

## 1 **Conserved cell types with divergent features between human and mouse cortex**

2 Rebecca D Hodge<sup>\*1</sup>, Trygve E Bakken<sup>\*1</sup>, Jeremy A Miller<sup>1</sup>, Kimberly A Smith<sup>1</sup>, Eliza R Barkan<sup>1</sup>,  
3 Lucas T Graybuck<sup>1</sup>, Jennie L Close<sup>1</sup>, Brian Long<sup>1</sup>, Osnat Penn<sup>1</sup>, Zizhen Yao<sup>1</sup>, Jeroen  
4 Eggermont<sup>2</sup>, Thomas Holtt<sup>2,3</sup>, Boaz P Levi<sup>1</sup>, Soraya I Shehata<sup>1</sup>, Brian Aevermann<sup>4</sup>, Allison  
5 Beller<sup>5</sup>, Darren Bertagnolli<sup>1</sup>, Krissy Brouner<sup>1</sup>, Tamara Casper<sup>1</sup>, Charles Cobbs<sup>6</sup>, Rachel Dalley<sup>1</sup>,  
6 Nick Dee<sup>1</sup>, Song-Lin Ding<sup>1</sup>, Richard G Ellenbogen<sup>7</sup>, Olivia Fong<sup>1</sup>, Emma Garren<sup>1</sup>, Jeff Goldy<sup>1</sup>,  
7 Ryder P Gwinn<sup>8</sup>, Daniel Hirschstein<sup>1</sup>, C Dirk Keene<sup>5</sup>, Mohamed Keshk<sup>4</sup>, Andrew L Ko<sup>7,9</sup>, Kanan  
8 Lathia<sup>1</sup>, Ahmed Mahfouz<sup>2,3</sup>, Zoe Maltzer<sup>1</sup>, Medea McGraw<sup>1</sup>, Thuc Nghi Nguyen<sup>1</sup>, Julie Nyhus<sup>1</sup>,  
9 Jeffrey G Ojemann<sup>7,9</sup>, Aaron Oldre<sup>1</sup>, Sheana Parry<sup>1</sup>, Shannon Reynolds<sup>1</sup>, Christine Rimorin<sup>1</sup>,  
10 Nadiya V Shapovalova<sup>1</sup>, Saroja Somasundaram<sup>1</sup>, Aaron Szafer<sup>1</sup>, Elliot R Thomsen<sup>1</sup>, Michael  
11 Tieu<sup>1</sup>, Richard H Scheuermann<sup>4,10</sup>, Rafael Yuste<sup>11</sup>, Susan M Sunkin<sup>1</sup>, Boudewijn Lelieveldt<sup>2,3</sup>,  
12 David Feng<sup>1</sup>, Lydia Ng<sup>1</sup>, Amy Bernard<sup>1</sup>, Michael Hawrylycz<sup>1</sup>, John W. Phillips<sup>1</sup>, Bosiljka Tasic<sup>1</sup>,  
13 Hongkui Zeng<sup>1</sup>, Allan R Jones<sup>1</sup>, Christof Koch<sup>1</sup>, Ed S Lein<sup>#1</sup>

14

15 <sup>1</sup> Allen Institute for Brain Science, Seattle, WA, USA

16 <sup>2</sup> Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands

17 <sup>3</sup> Department of Intelligent Systems, Delft University of Technology, Delft, the Netherlands

18 <sup>4</sup> J. Craig Venter Institute, La Jolla, CA, USA

19 <sup>5</sup> Department of Pathology, University of Washington, Seattle, WA, USA

20 <sup>6</sup> The Ben and Catherine Ivy Center for Advanced Brain Tumor Treatment, Swedish  
21 Neuroscience Institute, Seattle, WA, USA

22 <sup>7</sup> Department of Neurological Surgery, University of Washington School of Medicine, Seattle,  
23 WA

24 <sup>8</sup> Epilepsy Surgery and Functional Neurosurgery, Swedish Neuroscience Institute, Seattle, WA,  
25 USA

26 <sup>9</sup> Regional Epilepsy Center at Harborview Medical Center, Seattle, WA, USA

27 <sup>10</sup> Department of Pathology, University of California, San Diego, CA, USA

28 <sup>11</sup> Neurotechnology Center, Department of Biological Sciences, Columbia University, New York,  
29 NY, USA

30

31 \* Contributed equally

32 #Correspondence should be addressed to Ed S. Lein ([edl@alleninstitute.org](mailto:edl@alleninstitute.org))

## 33 Abstract

34 Elucidating the cellular architecture of the human neocortex is central to understanding our  
35 cognitive abilities and susceptibility to disease. Here we applied single nucleus RNA-  
36 sequencing to perform a comprehensive analysis of cell types in the middle temporal gyrus of  
37 human cerebral cortex. We identify a highly diverse set of excitatory and inhibitory neuronal  
38 types that are mostly sparse, with excitatory types being less layer-restricted than expected.  
39 Comparison to a similar mouse cortex single cell RNA-sequencing dataset revealed a  
40 surprisingly well-conserved cellular architecture that enables matching of homologous types and  
41 predictions of human cell type properties. Despite this general conservation, we also find  
42 extensive differences between homologous human and mouse cell types, including dramatic  
43 alterations in proportions, laminar distributions, gene expression, and morphology. These  
44 species-specific features emphasize the importance of directly studying human brain.

## 45 Introduction

46 The cerebral cortex, responsible for most of our higher cognitive abilities, is the most complex  
47 structure known to biology and is comprised of approximately 16 billion neurons and 61 billion  
48 non-neuronal cells organized into approximately 200 distinct anatomical or functional  
49 regions<sup>1,2,3,4</sup>. The human cortex is greatly expanded relative to the mouse, the dominant model  
50 organism in basic and translational research, with a 1200-fold increase in cortical neurons  
51 compared to only a 60-fold increase in sub-cortical neurons (excluding cerebellum)<sup>5,6</sup>. The  
52 general principles of neocortical development and the basic multilayered cellular  
53 cytoarchitecture of the neocortex appear relatively conserved across mammals<sup>7,8</sup>. However,  
54 whether the cellular and circuit architecture of cortex is fundamentally conserved across  
55 mammals, with a massive evolutionary areal expansion of a canonical columnar architecture in  
56 human, or is qualitatively and quantitatively specialized in human, remains an open question  
57 long debated in the field<sup>9,10</sup>. Addressing this question has been challenging due to a lack of  
58 tools to broadly characterize cell type diversity in complex brain regions, particularly in human  
59 brain tissues.

60 Prior studies have described differences in the cellular makeup of the cortex in human and  
61 specialized features of specific cell types<sup>11,12,13,14,15,16,17</sup>, although the literature is remarkably  
62 limited. For example, the supragranular layers of cortex, involved in cortico-cortical  
63 communication, are differentially expanded in mammalian evolution<sup>18</sup>. Furthermore, certain cell  
64 types show highly specialized features in human and non-human primate compared to mouse,  
65 such as the interlaminar astrocytes<sup>17</sup>, and the recently described rosehip cell<sup>19</sup>, a type of  
66 inhibitory interneuron in cortical layer 1 with distinctive morpho-electrical properties. All of these  
67 cellular properties are a function of the genes that are actively used in each cell type, and  
68 transcriptomic methods provide a powerful method to understand the molecular underpinnings  
69 of cellular phenotypes as well as a means for mechanistic understanding of species-specialized  
70 phenotypes. Indeed, a number of studies have shown significant differences in transcriptional  
71 regulation between mouse, non-human primate and human, including many genes associated  
72 with neuronal structure and function<sup>20,21,22,23</sup>.

73 Dramatic advances in single cell transcriptional profiling present a new approach for large-scale  
74 comprehensive molecular classification of cell types in complex tissues, and a metric for  
75 comparative analyses. The power of these methods is fueling ambitious new efforts to  
76 understand the complete cellular makeup of the mouse brain<sup>24</sup> and the even the whole human  
77 body<sup>25</sup>. Recent applications of single cell RNA-sequencing (scRNA-seq) methods in mouse  
78 cortex have demonstrated robust transcriptional signatures of neuronal and non-neuronal cell

79 types<sup>26,27,28</sup>, and suggest the presence of approximately 100 neuronal and non-neuronal cell  
80 types in any given cortical area. Similar application of scRNA-seq to human brain has been  
81 challenging due to the difficulty in dissociating intact cells from densely interconnected human  
82 tissue<sup>29</sup>. In contrast, single nucleus RNA-sequencing (snRNA-seq) methods allow for  
83 transcriptional profiling of intact neuronal nuclei that are relatively easy to isolate and enable use  
84 of frozen postmortem specimens from human brain repositories<sup>30,31,32</sup>. Importantly, it was  
85 recently shown that single nuclei contain sufficient gene expression information to distinguish  
86 closely related subtypes of cells at a similar resolution to scRNA-seq<sup>33,34</sup>, demonstrating that  
87 snRNA-seq is a viable method for surveying cell types that can be compared to scRNA-seq  
88 data. Early applications of snRNA-seq to human cortex demonstrated the feasibility of the  
89 approach but have not provided depth of coverage sufficient to achieve similar resolution to  
90 mouse studies<sup>35</sup>.

91 The current study aimed to establish a robust methodology for relatively unbiased cell type  
92 classification in human brain using snRNA-seq, and to perform the first comprehensive  
93 comparative analysis of cortical cell types to understand conserved and divergent features of  
94 human and mouse cerebral cortex. We first describe the cellular landscape of the human  
95 cortex, and then demonstrate a similar degree of cellular diversity between human and mouse  
96 and a conserved set of homologous cell types and subclasses. In contrast, we present evidence  
97 for extensive differences between homologous types, including evolutionary changes in relative  
98 proportions, laminar distributions, subtype diversity, gene expression and other cellular  
99 phenotypes.

## 100 Results

### 101 Transcriptomic taxonomy of cell types

102 A robust snRNA-seq methodology was established to analyze transcriptomically defined cell  
103 types in human cortex. We focused on the middle temporal gyrus (MTG), with samples largely  
104 derived from high-quality postmortem brain specimens. This region is frequently available  
105 through epilepsy surgery resections, permitting a comparison of postmortem versus acute  
106 neurosurgical tissues, as well as allowing future correlation with *in vitro* slice physiology  
107 experiments in MTG. Frozen tissue blocks were thawed, vibratome sectioned, and stained with  
108 fluorescent Nissl dye. Individual cortical layers were microdissected, tissues were homogenized  
109 to release nuclei, and nuclei were stained with an antibody against NeuN to differentiate  
110 neuronal (NeuN-positive) and non-neuronal (NeuN-negative) nuclei. Single nuclei were  
111 collected via fluorescence-activated cell sorting (FACS) (**Fig. 1A, Extended Data Figure 1A,**  
112 **Methods**). We sorted ~90% NeuN-positive and ~10% NeuN-negative nuclei across all cortical  
113 layers to enrich for neurons. The final dataset contained less than the targeted 10% non-  
114 neuronal nuclei because nearly 50% of NeuN-negative nuclei failed quality control criteria,  
115 potentially due to the lower RNA content of glia compared to neurons (**Methods**)<sup>27</sup>. SMART-  
116 Seqv4 (Takara Bio USA Inc.) was used to reverse transcribe mRNA and amplify cDNA.  
117 Sequencing libraries were generated using Nextera XT (Illumina), which were sequenced on a  
118 HiSeq 2500 at a median depth of 2.6 +/- 0.5 million reads/nucleus. Nuclei were collected from 8  
119 total human tissue donors (4 male, 4 female; 4 postmortem, 4 neurosurgical) ranging in age  
120 from 24 to 66 years (**Extended Data Table 1**). 15,206 nuclei were collected from postmortem  
121 tissue donors with no history of neuropathology or neuropsychiatric disorders, and 722 nuclei  
122 came from apparently histologically normal MTG distal to pathological tissue that was removed  
123 during surgical resections to treat epilepsy (**Methods**).

124 To evenly survey cell type diversity across cortical layers, nuclei were sampled based on the  
125 relative proportion of neurons in each layer reported in human temporal cortex<sup>36</sup>. Based on  
126 Monte Carlo simulations, we estimated that 14,000 neuronal nuclei were needed to target types  
127 as rare as 0.2% of the total neuron population (**Methods**). Using an initial subset of RNA-seq  
128 data, we observed more transcriptomic diversity in layers 1, 5, and 6 than in other layers so  
129 additional neuronal nuclei were sampled from those layers. In total, 15,928 nuclei passed quality  
130 control criteria and were split into three broad classes of cells (10,708 excitatory neurons, 4297  
131 inhibitory neurons, and 923 non-neuronal cells) based on NeuN staining and cell class marker  
132 gene expression (**Methods**).

133  
134 Nuclei from each broad class were iteratively clustered as described in<sup>33</sup>. Briefly, high variance  
135 genes were identified while accounting for gene dropouts, expression dimensionality was  
136 reduced with principal components analysis (PCA), and nuclei were clustered using Jaccard-  
137 Louvain community detection (**Methods**). On average, neuronal nuclei were larger than non-  
138 neuronal nuclei (**Extended Data Fig. 1B**), and median gene detection (**Extended Data Fig.**  
139 **1C,D**) was correspondingly higher for neurons (9046 genes) than for non-neuronal cells (6432  
140 genes), as previously reported for mouse<sup>26,27,28</sup>. Transcriptomic cell types were largely  
141 conserved across diverse individuals and tissue types (postmortem, neurosurgical), since all  
142 curated clusters contained nuclei derived from multiple donors, and nuclei from postmortem and  
143 neurosurgical tissue types clustered together (**Fig. 1B, Extended Data Fig. 2A**). However, a  
144 small, but consistent expression signature related to tissue type was apparent; for example,  
145 nuclei derived from neurosurgical tissues exhibited higher expression of some activity related  
146 genes (**Extended Data Fig. 2**). 325 nuclei were assigned to donor-specific or outlier clusters  
147 that contained marginal quality nuclei and were excluded from further analysis (**Methods**).

148 This analysis method defined 75 transcriptomically distinct cell types, including 45 inhibitory  
149 neuron types that express the canonical GABAergic interneuron marker *GAD1*, 24 excitatory  
150 neuron types that express the vesicular glutamate transporter *SLC17A7*, and 6 non-neuronal  
151 types that express the glutamate transporter *SLC1A3* (**Fig. 1C, D**). As expected based on prior  
152 studies<sup>26,27,28,31</sup>, the hierarchical relationships among types roughly mirrors the developmental  
153 origin of different cell types. We refer to the cell type clusters as cell *types*, intermediate order  
154 nodes as *subclasses*, and higher order nodes such as the interneurons derived from the caudal  
155 ganglionic eminence (CGE) as *classes*, and the broadest divisions such as excitatory neurons  
156 as *major classes*. Neuronal types split into two major classes representing cortical plate-derived  
157 glutamatergic excitatory neurons (n=10,525 nuclei) and ganglionic eminence-derived  
158 GABAergic inhibitory neurons (n=4164 nuclei). Non-neuronal types (n=914 nuclei) formed a  
159 separate main branch based on differential expression of many genes (**Fig. 1C**). We developed  
160 a principled nomenclature for clusters based on: 1) major cell class, 2) layer enrichment  
161 (including layers containing at least 10% of nuclei in that cluster), 3) a subclass marker gene  
162 (maximal expression of 14 manually-curated genes), and 4) a cluster-specific marker gene  
163 (maximal detection difference compared to all other clusters) (**Fig. 1D, Extended Data Fig. 3,**  
164 **Methods**). For example, the left-most inhibitory neuron type in **Figure 1D**, found in samples  
165 dissected from layers 1 and 2, and expressing the subclass marker *PAX6* and the specific  
166 marker *CDH12*, is named Inh L1-2 *PAX CDH12*. Additionally, we generated a searchable  
167 semantic representation of these cell type clusters that incorporates this accumulated  
168 knowledge about marker gene expression, layer enrichment, specimen source, and parent cell  
169 class to link them to existing anatomical and cell type ontologies<sup>37</sup> (**Supplementary Data**). We  
170 find broad correspondence to an earlier study<sup>31</sup>, but identify many additional types of excitatory  
171 and inhibitory neurons due to increased sampling and/or methodological differences (**Extended**  
172 **Data Fig. 4**). The majority of cell types were rare (<100 nuclei per cluster, <0.7% of cortical



173 neurons), including almost all interneuron types and deep layer excitatory neuron types. In  
174 contrast, the excitatory neurons of superficial layers 2-4 were dominated by a small number of  
175 relatively abundant types (>500 nuclei per cluster, >3.5% of neurons) (**Fig. 1C**). Both excitatory  
176 types and many interneuron types were restricted to a few layers, whereas non-neuronal nuclei  
177 were distributed across all layers, with the notable exception of one astrocyte type (**Fig. 1C**).

## 178 **Excitatory neurons often span multiple layers**

179 The 24 transcriptionally distinct excitatory neuron types broadly segregated by layer and  
180 expressed known laminar markers (**Fig. 2A-C**). In general, excitatory types were most similar to  
181 other types in the same or adjacent layers. Transcriptomic similarity by proximity for cortical  
182 layers has been described before, and interpreted as a developmental imprint of the inside-out  
183 generation of cortical layers<sup>38</sup>. Complex relationships between clusters are represented as  
184 constellation diagrams (**Fig. 2A, Methods**)<sup>26</sup>, where the circles represent core cells that were  
185 most transcriptionally similar to the cluster to which they were originally assigned, and indicate  
186 the size (proportional to circle area) and average laminar position of each cell type. The  
187 thickness of lines between cell clusters represents their similarity based on the number of nuclei  
188 whose assignment to a cluster switched upon reassignment (intermediate cells, **Methods**). This  
189 similarity by proximity is also apparent in the hierarchical dendrogram structure of cluster  
190 similarity in **Figure 2B**. One exception is the layer 5 Exc L5-6 *THEMIS C1QL3* type, which has a  
191 transcriptional signature similar to layer 2 and 3 types as well as several deep layer cell types  
192 (**Fig. 2A, B**). Two types, Exc L4-5 *FEZF2 SCN4B* and Exc L4-6 *FEZF2 IL26*, were so distinct  
193 that they occupied separate branches on the dendrogram and did not connect via intermediate  
194 cells to any other type (**Fig. 2A, B**).

195  
196 Each excitatory type showed selective expression of genes that can be used as cell type  
197 markers (**Fig. 2C**), although in general a small combinatorial profile (generally 2-3 genes per  
198 type) was necessary to distinguish each type from all other cortical cell types (**Fig. 2D**). The  
199 majority of these markers are novel as excitatory neuron markers, and belonged to diverse and  
200 functionally important gene families, such as BHLH transcription factors (*TWIST2*), collagens  
201 (*COL22A1*), and semaphorins (*SEMA3E*). Surprisingly, 16 out of 37 (41%) of these most  
202 specific marker genes were unannotated loci (LOCs), long non-coding RNAs (lincRNA),  
203 pseudogenes, and antisense transcripts. This may partially be a result of profiling nuclear RNA,  
204 as some of these transcripts have been shown to be enriched in the nucleus (**Fig. 2C**,  
205 **Extended Data Figs. 3, 5**)<sup>39</sup>.

206 Unexpectedly, most excitatory neuron types were present in multiple layers based on layer  
207 dissection information (**Fig. 2B**). Within the supragranular layers, three main types were  
208 enriched in layer 2 and 3 dissections. Additionally, ten *RORB*-expressing types were enriched in  
209 layer 3-6 dissections (**Fig. 2B, C**). Layers 5 and 6 contained 11 excitatory types: 4 types that  
210 expressed *THEMIS* (Thymocyte Selection Associated), 6 types that expressed *FEZF2*, and 1  
211 type that expressed the cytokine *IL15* (Interleukin 15). The majority of these types were similarly  
212 represented in layer 5 and 6 dissections (**Fig. 2B**). To clarify whether this crossing of layer  
213 boundaries was an artifact of dissection or a feature of MTG organization, we investigated the  
214 layer distribution of 10 types using multiplex fluorescence *in situ* hybridization (FISH) with  
215 combinatorial gene panels designed to discriminate clusters (**Fig. 2B, D, Extended Data Fig.**  
216 **6**). *In situ* distributions largely validated snRNA-seq predictions (**Fig. 2E**). Three types were  
217 mainly localized to layer 3c and the upper part of layer 4, defined as the dense band of granule  
218 cells visible in Nissl stained sections (**Fig. 2E**). Interestingly, one of these types (Exc L3-4 *RORB*  
219 *CARM1P1*) had large nuclei, suggesting that it may correspond to a subset of the giant  
220 pyramidal layer 3c neurons previously described in MTG<sup>40</sup> (**Fig. 2E, Extended Data Fig. 6**).

221 Two types were mostly restricted to layer 4 (Exc L3-5 *RORB ESR1*, Exc L4-5 *RORB DAPK2*),  
222 but the five other types examined all spanned multiple layers (**Fig. 2E**). Taken together, the  
223 snRNA-seq and *in situ* validation data indicate that transcriptomically defined excitatory neuron  
224 types are frequently not layer-specific, but rather spread across multiple anatomically defined  
225 layers.

## 226 **Heterogeneous expression within clusters**

227 A major evolutionary feature of human cortical architecture is the expansion of supragranular  
228 layers compared to other mammals, and morphological and physiological properties of  
229 pyramidal neurons vary across layers 2 and 3 of human temporal cortex<sup>40,41</sup>. In that light, it was  
230 surprising to find only three main excitatory clusters in human cortical layers 2 and 3. However,  
231 one cluster was very large (Exc L2-3 *LINC00507 FREM3*; n=2284 nuclei) and spanned layers 2  
232 and 3, posing the possibility that there is significant within-cluster heterogeneity. Indeed, we  
233 find continuous variation in gene expression in this cluster along the axis of cortical depth,  
234 illustrated well using two data visualization and mining tools built for this project to allow public  
235 access to this dataset. The Cytosplore MTG Viewer (<https://viewer.cytosplore.org>), is an  
236 extension of Cytosplore<sup>42</sup>, and presents a hierarchy of t-SNE maps of different subsets of MTG  
237 clusters<sup>43</sup>, with each map defined using informative marker genes (**Fig. 3A**). Layer dissection  
238 metadata overlaid onto the t-SNE map of Exc L2-3 *LINC00507 FREM3* revealed that nuclei in  
239 this type were ordered by layer, with nuclei sampled from layers 2 and 3 occupying relatively  
240 distinct locations in t-SNE space. Selecting nuclei at both ends of the cluster gradient in t-SNE  
241 space and computing differential expression between these nuclei revealed a set of genes with  
242 variable expression across this cluster (**Fig. 3A, Supplementary Movie 1**).

243 Examining this set of variable genes within Exc L2-3 *LINC00507 FREM3* using the RNA-Seq  
244 Data Navigator (<http://celltypes.brain-map.org/rnaseq/human>) showed gradient expression  
245 between layers 2 and 3 (**Fig. 3B**). Finally, single molecule FISH confirmed gradient expression  
246 of *LAMP5* and *COL5A2* across layers 2 and 3 in cells mapping to this cluster (**Fig. 3C,**  
247 **Extended Data Figs. 7, 8**). These results illustrate that there is additional diversity in human  
248 supragranular pyramidal neurons manifested as continuous variation in gene expression as a  
249 function of cortical depth that likely correlates with anatomical and functional heterogeneity of  
250 those cells.

## 251 **Inhibitory neuron diversity**

252 GABAergic inhibitory neurons split into two major branches, largely distinguished by expression  
253 of Adenosine Deaminase, RNA Specific B2 (*ADARB2*) and the transcription factor LIM  
254 Homeobox 6 (*LHX6*) (**Fig. 4A-F**). In mouse cortex, interneurons split into the same two major  
255 branches, also defined by expression of *Adarb2* and *Lhx6* and developmental origins in the  
256 caudal ganglionic eminence (CGE) and medial ganglionic eminence (MGE), respectively<sup>26</sup>.  
257 The *ADARB2* branch was further subdivided into the *LAMP5/PAX6* and *VIP* subclasses of  
258 interneurons, with likely developmental origins in the CGE. Surprisingly, the serotonin receptor  
259 subunit *HTR3A*, which marks CGE-derived interneurons in mouse<sup>44</sup>, was not a good marker of  
260 these types in human (**Fig. 4E**). The *LHX6* branch consisted of *PVALB* and *SST* subclasses of  
261 interneurons, likely originating in the medial ganglionic eminence MGE<sup>45,46</sup>. Consistent with  
262 mouse cortex<sup>26</sup>, the *ADARB2* branch showed a much higher degree of diversity in  
263 supragranular layers 1-3 compared to layers 4-6, whereas the opposite was true for the *LHX6*  
264 branch (**Fig. 4A, B**). As with the excitatory neuron taxonomy, many interneuron cluster specific  
265 markers were unannotated (LOC) genes, lincRNAs, pseudogenes, and antisense  
266 transcripts (**Fig. 4E, F**).

267 The *LAMP5/PAX6* subclass of interneurons included 6 transcriptomic types, many of which  
268 were enriched in layers 1 and 2 (**Fig. 4C**). Several types coexpressed *SST* (**Fig. 4E**), consistent  
269 with previous reports demonstrating *SST* expression in layer 1 of human MTG<sup>19</sup> and different  
270 from mouse *Lamp5* and *Pax6* interneurons<sup>26,27</sup>, which do not express *SST*. The Inh L1-4  
271 *LAMP5 LCP2* type expressed marker genes of rosehip cells, a type of interneuron with  
272 characteristic large axonal boutons that we described in a previous study of layer 1 MTG  
273 interneurons<sup>19</sup>. With whole cortex coverage, it is clear that this type is not restricted to layer 1  
274 but rather present across all cortical layers. Among *LAMP5/PAX6* types on the *ADARB2* (CGE-  
275 derived) branch, Inh L2-6 *LAMP5 CA1* cells uniquely expressed *LHX6*, suggesting possible  
276 developmental origins in the MGE, and appear similar to the *Lamp5 Lhx6* cells previously  
277 described in mouse cortex<sup>26,27</sup>.

278 *VIP* interneurons represented the most diverse subclass, containing 21 transcriptomic types  
279 (**Fig. 4A**), many of which were enriched in layers 2 and 3 (**Fig. 4C**). Several types in the *VIP*  
280 subclass (Inh L1 *SST CHRNA4* and Inh L1-2 *SST BAGE2*) appeared to be closely related to the  
281 L1 *SST NMBR* type of the *LAMP5/PAX6* subclass, as evidenced by intermediate cell  
282 connections between these types. Interestingly, these highly related types were all localized to  
283 layers 1 and 2. Furthermore, while both the Inh L1 *SST CHRNA4* and Inh L1-2 *SST*  
284 *BAGE2* were grouped into the *VIP* subclass, they appeared to lack expression of *VIP*. Rather,  
285 they expressed *SST*, consistent with expression of this gene in layer 1 and 2 interneurons as  
286 discussed above (**Fig. 4A, C, E**)<sup>19</sup>. The Inh L1-2 *GAD1 MC4R* type also lacked expression of  
287 *VIP* (**Fig. 4E**). Notably, this type specifically expresses the Melanocortin 4 Receptor, a gene  
288 linked to autosomal dominant obesity and previously shown to be expressed in a population of  
289 mouse hypothalamic neurons that regulate feeding behavior<sup>48,49</sup>.

