal 056M4834V 1:1000</p> <p>Mouse-anti-NR2F2 Abcam Ab41859 H7147 GR15505-8 1:700</p> <p>Mouse-anti-parvalbumin Swant 235 235 10-11(F) 1:1500</p>
-----------------	--

Rabbit-anti-nNOS Cayman Chemical 160870 polyclonal 189901 1:1000  
 Rabbit-anti-NPY Peninsula Laboratories T-4069 polyclonal A00721 1:300  
 Rat-anti-somatostatin Merck Millipore MAB354 YC7 2294201 1:50  
 Rabbit-anti-calbindin Swant CB-38a polyclonal 9.03 1:2000  
 Goat-anti-calretinin Swant CG1 polyclonal 15.1 1:700  
 Goat-anti-choline acetyltransferase Merck Millipore, Chemicon AB144P polyclonal 0608037072 1:100

#### Secondary antibody

Supplier name Catalog number Clone number Lot number Dilution  
 Alexa Fluor 488 AffiniPure Donkey Anti-Rabbit IgG Jackson ImmunoResearch 711-545-152 N.A. 119191 1:500  
 Alexa Fluor 488 AffiniPure Donkey Anti-Mouse IgG Jackson ImmunoResearch 715-485-150 N.A. 89132 1:500  
 Alexa Fluor 488 AffiniPure Donkey Anti-Rat IgG Jackson ImmunoResearch 712-485-150 N.A. 88670 1:500  
 Alexa Fluor 488 AffiniPure Donkey Anti-Goat IgG Jackson ImmunoResearch 705-485-003 N.A. 89133 1:500  
 Cy5 AffiniPure Donkey Anti-Mouse IgG Jackson ImmunoResearch 715-175-150 N.A. 105884 1:500  
 Cy5 AffiniPure Donkey Anti-Goat IgG Jackson ImmunoResearch 705-175-147 N.A. 114786 1:500

#### Validation

All antibodies were routinely evaluated by manufacturers using immunohistochemistry on rat cerebellum brain tissue.

## Eukaryotic cell lines

Policy information about [cell lines](#)

#### Cell line source(s)

No eukaryotic cell lines were used.

#### Authentication

No eukaryotic cell lines were used.

#### Mycoplasma contamination

No eukaryotic cell lines were used.

#### Commonly misidentified lines (See [ICLAC](#) register)

No commonly misidentified cell lines were used.

## Palaeontology

#### Specimen provenance

n/a

#### Specimen deposition

n/a

#### Dating methods

n/a

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

#### Laboratory animals

No animals and other organisms were used.

#### Wild animals

No animals and other organisms were used.

#### Field-collected samples

No animals and other organisms were used.

## Human research participants

Policy information about [studies involving human research participants](#)

#### Population characteristics

The specimens used in for single nucleus RNA sequencing were from two individual control Caucasian male donors, aged 50 and 54 years with no history of neuropsychiatric illness and no evidence of neuropathology. For electrophysiological recordings Caucasian donors were selected both genders (n=42, n=22 female and n=20 male, aged 49±18 years).

#### Recruitment

We used nonpathological neocortical tissue surgically removed from patients in a course of five years as part of the treatment protocol for aneurysms and brain tumors.

## ChIP-seq

### Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links  
*May remain private before publication.*

n/a

Files in database submission

n/a

Genome browser session  
(e.g. [UCSC](#))

n/a

## Methodology

Replicates

n/a

Sequencing depth

n/a

Antibodies

n/a

Peak calling parameters

n/a

Data quality

n/a

Software

n/a

## Flow Cytometry

### Plots

Confirm that:

- ☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☐ A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

Sample preparation

Microdissected sections of layer 1 from MTG were transferred into nuclei isolation medium containing 10mM Tris pH 8.0 (Ambion), 250mM sucrose, 25mM KCl (Ambion), 5mM MgCl<sub>2</sub> (Ambion) 0.1% Triton-X 100 (Sigma Aldrich), 1% RNasin Plus, 1X protease inhibitor (Promega), and 0.1mM DTT and placed into a 1ml dounce homogenizer (Wheaton). Tissue was homogenized to liberate nuclei using 10 strokes of the loose dounce pestle followed by 10 strokes of the tight pestle. Homogenate was strained through a 30µm cell strainer (Miltenyi Biotec) and centrifuged at 900xg for 10 min to pellet nuclei. Nuclei were then resuspended in staining buffer containing 1X PBS (Ambion), 0.8% nuclease-free BSA (Omni-Pur, EMD Millipore), and 0.5% RNasin Plus. Mouse monoclonal anti-NeuN antibody (EMD Millipore) was applied to nuclei preparations at a concentration of 1:1000 and samples were incubated for 30 min at 4°C. Control samples were incubated with mouse IgG1,k isotype control (BD Pharmingen). Samples were then centrifuged for 5 min at 500xg to pellet nuclei and pellets were resuspended in staining buffer as described above. Nuclei samples were incubated with secondary antibody (goat anti-mouse IgG, Alexa Fluor 594, ThermoFisher Scientific) for 30 min at 4°C, centrifuged for 5 min at 500xg, and resuspended in staining buffer. DAPI (4', 6-diamidino-2-phenylindole, ThermoFisher Scientific) was applied to nuclei samples at a concentration of 1µg/ml.

Instrument

BD FACSAria Fusion

Software

BD FACSDiva

Cell population abundance

10 nuclei pools and/or 10 pg and 1 pg total RNA

Gating strategy

Nuclei were first gated on DAPI and then passed through doublet discrimination gates prior to being gated on NeuN (Alexa Fluor 594) signal.

- ☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

## Magnetic resonance imaging

### Experimental design

Design type

n/a

Design specifications

n/a

Behavioral performance measures

n/a



## Acquisition

Imaging type(s)	n/a
Field strength	n/a
Sequence & imaging parameters	n/a
Area of acquisition	n/a
Diffusion MRI	<input type="checkbox"/> Used <input checked="" type="checkbox"/> Not used

## Preprocessing

Preprocessing software	n/a
Normalization	n/a
Normalization template	n/a
Noise and artifact removal	n/a
Volume censoring	n/a

## Statistical modeling &amp; inference

Model type and settings	n/a
Effect(s) tested	n/a
Specify type of analysis:	<input type="checkbox"/> Whole brain <input type="checkbox"/> ROI-based <input type="checkbox"/> Both
Statistic type for inference (See <a href="#">Eklund et al. 2016</a> )	n/a
Correction	n/a

## Models &amp; analysis

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Functional and/or effective connectivity
<input checked="" type="checkbox"/>	<input type="checkbox"/> Graph analysis
<input checked="" type="checkbox"/>	<input type="checkbox"/> Multivariate modeling or predictive analysis
Functional and/or effective connectivity	n/a
Graph analysis	n/a
Multivariate modeling and predictive analysis	n/a