290 The *SST* subclass consisted of 11 transcriptomic types, including one highly distinct type, Inh  
291 L3-6 *SST NPY*, that occupied its own discrete branch on the dendrogram and was not  
292 connected to other types in the *SST* constellation (**Fig. 4B, D**). Several *SST* types displayed  
293 laminar enrichments, with Inh L5-6 *SST TH* cells being a particularly restricted type, found only  
294 in layers 5 and 6. We further validated marker gene expression and the spatial distribution of the  
295 Inh L3-6 *SST NPY* and Inh L5-6 *SST TH* types using ISH from the Allen Human Brain Atlas  
296 (<http://human.brain-map.org/>; **Fig. 4G**). ISH for *TH* confirmed that expression of this gene is  
297 sparse and restricted to layers 5-6; interestingly, *Th* ISH in mouse temporal association area  
298 (TEa; the closest homolog to human MTG) showed similar sparse labeling restricted to layers 5  
299 and 6, suggesting that this gene may mark similar cell types in human and mouse  
300 (<http://mouse.brain-map.org/>; **Fig. 4G**). In contrast, the well-known interneuron marker  
301 neuropeptide Y (*Npy*) was broadly expressed in a scattered pattern throughout all layers in  
302 mouse TEa, whereas, in human MTG, *NPY* labeled only a single interneuron type whose  
303 sparsity was confirmed by ISH (**Fig. 4G**), indicating that this heavily-studied marker labels a  
304 different cohort of cell types in human and mouse<sup>50,51</sup>.

305  
306 The *PVALB* subclass comprised 7 clusters, including two types that were grouped into this  
307 branch but did not appear to express *PVALB* (**Fig. 4F**). One of these types, Inh L5-6 *SST*  
308 *MIR548F2*, had low expression of *SST*, whereas the other type, Inh L5-6 *GAD1 GLP1R*, did not  
309 express any canonical interneuron subclass markers. Intermediate cells connected the Inh L5-6  
310 *SST MIR548F2* type in the *PVALB* constellation to the Inh L5-6 *SST TH* type in the *SST*  
311 constellation. Two other connections between the *SST* and *PVALB* constellations were  
312 apparent, both of which included the Inh L2-4 *SST FRZB* cluster (**Fig. 4B**). One highly  
313 distinctive *PVALB* type (Inh L2-5 *PVALB SCUBE3*) (**Fig. 4B, D**) likely corresponds to chandelier  
314 (axo-axonic) cells as it expresses *UNC5B*, a marker of chandelier (axo-axonic) cells in mouse<sup>52</sup>  
315 (**Fig. 4H**). Multiplex FISH (RNAscope, **Methods**) validated expression of several novel marker

316 genes (*NOG*, *COL15A1*, **Fig. 4H**) and showed enrichment of these cells mainly in layers 2-4,  
317 consistent with the pattern observed in the snRNA-seq data (**Fig. 4D, H**).

### 318 **Diverse morphology of astrocyte types**

319 Although non-neuronal (NeuN-) cells were not sampled as deeply as neurons, all major glial  
320 types - astrocytes, oligodendrocytes, endothelial cells, and microglia - were identified (**Fig. 5A**).  
321 In contrast to studies of mouse cortex where non-neuronal cells were more extensively sampled  
322 or selectively targeted with Cre lines<sup>26,28,53</sup>, we did not find other types of immune or vascular  
323 cells. This decreased diversity is likely largely due to more limited non-neuronal sampling, but  
324 may also reflect the age of tissue analyzed. For example, previous reports showed that adult  
325 mouse cortex contains mainly oligodendrocyte progenitor cells (OPCs) and mature  
326 oligodendrocytes, but few immature and myelinating oligodendrocyte types<sup>28,53</sup>, similarly, we  
327 found only two oligodendrocyte types, one of which expressed markers of oligodendrocyte  
328 progenitor cells (OPCs) (e.g. *PDGFRA*, *OLIG2*) and another that expressed mature  
329 oligodendrocyte markers (e.g. *OPALIN*, *MAG*) (**Fig. 5A, B**).

330  
331 Astrocytes in human cortex are both functionally<sup>54</sup> and morphologically<sup>17</sup> specialized in  
332 comparison to rodent astrocytes, with distinct morphological types residing in different layers of  
333 human cortex (**Fig. 5C**). Interlaminar astrocytes, described only in primates to date, reside in  
334 layer 1 and extend long processes into lower layers, whereas protoplasmic astrocytes are found  
335 throughout cortical layers 2-6<sup>17</sup> (**Fig.5C**). Similarly, we find two astrocyte clusters with different  
336 laminar distributions. Astro L1-2 *FGFR3 GFAP* originated mostly from layer 1 and 2 dissections,  
337 whereas the Astro L1-6 *FGFR3 SLC14A1* type was found in all layers (**Fig.5A**). The two  
338 astrocyte types we identified were distinguished by expression of the specific marker  
339 gene *ID3* along with higher expression of *GFAP* and *AQP4* in the Astro L1-2 *FGFR3 GFAP* type  
340 than in the Astro L1-6 *FGFR3 SLC14A1* type (**Fig. 5B, D**). To determine if these two  
341 transcriptomic types correspond to distinct morphological types, we labeled cells with a  
342 combination of multiplex FISH and immunohistochemistry for GFAP protein. Cells with  
343 high *GFAP* and *AQP4* expression, characteristic of the Astro L1-2 *FGFR3 GFAP* type and  
344 consistent with previous reports of interlaminar astrocytes<sup>55</sup>, were present predominantly in the  
345 upper half of layer 1 (**Fig. 5E**). Coexpression of *AQP4* and *ID3* was apparent in layer 1 cells  
346 that had extensive, long-ranging GFAP-positive processes characteristic of interlaminar  
347 astrocytes (**Fig. 5E**). In contrast, GFAP-positive cells with protoplasmic astrocyte morphology  
348 lacked expression of *ID3*, consistent with the Astro L1-6 *FGFR3 SLC14A1* type (**Fig. 5E**).

349  
350 Interestingly, while most nuclei contributing to the Astro L1-6 *FGFR3 GFAP* cluster came from  
351 layer 1 and 2 dissections, seven nuclei were from layer 5 and 6 dissections and expressed *ID3*  
352 as well as a distinct set of marker genes (**Fig. 5D**). Based on their laminar origin, we  
353 hypothesized that these nuclei may correspond to fibrous astrocytes, which are enriched in  
354 white matter<sup>17</sup> (**Fig. 5C**). Indeed, astrocytes at the border of layer 6 and the underlying white  
355 matter coexpressed *ID3* and *AQP4* and had relatively thick, straight GFAP-positive processes  
356 characteristic of fibrous astrocytes (**Fig. 5E**), suggesting that the Astro L1-6 *FGFR3*  
357 *GFAP* cluster contains a mixture of two different morphological astrocyte types. Given that  
358 nuclei corresponding to fibrous astrocytes express distinct marker genes from interlaminar  
359 astrocytes (**Fig. 5D**), it is likely that fibrous astrocytes will form a separate transcriptomic type  
360 with increased sampling.



## 361 Human and mouse cell type homology

362 Single cell transcriptomics not only provides a new method for comprehensive analysis of  
363 species-specific cellular diversity, but also a quantitative metric for comparative analysis  
364 between species. Furthermore, identification of homologous cell types or classes allows  
365 inference of cellular properties from much more heavily studied model organisms. The  
366 availability of densely sampled single cell or single nucleus RNA-seq datasets in human  
367 (described here) and mouse<sup>26</sup> cortex using the same RNA-seq profiling platform allowed a  
368 direct comparison of transcriptomic cell types. The success of such a comparison is predicated  
369 on the idea of conserved transcriptional patterning. As a starting point, we asked whether the  
370 same types of genes discriminate human interneuron cell types as those reported for mouse  
371 interneuron types<sup>52</sup>. Indeed, we find the same sets of genes (mean = 21 genes/set) best  
372 discriminate human interneuron types (**Fig.6A**), including genes central to neuronal connectivity  
373 and signaling. Similar functional classes of genes also discriminate human and mouse  
374 excitatory neuron types (although with less conservation for classes of genes that discriminate  
375 non-neuronal cell types; **Extended Data Fig.9A**), indicating that shared expression patterns  
376 between species may facilitate matching cell types.

377 Simply combining expression data for inhibitory neuron nuclei from human MTG and for cells  
378 from mouse V1 was not sufficient for identification of homologous cell types. PCA analysis  
379 resulted in samples clearly separated by species along the first principal component that  
380 explained almost 20% of expression variation (**Fig.6B**, **Extended Data Fig.9B**). Recent work  
381 has demonstrated the power of canonical correlation analysis (CCA) to align single cell RNA-  
382 seq data from human and mouse based on shared co-expression patterns<sup>56</sup>. Application of  
383 CCA and graph-based clustering to human and mouse cortical samples was much more  
384 successful (**Fig.6B**), and allowed matching of human and mouse types based on shared CCA  
385 cluster membership for inhibitory neurons (**Fig.6C**, **Extended Data Fig.9E**), excitatory neurons  
386 (**Fig.6D**, **Extended Data Fig.9F**) and non-neuronal cells (**Fig.6E**, **Extended Data Fig.9G**).

387 Remarkably, shared co-expression between mouse V1 and human MTG enabled the  
388 identification of homologous types at approximately half the resolution of the full human  
389 classification (38 types versus 75 types). Combining the CCA results allowed generation of a  
390 hierarchical taxonomy including 34 neuronal and 4 non-neuronal cell types and subclasses (**Fig.**  
391 **6F**). A hybrid nomenclature from human and mouse<sup>27</sup> was used to describe these homologous  
392 types. Ten cell types were matched one-to-one between species, whereas other types were  
393 matched at a subclass resolution. Transcriptomically distinct cell types more often had one-to-  
394 one matches, likely because more redundant marker genes compensated for divergent  
395 expression patterns, and we find even most rare types had homologous types in mouse and  
396 human.  
397

398 This homology alignment enabled prediction of the anatomical, functional, and connectional  
399 properties of human cell types based on the much larger mouse literature for homologous cell  
400 types. For example, the human cluster Inh L2-5 *PVALB SCUBE3* described above matches  
401 one-to-one with the mouse chandelier (or axo-axonic) cell type *Pvalb Vipr2*, suggesting that this  
402 cell type selectively innervates the axon initial segment of excitatory neurons. Also, the human  
403 cluster Inh L3-6 *SST NPY* matches the mouse *Sst Chodl* type and is therefore predicted to have  
404 long-range projections and contribute to sleep regulation<sup>26,57,58</sup>. Many other anatomically  
405 defined interneuron types could be similarly inferred, including basket, Martinotti, bipolar,  
406 neurogliaform, and single-bouquet cells (**Fig. 6C**), although future experiments will be  
407 necessary to confirm these predictions.  
408



409  
410 The long-range projection targets of human glutamatergic neurons (e.g. intratelencephalic (IT),  
411 pyramidal tract (PT), and corticothalamic (CT)) that would otherwise be experimentally  
412 inaccessible can also be inferred based on their best transcriptomic match to mouse cell types;  
413 for example, the human Exc L4-5 *FEZF2 SCN4B* type corresponds to the PT sub-cortically  
414 projecting layer 5 pyramidal cells (**Fig. 6D**). The Exc L4-6 *FEZF2 IL26* matches two mouse layer  
415 5 types (L5 NP *Slc17a8* and L5 NP *Rapgef3*) that lack long-range projections<sup>26,59</sup>. Finally, layer  
416 6b (subplate) types can be identified by homology, and among human layer 6b types, Exc  
417 L6 *FEZF2 OR2T8* has much larger nuclei (**Extended Data Fig. 1B**) and corresponds to the  
418 mouse L6b *Rprm* type that selectively projects to thalamus rather than cortex.

419  
420 Four of five human non-neuronal cell types matched mouse cell types (**Fig. 6E**), while  
421 endothelial cells had such divergent global expression patterns between species that they could  
422 not be matched by CCA despite the expression of conserved canonical marker genes  
423 (e.g. *EMCN* and *NOSTRIN*). The mouse Oligo *Enpp6* cluster partially overlapped nuclei from  
424 human OPC and mature oligodendrocyte clusters and appears to represent an immature  
425 oligodendrocyte type<sup>26,53</sup> that is rare or not present in adult human cortex. The morphologically  
426 distinct human layer 1 astrocyte type, Astro L1-2 *FGFR3 GFAP*, did not match any clusters  
427 from<sup>26</sup>, although a layer 1 enriched astrocyte with shared marker gene expression was  
428 previously reported in mouse<sup>28</sup>. Finally, while the majority of human microglia clustered with  
429 mouse microglia, two nuclei clustered with mouse perivascular macrophages (**Extended Data**  
430 **Fig.9D**), suggesting that this rare type was likely undersampled in human.

431  
432 Only three mouse neuronal types and two human interneuron types lacked homologous types,  
433 although all three mouse types are very rare and may not have been sampled in human. The  
434 mouse *Meis2* inhibitory type, which is primarily restricted to white matter and has an embryonic  
435 origin outside of the ganglionic eminence<sup>26</sup>, may have been missed due to limited sampling of  
436 layer 6b and underlying white matter. Mouse Cajal-Retzius cells are glutamatergic neurons in  
437 layer 1. These cells are exceedingly rare (less than 0.1% of layer 1 neurons) in adult human  
438 cortex<sup>60,61</sup> and were not expected to be sampled. Finally, the mouse layer 5 pyramidal tract type  
439 L5 PT *Chma6*, a rare excitatory neuron type with strong projections to superior colliculus<sup>59</sup>, has  
440 no matching human cluster. However, 2 of 25 nuclei from the human pyramidal tract (PT)-like  
441 cluster Exc L4-5 *FEZF2 SCN4B* are more similar to this distinct mouse PT type than to other  
442 mouse PT types (**Extended Data Fig.9C**), suggesting this mismatch is also due to  
443 undersampling in human. Interestingly, both human interneuron types that lack closely matched  
444 mouse homologues (Inh L1 *SST CHRNA4* and Inh L1-2 *GAD1 MC4R*) are highly enriched in  
445 layer 1. Along with the phenotypic specialization of the layer 1 rosehip neuron<sup>19</sup>, it appears that  
446 layer 1 may be a hotspot of evolutionary change at the level of inhibitory cell types.

447 While many homologous subclasses had comparable diversity between species, some  
448 subclasses had expanded diversity in human or mouse. Human layer 4 excitatory neurons are  
449 more diverse than those of mouse (**Fig. 6D**), contributing to increased diversity of supragranular  
450 layers due to mixing into layer 3 as described above. Mouse layer 5 PT types are much more  
451 diverse than those in human, which may reflect either a true species difference or  
452 undersampling, as they make up <1% of layer 5 excitatory neurons in human MTG. Layer 6 CT  
453 types also show greater diversity in mouse V1 than human MTG; however, this difference may  
454 reflect an areal difference between a primary sensory area that has strong, reciprocal  
455 connections with the thalamus and an area of association cortex. Indeed, we find increased  
456 diversity of cell types in human visual cortex that match mouse layer 6 CT types (data not  
457 shown).

## 458 **Divergent proportions of cell types**

459 Alterations in the relative proportions of cell types could have profound consequences for  
460 cortical circuit function. snRNA-seq data predicted a significant species difference in the  
461 proportions of interneuron classes. Human MTG showed similar proportions of MGE-derived  
462 (44% *LHX6+* nuclei) and CGE-derived (50% *ADARB2+* nuclei) interneurons, whereas in mouse  
463 cortex roughly 70% of interneurons are MGE-derived and ~30% are CGE-derived<sup>44,62</sup>. To  
464 validate these differences, we applied multiplex FISH to quantify the proportions of CGE  
465 (*ADARB2+*) and MGE (*LHX6+*) interneurons in human MTG and mouse TEa (**Fig. 7, Extended  
466 Data Fig. 10**). Interneurons that co-expressed *ADARB2* and *LHX6*, corresponding to the human  
467 Inh L2-6 *LAMP5 CA1* and mouse *Lamp5 Lhx6* types (**Figs. 1, 4**), were considered separately.  
468 Consistent with the snRNA-seq data, we found similar proportions of MGE (50.2 ± 2.3%) and  
469 CGE (44.2 ± 2.4%) interneurons in human MTG, whereas we found more than twice as many  
470 MGE (67.8 ± 0.9%) than CGE (30.8 ± 1.2%) interneurons in mouse TEa. The increased  
471 proportion of CGE-derived interneurons in human was greatest in layer 4, whereas the  
472 decreased proportion of MGE interneurons in human was greatest in layers 4-6 (**Fig. 7A**).  
473 Interestingly, both the snRNA-seq data (6.1% of *GAD1+* cells) and *in situ* cell counts (5.6 ±  
474 0.3% of *GAD1+* cells) confirmed a significant increase in the proportion of the Inh L2-6 *LAMP5  
475 CA1* type in human MTG versus the *Lamp5 Lhx6* type in mouse TEa (1.4 ± 0.2% of *GAD1+*  
476 cells), most notably in layer 6 (**Fig. 7A**).

477  
478 Another major predicted mismatch was seen for the sub-cortically projecting PT neurons, which  
479 comprise approximately 20% of layer 5 excitatory neurons in mouse but less than 1% in human  
480 based on single cell<sup>26</sup> and single nucleus RNA-seq sampling. To directly compare the spatial  
481 distribution and abundance of PT types between species, we performed ISH for a pan-layer 5  
482 PT marker (*Fam84b*)<sup>26</sup> in mouse TEa and for markers of the homologous layer 5 PT type Exc  
483 L4-5 *FEZF2 SCN4B* in human MTG. In mouse TEa, *Fam84b* was expressed in many neurons  
484 in superficial layer 5 (**Fig. 7B**). To unambiguously label PT neurons in human MTG, we  
485 performed triple FISH with the pan-excitatory marker *SLC17A7*, the PT markers *FAM84B*  
486 or *POU3F1*, and *NPTX1*, which labels most *SLC17A7*-positive layer 5 neurons but not PT cells  
487 (**Fig. 7B, Extended Data Fig. 11**). In MTG, *SLC17A7*/*NPTX1*- cells co-labeled  
488 with *FAM84B* or *POU3F1* were sparsely distributed predominantly in superficial layer 5 and  
489 were large with a prominent, thick apical dendrite (**Fig. 7B, Extended Data Fig. 11**). Thus, PT  
490 cells have a similar distribution within layer 5 in human and mouse but are much less abundant  
491 in human, likely reflecting an evolutionary scaling constraint as discussed below.

## 492 **Divergent expression between homologous types**

493 The identification of homologous or consensus cell types or classes allows direct analysis of the  
494 conservation and divergence of gene expression patterns across these types. For each pair of  
495 homologous cell types, we compared expression levels of 14,414 orthologous genes between  
496 human and mouse. Nuclear expression levels were estimated based on intronic reads to better  
497 compare human single nucleus and mouse single cell RNA-seq data. The Exc L3c/L5a type  
498 (Exc L3-4 *RORB CARM1P1* in human) has the most conserved expression ( $r = 0.78$ ) of all  
499 types, and yet 12% of genes have highly divergent expression (defined as >10-fold difference),  
500 including many specific markers (orange dots, **Fig. 8A**) for this cell type. Microglia had the least  
501 conserved expression ( $r = 0.60$ ), and more than 20% of genes were highly divergent (**Fig.  
502 8B**). Surprisingly, the Exc L3c/L5a consensus type shows a striking shift in layer position  
503 between human, where Exc L3-4 *RORB CARM1P1* is highly enriched in layer 3c of MTG, and  
504 mouse, where the homologous type L5 *Endou* is enriched in layer 5a of mouse V1 (**Fig.  
505 8A**). This laminar shift of a homologous cell type helps explain the reported expression shift of

506 several genes from layer 5 in mouse to layer 3 in human <sup>20</sup>, including two genes (*BEND5*  
507 and *PRSS12*) expressed in Exc L3-4 *RORB CARM1P1* but not in layer 3 of mouse TEa.

508  
509 Over half of all genes analyzed (8222, or 57%) had highly divergent expression in at least one  
510 of the 38 homologous types, and many genes had divergent expression restricted to a specific  
511 cell type or broad class (**Fig. 8C**). Non-neuronal cell types had the most highly divergent  
512 expression including 2025 genes with >10-fold species difference, supporting increased  
513 evolutionary divergence of non-neuronal expression patterns between human and mouse brain  
514 described previously <sup>22</sup>.

515 Most genes had divergent expression in a subset of types rather than all types, and this resulted  
516 in a shift in the cell type specificity or patterning of genes. These expression pattern changes  
517 were quantified as the beta score of log-fold differences across cell types  
518 (**Methods, Supplementary Table 2**), and scores were approximately log-normally distributed  
519 with a long tail of highly divergent genes (**Fig. 8D**). Cell type marker genes tended to be less  
520 conserved than more commonly expressed genes (**Fig. 8E**). In many cases, the most defining  
521 markers for cell types were not shared between human and mouse. For example, chandelier  
522 interneurons selectively express *Vipr2* in mouse but *COL15A1* and *NOG* in human (**Fig. 4H**).  
523 Interestingly, the functional classes of genes that best differentiate cell types within a species  
524 (**Fig. 6A**) are the same functional classes that show the most divergent expression patterns  
525 between species (**Fig. 8F**). In other words, the same gene families show cell type specificity in  
526 both species, but their patterning across cell types frequently differs.

527  
528 The top 20 most divergent gene families between human and mouse (i.e. highest median  
529 pattern change) include neurotransmitter receptors (serotonin, adrenergic, glutamate, peptides,  
530 and glycine), ion channels (chloride), and cell adhesion molecules involved in axonal pathfinding  
531 (netrins and cadherins). Among the top 3% most divergent genes (see **Supplementary Table 2**  
532 for full list), the extracellular matrix collagens *COL24A1* and *COL12A1* and the glutamate  
533 receptor subunits *GRIK1* and *GRIN3A* were expressed in different cell types between species  
534 and were validated by ISH to have different laminar distributions in human MTG and mouse TEa  
535 (**Fig. 8G**). The cumulative effect of so many differences in the cellular patterning of genes with  
536 well characterized roles in neuronal signaling and connectivity is certain to cause many  
537 differences in human cortical circuit function.

## 538 Discussion

539 Single cell transcriptomics provides a powerful tool to systematically characterize the cellular  
540 diversity of complex brain tissues, allowing a paradigm shift in neuroscience from the historical  
541 emphasis on cellular anatomy to a molecular classification of cell types and the genetic  
542 blueprints underlying the properties of each cell type. Echoing early anatomical studies <sup>10</sup>,  
543 recent studies of mouse neocortex have shown a great diversity of cell types <sup>26,28</sup>. Similar  
544 studies of human cortex <sup>35,31,32</sup> have shown the same broad classes of cells but much less  
545 subtype diversity (**Extended Data Fig. 4**), likely resulting from technical differences, such as  
546 fewer nuclei sampled or reduced gene detection. A recent study showed a high degree of  
547 cellular diversity in human cortical layer 1 <sup>19</sup> by densely sampling high-quality postmortem  
548 human tissue with snRNA-seq and including intronic sequence to capture signal in nuclear  
549 transcripts <sup>33</sup>. The current study takes a similar dense sampling approach by sequencing  
550 approximately 16,000 single nuclei spanning all cortical layers of MTG, and defines 75 cell types  
551 representing non-neuronal (6), excitatory (24) and inhibitory (45) neuronal types. Importantly,

552 robust cell typing could be achieved despite the increased biological and technical variability  
553 between human individuals. Nuclei from postmortem and acute surgically resected samples  
554 clustered together, and all clusters described contained nuclei from multiple  
555 individuals. Importantly, the ability to use these methods to study the fine cellular architecture of  
556 the human brain and to identify homologous cell types based on gene expression allows  
557 inference of cellular phenotypes across species as well. In particular, since so much knowledge  
558 has been accumulated about the cellular makeup of rodent cortex based on transcriptomics,  
559 physiology, anatomy and connectivity, this approach immediately allows strong predictions  
560 about such features as well as others that are not currently possible to measure in human such  
561 as developmental origins and long-range projection targets.

562  
563 This molecular paradigm can help unify the field and increase the cellular resolution of many  
564 studies but has several consequences and challenges. Unambiguous definition of  
565 transcriptomic cell types *in situ* typically requires the detection of two or more markers with  
566 multiplexed molecular methods, demonstrating the need to further develop spatial  
567 transcriptomics methods<sup>63</sup>. Developing consistent nomenclature will also be challenging,  
568 particularly when marker genes are not conserved across species. Establishing cell type  
569 homologies across species can generate hypotheses about conserved and divergent cell  
570 features, and facilitates the larger, open access efforts to profile single cells across the brain  
571 underway in mouse, monkey, and human through the BRAIN Initiative<sup>24</sup> and the Human Cell  
572 Atlas<sup>25</sup>. The current data are made publicly available with two new viewer applications to mine  
573 expression data across transcriptomic cell types in both human and mouse cortex ([www.brain-](http://www.brain-map.org)  
574 [map.org](http://www.brain-map.org); [viewer.cytosplore.org](http://viewer.cytosplore.org)).

575 Interestingly, whereas excitatory neuron types are traditionally referred to as being confined to a  
576 single cortical layer, we find instead that many transcriptomically-defined excitatory types are  
577 represented in multiple layers. In part, this may reflect indistinct laminar boundaries in MTG; for  
578 example, von Economo<sup>40</sup> noted intermixing of granule and pyramidal neurons in layer 4 along  
579 with blending of layer 4 pyramidal neurons into adjacent layers 3 and 5 in MTG. However, we  
580 find several types with broad spatial distributions across multiple layer boundaries, suggesting  
581 that indistinct laminar boundaries do not fully account for this lack of strict laminar segregation.  
582 Examination of the spatial distribution of excitatory neuron types in additional cortical areas will  
583 be necessary to determine if this is a particular feature of MTG or a more widespread  
584 phenomenon in human cortex.

585  
586 The transcriptomic cellular organization and diversity in human MTG are surprisingly similar to  
587 those of mouse V1<sup>26</sup>, despite many differences in these data sets. First, mouse scRNA-seq  
588 was compared to human snRNA-seq, and to mitigate this, expression levels were estimated  
589 using intronic sequence that should be almost exclusively retained in the nucleus<sup>33</sup>. Second,  
590 young adult (~8-week-old) mice were compared to older (24-66 years) human specimens;  
591 however, prior transcriptomic studies demonstrated stable gene expression throughout  
592 adulthood in human<sup>64,65</sup>. Third, MTG in human was compared to V1 in mouse. This areal  
593 difference is expected to primarily affect comparison of excitatory neurons that vary more  
594 between regions than inhibitory neurons or glia<sup>26</sup>. Finally, scRNA-seq introduces significant  
595 biases due to differential survival of cell types during dissociation, necessitating the use of Cre-  
596 lines to enrich for under-sampled and rare cell types in mouse cortex<sup>26</sup>. In contrast, we found  
597 that snRNA-seq provides more unbiased sampling and estimates of cell type proportions.  
598 Despite these differences, the human and mouse cell type taxonomies could be matched at  
599 high resolution and reveal a “canonical” cellular architecture that is conserved between cortical  
600 areas and species. Beyond similarities in overall diversity and hierarchical organization, 10 cell



601 types could be unambiguously mapped one-to-one between species, and 28 additional  
602 subclasses could be mapped at a higher level in the taxonomic tree. One-to-one matches were  
603 highly distinctive cell types, including several non-neuronal and neuronal types, such as  
604 chandelier cells. Comparison of absolute numbers of types between studies is challenging, but  
605 no major classes have missing homologous types other than exceedingly rare types that were  
606 likely undersampled in human, such as Cajal-Retzius cells.

607 A striking feature of cortical evolution is the relative expansion of the supragranular layers  
608 involved in cortico-cortical communication<sup>18</sup>. Consistent with this expansion, we find increased  
609 diversity of excitatory neurons in layers 2-4 in human compared to mouse. Layers 2 and 3 are  
610 dominated by three major types, but the most common layer 2/3 type exhibits considerable  
611 transcriptomic heterogeneity in the form of gene expression gradients, which would be expected  
612 to correlate with other cellular phenotypes. We also find expanded diversity of excitatory types  
613 in deep layer 3, along with a surprising increase in diversity in human layer 4 compared to  
614 mouse.

615 We observed several other evolutionary changes in cell type proportions and diversity that  
616 substantially alter the human cortical microcircuit. The relative proportions of major classes of  
617 GABAergic interneurons vary between human MTG and mouse V1, with human MTG having  
618 fewer *PVALB*- and *SST*-expressing interneurons and more *LAMP5/PAX6*- and *VIP*-expressing  
619 interneurons. Since these interneuron classes are derived from the MGE and CGE,  
620 respectively, in mouse, this difference is consistent with increased generation of CGE-derived  
621 interneurons in human<sup>45</sup>. Another major species difference is seen for human layer 5 excitatory  
622 neurons that are homologous to mouse sub-cortically projecting (PT) neurons. Both the  
623 frequency (<1% in human versus approximately 20% in mouse) and diversity (1 type in human  
624 versus 5 types in mouse)<sup>26</sup> of PT neurons are markedly reduced in human, although reduced  
625 diversity may be an artifact of limited sampling in human. The sparsity of this type was  
626 confirmed *in situ* and was not a technical artifact of tissue processing. Rather, this sparsity likely  
627 reflects the 1200-fold expansion of human cortex relative to mouse compared to only 60-fold  
628 expansion of sub-cortical regions that are targets of these neurons<sup>4,5</sup>. If the number of PT  
629 neurons scales with the number of their sub-cortical projection targets, then the 20-fold greater  
630 expansion of cortical neurons would lead to a 20-fold dilution of PT neuron frequency as we  
631 observed. Indeed, the number of human corticospinal neurons, a subset of sub-cortically  
632 projecting neurons, has scaled linearly with the number of target neurons in the spinal cord,  
633 both increasing 40-fold compared to mouse<sup>66,67,68</sup>. Thus, this striking difference in cell type  
634 frequency may be a natural consequence of allometric scaling of the mammalian brain<sup>69</sup>.

635 Our results demonstrate striking species divergence of gene expression between homologous  
636 cell types, as observed in prior studies at the single gene<sup>20</sup> or gross structural level<sup>21</sup>. We find  
637 more than half of all orthologous genes show a major (>10-fold) difference in expression in at  
638 least one of the 38 consensus cell types, and up to 20% of genes in any given cell type showing  
639 such major divergent expression. Several cell types, including endothelial cells, had such  
640 substantial expression divergence that they could not be matched across species using the  
641 methods employed here. These gene expression differences are likely to be functionally  
642 relevant, as divergent genes are associated with neuronal connectivity and signaling, signaling,  
643 including axon guidance genes, ion channels, and neuropeptide signaling. Surprisingly,  
644 serotonin receptors are the most divergent gene family, challenging the use of mouse models  
645 for the many neuropsychiatric disorders involving serotonin signaling<sup>70</sup>. Finally, the more  
646 selectively expressed a gene is in one species the less likely its pattern is to be conserved, and  
647 many well-known markers of specific cell types do not have conserved patterns.



648  
649 Homologous cell types can have highly divergent features in concert with divergent gene  
650 expression. Here, we show that the interlaminar astrocyte, which has dramatic morphological  
651 specialization in primates including human, corresponds to one of two transcriptomic astrocyte  
652 types. A recent scRNA-seq analysis of mouse cortex also found 2 types, with one enriched in  
653 layer 1<sup>28</sup>. However, this mouse astrocyte type had less complex morphology and did not extend  
654 the long-range processes characteristic of interlaminar astrocytes. Thus, a 10-fold increase in  
655 size, the formation of a long process, and other phenotypic differences<sup>17,55,54</sup> are evolutionary  
656 variations on a conserved genetically defined cell type. Similarly, a recent study identified the  
657 rosehip interneuron in human layer 1<sup>19</sup>, which showed species differences in anatomy,  
658 physiology and marker gene profiles suggesting that it is a novel type of interneuron in human  
659 cortex. In fact, we now find that this rosehip type can be mapped to a mouse neurogliaform  
660 interneuron type. Thus, phenotypic differences large enough to define cell types with  
661 conventional criteria represent relatively minor variation on a conserved genetic blueprint for  
662 neurons as well.

663 Together these observations quantitatively frame the debate of whether human cortex is  
664 different from that of other mammals<sup>9,10</sup>, revealing a basic transcriptomic similarity of cell types  
665 punctuated by differences in proportions and gene expression between species that could  
666 greatly influence microcircuit function. The current results help to resolve the seeming paradox  
667 of conserved structure across mammals but failures in the use of mouse for pre-clinical studies  
668<sup>71,70</sup>, and they highlight the need to analyze the human brain in addition to model organisms.  
669 The magnitude of differences between human and mouse suggest that similar profiling of more  
670 closely related non-human primates will be necessary to study many aspects of human brain  
671 structure and function. The enhanced resolution afforded by these molecular technologies also  
672 has great promise for accelerating a mechanistic understanding of brain evolution and disease.

## 673 **Acknowledgements**

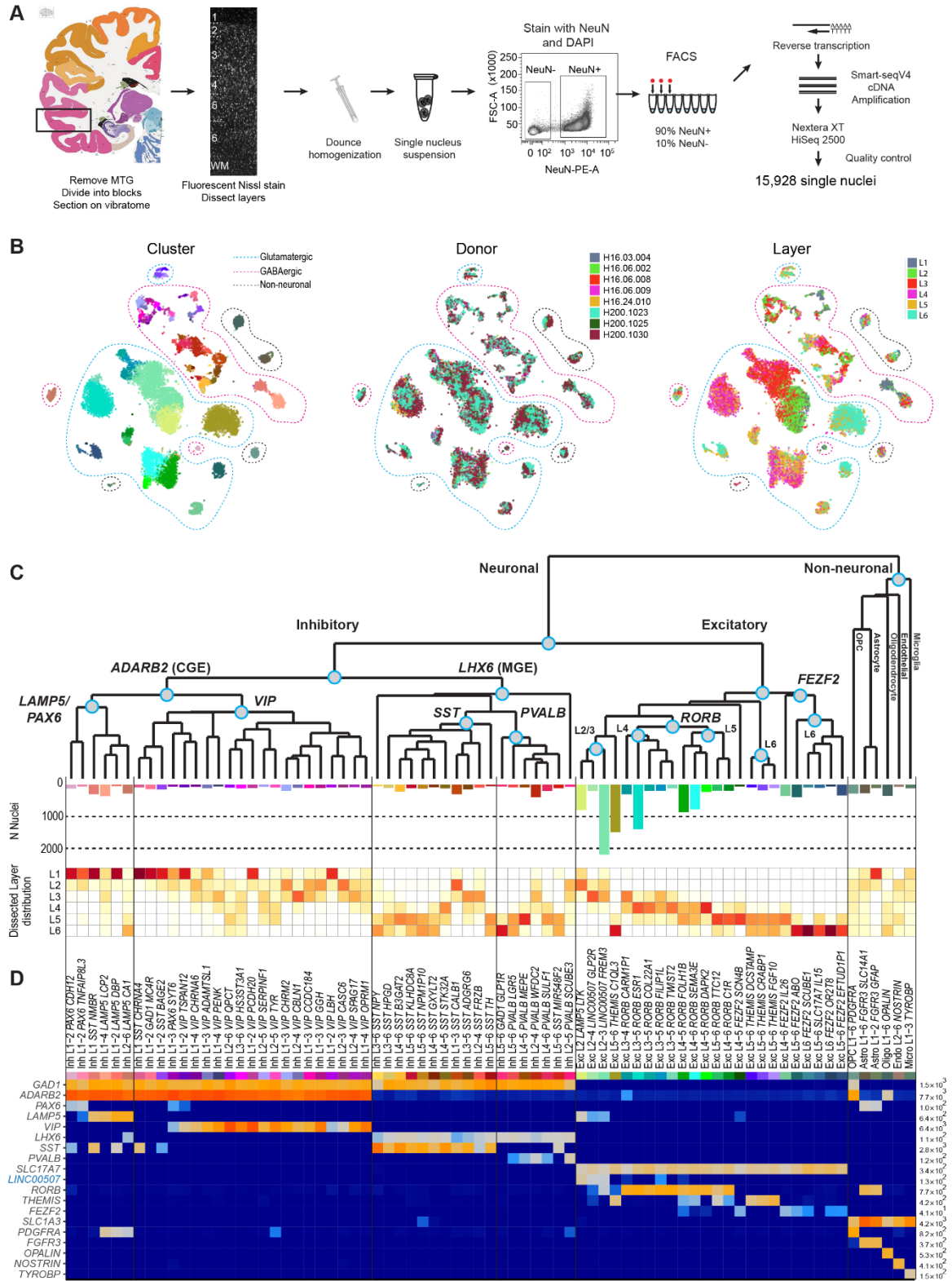
674 We would like to thank the Tissue Procurement, Tissue Processing, and Facilities teams at the  
675 Allen Institute for Brain Science for assistance with the transport and processing of postmortem  
676 and neurosurgical brain specimens. We thank the Technology team at the Allen Institute for  
677 assistance with data management. We gratefully acknowledge our collaborators at local  
678 hospitals (Swedish Medical Center, Harborview Medical Center/UW Medicine, and University of  
679 Washington Medical Center) for help with the coordination of human neurosurgical tissue  
680 collections. We thank Joe Davis and the San Diego Medical Examiner's Office for assistance  
681 with postmortem tissue donations. We acknowledge the Molecular Biology, Histology, and  
682 Imaging teams at the Allen Institute for Brain Science for performing chromogenic *in situ*  
683 hybridization experiments. This work was funded by the Allen Institute for Brain Science, and by  
684 US National Institutes of Health grant 5 U01 MH114812-02 to E.S.L. Funding from NWO-AES  
685 projects 12721: 'Genes in Space' and 12720: 'VANPIRE' (P.I. Anna Vilanova) for development  
686 of the Cytosplore MTG Viewer is gratefully acknowledged. We thank Baldur van Lew for  
687 scripting and narration of Cytosplore instructional and use case videos. The authors thank the  
688 Allen Institute founder, Paul G. Allen, for his vision, encouragement, and support.

## 689 **Author Contributions**

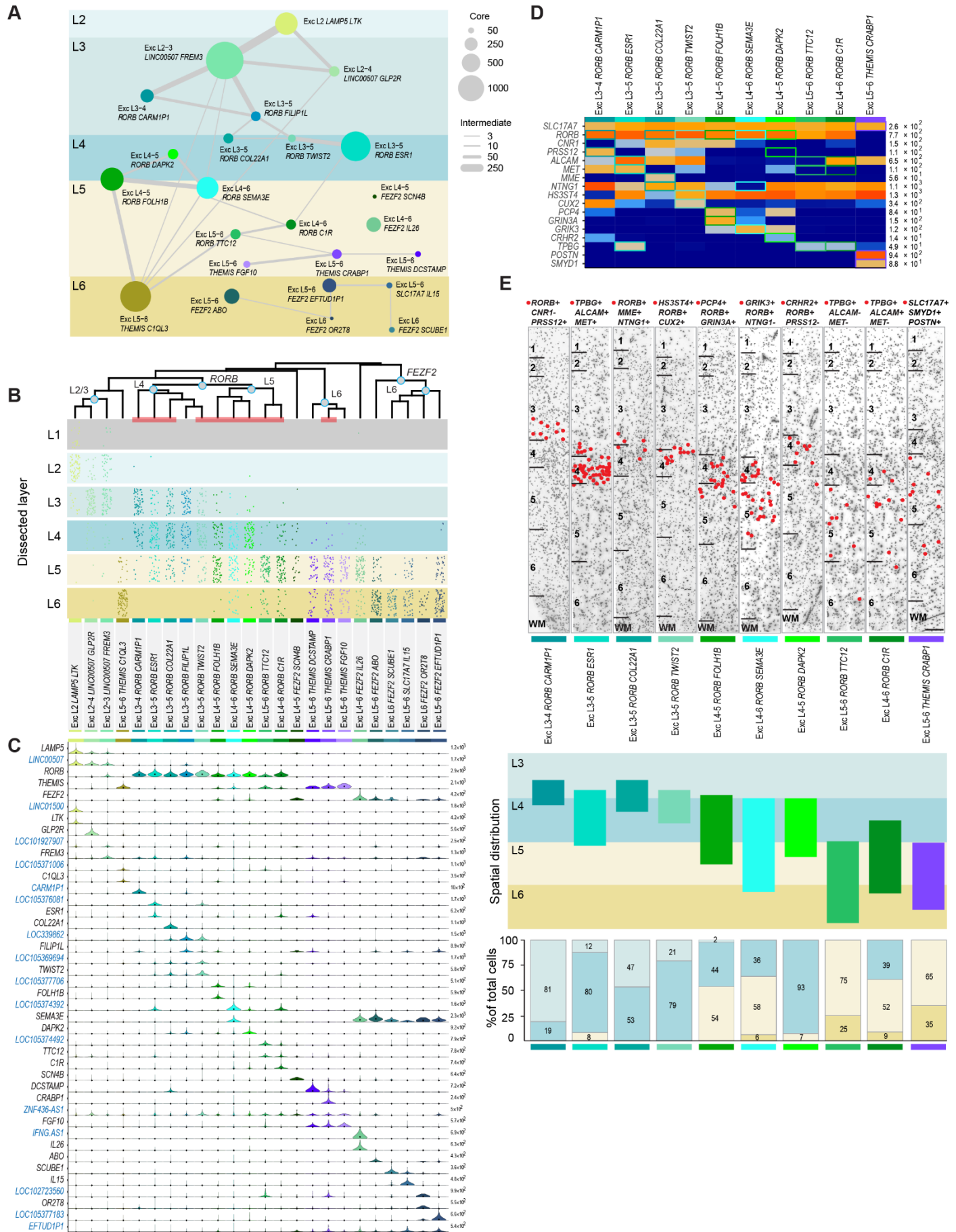
690 E.S.L conceptualized and supervised the study. E.S.L. and R.Y. conceptualized the Human Cell  
691 Types Program. R.D.H and T.E.B. designed experiments. R.D.H., E.R.B., B. Long., J.L.C.,  
692 B.P.L., S.I.S., K.B, J.G., D.H., S.L.D., M.M., S.P., E.R.T, N.V.S., and Z.M. contributed to nuclei  
693 isolation and/or validation experiments. T.E.B., J.A.M., O.P., Z.Y., O.F., J.G., S.S., and M.H.

694 contributed to computational analyses. K.A.S. and B.T. managed the single-nucleus RNA-seq  
695 pipeline. L.T.G. developed data visualization tools. B.T. and H.Z. provided the mouse cortex  
696 transcriptomic cell type taxonomy for the cross-species comparative study. D.B., K.L., C.R. and  
697 M.T. performed single-nucleus RNA-seq. A. Bernard and J.P. managed establishment of single-  
698 nucleus RNA-seq pipeline. A. Bernard and M.M contributed to the development and  
699 management of histological methods and data generation. K.B. performed  
700 immunohistochemistry experiments. R.D., N.D., T.C., J.N., A.O. processed postmortem brain  
701 tissues. A. Bernard and N.D. managed acquisition of postmortem and neurosurgical tissues. A.  
702 Beller, C.D.K, C.C., R.G.E., R.P.G., A.L.K, and J.G.O. contributed to neurosurgical tissue  
703 collections. B.A., M.K., and R.H.S. developed the semantic representation of clusters. J.E., T.H.,  
704 A.M., and B. Lelieveldt developed the Cytosplore MTG Viewer. L.T.G., J.A.M., D.F., L.N, and A.  
705 Bernard contributed to the development of the RNA-Seq Data Navigator. S.R., A.S., and S.M.S.  
706 provided program management and/or regulatory compliance support. C.K. and A.R.J. provided  
707 institutional support and project oversight. E.S.L. and H.Z. directed the Allen Institute Cell Types  
708 Program. R.D.H., T.E.B., and E.S.L. wrote the paper with contributions from J.A.M and J.L.C.,  
709 and in consultation with all authors.

710 **Figures**

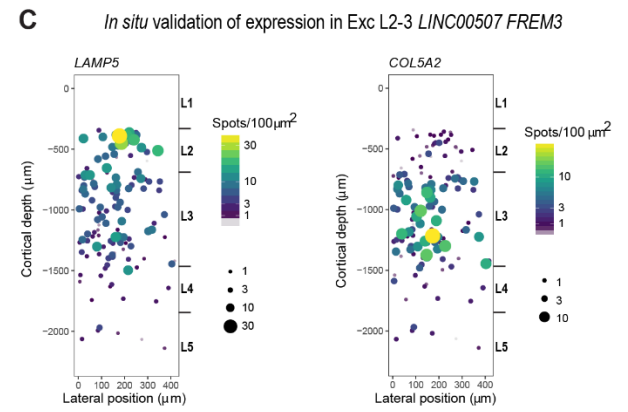
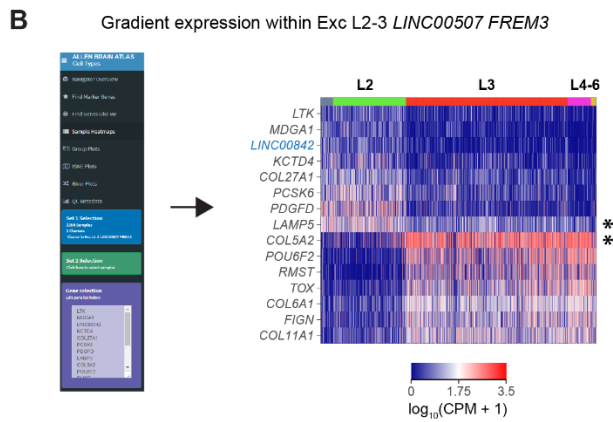
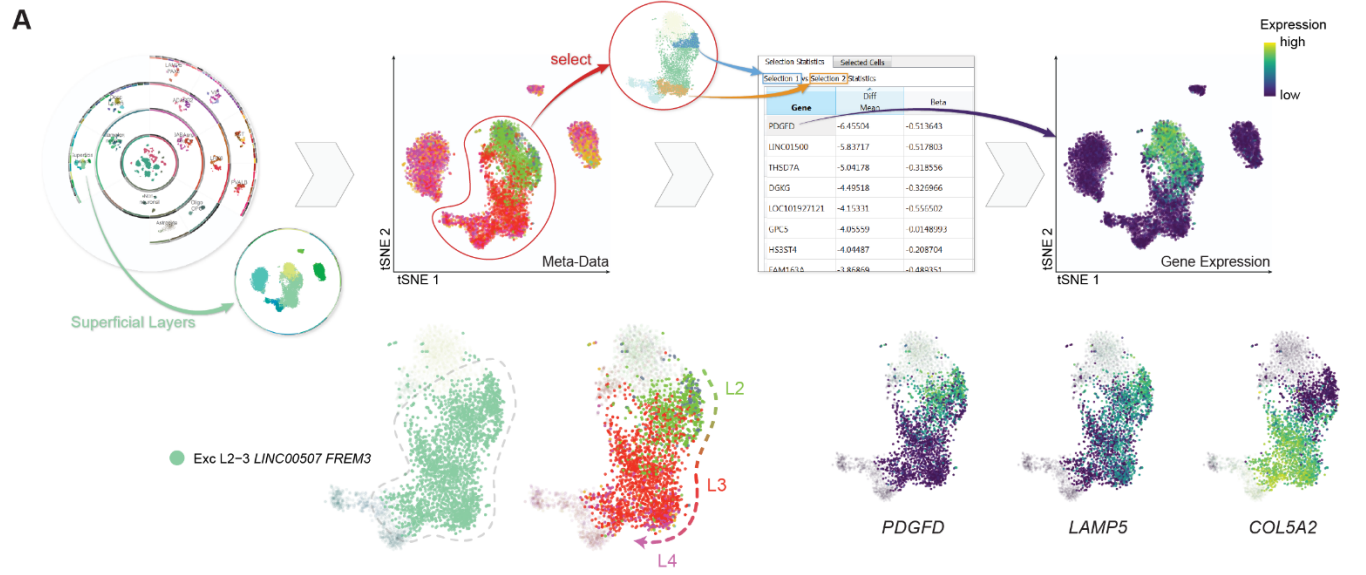


712 **Figure 1. Cell type taxonomy in human middle temporal gyrus (MTG).** (A) Schematic  
713 diagram illustrating nuclei isolation from frozen MTG specimens by vibratome sectioning,  
714 fluorescent Nissl staining and dissection of specific cortical layers. Single neuronal (NeuN+) and  
715 non-neuronal (NeuN-) nuclei were collected by fluorescence-activated cell sorting (FACS), and  
716 RNA-sequencing of single nuclei used SMART-seqv4, Nextera XT, and HiSeq2500  
717 sequencing. (B) Overview of transcriptomic cell type clusters visualized using t-distributed  
718 stochastic neighbor embedding (t-SNE). On the left t-SNE map, each dot corresponding to one  
719 of 15,928 nuclei has a cell-type specific color that is used throughout the remainder of the  
720 manuscript. In the middle, donor metadata is overlaid on the t-SNE map to illustrate the  
721 contribution of nuclei from different individuals to each cluster. In the list of specimens,  
722 H16.03.004-H16.06.009 are neurosurgical tissue donors and H16.24.010-H200.1030 are  
723 postmortem donors. On the right, layer metadata is overlaid on the t-SNE map to illustrate the  
724 laminar composition of each cluster. (C) Hierarchical taxonomy of cell types based on median  
725 cluster expression consisting of 69 neuronal (45 inhibitory, 24 excitatory) and 6 non-neuronal  
726 transcriptomic cells types. Major cell classes are labeled at branch points in the dendrogram.  
727 The bar plot below the dendrogram represents the number of nuclei within each cluster. The  
728 laminar distributions of clusters are shown in the plot that follows. For each cluster, the  
729 proportion of nuclei in each layer is depicted using a scale from white (low) to dark red  
730 (high). (D) Heatmap showing the expression of cell class marker genes (blue, non-coding)  
731 across clusters. Maximum expression values for each gene are listed on the far-right hand side.  
732 Gene expression values are quantified as counts per million of intronic plus exonic reads and  
733 displayed on a  $\log_{10}$  scale.

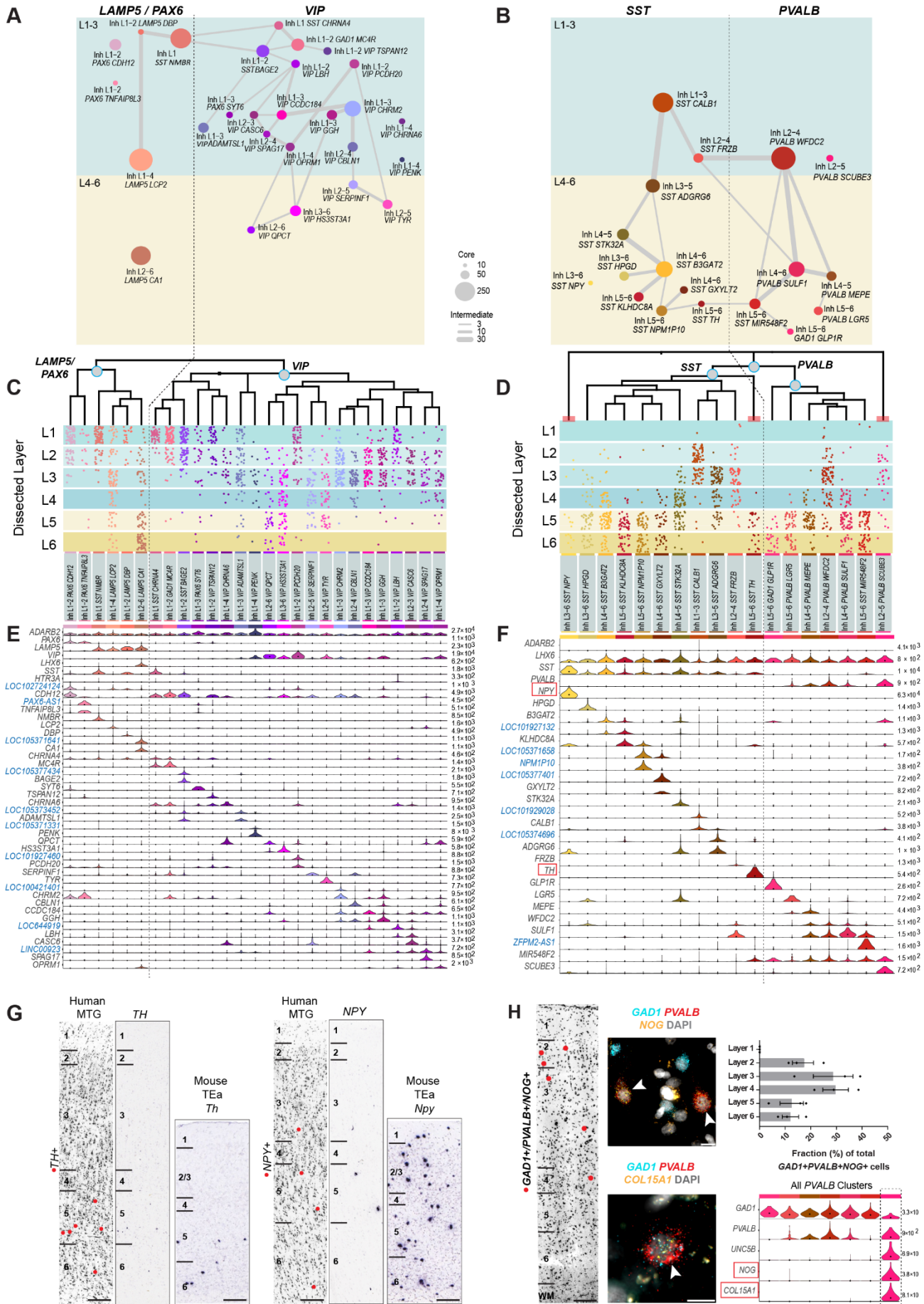




735 **Figure 2. Excitatory neuron diversity and marker gene expression. (A)** Constellation  
736 diagram for excitatory cell types. The number of cells that could be unambiguously assigned to  
737 each cluster (core cells) is represented by disc area and the number of cells with uncertain  
738 membership between each pair of clusters (intermediate cells) is represented by line  
739 thickness. **(B)** Dendrogram illustrating overall gene expression similarity between cell types.  
740 Layer distributions of cell types are shown as dot plots where each dot represents a single  
741 nucleus from a layer-specific dissection. Note that incidental capture of some layer 2 excitatory  
742 neurons occurred in layer 1 dissections and is reflected in the dot plots. Clusters marked by a  
743 red bar at the base of the dendrogram are examined using fluorescent in situ hybridization  
744 (FISH) in (D-E). **(C)** Violin plot showing marker gene (blue, non-coding) expression distributions  
745 across clusters. Each row represents a gene, black dots show median gene expression within  
746 clusters, and the maximum expression value for each gene is shown on the right-hand side of  
747 each row. Gene expression values are shown on a linear scale. **(D)** Heatmap summarizing  
748 combinatorial 3-gene panels used for multiplex fluorescent in situ hybridization assays to explore  
749 the spatial distribution of 10 excitatory clusters. Gene combinations for each cluster are  
750 indicated by colored boxes on the heatmap. **(E)** Representative inverted images of DAPI-  
751 stained cortical columns spanning layers 1-6 for each marker gene panel. Red dots depict the  
752 locations of cells positive for the specific marker gene combinations for each cluster. Marker  
753 gene combinations are listed at the top of each image. Cluster names along with color coded  
754 cluster-specific bars are beneath each panel. Scale bar, 250 $\mu$ m. Below the DAPI images, a  
755 schematic diagram of the spatial distribution (i.e. the laminar extent) of each cluster examined.  
756 The schematic is based on the observed positions of labeled cells across n=3-4 sections per  
757 cell type and n=2-3 donors per cell type. Bar plots below summarize counts of the percentage of  
758 labeled cells per layer, expressed as a fraction of the total number of labeled cells for each type.  
759 Bars are color coded to represent different cortical layers using the scheme shown in (A). The  
760 cluster represented by each bar is indicated by the colored bar at the bottom of the plot. Cell  
761 counts are cumulative values from n=2-3 subjects for each cell type.

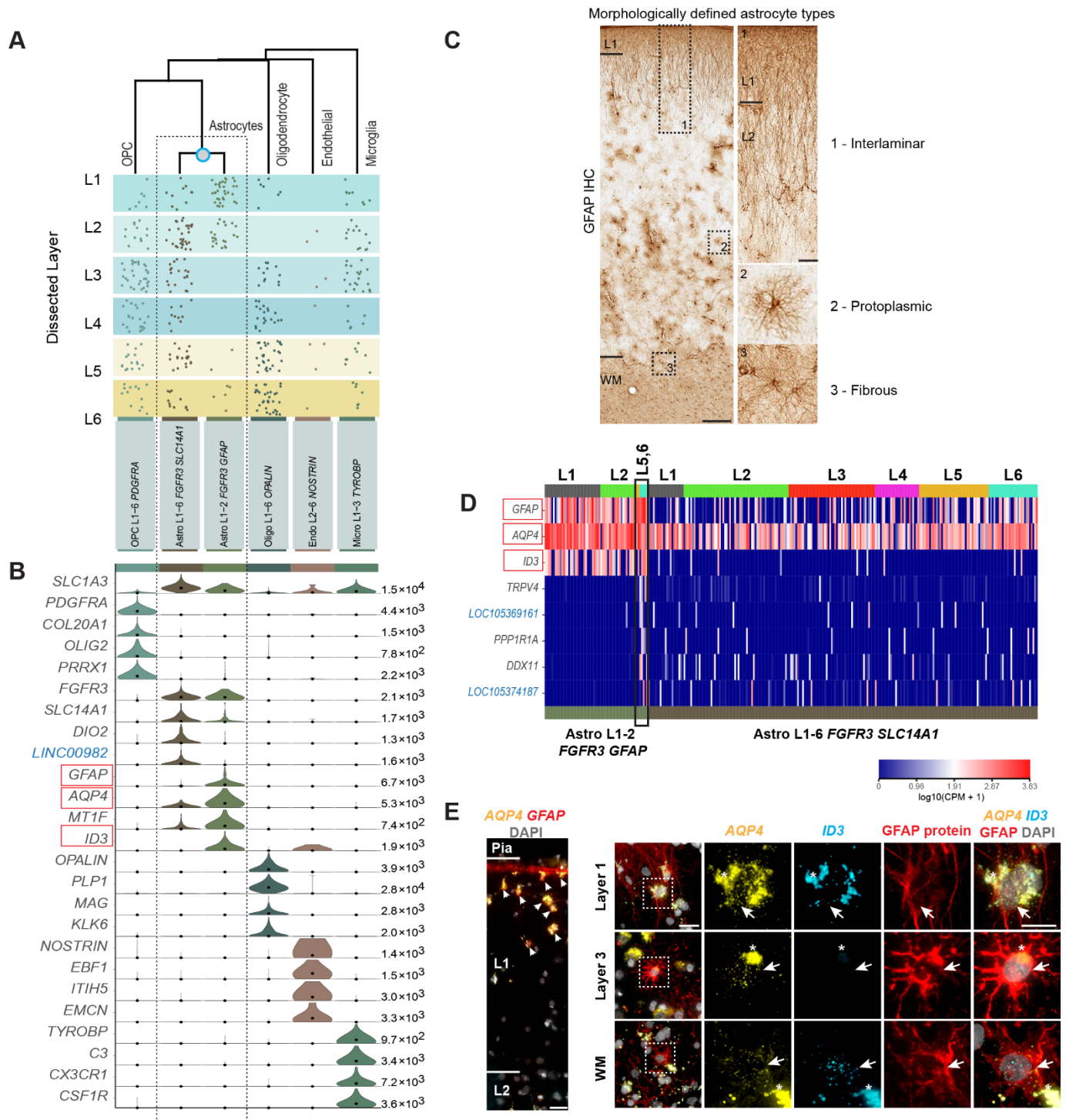


763 **Figure 3. Gene expression heterogeneity within the Exc L2-3 *LINC00507* *FREM3* cell**  
764 **type. (A, B)** Transcriptomics data visualization tools for exploring gene expression gradients in  
765 human cortical neurons. **(A)** Cytosplore MTG Viewer. Top panels, left to right: the hierarchy  
766 viewer shows an overview of the t-SNE map of all clusters. Zooming in allows for visualization  
767 and selection of superficial layer excitatory neurons on the t-SNE map. Overlaying layer  
768 metadata on the t-SNE map shows that nuclei within the EXC L2-3 *LINC00507* *FREM3* cell type  
769 are sorted by cortical layer. Differential expression analysis, computed by selecting nuclei on  
770 opposite ends of the cluster, reveals gene expression gradients organized along the layer  
771 structure of the cluster. Bottom panels, left to right: t-SNE map showing the EXC L2-3  
772 *LINC00507* *FREM3* cluster outlined by dashed gray line. Overlaying layer metadata on the  
773 cluster highlights its layer structure. Examples of genes that exhibit expression heterogeneity  
774 across the layer structure of the cluster are shown to the right. **(B)** RNA-Seq Data Navigator.  
775 Selection of the sample heatmaps option in the browser allows for visualization of gene  
776 expression patterns in the EXC L2-3 *LINC00507* *FREM3* cluster. Each row in the heatmap  
777 represents a gene (blue, non-coding), and nuclei in the cluster are ordered by layer (colored bar  
778 at the top of the heatmap). The selected genes illustrate opposing gene expression gradients  
779 across the layer structure of the cluster. Genes marked with an asterisk were included in the  
780 validation experiments in (C). **(C)** Single molecule fluorescent in situ hybridization (smFISH)  
781 validation of gene expression heterogeneity. Panels show quantification of *LAMP5* (left)  
782 and *COL5A2* (right) expression in cells located in layers 2-3. Each circle represents a cell, the  
783 size of each circle is proportional to the number of smFISH spots per cell, and circles are color-  
784 coded per the scale shown to the right of each panel. Consistent with the RNA-seq data shown  
785 in panels A and B, smFISH analysis demonstrates that these genes exhibit opposing expression  
786 gradients across cortical layers 2 and 3.



788 **Figure 4. Inhibitory neuron diversity and marker gene expression. (A, B)** Constellation  
789 diagrams for *LAMP5/PAX6* and *VIP* (A) and *SST/PVALB* (B) subclasses. The number of core  
790 cells within each cluster is represented by disc area and the number of intermediate cells by  
791 weighted lines. **(C, D)** Dendrograms illustrate gene expression similarity between cell types.  
792 Below each dendrogram, the spatial distribution of each type is shown. Each dot represents a  
793 single nucleus derived from a layer-specific dissection. Red bars at the base of the dendrogram  
794 in (D) indicate clusters examined using *in situ* hybridization (ISH) in (G-H). **(E, F)** Violin plots of  
795 marker gene expression distributions across clusters. Rows are genes (blue, non-coding  
796 transcripts), black dots in each violin represent median gene expression within clusters, and the  
797 maximum expression value for each gene is shown on the right-hand side of each row. Gene  
798 expression values are shown on a linear scale. Genes shown in (G) are outlined by red boxes in  
799 (F). **(G)** Chromogenic single gene ISH for *TH* (left), a marker of Inh L5–6 *SST TH*, and *NPY*  
800 (right), a marker of Inh L3–6 *SST NPY*, from the Allen Human Brain Atlas. Left columns show  
801 grayscale images of the Nissl stained section nearest the ISH stained section shown in the right  
802 panel for each gene. Red dots overlaid on the Nissl section show the laminar positions of cells  
803 positive for the gene assayed by ISH. Chromogenic ISH for *Th* and *Npy* in mouse temporal  
804 association cortex (TEa) from the Allen Mouse Brain Atlas are to the right of the human ISH  
805 images. Scale bars: human (250 $\mu$ m), mouse (100 $\mu$ m). **(H)** RNAscope multiplex fluorescent ISH  
806 for markers of putative chandelier cell cluster Inh L2–5 *PVALB SCUBE3*. Left panel -  
807 representative inverted DAPI-stained cortical column with red dots marking the position of cells  
808 positive for the genes *GAD1*, *PVALB*, and *NOG* (scale bar, 250 $\mu$ m). Middle - images of cells  
809 positive for *GAD1*, *PVALB*, and the specific marker genes *NOG* (top, scale bar 10 $\mu$ m)  
810 and *COL15A1* (bottom, scale bar 10 $\mu$ m). White arrows mark triple positive cells. Right - bar plot  
811 summarizes counts of *GAD1+*, *PVALB+*, *NOG+* cells across layers (expressed as percentage of  
812 total triple positive cells). Bars show the mean, error bars represent the standard error of the  
813 mean (SEM), and dots represent data points for individual specimens (n=3 subjects). Violin plot  
814 shows gene expression distributions across clusters in the *PVALB* subclass for the chandelier  
815 cell marker *UNC5B* and the Inh L2–5 *PVALB SCUBE3* cluster markers *NOG* and *COL15A1*.

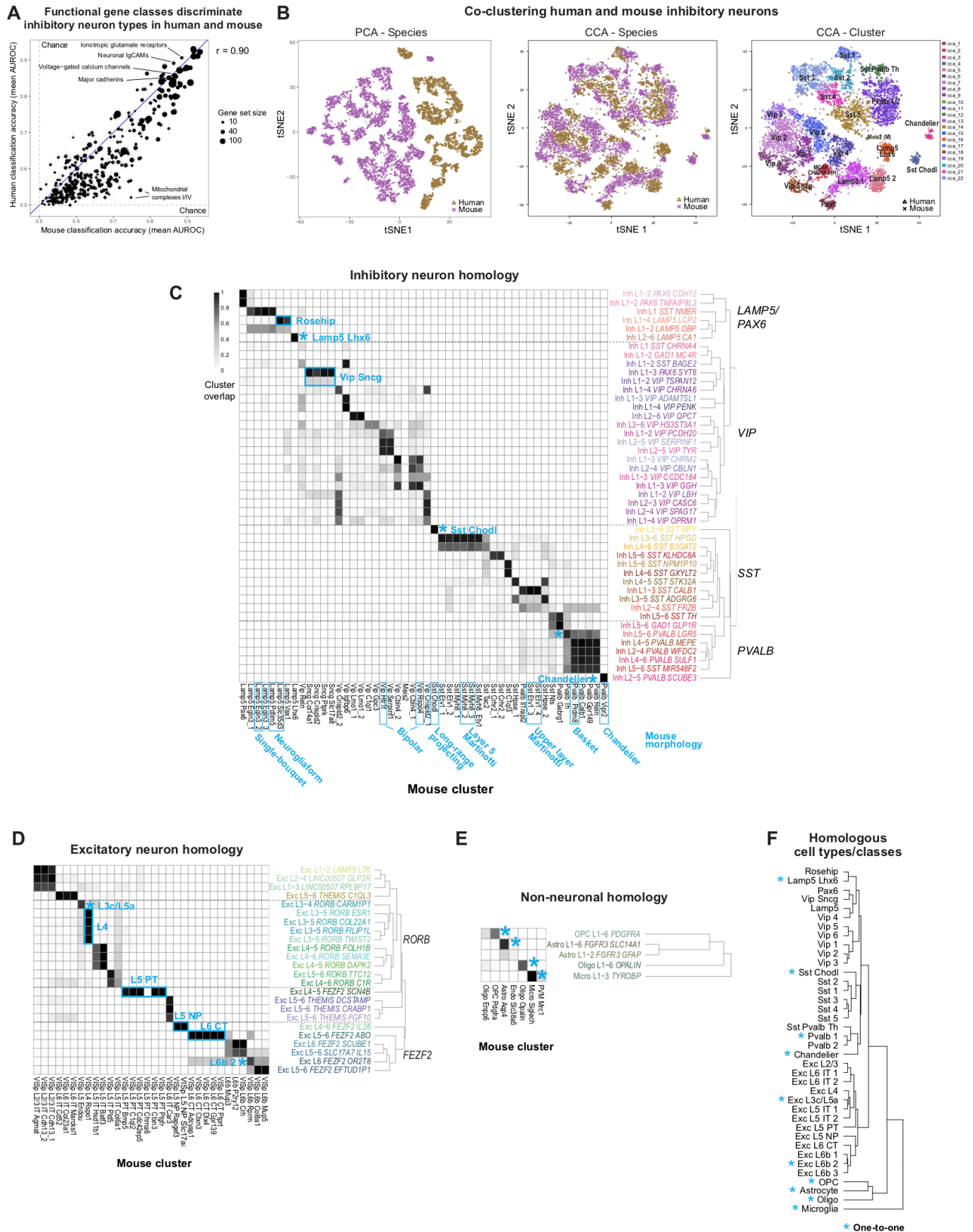




816

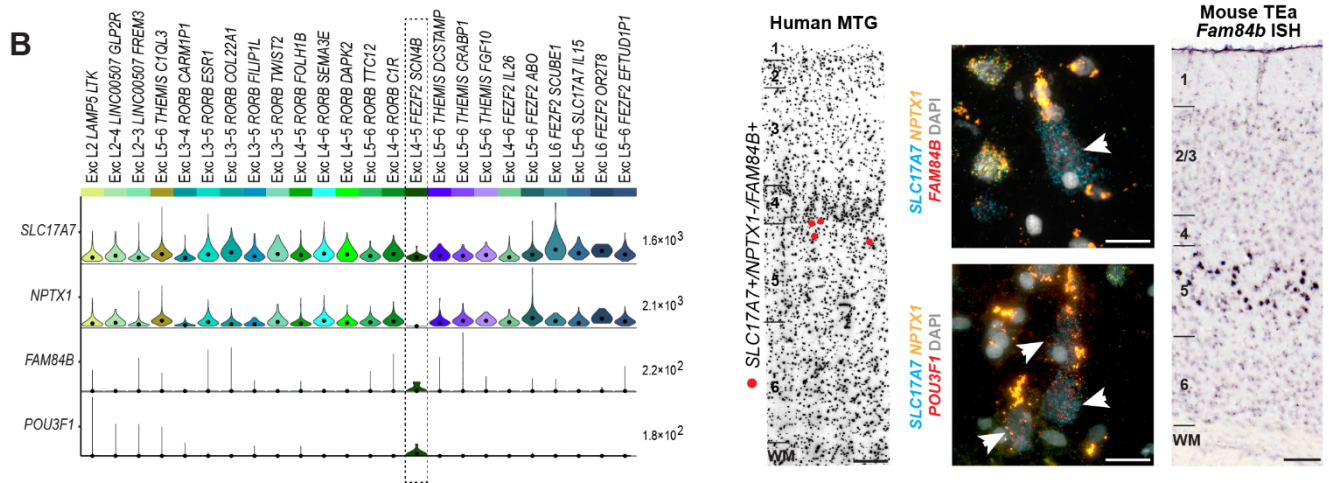
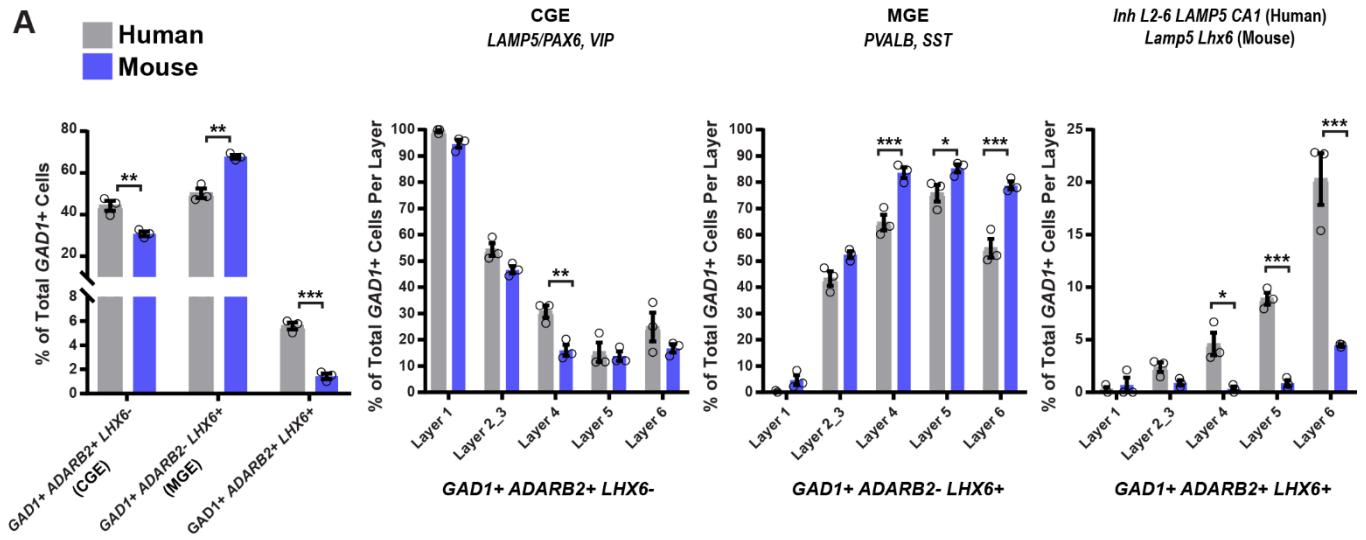
817

818 **Figure 5. Non-neuronal cell type diversity and marker gene expression.** (A) Dendrogram  
819 illustrating overall gene expression similarity between non-neuronal cell types, with the spatial  
820 distribution of types shown beneath the dendrogram. Each dot represents a single nucleus from  
821 a layer-specific dissection. (B) Violin plots show expression distributions of marker genes across  
822 clusters. Each row represents a gene (blue, non-coding), black dots represent median gene  
823 expression within clusters, and the maximum expression value for each gene is shown on the  
824 right-hand side of each row. Gene expression values are shown on a linear  
825 scale. (C) Immunohistochemistry (IHC) for GFAP in human MTG illustrates the features of  
826 morphologically-defined astrocyte types. Black boxes on the left panel indicate regions shown at  
827 higher magnification on the right. Scale bars: low mag (250µm), high mag (50µm). (D) Heatmap  
828 illustrating marker gene expression in the Astro L1-2 *FGFR3 GFAP* and Astro L1-6 *FGFR3*  
829 *SLC14A1* clusters. Each row is a gene, each column a single nucleus, and the heatmap is  
830 ordered per the layers that nuclei were dissected from. A minority of nuclei in the Astro  
831 L1-2 *FGFR3 GFAP* cluster came from deep layers (black box on heatmap) and express marker  
832 genes distinct from the other nuclei in the cluster. Red boxes in (B, D) are genes examined in  
833 (E). (E) RNAscope multiplex fluorescent *in situ* hybridization (FISH) for astrocyte markers. Left -  
834 expression of *AQP4* and *GFAP* in layer 1 (scale bar, 25µm). Cells expressing high levels  
835 of *AQP4* and *GFAP*, consistent with the Astro L1-2 *FGFR3 GFAP* cluster, are localized to the  
836 top half of layer 1 (white arrowheads). Right - FISH for *AQP4* and *ID3* combined with GFAP  
837 immunohistochemistry. White box indicates area shown at higher magnification to the right.  
838 Scale bars: low mag (25µm), high mag (15µm). Asterisks mark lipofuscin autofluorescence. Top  
839 row: *AQP4* expressing cells in layer 1 coexpress *ID3* and have long, GFAP-labeled processes  
840 that span layer 1. Middle row: protoplasmic astrocyte located in layer 3 lacks expression of *ID3*,  
841 consistent with the Astro L1-6 *FGFR3 SLC14A1* type. Bottom row: fibrous astrocyte at the  
842 white matter (WM)/layer 6 boundary triple positive for *AQP4*, *ID3*, and GFAP protein.



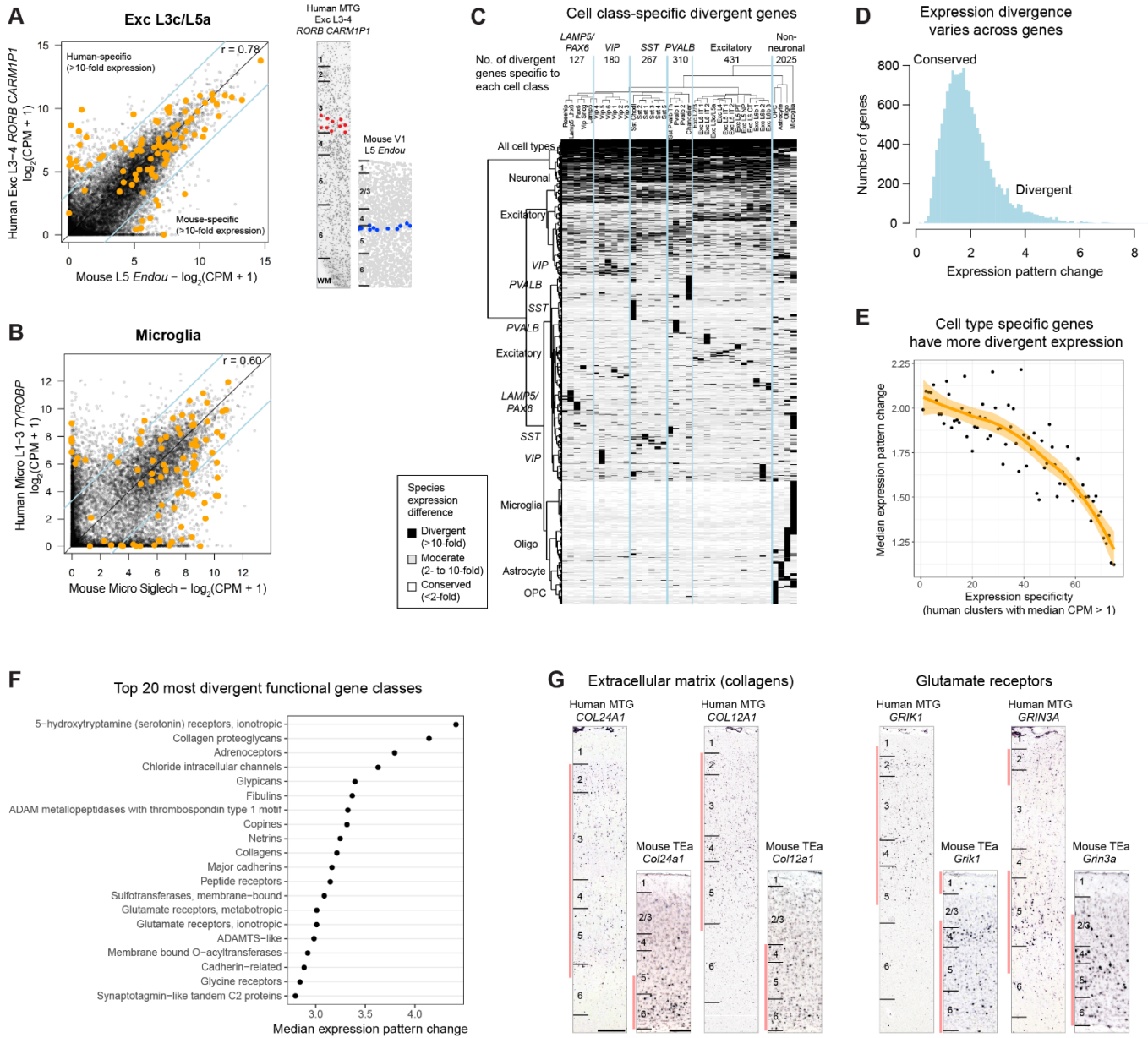
844 **Figure 6. Evolutionary conservation of cell types between human and mouse. (A)**  
845 Inhibitory neurons can be assigned to cell types based on expression patterns of functional  
846 gene families (n = 5 to 261 genes). Classification performance (average MetaNeighbor AUROC  
847 score across clusters) varies among functional classes of genes and is highly correlated (r =  
848 0.90) between human and mouse. Error bars correspond to the standard deviation of average  
849 AUROC scores across ten sub-sampled iterations. **(B)** Human (gold) and mouse (purple)  
850 inhibitory neurons were aligned with principal components analysis (PCA; left) and canonical  
851 correlation analysis (CCA; middle), and the first 30 basis vectors were represented using t-SNE.  
852 Right: CCA clusters were identified by the Louvain algorithm using 30 nearest neighbors and  
853 annotated based on cluster labels from this study and mouse. Clusters labeled with (M) or (H)  
854 contain only mouse cells or human nuclei, respectively. **(C-E)** Human and mouse cell type  
855 homologies for inhibitory neurons (C), excitatory neurons (D), and non-neuronal cells (E) were  
856 predicted based on shared CCA cluster membership. Greyscale indicates, for each pair of  
857 human (rows) and mouse (columns) clusters, the minimum proportion of human nuclei or  
858 mouse cells that co-cluster using CCA. Note that rows and columns need not sum to one  
859 because clusters can partially overlap. One-to-one matches are indicated by an asterisk. Known  
860 morphologies are indicated for mouse inhibitory types and known projection targets are given  
861 for excitatory types (IT - intratelencephalic, PT - pyramidal tract/sub-cortical, NP - near-  
862 projecting, CT - corticothalamic). Note that human endothelial nuclei could not be aligned by  
863 CCA and were excluded from the analysis. **(F)** Hierarchical taxonomy of 34 neuronal and 4 non-  
864 neuronal homologous cell types and cell classes, including 10 cell types that match one-to-one  
865 between human and mouse.





866

867 **Figure 7. Frequency differences in cell classes and types between human and mouse.**  
868 **(A)** Quantification of broad interneuron classes in human MTG and mouse temporal association  
869 area (TEa) based on counts of cells labeled using RNAscope multiplex fluorescent in situ  
870 hybridization (FISH). Sections were labeled with the gene panel *GAD1/Gad1*, *LHX6/Lhx6*,  
871 and *ADARB2/Adarb2* (human/mouse). Bar plots from left to right: (1) Percentage of each major  
872 cell class of total *GAD1+* cells. (2) Percentage of *GAD1+/ADARB2+/LHX6-* cells of total *GAD1+*  
873 cells per layer, representing *LAMP5/PAX6*, and *VIP* types. (3) Percentage of *GAD1+/ADARB2-*  
874 */LHX6+* cells of total *GAD1+* cells per layer, representing all *PVALB* and *SST* types. (4)  
875 Percentage of *GAD1+/ADARB2+/LHX6+* cells of total *GAD1+* cells per layer, representing the  
876 Inh L2–6 *LAMP5 CA1* type (human) or *Lamp5 Lhx6* type (mouse). Bars represent the mean,  
877 error bars the standard error of the mean, and circles show individual data points for each  
878 specimen (n=3 specimens for both human and mouse; t-test with Holm-Sidak correction for  
879 multiple comparisons, \*p<0.05 \*\*p<0.01, \*\*\*p<0.001). **(B)** Left to right: violin plot showing  
880 expression of specific markers of the putative pyramidal tract (PT) EXC L4-5 *FEZF2 SCN4B*  
881 cluster (black box) and *NPTX1*, a gene expressed by all non-PT excitatory neurons. Each row  
882 represents a gene, the black dots in each violin represent median gene expression within  
883 clusters, and the maximum expression value for each gene is shown on the right-hand side of  
884 each row. Gene expression values are shown on a linear scale. A representative inverted  
885 DAPI-stained cortical column (scale bar, 200µm) with red dots marking the position of cells  
886 positive for the genes *SLC17A7* and *FAM84B* and negative for *NPTX1* illustrates the relative  
887 abundance of the EXC L4-5 *FEZF2 SCN4B* type in human MTG. Representative examples of  
888 RNAscope multiplex FISH stained sections from human MTG showing *FAM84B* (top, white  
889 arrows, scale bar, 25µm) and *POU3F1*-expressing cells (bottom, white arrows, scale bar,  
890 25µm). Expression of *Fam84b* in mouse TEa (scale bar, 75µm) is shown in the adjacent panel.



891

892 **Figure 8. Divergent cell type expression between human and mouse. (A)** Left: Comparison  
893 of expression levels of 14,414 orthologous genes between human and mouse for the most  
894 highly conserved one-to-one homologous type, Exc L3c/L5a. Genes outside of the blue lines  
895 have highly divergent expression (>10-fold change) between human and mouse. Approximately  
896 100 genes (orange dots) are relatively specific markers in human and/or mouse. Right: ISH  
897 validation of layer distributions in human MTG and mouse primary visual cortex (data from Tasic  
898 et al., 2017). Cells are labeled based on expression of cluster marker genes in human  
899 (*RORB+/CNR1-/PRSS12+*) and mouse (*Scnn1a+/Hsd11b1+*). **(B)** Comparison of expression  
900 between human and mouse microglia, the least conserved homologous type. **(C)** Patterns of  
901 expression divergence between human and mouse for 8222 genes (57% of orthologous genes)  
902 with at least 10-fold expression change in one or more homologous cell types. Genes were  
903 hierarchically clustered and groups of genes that have similar patterns of expression divergence  
904 are labeled by the affected cell class. Top row: number of genes with expression divergence  
905 restricted to each broad class of cell types. **(D)** For each gene, the expression pattern change  
906 was quantified by the beta score (see Methods) of the absolute log fold change in expression  
907 between human and mouse. Genes with divergent patterns have large expression changes  
908 among a subset of homologous cell types. Genes with conserved patterns have similar  
909 expression levels in human and mouse or have a similar expression level change in all types.  
910 Pattern changes are approximately log-normally distributed, and a minority of genes have highly  
911 divergent patterns. **(E)** Genes expressed in fewer human cell types tended to have greater  
912 evolutionary divergence than more ubiquitously expressed genes. A loess curve and standard  
913 error was fit to median expression pattern changes across genes binned by numbers of clusters  
914 with expression (median CPM > 1). **(F)** Gene families with the most divergent expression  
915 patterns (highest median pattern change) include neurotransmitter receptors, ion channels, and  
916 cell adhesion molecules. **(G)** Genes estimated to have highly divergent expression patterns  
917 have different laminar expression validated by ISH in human and mouse. Red bars highlight  
918 layers with enriched expression. Scale bars: human (250µm), mouse (100µm).



## 919 **Methods**

### 920 **Post-mortem tissue donors**

921 Males and females 18 – 68 years of age with no known history of neuropsychiatric or  
922 neurological conditions ('control' cases) were considered for inclusion in this study (**Extended**  
923 **Data Table 1**). De-identified postmortem human brain tissue was collected after obtaining  
924 permission from decedent next-of-kin. The Western Institutional Review Board (WIRB) reviewed  
925 the use of de-identified postmortem brain tissue for research purposes and determined that, in  
926 accordance with federal regulation 45 CFR 46 and associated guidance, the use of and  
927 generation of data from de-identified specimens from deceased individuals did not constitute  
928 human subjects research requiring IRB review. Postmortem tissue collection was performed in  
929 accordance with the provisions of the Uniform Anatomical Gift Act described in Health and  
930 Safety Code §§ 7150, et seq., and other applicable state and federal laws and regulations.  
931 Routine serological screening for infectious disease (HIV, Hepatitis B, and Hepatitis C) was  
932 conducted using donor blood samples and only donors negative for all three tests were  
933 considered for inclusion in the study. Tissue RNA quality was assessed using an Agilent  
934 Bioanalyzer-generated RNA Integrity Number (RIN) and Agilent Bioanalyzer electropherograms  
935 for 18S/28S ratios. Specimens with RIN values  $\geq 7.0$  were considered for inclusion in the study  
936 (**Extended Data Table 1**).

### 937 **Processing of whole brain postmortem specimens**

938 Whole postmortem brain specimens were transported to the Allen Institute on ice. Standard  
939 processing of whole brain specimens involved bisecting the brain through the midline and  
940 embedding of individual hemispheres in Cavex Impressional Alginate for slabbing. Coronal brain  
941 slabs were cut at 1cm intervals through each hemisphere and individual slabs were frozen in a  
942 slurry of dry ice and isopentane. Slabs were then vacuum sealed and stored at  $-80^{\circ}\text{C}$  until the  
943 time of further use.

944  
945 Middle temporal gyrus (MTG) was identified on and removed from frozen slabs of interest, and  
946 subdivided into smaller blocks for further sectioning. Individual tissue blocks were processed by  
947 thawing in PBS supplemented with 10mM DL-Dithiothreitol (DTT, Sigma Aldrich), mounting on a  
948 vibratome (Leica), and sectioning at  $500\mu\text{m}$  in the coronal plane. Sections were placed in  
949 fluorescent Nissl staining solution (Neurotrace 500/525, ThermoFisher Scientific) prepared in  
950 PBS with 10mM DTT and 0.5% RNasin Plus RNase inhibitor (Promega) and stained for 5 min  
951 on ice. After staining, sections were visualized on a fluorescence dissecting microscope (Leica)  
952 and cortical layers were individually microdissected using a needle blade micro-knife (Fine  
953 Science Tools).

### 954 **Neurosurgical tissue donors**

955 Tissue procurement from neurosurgical donors was performed outside of the supervision of the  
956 Allen Institute at local hospitals, and tissue was provided to the Allen Institute under the  
957 authority of the IRB of each participating hospital. A hospital-appointed case coordinator  
958 obtained informed consent from donors prior to surgery. Tissue specimens were de-identified  
959 prior to receipt by Allen Institute personnel. The specimens collected for this study were  
960 apparently non-pathological tissues removed during the normal course of surgery to access  
961 underlying pathological tissues. Tissue specimens collected were determined to be non-  
962 essential for diagnostic purposes by medical staff and would have otherwise been discarded.

## 963 **Processing of neurosurgical tissue samples**

964 Neurosurgical tissue was transported to the Allen Institute in chilled, oxygenated artificial  
965 cerebrospinal fluid (ACSF) consisting of the following: 0.5 mM calcium chloride (dehydrate), 25  
966 mM D-glucose, 20 mM HEPES, 10 mM magnesium sulfate, 1.2 mM sodium phosphate  
967 monobasic monohydrate, 92 mM N-methyl-d-glucamine chloride (NMDG-Cl), 2.5 mM potassium  
968 chloride, 30 mM sodium bicarbonate, 5 mM sodium L-ascorbate, 3 mM sodium pyruvate, and 2  
969 mM thiourea. The osmolality of the solution was 295-305 mOsm/kg and the pH was 7.3. Slices  
970 were prepared using a Compresstome VF-200 or VF-300 vibratome (Precisionary Instruments).  
971 After sectioning, slices were recovered in ACSF containing 2 mM calcium chloride (dehydrate),  
972 25 mM D-glucose, 20 mM HEPES, 2 mM magnesium sulfate, 1.2 mM sodium phosphate  
973 monobasic monohydrate, 2.5 mM potassium chloride, 30 mM sodium bicarbonate, 92 mM  
974 sodium chloride, 5 mM sodium L-ascorbate, 3 mM sodium pyruvate, and 2 mM thiourea at room  
975 temperature for at least 1 hour. After the recovery period, slices were transferred to RNase-free  
976 microcentrifuge tubes, snap frozen, and stored at -80°C until the time of use. Microdissection of  
977 cortical layers was carried out on tissue slices that were thawed and stained as described above  
978 for postmortem tissue.

## 979 **Nucleus sampling plan**

980 We estimated that 16 cells were required to reliably discriminate two closely related *Sst+*  
981 interneuron types reported by<sup>27</sup>. Monte Carlo simulations were used to estimate the sampling  
982 depth  $N$  needed to be 95% confident that at least 16 nuclei of frequency  $f$  have been selected  
983 from the population. Calculating  $N$  for a range of  $f$  revealed a simple linear approximation:  $N =$   
984  $28 / f$ . Subtypes of mouse cortical layer 5 projection neurons can be rarer than 1% of the  
985 population<sup>72</sup>, so we targeted neuron types as rare as 0.2% of all cortical neurons. We initially  
986 sampled 14,000 neuronal nuclei distributed across cortical layers relative to the proportion of  
987 neurons reported in each layer<sup>36</sup>. We sampled approximately 1000 additional neuronal nuclei  
988 from layers with increased diversity observed based on RNA-seq data. We also targeted 1500  
989 (10%) non-neuronal (NeuN-) nuclei and obtained approximately 1000 nuclei that passed QC,  
990 and we expected to capture types as rare as 3% of the non-neuronal population.

## 991 **Nucleus isolation and sorting**

992 Microdissected tissue pieces were placed in into nuclei isolation medium containing 10mM Tris  
993 pH 8.0 (Ambion), 250mM sucrose, 25mM KCl (Ambion), 5mM MgCl<sub>2</sub> (Ambion) 0.1% Triton-X  
994 100 (Sigma Aldrich), 1% RNasin Plus, 1X protease inhibitor (Promega), and 0.1mM DTT in 1ml  
995 dounce homogenizer (Wheaton). Tissue was homogenized using 10 strokes of the loose  
996 dounce pestle followed by 10 strokes of the tight pestle and the resulting homogenate was  
997 passed through 30µm cell strainer (Miltenyi Biotech) and centrifuged at 900xg for 10 min to  
998 pellet nuclei. Nuclei were resuspended in buffer containing 1X PBS (Ambion), 0.8% nuclease-  
999 free BSA (Omni-Pur, EMD Millipore), and 0.5% RNasin Plus. Mouse anti-NeuN conjugated to  
1000 PE (EMD Millipore) was added to preparations at a dilution of 1:500 and samples were  
1001 incubated for 30 min at 4°C. Control samples were incubated with mouse IgG1,k-PE Isotype  
1002 control (BD Pharmingen). Samples were then centrifuged for 5 min at 400xg to pellet nuclei and  
1003 pellets were resuspended in 1X PBS, 0.8% BSA, and 0.5% RNasin Plus. DAPI (4', 6-diamidino-  
1004 2-phenylindole, ThermoFisher Scientific) was applied to nuclei samples at a concentration of  
1005 0.1µg/ml.

1006 Single nucleus sorting was carried out on either a BD FACSAria II SORP or BD FACSAria  
1007 Fusion instrument (BD Biosciences) using a 130µm nozzle. A standard gating strategy was  
1008 applied to all samples. First, nuclei were gated on their size and scatter properties and then on

1009 DAPI signal. Doublet discrimination gates were used to exclude nuclei aggregates. Lastly,  
1010 nuclei were gated on NeuN signal (PE). Ten percent of nuclei were intentionally sorted as  
1011 NeuN-negative and the remaining 90% of nuclei were NeuN-positive. Single nuclei were sorted  
1012 into 8-well strip tubes containing 11.5µl of SMART-seq v4 collection buffer (Takara)  
1013 supplemented with ERCC MIX1 spike-in synthetic RNAs at a final dilution of 1x10<sup>-8</sup> (Ambion).  
1014 Strip tubes containing sorted nuclei were briefly centrifuged and stored at -80°C until the time of  
1015 further processing. Index sorting was carried out for most samples to allow properties of nuclei  
1016 detected during sorting to be connected with the cell type identity revealed by subsequent  
1017 snRNA-seq.

## 1018 **RNA-sequencing**

1019 We used the SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (Takara #634894) per the  
1020 manufacturer's instructions for reverse transcription of RNA and subsequent cDNA  
1021 amplification. Standard controls were processed alongside each batch of experimental  
1022 samples. Control strips included: 2 wells without cells, 2 wells without cells or ERCCs (i.e. no  
1023 template controls), and either 4 wells of 10 pg of Human Universal Reference Total RNA  
1024 (Takara 636538) or 2 wells of 10 pg of Human Universal Reference and 2 wells of 10 pg Control  
1025 RNA provided in the Clontech kit. cDNA was amplified with 21 PCR cycles after the reverse  
1026 transcription step. AMPure XP Bead (Beckman Coulter A63881) purification was done using an  
1027 Agilent Bravo NGS Option A instrument with a bead ratio of 1x, and purified cDNA was eluted in  
1028 17 µl elution buffer provided by Takara. All samples were quantitated using PicoGreen®  
1029 (ThermoFisher Scientific) on a Molecular Dynamics M2 SpectraMax instrument. cDNA libraries  
1030 were examined on either an Agilent Bioanalyzer 2100 using High Sensitivity DNA chips or an  
1031 Advanced Analytics Fragment Analyzer (96) using the High Sensitivity NGS Fragment Analysis  
1032 Kit (1bp-6000bp). Purified cDNA was stored in 96-well plates at -20°C until library preparation.

1033 The NexteraXT DNA Library Preparation (Illumina FC-131-1096) kit with NexteraXT Index Kit  
1034 V2 Sets A-D (FC-131-2001, 2002, 2003, or 2004) was used for sequencing library preparation.  
1035 NexteraXT DNA Library prep was done at either 0.5x volume manually or 0.4x volume on the  
1036 Mantis instrument (Formulatrix). Three different cDNA input amounts were used in generating  
1037 the libraries: 75pg, 100pg, and 125pg. AMPure XP bead purification was done using the Agilent  
1038 Bravo NGS Option A instrument with a bead ratio of 0.9x and all samples were eluted in 22 µl of  
1039 Resuspension Buffer (Illumina). Samples were quantitated using PicoGreen on a Molecular  
1040 Dynamics M2 SpectraMax instrument. Sequencing libraries were assessed using either an  
1041 Agilent Bioanalyzer 2100 with High Sensitivity DNA chips or an Advanced Analytics Fragment  
1042 Analyzer with the High Sensitivity NGS Fragment Analysis Kit for sizing. Molarity was  
1043 calculated for each sample using average size as reported by Bioanalyzer or Fragment  
1044 Analyzer and pg/µl concentration as determined by PicoGreen. Samples were normalized to 2-  
1045 10 nM with Nuclease-free Water (Ambion). Libraries were multiplexed at 96 samples/lane and  
1046 sequenced on an Illumina HiSeq 2500 instrument using Illumina High Output V4 chemistry.

## 1047 **RNA-seq gene expression quantification**

1048 Raw read (fastq) files were aligned to the GRCh38 human genome sequence (Genome  
1049 Reference Consortium, 2011) with the RefSeq transcriptome version GRCh38.p2 (current as of  
1050 4/13/2015) and updated by removing duplicate Entrez gene entries from the gtf reference file for  
1051 STAR processing. For alignment, Illumina sequencing adapters were clipped from the reads  
1052 using the fastqMCF program<sup>73</sup>. After clipping, the paired-end reads were mapped using Spliced  
1053 Transcripts Alignment to a Reference (STAR)<sup>74</sup> using default settings. STAR uses and builds its  
1054 own suffix array index which considerably accelerates the alignment step while improving on  
1055 sensitivity and specificity, due to its identification of alternative splice junctions. Reads that did

1056 not map to the genome were then aligned to synthetic constructs (i.e. ERCC) sequences and  
1057 the E. coli genome (version ASM584v2). The final results files included quantification of the  
1058 mapped reads (raw exon and intron counts for the transcriptome-mapped reads). Also, part of  
1059 the final results files are the percentages of reads mapped to the RefSeq transcriptome, to  
1060 ERCC spike-in controls, and to E. coli. Quantification was performed using `summerizeOverlaps`  
1061 from the R package `GenomicAlignments`<sup>75</sup>. Read alignments to the genome (exonic, intronic,  
1062 and intergenic counts) were visualized as beeswarm plots using the R package `beeswarm`.

1063  
1064 Expression levels were calculated as counts per million (CPM) of exonic plus intronic reads, and  
1065  $\log_2(\text{CPM} + 1)$  transformed values were used for a subset of analyses as described below.  
1066 Gene detection was calculated as the number of genes expressed in each sample with CPM >  
1067 0. CPM values reflected absolute transcript number and gene length, i.e. short and abundant  
1068 transcripts may have the same apparent expression level as long but rarer transcripts. Intron  
1069 retention varied across genes so no reliable estimates of effective gene lengths were available  
1070 for expression normalization. Instead, absolute expression levels were estimated as fragments  
1071 per kilobase per million (FPKM) using only exonic reads so that annotated transcript lengths  
1072 could be used.

### 1073 **Quality control of RNA-seq data**

1074 Nuclei were included for clustering analysis if they passed all of the following quality control  
1075 (QC) thresholds:

- 1076 • >30% cDNA longer than 400 base pairs
- 1077 • >500,000 reads aligned to exonic or intronic sequence
- 1078 • >40% of total reads aligned
- 1079 • >50% unique reads
- 1080 • TA nucleotide ratio > 0.7

1081 After clustering (see below), clusters were identified as outliers if more than half of nuclei co-  
1082 expressed markers of inhibitory (*GAD1*, *GAD2*) and excitatory (*SLC17A7*) neurons or were  
1083 NeuN+ but did not express the pan-neuronal marker *SNAP25*. Median values of QC metrics  
1084 listed above were calculated for each cluster and used to compute the median and inter-quartile  
1085 range (IQR) of all cluster medians. Clusters were also identified as outliers if the cluster median  
1086 QC metrics deviated by more than three times the IQRs from the median of all clusters.

1087  
1088 Clusters were identified as donor-specific if they included fewer nuclei sampled from donors  
1089 than expected by chance. For each cluster, the expected proportion of nuclei from each donor  
1090 was calculated based on the laminar composition of the cluster and laminar sampling of the  
1091 donor. For example, if 30% of layer 3 nuclei were sampled from a donor, then a layer 3-enriched  
1092 cluster should contain approximately 30% of nuclei from this donor. In contrast, if only layer 5  
1093 were sampled from a donor, then the expected sampling from this donor for a layer 1-enriched  
1094 cluster was zero. If the difference between the observed and expected sampling was greater  
1095 than 50% of the number of nuclei in the cluster, then the cluster was flagged as donor-specific  
1096 and excluded.



1097

1098 To confirm exclusion, clusters automatically flagged as outliers or donor-specific were manually  
1099 inspected for expression of broad cell class marker genes, mitochondrial genes related to  
1100 quality, and known activity-dependent genes.

### 1101 **Clustering RNA-seq data**

1102 Nuclei and cells were grouped into transcriptomic cell types using an iterative clustering  
1103 procedure based on community detection in a nearest neighbor graph as described in <sup>33</sup>. Briefly,  
1104 intronic and exonic read counts were summed, and log<sub>2</sub>-transformed expression (CPM + 1) was  
1105 centered and scaled across nuclei. X- and Y-chromosome were excluded to avoid nuclei  
1106 clustering based on sex. Many mitochondrial genes had expression that was correlated with  
1107 RNA-seq data quality, so nuclear and mitochondrial genes downloaded from Human  
1108 MitoCarta2.0 <sup>76</sup> were excluded. Differentially expressed genes were selected while accounting  
1109 for gene dropouts, and principal components analysis (PCA) was used to reduce dimensionality.  
1110 Nearest-neighbor distances between nuclei were calculated using up to 20 principal  
1111 components, Jaccard similarity coefficients were computed, and Louvain community detection  
1112 was used to cluster this graph with 15 nearest neighbors. Marker genes were defined for all  
1113 cluster pairs using two criteria: 1) significant differential expression (Benjamini-Hochberg false  
1114 discovery rate < 0.05) using the R package *limma* and 2) either binary expression (CPM > 1 in  
1115 >50% nuclei in one cluster and <10% in the second cluster) or >100-fold difference in  
1116 expression. Pairs of clusters were merged if either cluster lacked at least one marker gene.  
1117 Clustering was then applied iteratively to each sub-cluster until the occurrence of one of four  
1118 stop criteria: 1) fewer than six nuclei (due to a minimum cluster size of three); 2) no significantly  
1119 variable genes; 3) no significantly variable PCs; 4) no significant clusters.

1120

1121 To assess the robustness of clusters, the iterative clustering procedure described above was  
1122 repeated 100 times for random subsets of 80% of nuclei. A co-clustering matrix was generated  
1123 that represented the proportion of clustering iterations that each pair of nuclei were assigned to  
1124 the same cluster. We defined consensus clusters by iteratively splitting the co-clustering matrix  
1125 as described in <sup>26</sup>. We used the co-clustering matrix as the similarity matrix and clustered using  
1126 either Louvain (>= 4000 nuclei) or Ward's algorithm (< 4000 nuclei). We defined  $N_{k,l}$  as the  
1127 average probabilities of nuclei within cluster  $k$  to co-cluster with nuclei within cluster  $l$ . We  
1128 merged clusters  $k$  and  $l$  if  $N_{k,l} > \max(N_{k,k}, N_{l,l}) - 0.25$  or if the sum of  $-\log_{10}$ (adjusted P-value) of  
1129 differentially expressed genes between clusters  $k$  and  $l$  was less than 150. Finally, we refined  
1130 cluster membership by reassigning each nucleus to the cluster to which it had maximal average  
1131 co-clustering. We repeated this process until cluster membership converged.

1132 Cluster names were defined using an automated strategy which combined molecular  
1133 information (marker genes) and anatomical information (layer of dissection). Clusters were  
1134 assigned a broad class of interneuron, excitatory neuron, microglia, astrocyte, oligodendrocyte  
1135 precursor, oligodendrocyte, or endothelial cell based on maximal median cluster CPM of *GAD1*,  
1136 *SLC17A7*, *TYROBP*, *AQP4*, *PDGFRA*, *OPALIN*, or *NOSTRIN*, respectively. Enriched layers  
1137 were defined as the range of layers which contained at least 10% of the total cells from that  
1138 cluster. Clusters in were then assigned a broad marker, defined by maximal median CPM  
1139 of *PAX6*, *LAMP5*, *VIP*, *SST*, *PVALB*, *LINC00507*, *RORB*, *THEMIS*, *FEZF2*, *TYROBP*, *FGFR3*,  
1140 *PDGFRA*, *OPALIN*, or *NOSTRIN*. Finally, clusters in all broad classes with more than one  
1141 cluster (e.g., interneuron, excitatory neuron, and astrocyte) were assigned a gene showing the  
1142 most specific expression in that cluster. These marker genes had the greatest difference in the  
1143 proportion of expression (CPM > 1) with a cluster compared to all other clusters regardless of  
1144 mean expression level.

## 1145 **Scoring cluster marker genes**

1146 Many genes were expressed in the majority of nuclei in a subset of clusters. A marker score  
1147 (beta) was defined for all genes to measure how binary expression was among clusters,  
1148 independent of the number of clusters labeled (**Supplementary Table 2**). First, the proportion  
1149 ( $x_i$ ) of nuclei in each cluster that expressed a gene above background level (CPM > 1) was  
1150 calculated. Then, scores were defined as the squared differences in proportions normalized by  
1151 the sum of absolute differences plus a small constant ( $\epsilon$ ) to avoid division by zero. Scores  
1152 ranged from 0 to 1, and a perfectly binary marker had a score equal to 1.

$$1153 \quad \beta = \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - x_j)^2}{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j| + \epsilon}.$$

1154  
1155

## 1156 **Assigning core and intermediate cells**

1157 We defined core and intermediate cells as described<sup>26</sup>. Specifically, we used a nearest-centroid  
1158 classifier, which assigns a cell to the cluster whose centroid has the highest Pearson's  
1159 correlation with the cell. Here, the cluster centroid is defined as the median expression of the  
1160 1200 marker genes with the highest beta score. To define core vs. intermediate cells, we  
1161 performed 5-fold cross-validation 100 times. In each round, the cells were randomly partitioned  
1162 into 5 groups, and cells in each group of 20% of the cells were classified by a nearest centroid  
1163 classifier trained using the other 80% of the cells. A cell classified to the same cluster as its  
1164 original cluster assignment more than 90 times was defined as a core cell, the others were  
1165 designated intermediate cells. We define 14,204 core cells and 1,399 intermediate cells, which  
1166 in most cases classify to only 2 clusters (1,345 out of 1,399, 96.1%). Most cells are defined as  
1167 intermediate because they are confidently assigned to a different cluster from the one originally  
1168 assigned (1,220 out of 1,399, 87.2%) rather than because they are not confidently assigned to  
1169 any cluster.

## 1170 **Cluster dendrograms**

1171 Clusters were arranged by transcriptomic similarity based on hierarchical clustering. First, the  
1172 average expression level of the top 1200 marker genes (highest beta scores, as above) was  
1173 calculated for each cluster. A correlation-based distance matrix ( $D_{xy} = \frac{1-\rho(x,y)}{2}$ ) was calculated,  
1174 and complete-linkage hierarchical clustering was performed using the "hclust" R function with  
1175 default parameters. The resulting dendrogram branches were reordered to show inhibitory  
1176 clusters followed by excitatory clusters, with larger clusters first, while retaining the tree  
1177 structure. Note that this measure of cluster similarity is complementary to the co-clustering  
1178 separation described above. For example, two clusters with similar gene expression patterns  
1179 but a few binary marker genes may be close on the tree but highly distinct based on co-  
1180 clustering.

## 1181 **Mapping cell types to reported clusters**

1182 69 neuronal clusters in MTG were matched to 16 neuronal clusters reported by<sup>31</sup> using nearest-  
1183 centroid classifier of expression signatures. Specifically, single nucleus expression data was  
1184 downloaded for 3042 single cells and 25,051 genes. 1359 marker genes (beta score > 0.4) of  
1185 MTG clusters that had a matching gene in the Lake et al. dataset were selected, and the  
1186 median expression for these genes was calculated for all MTG clusters. Next, Pearson's

1187 correlations were calculated between each nucleus in the Lake et al. dataset and all 69 MTG  
1188 clusters based on these 1359 genes. Nuclei were assigned to the cluster with the maximum  
1189 correlation. A confusion matrix was generated to compare the cluster membership of nuclei  
1190 reported by Lake et al. and assigned MTG cluster. The proportion of nuclei in each MTG cluster  
1191 that were members of each of the 16 Lake et al. clusters were visualized as a dot plot with circle  
1192 sizes proportional to frequency and colored by MTG cluster color.

### 1193 **Colorimetric *in situ* hybridization**

1194 *In situ* hybridization (ISH) data for human and mouse cortex was from the Allen Human Brain  
1195 Atlas and Allen Mouse Brain Atlas. All ISH data is publicly accessible at: [www.brain-map.org](http://www.brain-map.org).  
1196 Data was generated using a semi-automated technology platform as described <sup>77</sup>, with  
1197 modifications for postmortem human tissues as previously described <sup>20</sup>. Digoxigenin-labeled  
1198 riboprobes were generated for each human gene such that they would have >50% overlap with  
1199 the orthologous mouse gene in the Allen Mouse Brain Atlas <sup>77</sup>.

### 1200 **GFAP immunohistochemistry**

1201 Tissue slices (350 µm) from neurosurgical specimens were fixed for 2-4 days in 4%  
1202 paraformaldehyde in PBS at 4°C, washed in PBS, and cryoprotected in 30% sucrose.  
1203 Cryoprotected slices were frozen and re-sectioned at 30 µm using a sliding microtome (Leica  
1204 SM2000R). Free floating sections were mounted onto gelatin coated slides and dried overnight  
1205 at 37 °C. Slides were washed in 1X tris buffered saline (TBS), followed by incubation in 3%  
1206 hydrogen peroxide in 1X TBS. Slides were then heated in sodium citrate (pH 6.0) for 20 minutes  
1207 at 98 °C. After cooling, slides were rinsed in MilliQ water followed by 1X TBS. Primary antibody  
1208 (mouse anti-GFAP, EMD Millipore, #MAB360, clone GA5, 1:1500) was diluted in Renaissance  
1209 Background Reducing Diluent (Biocare #PD905L). Slides were processed using a Biocare  
1210 IntelliPATH FLX Automated Slide Stainer. After primary antibody incubation, slides were  
1211 incubated in Mouse Secondary Reagent (Biocare #IPSC5001G20), rinsed with 1X TBS,  
1212 incubated in Universal HRP Tertiary Reagent (Biocare #IPT5002G20), rinsed in 1X TBS, and  
1213 incubated in IP FLXDAB (Biocare Buffer #IPBF5009G20), and DAB chromogen (Biocare  
1214 Chromogen #IPC5008G3). Slides were then rinsed in 1X TBS, incubated in DAB sparkle  
1215 (Biocare #DSB830M), washed in MilliQ water, dehydrated through a series of graded alcohols,  
1216 cleared with Formula 83, and coverslipped with DPX. Slides were imaged using an Aperio  
1217 ScanScope XT slide scanner (Leica).

### 1218 **Multiplex fluorescence *in situ* hybridization (FISH)**

1219 Genes were selected for multiplex FISH experiments that discriminated cell types and broader  
1220 classes by visual inspection of differentially expressed genes that had relatively binary  
1221 expression in the targeted types.

### 1222 **Single molecule FISH (smFISH)**

1223 Fresh-frozen human brain tissue from the MTG was sectioned at 10µm onto Poly-L-lysine  
1224 coated coverslips as described previously <sup>78</sup>, let dry for 10 min at room temperature, then fixed  
1225 for 15 min at 4 C in 4% PFA. Sections were washed 3 x 10 min in PBS, then permeabilized and  
1226 dehydrated with 100% isopropanol at room temperature for 3 min and allowed to dry. Sections  
1227 were stored at -80 C until use. Frozen sections were rehydrated in 2XSSC (Sigma Aldrich  
1228 20XSSC, 15557036) for 5 min, then treated 2 X 5 min with 4%SDS (Sigma Aldrich, 724255) and  
1229 200mM boric acid (Sigma Aldrich, cat# B6768) pH 8.5 at room temperature. Sections were  
1230 washed 3 times in 2X SSC, then once in TE pH 8 (Sigma Aldrich, 93283). Sections were  
1231 heatshocked at 70 C for 10 min in TE pH 8, followed by 2XSSC wash at room temperature.

1232 Sections were then incubated in hybridization buffer (10% Formamide (v/v, Sigma Aldrich  
1233 4650), 10% Dextran Sulfate (w/v, Sigma Aldrich D8906), 200ug/mL BSA (Ambion AM2616), 2  
1234 mM Ribonucleoside vanadyl complex (New England Biolabs, S1402S), 1mg/ml tRNA (Sigma  
1235 10109541001) in 2XSSC) for 5 min at 38.5 C. Probes were diluted in hybridization buffer at a  
1236 concentration of 250 nM and hybridized at 38.5 C for 2 h. Following hybridization, sections were  
1237 washed 2 X 15 min at 38.5 C in wash buffer (2XSSC, 20% Formamide), and 1 X 15 min in wash  
1238 buffer with 5 ug/ml DAPI (Sigma Aldrich, 32670). Sections are then imaged in Imaging buffer  
1239 (20 mM Tris-HCl pH 8, 50 mM NaCl, 0.8% Glucose (Sigma Aldrich, G8270), 3 U/ml Glucose  
1240 Oxidase (Sigma Aldrich, G2133), 90 U/ml Catalase (Sigma Aldrich, C3515). Following imaging,  
1241 sections were incubated 3 X 10 min in stripping buffer (65% Formamide, 2X SSC) at 30 C to  
1242 remove hybridization probes from the first round. Sections were then washed in 2X SSC for 3 X  
1243 5 min at room temperature prior to repeating the hybridization procedure.

#### 1244 ***RNAscope multiplex FISH***

1245 Human tissue specimens used for RNAscope multiplex FISH came from either neurosurgical  
1246 resections or postmortem brain specimens. Mouse tissue for RNAscope experiments was from  
1247 adult (P56 +/- 3 days) wildtype C57Bl/6J mice. All animal procedures were approved by the  
1248 Institutional Animal Care and Use Committee at the Allen Institute for Brain Science (Protocol  
1249 No. 1511). Mice were provided food and water ad libitum, maintained on a regular 12-h  
1250 day/night cycle, and housed in cages with various enrichment materials added, including  
1251 nesting materials, gnawing materials, and plastic shelters. Mice were anesthetized with 5%  
1252 isoflurane and intracardially perfused with either 25 or 50 ml of ice cold, oxygenated artificial  
1253 cerebral spinal fluid (0.5mM CaCl<sub>2</sub>, 25mM D-Glucose, 98mM HCl, 20mM HEPES, 10mM  
1254 MgSO<sub>4</sub>, 1.25mM NaH<sub>2</sub>PO<sub>4</sub>, 3mM Myo-inositol, 12mM N-acetylcysteine, 96mM N-methyl-D-  
1255 glucamine, 2.5mM KCl, 25mM NaHCO<sub>3</sub>, 5mM sodium L-Ascorbate, 3mM sodium pyruvate,  
1256 0.01mM Taurine, and 2mM Thiourea). The brain was then rapidly dissected, embedded in  
1257 optimal cutting temperature (O.C.T.) medium, and frozen in a slurry of dry ice and ethanol.  
1258 Tissues were stored at -80C until for later cryosectioning.

1259 Fresh-frozen mouse or human tissues were sectioned at 14-16um onto Superfrost Plus glass  
1260 slides (Fisher Scientific). Sections were dried for 20 minutes at -20C and then vacuum sealed  
1261 and stored at -80C until use. The RNAscope multiplex fluorescent v1 kit was used according to  
1262 the manufacturer's instructions for fresh-frozen tissue sections (ACD Bio), with the following  
1263 minor modifications: (1) fixation was performed for 60 minutes in 4% paraformaldehyde in 1X  
1264 PBS at 4 °C, and (2) the protease treatment step was shortened to 15 min. Sections were  
1265 imaged using either a 40X or 60X oil immersion lens on a Nikon TiE fluorescent microscope  
1266 equipped with NIS-Elements Advanced Research imaging software (version 4.20).

#### 1267 ***RNAscope multiplex FISH with GFAP immunohistochemistry***

1268 Tissue sections were processed for RNAscope multiplex FISH detection of *ID3* (ACD Bio,  
1269 #492181-C3, NM\_002167.4 ) and *AQP4* (ACD Bio, #482441, NM\_001650.5 ) exactly as  
1270 described above. At the end of the RNAscope protocol, sections were fixed in 4%  
1271 paraformaldehyde for 15 minutes at room temperature and then washed twice in 1X PBS for 5  
1272 minutes. Sections were incubated in blocking solution (10% normal donkey serum, 0.1% triton-x  
1273 100 in 1X PBS) for 30 minutes at room temperature and then incubated in primary antibody  
1274 diluted 1:100 in blocking solution (mouse anti-GFAP, Sigma-Aldrich, #G3893, clone G-A-5) for  
1275 18 hours at 4C. Sections were then washed 3 times for 5 minutes each in 1X PBS, incubated  
1276 with secondary antibody (goat anti-mouse IgG(H+L) Alexa Fluor 568 conjugate, ThermoFisher  
1277 Scientific, #A-11004) for 30 minutes at room temperature, rinsed in 1X PBS 3 times for 5  
1278 minutes each, counterstained with DAPI (1ug/ml), and mounted with ProLong Gold mounting



1279 medium (ThermoFisher Scientific). Sections were imaged using either a 40X or 60X oil  
1280 immersion lens on a Nikon TiE fluorescent microscope equipped with NIS-Elements Advanced  
1281 Research imaging software (version 4.20).

### 1282 ***In situ* validation of excitatory types**

1283 To validate excitatory neuron types, clusters were labeled with cell type specific combinatorial  
1284 gene panels. For each gene panel, positive cells were manually called by visual assessment of  
1285 RNA spots for each gene. The total number of positive cells was quantified for each section.  
1286 Cells were counted on at least three sections derived from at least two donors for each probe  
1287 combination. DAPI staining was used to determine the boundaries of cortical layers within each  
1288 tissue section and the laminar position of each positive cell was recorded. The percentage of  
1289 labeled cells per layer, expressed as a fraction of the total number of labeled cells summed  
1290 across all layers, was calculated for each type. Probes used in these experiments were as  
1291 follows (all from ACD Bio): *SLC17A7* (#415611, NM\_020309.3), *RORB* (#446061, #446061-  
1292 C2, NM\_006914.3), *CNR1* (#591521-C2, NM\_001160226.1), *PRSS12* (#493931-  
1293 C3, NM\_003619.3), *ALCAM* (#415731-C2, NM\_001243283.1), *MET*  
1294 (#431021, NM\_001127500.1), *MME* (#410891-C2, NM\_007289.2), *NTNG1* (#446101-  
1295 C3, NM\_001113226.1), *HS3ST4* (#506181, NM\_006040.2), *CUX2* (#425581-  
1296 C3, NM\_015267.3), *PCP4* (#446111, NM\_006198.2), *GRIN3A* (#534841-C3, NM\_133445.2),  
1297 *GRIK3* (#493981, NM\_000831.3), *CRHR2* (#469621, NM\_001883.4), *TPBG*  
1298 (#405481, NM\_006670.4), *POSTN* (#409181-C3, NM\_006475.2), *SMYD1* (#493951-  
1299 C2, NM\_001330364.1)

### 1300 ***In situ* validation of putative chandelier cells**

1301 Tissue sections were labeled with the gene panel *GAD1*, *PVALB*, and *NOG*, or *COL15A1*,  
1302 specific markers of the Inh L2-5 *PVALB SCUBE3* putative chandelier cell cluster. Probes were  
1303 as follows (all from ACD Bio): *GAD1* (#404031-C3, NM\_000817.2), *PVALB* (#422181-  
1304 C2, NM\_002854.2), *NOG* (#416521, NM\_005450.4), *COL15A1* (#484001, NM\_001855.4).  
1305 Counts were conducted on sections from 3 human tissue donors. For each donor, the total  
1306 number of *GAD1+*, *PVALB+* and *NOG+* cells was summed across multiple sections. The  
1307 laminar position of each cell, based on boundaries defined by assessing DAPI staining patterns  
1308 in each tissue section, was recorded. The proportion of chandelier cells in each layer was  
1309 calculated as a fraction of the total number of *GAD1+/PVALB+/NOG+* cells summed across all  
1310 layers for each specimen.

### 1311 **Cell counts of broad interneuron classes**

1312 Tissue sections were labeled with the RNAscope Multiplex Fluorescent kit (ACD Bio) as  
1313 described above. For human tissue sections, the following probes (all from ACD Bio) were  
1314 used: *GAD1* (#404031, NM\_000817.2); *ADARB2* (#511651-C3, NM\_018702.3); *LHX6*  
1315 (#460051-C2, NM\_014368.4). For mouse tissue sections, the following probes were used: *Gad1*  
1316 (#400951, NM\_008077.4); *Adarb2* (#519971-C3, NM\_052977.5); *Lhx6* (#422791-C2,  
1317 NM\_001083127.1). The expression of each gene was assessed by manual examination of  
1318 corresponding RNA spots. Cell counts were conducted on sections from 3 human tissue donors:  
1319 2 neurosurgical and 1 postmortem. For mouse, 3 independent specimens were used. For both  
1320 human and mouse, >500 total *GAD1+* cells per specimen were counted (Human, n=2706, 1553,  
1321 and 3476 *GAD1+* cells per donor, respectively; Mouse, n=1897, 2587, and 708 *GAD1+* cells per  
1322 specimen, respectively). Expression of *ADARB2/Adarb2* and *LHX6/Lhx6* was manually  
1323 assessed in each *GAD1+* cell and cells were scored as being positive or negative for each  
1324 gene. At the same time, the laminar position of each *GAD1+* cell was recorded. Cell density,

1325 highlighted by DAPI staining, was used to determine laminar boundaries. The percentage of  
1326 each cell class expressed as a fraction of total *GAD1*+ cells and the percentage of each cell  
1327 class per layer, expressed as a fraction of the total number of *GAD1*+ cells per layer, were  
1328 calculated for each specimen. Statistical comparisons between human and mouse were done  
1329 using unpaired two-tailed t-tests with Holm-Sidak correction for multiple comparisons.

### 1330 **Imaging and quantification of smFISH expression**

1331 smFISH images were collected using an inverted microscope in an epifluorescence  
1332 configuration (Zeiss Axio Observer.Z1) with a 63x oil immersion objective with numerical  
1333 aperture 1.4. The sample was positioned in x, y and z with a motorized x, y stage with linear  
1334 encoders and z piezo top-plate (Applied Scientific Instruments MS 2000-500) and z stacks with  
1335 300 nm plane spacing were collected in each color at each stage position through the entire z  
1336 depth of the sample. Fluorescence emission was filtered using a high-speed filterwheel (Zeiss)  
1337 directly below the dichroic turret and imaged onto a sCMOS camera (Hamamatsu ORCA  
1338 Flash4.0) with a final pixel size of 100 nm. Images were collected after each round of  
1339 hybridization using the same configuration of x, y tile locations, aligned manually before each  
1340 acquisition based on DAPI fluorescence. smFISH signal was observed as diffraction-limited  
1341 spots which were localized in 3D image stacks by finding local maxima after spatial bandpass  
1342 filtering. These maxima were filtered for total intensity and radius to eliminate dim background  
1343 and large, bright lipofuscin granules. Outlines of cells and cortical layers were manually  
1344 annotated on images of *GAD*, *SLC17A7* and DAPI as 2D polygons using FIJI. The number of  
1345 mRNA molecules in each cell for each gene was then calculated and converted to densities  
1346 (spots per 100 $\mu\text{m}^2$ ).

1347  
1348 Background expression of the excitatory neuron marker *SLC17A7* was defined as the 95th  
1349 quantile of *SLC17A7* spot density among cells in cortical layer 1, since no excitatory cells should  
1350 be present in layer 1. Excitatory neurons were defined as any cell with *SLC17A7* spot density  
1351 greater than this threshold. To map excitatory cells to MTG reference clusters, spot counts were  
1352 log-transformed and scaled so that the 90th quantile of expression for each gene in smFISH  
1353 matched the maximum median cluster expression of that gene among the reference clusters.  
1354 Reference clusters that could not be discriminated based on the smFISH panel of nine genes  
1355 were merged and all comparisons between smFISH and RNA-seq cluster classes were  
1356 performed using these cluster groups. Scaled spot densities for each cell were then compared  
1357 to median expression levels of each reference cluster using Pearson correlation, and each cell  
1358 was assigned to the cluster with the highest correlation. For cells that mapped to the Exc L2-  
1359 3 *LINC00507* *FREM3* cluster, *LAMP5* and *COL5A2* expression was plotted as a dot plot where  
1360 the size and color of dots corresponded to probe spot density and the location corresponded to  
1361 the *in situ* location.

### 1362 **MetaNeighbor analysis**

1363 To compare the ability of different gene sets to distinguish cell types in mouse versus human  
1364 cortex, we performed a modified supervised MetaNeighbor analysis<sup>79</sup> independently for both  
1365 species. First, we divided our data sets into two artificial experiments, selecting random groups  
1366 of equal size up to a maximum of 10 cells per cluster for each experiment. We next ran  
1367 MetaNeighbor separately for clusters from each broad class (GABAergic, glutamatergic, and  
1368 non-neuronal) using the R function “run\_MetaNeighbor” where “experiment\_labels” are 1 or 2  
1369 corresponding to the two artificial experiments, “celltype\_labels” are 2 for cells in the targeted  
1370 cluster and 1 for cells in all other clusters of the same broad class, and “genesets” were all of  
1371 the HGNC gene sets included in Table S3 of<sup>52</sup>. Mean AUROC scores for each gene set were

1372 then calculated by averaging the reported AUROC scores for a gene set across all clusters  
1373 within a given broad class. This processes was repeated for 10 divisions of the human and  
1374 mouse data into random experimental groups. Means and standard deviations of these mean  
1375 AUROC scores for human and mouse GABAergic cell types are compared in **Fig 5**.

### 1376 **Estimation of cell type homology**

1377 We aligned single nucleus and single cell RNA-seq data from human MTG and mouse primary  
1378 visual cortex by applying canonical correlation analysis (CCA) as implemented in the Seurat R  
1379 package <sup>56</sup>. We used log<sub>2</sub>-transformed CPM of intronic plus exonic reads for both datasets.  
1380 Including exonic reads increased experimental differences due to measuring whole cell versus  
1381 nuclear transcripts, but this was out-weighed by improved gene detection. We separated each  
1382 of the datasets into three broad cell classes: GABAergic, glutamatergic, and non-neuronal,  
1383 based on their assigned clusters, and selected up to 200 cells from each cluster. We included  
1384 mouse non-neuronal cells from cell types that we had captured in our human survey, including  
1385 astrocytes, oligodendrocyte precursors, oligodendrocytes, endothelial cells, and microglia. For  
1386 each of these datasets, we selected the union of the top 2,000 genes with the highest  
1387 dispersion for human and mouse and calculated 40 canonical correlates with diagonal CCA.  
1388 Following this step, we removed 88 nuclei or cells for which the variance explained by CCA was  
1389 less than half of the variance explained by PCA, and aligned the canonical basis vectors to  
1390 allow integrated analysis. In particular, all human endothelial nuclei and over half of human  
1391 microglial nuclei were removed along with mouse Cajal-Retzius cells.

1392 We defined homologous cell types by clustering canonical correlates and identifying human and  
1393 mouse samples that co-clustered. Initially, the first 10 canonical correlates were selected, and a  
1394 weighted graph was constructed based on the Jaccard similarity of the 10 nearest neighbors of  
1395 each sample. Louvain community detection was run to identify clusters that optimized the global  
1396 modularity of the partitioned graph. For each pair of human and mouse clusters, the overlap  
1397 was defined as the sum of the minimum proportion of samples in each cluster that overlapped  
1398 within each CCA cluster. This approach identified pairs of human and mouse clusters that  
1399 consistently co-clustered within one or more CCA clusters. Cluster overlaps varied from 0 to 1  
1400 and were visualized as a heatmap with human clusters as rows and mouse clusters as columns.  
1401 Cell type homologies were identified as one-to-one, one-to-many, or many-to-many based on  
1402 the pattern of overlap between clusters. A quality score was calculated for the homology  
1403 mapping that rewarded overlaps greater than 0.6 (0.2 for non-neuronal clusters) and penalized  
1404 for clusters lacking any overlaps. For each human cluster, the inverse of the sum of the number  
1405 of overlapping mouse clusters was calculated, and this value was set to -1 if no overlapping  
1406 clusters were found. The quality score was defined as the sum of the scores for the individual  
1407 clusters and could range from -38 (no overlap) to 38 (all one-to-one matches). Including more  
1408 canonical correlates or fewer nearest neighbors increased the number of cell types that could  
1409 be discriminated within each species (increasing the quality score) but also resulted in more  
1410 species-specific clusters (decreasing the quality score). A grid based search was used to select  
1411 the number of canonical correlates and nearest neighbors that maximized the quality score.  
1412 Detection of homologous cell types was confirmed by visual inspection (**Supplementary Table**  
1413 **1**).

### 1415 **Quantification of expression divergence**

1416 For each pair of 38 homologous human and mouse cell types, the average expression of 14,414  
1417 orthologous genes was calculated as the average counts per million of intronic reads. Only  
1418 intronic reads were used to better compare these single nucleus (human) and single cell

1419 (mouse) datasets. Average expression values were log<sub>2</sub>-transformed and scatter plots and  
1420 Pearson's correlations were calculated to compare human and mouse. Genes were ranked  
1421 based on their cell type-specificity in human and mouse using a tau score defined in <sup>80</sup>, and the  
1422 union of the top 50 markers in human and mouse were highlighted in the scatter plots. The fold  
1423 difference in expression between human and mouse was calculated for all genes and  
1424 homologous cell types and thresholded to identify large (>10-fold), moderate (2- to 10-fold), and  
1425 small (<2-fold) differences. A heatmap was generated showing expression differences across  
1426 cell types, and hierarchical clustering using Ward's method was applied to group genes with  
1427 similar patterns of expression change. For each of 6 major classes of cell types  
1428 (*LAMP5/PAX6*, *VIP*, *SST*, *PVALB*, excitatory, non-neuronal), the number of genes was  
1429 quantified that had >10-fold change in at least one cell type in that class and <10-fold change in  
1430 all cell types in the other 5 classes. The expression pattern change of 14,414 genes was  
1431 quantified as the beta score (see marker score methods above) of log<sub>2</sub>-expression differences  
1432 across 38 homologous cell types (**Supplementary Table 2**). Genes with high scores have a  
1433 large fold-change in expression in one or more (but not all) cell types. For each gene, the  
1434 number of clusters with median expression (CPM) > 1 was compared to the median pattern  
1435 change of those genes. A loess curve and standard error were fit using the R package *ggplot*.  
1436 Finally, the median pattern change was calculated for the functional gene families used in the  
1437 MetaNeighbor analysis described above.

#### 1438 **Data and Code Availability**

1439 Data and code used to produce figures will be available  
1440 from [https://github.com/AllenInstitute/MTG\\_celltypes](https://github.com/AllenInstitute/MTG_celltypes). RNA-seq data from this study is publicly  
1441 available and can be downloaded at <http://celltypes.brain-map.org/>, and data can be visualized  
1442 and analyzed using two complementary viewers at [http://celltypes.brain-](http://celltypes.brain-map.org/mnaseq/human)  
1443 [map.org/mnaseq/human](http://celltypes.brain-map.org/mnaseq/human) and <https://viewer.cytosplore.org/>.

#### 1444 **References**

- 1445 1. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–  
1446 178 (2016).
- 1447 2. Nieuwenhuys, R. The myeloarchitectonic studies on the human cerebral cortex of the Vogt-  
1448 Vogt school, and their significance for the interpretation of functional neuroimaging data. *Brain*  
1449 *Struct Funct* **218**, 303–52 (2013).
- 1450 3. Essen, D. C. V., Glasser, M. F., Dierker, D. L., Harwell, J. & Coalson, T. Parcellations and  
1451 Hemispheric Asymmetries of Human Cerebral Cortex Analyzed on Surface-Based Atlases.  
1452 *Cerebral Cortex* **22**, 2241–2262 (2011).
- 1453 4. Azevedo, F. A. C. *et al.* Equal numbers of neuronal and nonneuronal cells make the human  
1454 brain an isometrically scaled-up primate brain. *The Journal of Comparative Neurology* **513**,  
1455 532–541 (2009).
- 1456 5. Herculano-Houzel, S., Mota, B. & Lent, R. Cellular scaling rules for rodent brains. *Proc Natl*  
1457 *Acad Sci U S A* **103**, 12138–43 (2006).
- 1458 6. Azevedo, F. A. *et al.* Equal numbers of neuronal and nonneuronal cells make the human brain  
1459 an isometrically scaled-up primate brain. *J Comp Neurol* **513**, 532–41 (2009).



- 1460 7.Geschwind, D. H. & Rakic, P. Cortical evolution: judge the brain by its cover. *Neuron* **80**, 633–  
1461 47 (2013).
- 1462 8.DeFelipe, J. The evolution of the brain, the human nature of cortical circuits, and intellectual  
1463 creativity. *Front Neuroanat* **5**, 29 (2011).
- 1464 9.Cajal, S. Ramón y. *La Textura del Sistema Nerviosa del Hombre y los Vertebrados*. (1904).
- 1465 10.Nó, R. Lorente de. La corteza cerebral del ratón. *Trab. Lab. Invest. Bio. (Madrid)* **20**, (1922).
- 1466 11.Poorthuis, R. B. *et al.* Rapid Neuromodulation of Layer 1 Interneurons in Human Neocortex.  
1467 *Cell Rep* **23**, 951–958 (2018).
- 1468 12.Eyal, G. *et al.* Unique membrane properties and enhanced signal processing in human  
1469 neocortical neurons. *Elife* **5**, (2016).
- 1470 13.Szegedi, V. *et al.* Plasticity in Single Axon Glutamatergic Connection to GABAergic  
1471 Interneurons Regulates Complex Events in the Human Neocortex. *PLoS Biol* **14**, e2000237  
1472 (2016).
- 1473 14.DeFelipe, J. Types of neurons, synaptic connections and chemical characteristics of cells  
1474 immunoreactive for calbindin-D28K, parvalbumin and calretinin in the neocortex. *J Chem*  
1475 *Neuroanat* **14**, 1–19 (1997).
- 1476 15.Benavides-Piccione, R., Ballesteros-Yáñez, I., DeFelipe, J. & Yuste, R. Cortical area and  
1477 species differences in dendritic spine morphology. *J Neurocytol* **31**, 337–46 (2002).
- 1478 16.Gabbott, P. L. Subpial Fan Cell - A Class of Calretinin Neuron in Layer 1 of Adult Monkey  
1479 Prefrontal Cortex. *Front Neuroanat* **10**, 28 (2016).
- 1480 17.Oberheim, N. A. *et al.* Uniquely hominid features of adult human astrocytes. *J Neurosci* **29**,  
1481 3276–87 (2009).
- 1482 18.Hill, R. S. & Walsh, C. A. Molecular insights into human brain evolution. *Nature* **437**, 64–7  
1483 (2005).
- 1484 19.Boldog, E. *et al.* Transcriptomic and morphophysiological evidence for a specialized human  
1485 cortical GABAergic cell type. *bioRxiv* (2017). doi:10.1101/216085
- 1486 20.Zeng, H. *et al.* Large-scale cellular-resolution gene profiling in human neocortex reveals  
1487 species-specific molecular signatures. *Cell* **149**, 483–96 (2012).
- 1488 21.Bakken, T. E. *et al.* A comprehensive transcriptional map of primate brain development.  
1489 *Nature* **535**, 367–75 (2016).
- 1490 22.Hawrylycz, M. *et al.* Canonical genetic signatures of the adult human brain. *Nat Neurosci* **18**,  
1491 1832–44 (2015).
- 1492 23.Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–  
1493 206 (2014).
- 1494 24.Ecker, J. R. *et al.* The BRAIN Initiative Cell Census Consortium: Lessons Learned toward  
1495 Generating a Comprehensive Brain Cell Atlas. *Neuron* **96**, 542–557 (2017).
- 1496 25.Regev, A. *et al.* The Human Cell Atlas. *Elife* **6**, (2017).



- 1497 26.Tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *bioRxiv*  
1498 (2017). doi:10.1101/229542
- 1499 27.Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics.  
1500 *Nat Neurosci* **19**, 335–46 (2016).
- 1501 28.Zeisel, A. *et al.* Brain structure. Cell types in the mouse cortex and hippocampus revealed by  
1502 single-cell RNA-seq. *Science* **347**, 1138–42 (2015).
- 1503 29.Darmanis, S. *et al.* A survey of human brain transcriptome diversity at the single cell level.  
1504 *Proc Natl Acad Sci U S A* **112**, 7285–90 (2015).
- 1505 30.Krishnaswami, S. R. *et al.* Using single nuclei for RNA-seq to capture the transcriptome of  
1506 postmortem neurons. *Nat Protoc* **11**, 499–524 (2016).
- 1507 31.Lake, B. B. *et al.* Neuronal subtypes and diversity revealed by single-nucleus RNA  
1508 sequencing of the human brain. *Science* **352**, 1586–90 (2016).
- 1509 32.Habib, N. *et al.* Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat Methods* **14**,  
1510 955–958 (2017).
- 1511 33.Bakken, T. E. *et al.* Equivalent high-resolution identification of neuronal cell types with single-  
1512 nucleus and single-cell RNA-sequencing. *bioRxiv* (2017). doi:10.1101/239749
- 1513 34.Lake, B. B. *et al.* A comparative strategy for single-nucleus and single-cell transcriptomes  
1514 confirms accuracy in predicted cell-type expression from nuclear RNA. *Sci Rep* **7**, 6031 (2017).
- 1515 35.Lake, B. B. *et al.* Integrative single-cell analysis of transcriptional and epigenetic states in the  
1516 human adult brain. *Nat Biotechnol* **36**, 70–80 (2018).
- 1517 36.DeFelipe, J., Alonso-Nanclares, L. & Arellano, J. I. Microstructure of the neocortex:  
1518 comparative aspects. *J Neurocytol* **31**, 299–316 (2002).
- 1519 37.Bakken, T. *et al.* Cell type discovery and representation in the era of high-content single cell  
1520 phenotyping. *BMC Bioinformatics* **18**, 559 (2017).
- 1521 38.Bakken, T. E. *et al.* Spatiotemporal dynamics of the postnatal developing primate brain  
1522 transcriptome. *Human Molecular Genetics* **24**, 4327–4339 (2015).
- 1523 39.Werner, M. S. *et al.* Chromatin-enriched lncRNAs can act as cell-type specific activators of  
1524 proximal gene transcription. *Nat Struct Mol Biol* **24**, 596–603 (2017).
- 1525 40.Von Economo, C. *Cellular structure of the human cerebral cortex*. (Karger Medical and  
1526 Scientific Publishers, 2009).
- 1527 41.Kalmbach, B. *et al.* h-channels contribute to divergent electrophysiological properties of  
1528 supragranular pyramidal neurons in human versus mouse cerebral cortex. *bioRxiv* (2018).  
1529 doi:10.1101/312298
- 1530 42.Cytosplore: Interactive Immune Cell Phenotyping for Large Single-Cell Datasets. *Computer*  
1531 *Graphics Forum* **35**, (2016).
- 1532 43.Hollt, T. *et al.* CyteGuide: Visual Guidance for Hierarchical Single-Cell Analysis. *IEEE Trans*  
1533 *Vis Comput Graph* **24**, 739–748 (2018).

- 1534 44.Lee, S., Hjerling-Leffler, J., Zagha, E., Fishell, G. & Rudy, B. The largest group of superficial  
1535 neocortical GABAergic interneurons expresses ionotropic serotonin receptors. *J Neurosci* **30**,  
1536 16796–808 (2010).
- 1537 45.Hansen, D. V. *et al.* Non-epithelial stem cells and cortical interneuron production in the  
1538 human ganglionic eminences. *Nat Neurosci* **16**, 1576–87 (2013).
- 1539 46.Ma, T. *et al.* Subcortical origins of human and monkey neocortical interneurons. *Nat*  
1540 *Neurosci* **16**, 1588–97 (2013).
- 1541 47.Tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *bioRxiv*  
1542 (2017). doi:10.1101/229542
- 1543 48.Shah, B. P. *et al.* MC4R-expressing glutamatergic neurons in the paraventricular  
1544 hypothalamus regulate feeding and are synaptically connected to the parabrachial nucleus.  
1545 *Proc Natl Acad Sci U S A* **111**, 13193–8 (2014).
- 1546 49.Horstmann, A. *et al.* Common genetic variation near MC4R has a sex-specific impact on  
1547 human brain structure and eating behavior. *PLoS One* **8**, e74362 (2013).
- 1548 50.Raghanti, M. A. *et al.* Neuropeptide Y-immunoreactive neurons in the cerebral cortex of  
1549 humans and other haplorrhine primates. *Am J Primatol* **75**, 415–24 (2013).
- 1550 51.Xu, X., Roby, K. D. & Callaway, E. M. Immunochemical characterization of inhibitory mouse  
1551 cortical neurons: three chemically distinct classes of inhibitory cells. *J Comp Neurol* **518**, 389–  
1552 404 (2010).
- 1553 52.Paul, A. *et al.* Transcriptional Architecture of Synaptic Communication Delineates GABAergic  
1554 Neuron Identity. *Cell* **171**, 522–539.e20 (2017).
- 1555 53.Marques, S. *et al.* Oligodendrocyte heterogeneity in the mouse juvenile and adult central  
1556 nervous system. *Science* **352**, 1326–1329 (2016).
- 1557 54.Zhang, Y. *et al.* Purification and Characterization of Progenitor and Mature Human  
1558 Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron* **89**, 37–53  
1559 (2016).
- 1560 55.Sosunov, A. A. *et al.* Phenotypic heterogeneity and plasticity of isocortical and hippocampal  
1561 astrocytes in the human brain. *J Neurosci* **34**, 2285–98 (2014).
- 1562 56.Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell  
1563 transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**,  
1564 411–420 (2018).
- 1565 57.Kilduff, T. S., Cauli, B. & Gerashchenko, D. Activation of cortical interneurons during sleep:  
1566 an anatomical link to homeostatic sleep regulation? *Trends Neurosci* **34**, 10–9 (2011).
- 1567 58.He, M. *et al.* Strategies and Tools for Combinatorial Targeting of GABAergic Neurons in  
1568 Mouse Cerebral Cortex. *Neuron* **92**, 555 (2016).
- 1569 59.Sorensen, S. A. *et al.* Correlated Gene Expression and Target Specificity Demonstrate  
1570 Excitatory Projection Neuron Diversity. *Cerebral Cortex* **25**, 433–449 (2013).
- 1571 60.Belichenko, P. V., Vogt, W. D. M., Myklóssy, J. & Celio, M. R. Calretinin-positive Cajal-  
1572 Retzius cells persist in the adult human neocortex. *Neuroreport* **6**, 1869–74 (1995).

- 1573 61. Glezer, I. I., Hof, P. R. & Morgane, P. J. Calretinin-immunoreactive neurons in the primary  
1574 visual cortex of dolphin and human brains. *Brain Res* **595**, 181–8 (1992).
- 1575 62. Miyoshi, G. *et al.* Genetic fate mapping reveals that the caudal ganglionic eminence  
1576 produces a large and diverse population of superficial cortical interneurons. *J Neurosci* **30**,  
1577 1582–94 (2010).
- 1578 63. Lein, E., Borm, L. E. & Linnarsson, S. The promise of spatial transcriptomics for  
1579 neuroscience in the era of molecular cell typing. *Science* **358**, 64–69 (2017).
- 1580 64. Colantuoni, C. *et al.* Temporal dynamics and genetic control of transcription in the human  
1581 prefrontal cortex. *Nature* **478**, 519–23 (2011).
- 1582 65. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489  
1583 (2011).
- 1584 66. Bahney, J. & Bartheld, C. S. von. The Cellular Composition and Glia-Neuron Ratio in the  
1585 Spinal Cord of a Human and a Nonhuman Primate: Comparison with Other Species and Brain  
1586 Regions. *The Anatomical Record* **301**, 697–710 (2017).
- 1587 67. Bjugn, R. The use of the optical disector to estimate the number of neurons, glial and  
1588 endothelial cells in the spinal cord of the mouse—with a comparative note on the rat spinal cord.  
1589 *Brain Res* **627**, 25–33 (1993).
- 1590 68. Lassek, A. M. & Rassmussen, G. L. The human pyramidal tract: A fiber and numerical  
1591 analysis. *Archives of Neurology & Psychiatry* **42**, 872–876 (1939).
- 1592 69. Finlay, B. & Darlington, R. Linked regularities in the development and evolution of  
1593 mammalian brains. *Science* **268**, 1578–1584 (1995).
- 1594 70. Markou, A., Chiamulera, C., Geyer, M. A., Tricklebank, M. & Steckler, T. Removing obstacles  
1595 in neuroscience drug discovery: the future path for animal models. *Neuropsychopharmacology*  
1596 **34**, 74–89 (2009).
- 1597 71. Nestler, E. J. & Hyman, S. E. Animal models of neuropsychiatric disorders. *Nature*  
1598 *Neuroscience* **13**, 1161–1169 (2010).
- 1599 72. Sorensen, S. A. *et al.* Correlated Gene Expression and Target Specificity Demonstrate  
1600 Excitatory Projection Neuron Diversity. *Cerebral Cortex* **25**, 433–449 (2013).
- 1601 73. Aronesty, E. Comparison of Sequencing Utility Programs. *The Open Bioinformatics Journal*  
1602 **7**, 1–8 (2013).
- 1603 74. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2012).
- 1604 75. Lawrence, M. *et al.* Software for Computing and Annotating Genomic Ranges. *PLoS*  
1605 *Computational Biology* **9**, e1003118 (2013).
- 1606 76. Calvo, S. E., Clauser, K. R. & Mootha, V. K. MitoCarta2.0: an updated inventory of  
1607 mammalian mitochondrial proteins. *Nucleic Acids Res* **44**, D1251–7 (2016).
- 1608 77. Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature*  
1609 **445**, 168–76 (2007).
- 1610 78. Lyubimova, A. *et al.* Single-molecule mRNA detection and counting in mammalian tissue.  
1611 *Nat Protoc* **8**, 1743–58 (2013).

- 1612 79.Crow, M., Paul, A., Ballouz, S., Huang, Z. J. & Gillis, J. Characterizing the replicability of cell  
1613 types defined by single cell RNA-sequencing data using MetaNeighbor. *Nat Commun* **9**, 884  
1614 (2018).
- 1615 80.Yanai, I. *et al.* Genome-wide midrange transcription profiles reveal expression level  
1616 relationships in human tissue specification. *Bioinformatics* **21**, 650–9 (2005).

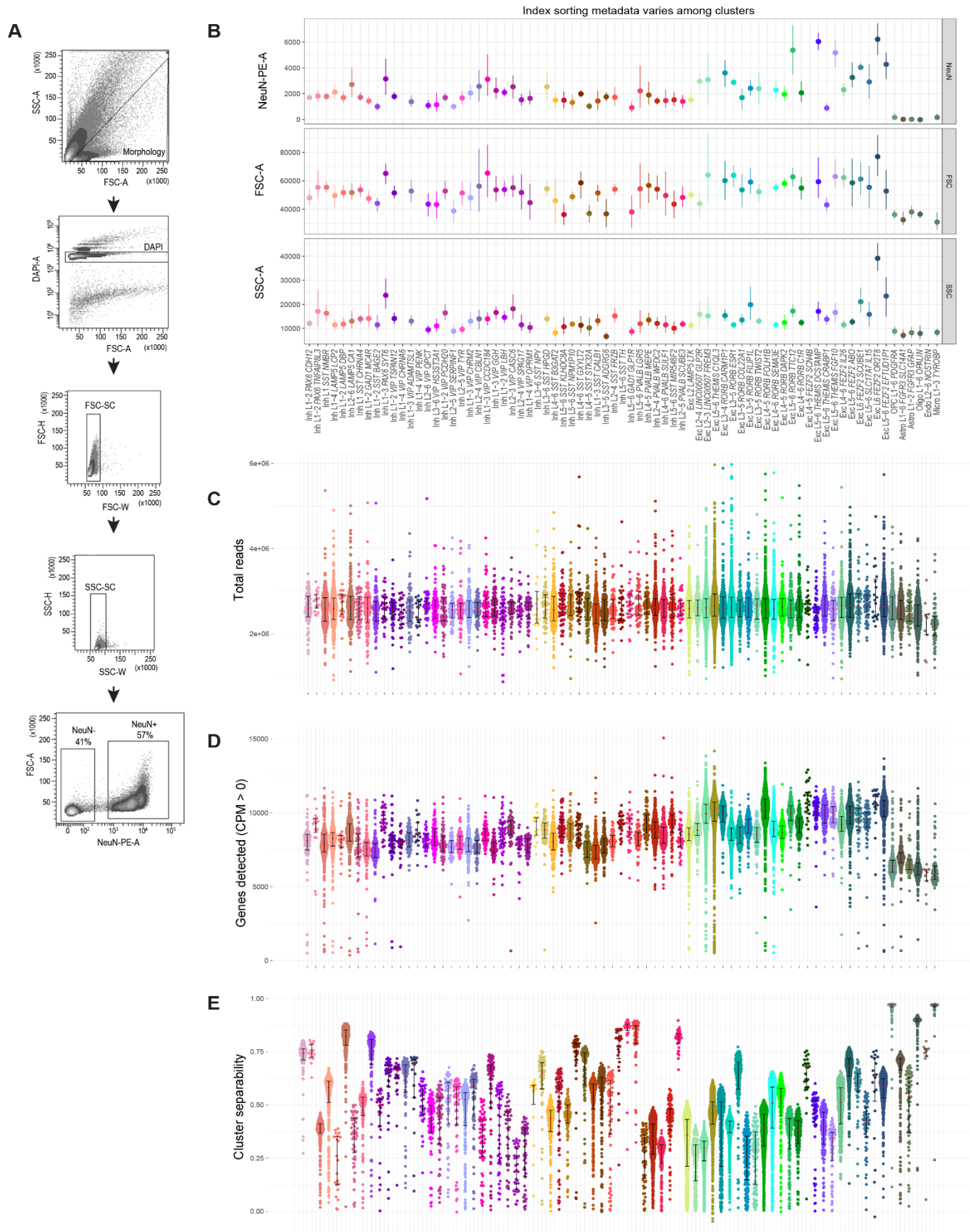
1617

Specimen ID	Tissue Type	Age	Sex	Race	Cause of Death	PMI (hr)	Tissue RIN	Hemisphere Sampled	Reason for Surgery	Number of nuclei sampled
H200.1023	P	43	F	Iranian descent	Mitral valve prolapse	18.5	7.4 ± 0.7	L	N/A	6170
H200.1025	P	50	M	Caucasian	CV	24.5	7.6 ± 1.0	L	N/A	1334
H200.1030	P	54	M	Caucasian	CV	25	7.7 ± 0.8	L	N/A	7331
H16.24.010	P	66	M	Caucasian	CV	21	7.2 ± 1.4	L	N/A	371
H16.06.002	N	35	F	Caucasian	N/A	N/A	7.1 ± 0.9	R	Epilepsy	97
H16.06.008	N	24	F	Hispanic	N/A	N/A	8.1 ± 0.8	L	Epilepsy	197
H16.06.009	N	48	F	Caucasian	N/A	N/A	7.1	L	Epilepsy	220
H16.03.004	N	25	M	Not noted	N/A	N/A	8.2 ± 0.8	R	Tumor removal, epilepsy	208

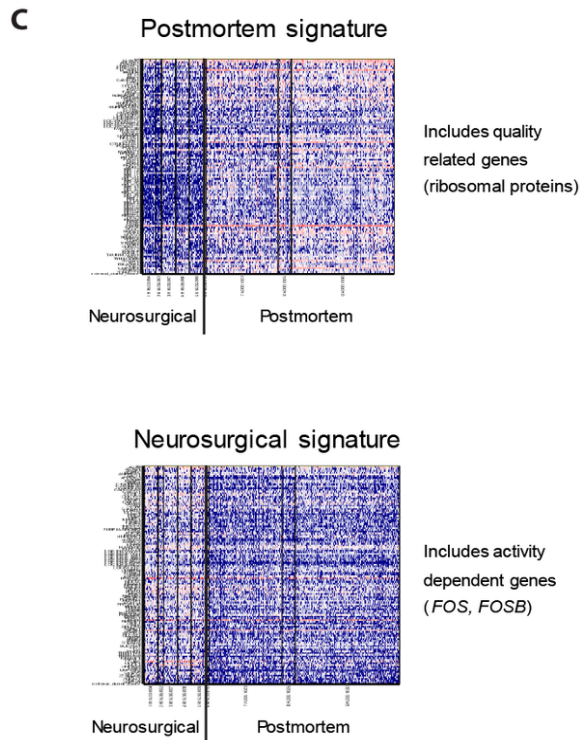
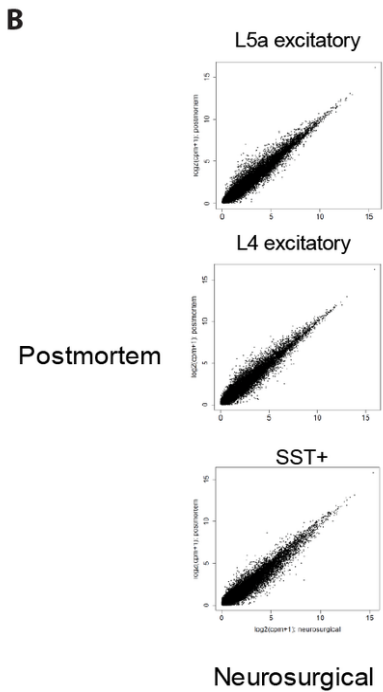
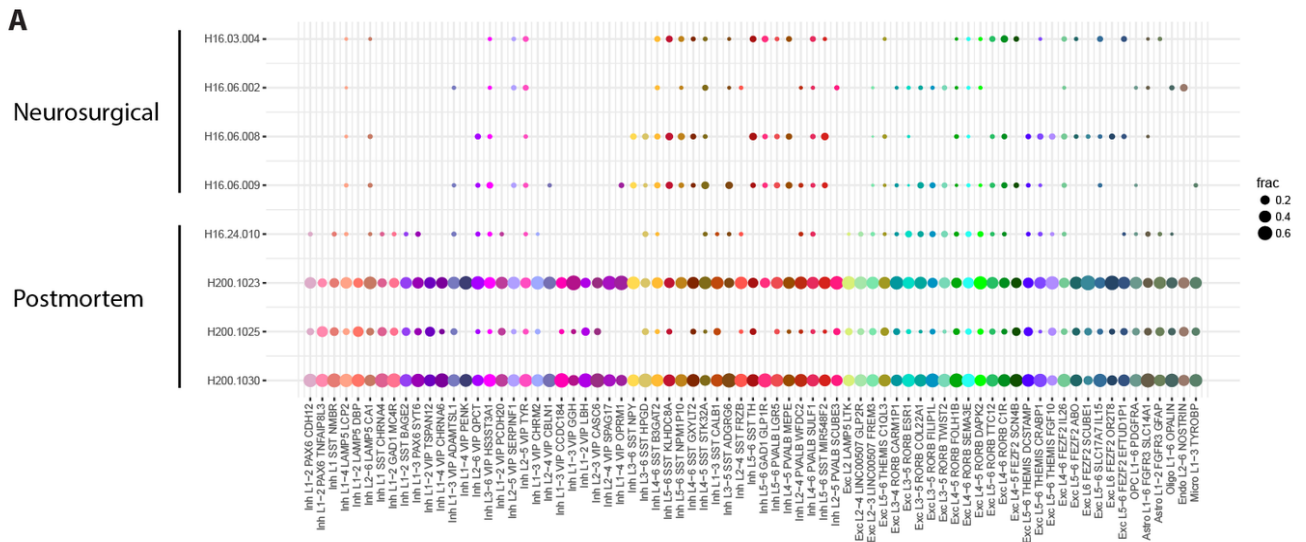
1618

1619 **Extended Data Table 1. Summary of human tissue donor information.** Tissue types - P,  
 1620 postmortem, N - neurosurgical. Cause of death - CV, cardiovascular, N/A, not applicable. PMI -  
 1621 postmortem interval. RIN - RNA Integrity Number.





1623 **Extended Data Figure 1. Nuclei metadata summarized by cluster. (A)** FACS gating scheme  
1624 for nuclei sorts. **(B)** FACS metadata for index sorted single nuclei shows significant variability in  
1625 NeuN fluorescence intensity (NeuN-PE-A), size (forward-scatter area, FSC-A), and granularity  
1626 (side-scatter area, SSC-A) across clusters. As expected, non-neuronal nuclei have almost no  
1627 NeuN staining and are smaller (as inferred by lower FSC values). **(C-E)** Scatter plots plus  
1628 median and interquartile interval of three QC metrics grouped and colored by  
1629 cluster. **(C)** Median total reads were approximately 2.6 million for all cell types, although slightly  
1630 lower for non-neuronal nuclei. **(D)** Median gene detection was highest among excitatory neuron  
1631 types in layers 5 and 6 and a subset of types in layer 3, lower among inhibitory neuron types,  
1632 and significantly lower among non-neuronal types. **(E)** Cluster separability varied substantially  
1633 among cell types, with a subset of neuronal types and all non-neuronal types being highly  
1634 discrete.

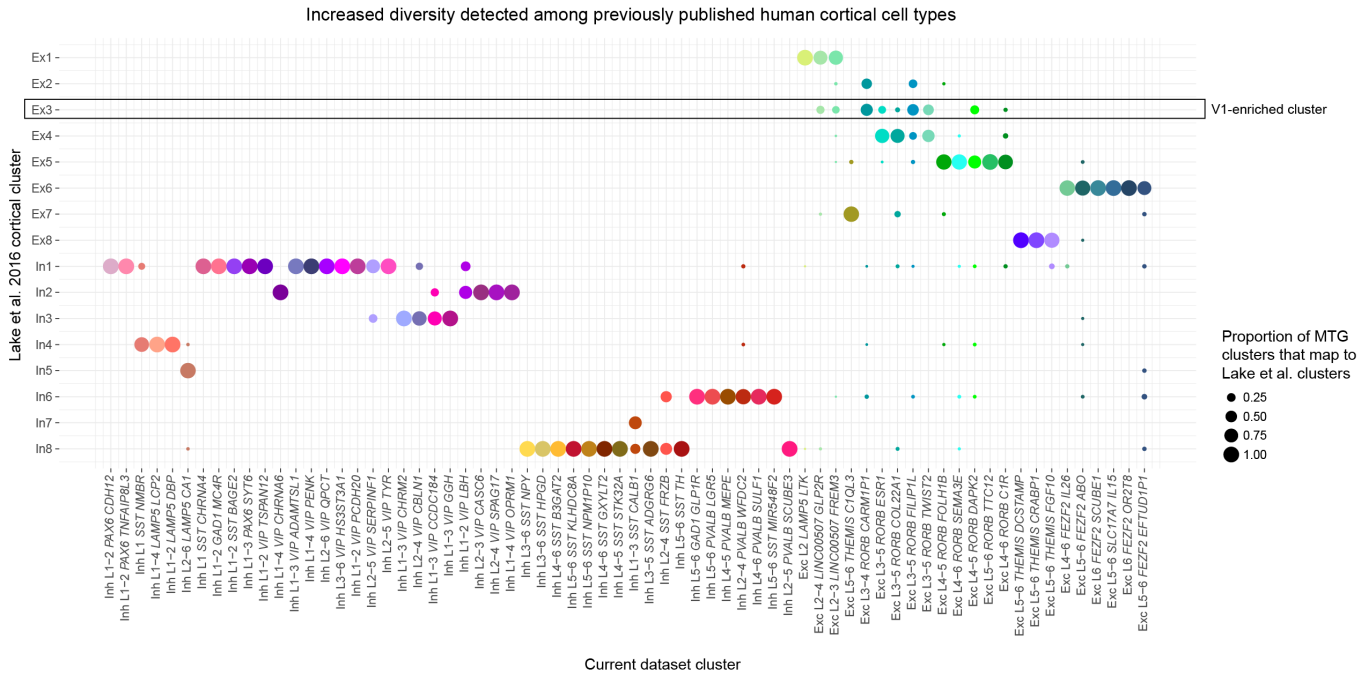


1636 **Extended Data Figure 2. Small but consistent expression signature of donor tissue**  
1637 **source. (A)** Dot plot showing the proportion of nuclei isolated from neurosurgical and  
1638 postmortem donors among human MTG clusters. Note that most nuclei from neurosurgical  
1639 donors were isolated only from layer 5 so clusters enriched in other layers, such as layer 1  
1640 interneurons, have low representation of these donors. **(B)** Highly correlated expression  
1641 between nuclei from postmortem and neurosurgical donors among two classes of excitatory  
1642 neurons and one class of inhibitory neurons. Nuclei were pooled and compared within these  
1643 broad classes due to the low sampling of individual clusters from neurosurgical  
1644 donors. **(C)** Expression ( $\log_{10}(\text{CPM} + 1)$ ) heatmaps of genes that are weakly but consistently up-  
1645 regulated in nuclei from postmortem or neurosurgical donors including ribosomal genes and  
1646 activity-dependent genes, respectively.



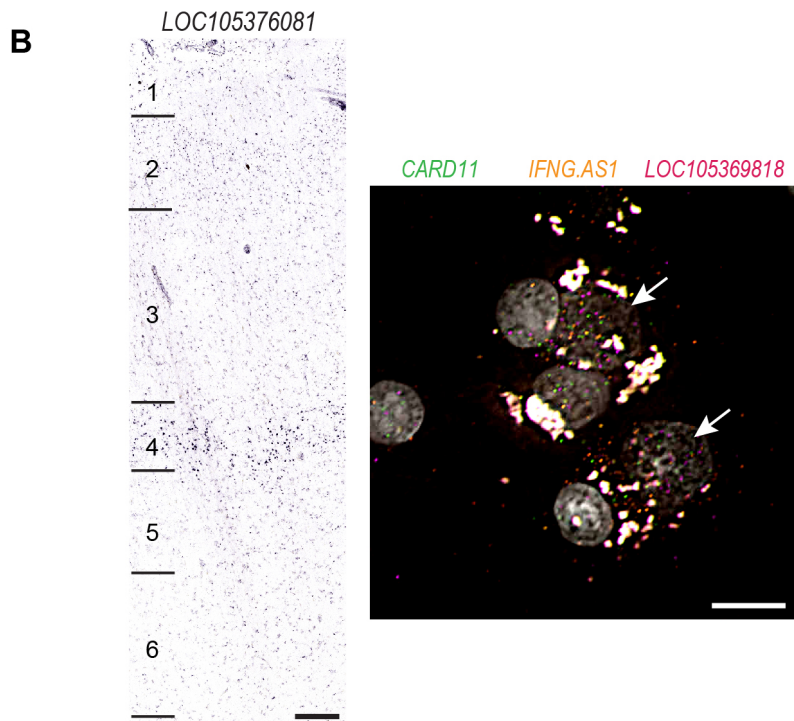
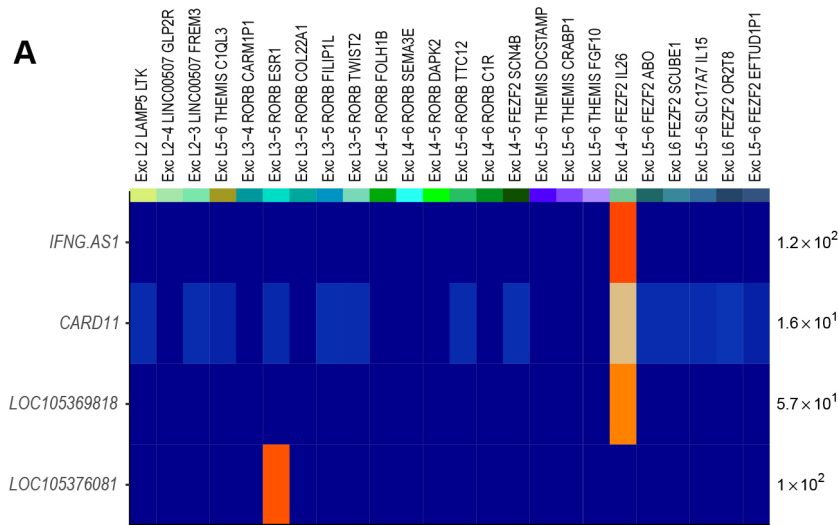


1648 **Extended Data Figure 3. Expression of cell type specific markers.** Violin plots of the best  
1649 cell type markers include many non-coding genes (blue symbols): lncRNAs, antisense  
1650 transcripts, and unnamed (LOC) genes. Expression values are on a linear scale and dots  
1651 indicate median expression. Note that LOC genes were excluded from cluster names, and the  
1652 best non-LOC marker genes were used instead.



1653

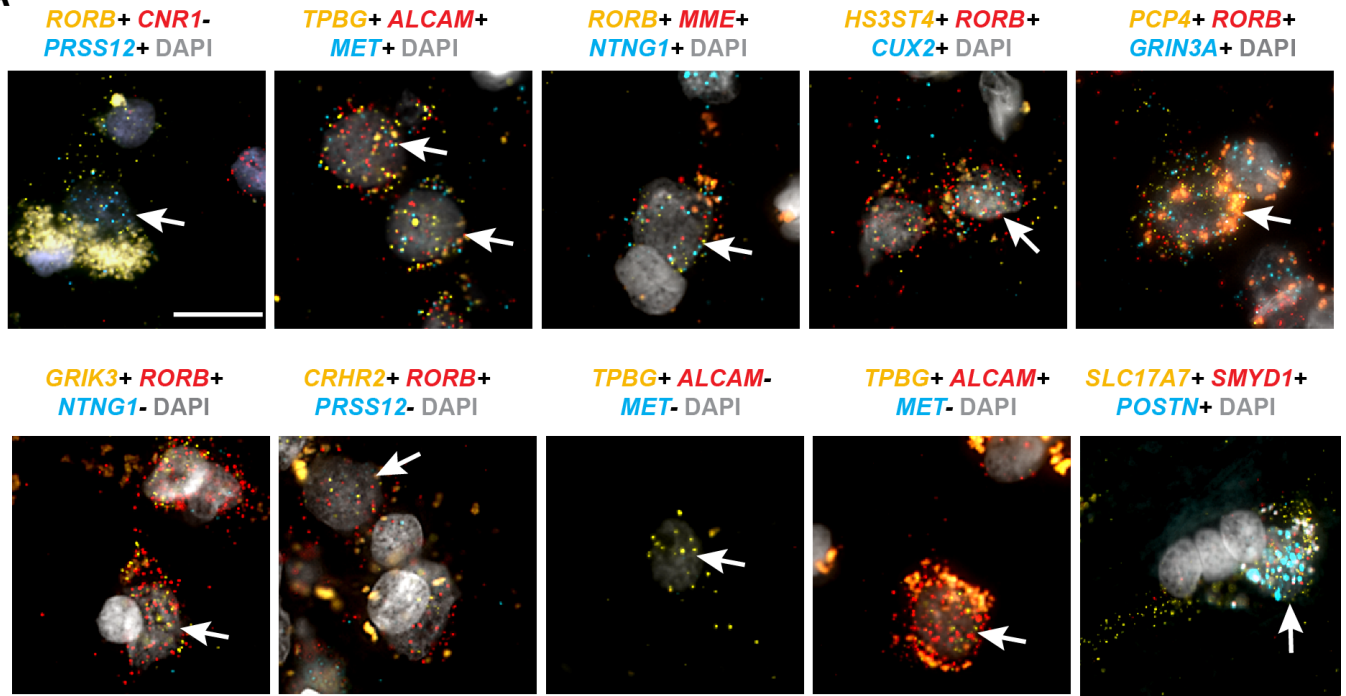
1654 **Extended Data Figure 4. Matching MTG clusters to reported human cortical cell**  
 1655 **types.** Dot plot showing the proportion of each MTG cluster that matches 16 clusters reported  
 1656 by<sup>29</sup> based on a centroid expression classifier. Ex3 was highly enriched in visual cortex and not  
 1657 detected in temporal cortex by Lake et al.



1658

1659 **Extended Data Figure 5. In situ validation of LOC and antisense transcripts as cell type**  
 1660 **specific markers. (A)** Heatmap illustrating cell type specific expression of several LOC genes  
 1661 and one antisense transcript (IFNG-AS1). **(B)** Left - chromogenic in situ hybridization  
 1662 for LOC105376081, a specific marker of the Exc L3-5 *RORB ESR1* type shows expression of  
 1663 this gene predominantly in layer 4, consistent with the anatomical location of this cell type. Scale  
 1664 bar, 100µm. Right - triple RNAscope FISH for markers of the Exc L4-6 *FEZF2 IL26* type.  
 1665 Coexpression of the protein coding gene *CARD11* with *IFNG-AS1*, an antisense transcript, and  
 1666 *LOC105369818* is apparent within several DAPI-labeled nuclei (white arrows). Scale bar,  
 1667 15µm.

A



1668

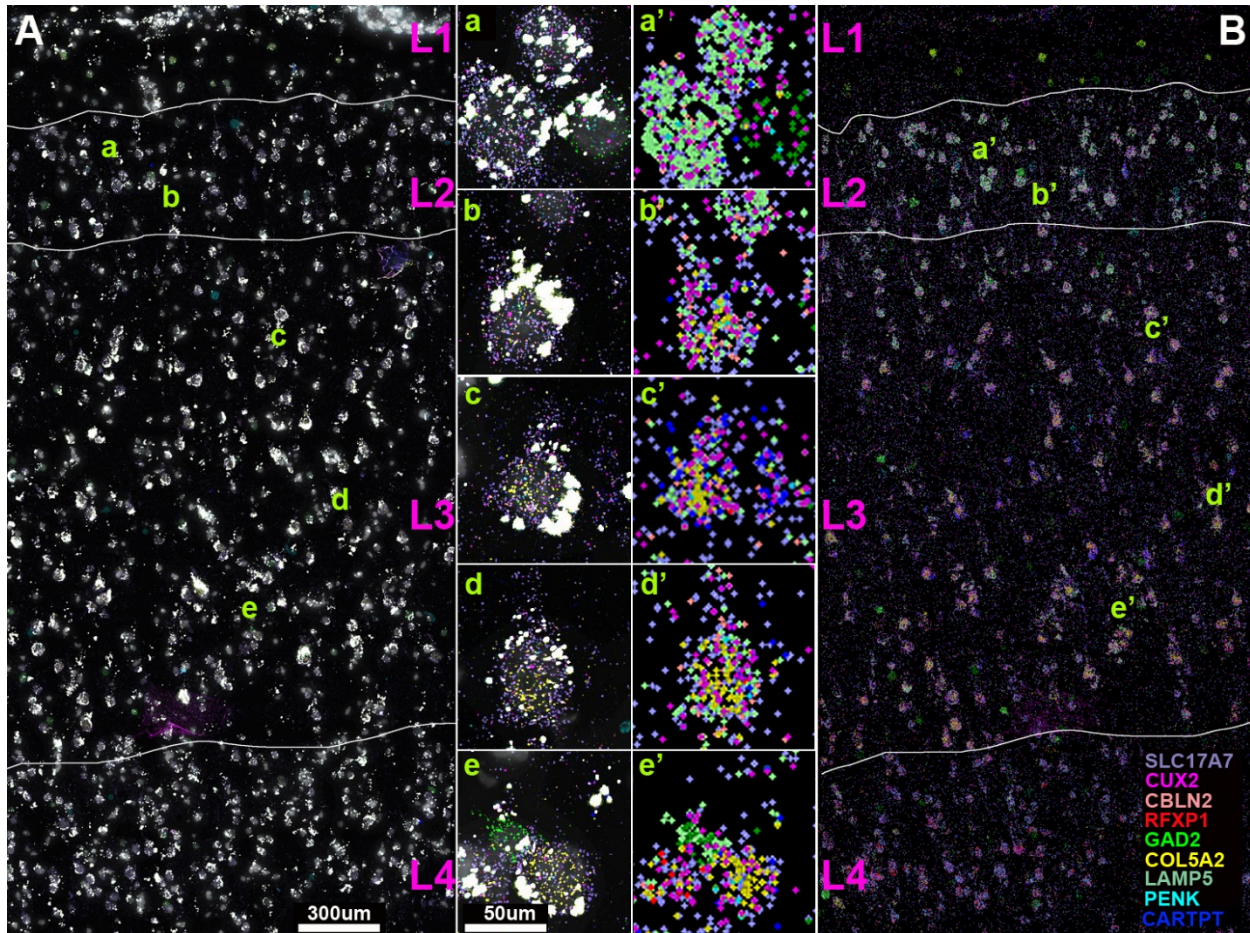
1669

1670

1671

**Extended Data Figure 6. RNAscope multiplex FISH validation of 10 excitatory neuron types.** Gene combinations probed are listed above each image. Labeled cells are indicated by white arrows. Scale bar, 20 $\mu$ m.

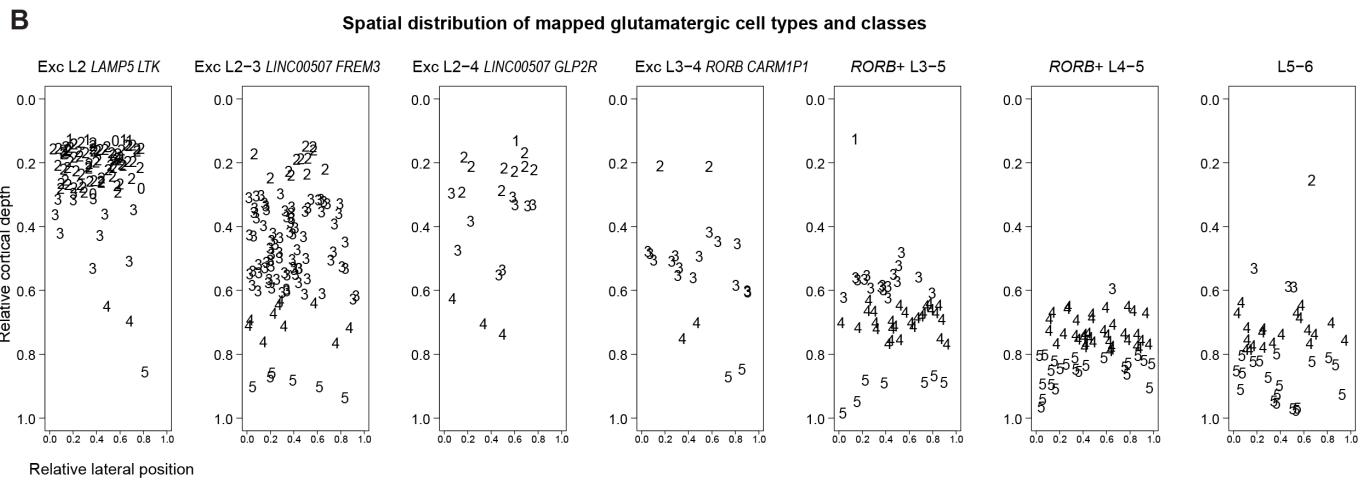
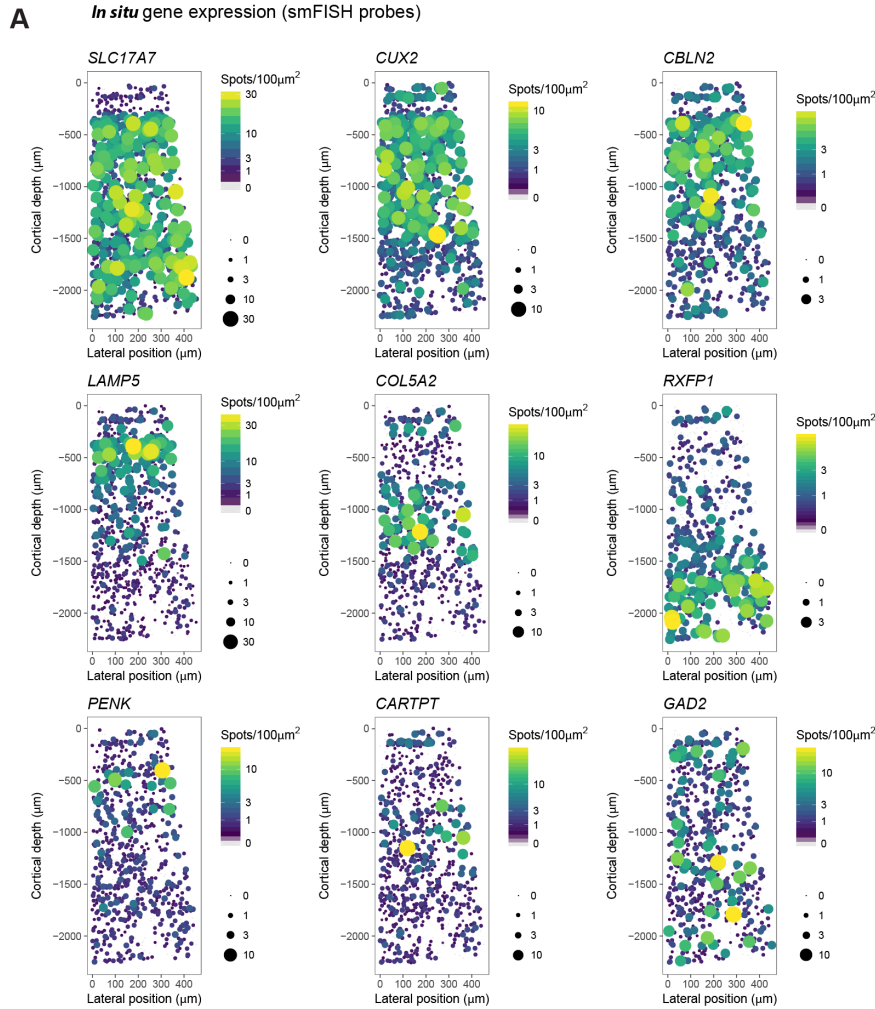




1672

1673 **Extended Data Figure 7. Single molecule (sm)FISH analysis of gene expression levels in**  
1674 **human MTG layers 2 and 3.** smFISH was performed with probes against *SLC17A7*, *CUX2*,  
1675 *CBLN2*, *RFXP1*, *GAD2*, *COL5A2*, *LAMP5*, *PENK*, and *CARTPT* mRNA. (A) smFISH image  
1676 (100x). Spots for each gene are pseudocolored as indicated in the bottom right legend. Layer  
1677 demarcations are indicated in magenta. Scale bar = 300 um. B) Spot indications for each gene,  
1678 pseudocolored as indicated in the bottom right legend, as in A. a,a') Superficial layer 2 cells  
1679 express *SLC17A7*(lavender), *CUX2* (magenta), and *LAMP5* (mint). b,b') At deeper locations in  
1680 layer 2, an example of an *SLC17A7*-expressing cell with *CUX2*, *LAMP5* and *COL5A2*  
1681 expression. Note that *LAMP5* expression (mint) decreases in *CUX2/SLC17A7*-expressing cells,  
1682 while *COL5A2/CUX2*-expressing cells increase with depth along Layers 2 and 3 (see, c,c'; d,d';  
1683 e,e').

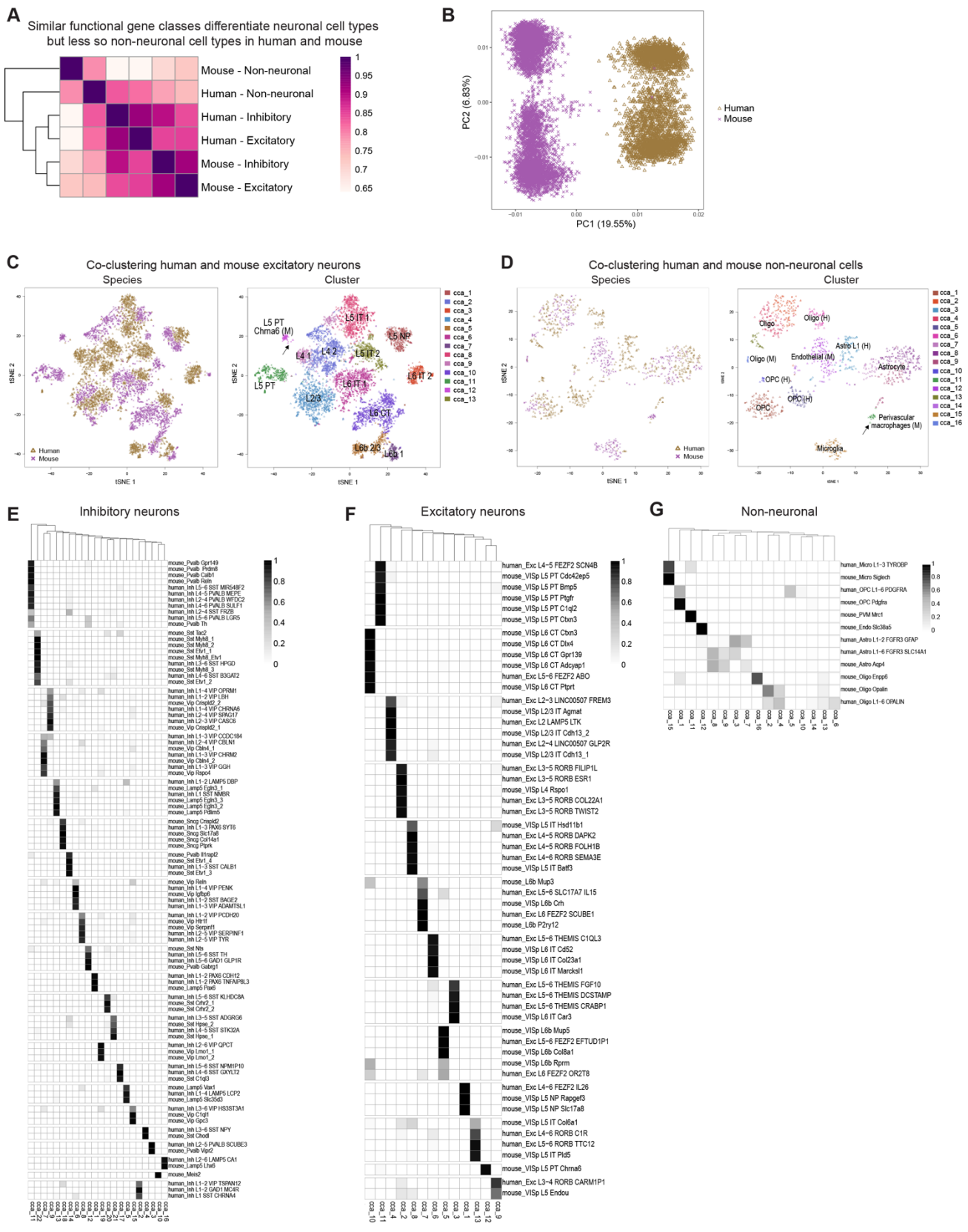




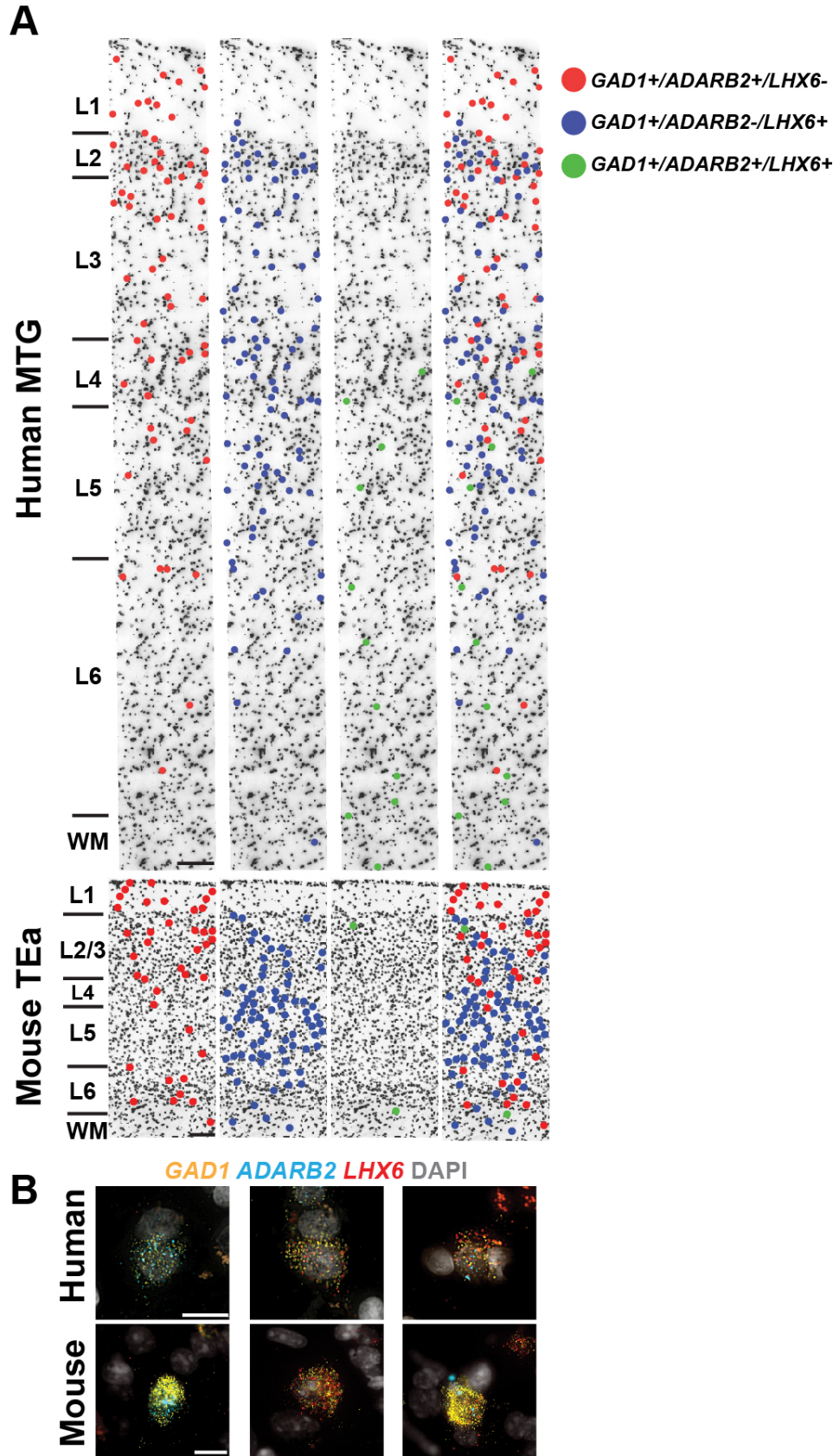
1684

1685

1686 **Extended Data Figure 8. Laminar distribution of superficial excitatory neuron types**  
1687 **validated by smFISH. (A)** Probe density (spots per 100 $\mu$ m<sup>2</sup>) for 9 genes assayed across layers  
1688 1-4 (and partially layer 5) of human MTG. The cortical slice was approximately 0.5mm wide and  
1689 2mm deep. Points correspond to cellular locations *in situ* where the y-axis is the cortical depth  
1690 from the pial surface and the x-axis is the lateral position. Point size and color correspond to  
1691 probe density. Cells that lack probe expression are shown as small grey points. **(B)** *In situ*  
1692 location of cells mapped to indicated cell types and classes (different panels) based on  
1693 expression levels of 9 genes shown in (A). Numbers indicate qualitative calls of the layer to  
1694 which each cell belongs based on cytoarchitecture. 0 indicates that the cell was not annotated.



1696 **Extended Data Figure 9. Quantifying human and mouse cell type homology. (A)** Heatmap  
1697 of Pearson's correlations between average MetaNeighbor AUROC scores for three broad  
1698 classes of human and mouse cortical cell types. Rows and columns are ordered by average-  
1699 linkage hierarchical clustering. **(B)** Human (gold) and mouse (purple) GABAergic neurons  
1700 projected on the first two principal components of a PCA combining expression data from both  
1701 species. Almost 20% of expression differences are explained by species, while 6% are  
1702 explained by major subclasses of interneurons. **(C)** t-SNE plots of first 30 basis vectors from a  
1703 CCA of human and mouse glutamatergic neurons colored by species and CCA Cluster labeled  
1704 with (M) contains only mouse cells. cluster. Arrow highlights two human nuclei that cluster with  
1705 the mouse-specific (M) L5 PT Chrna6 cluster. **(D)** t-SNE plots of first 10 basis vectors from a  
1706 CCA of human and mouse non-neuronal cells colored by species and CCA Clusters labeled  
1707 with (M) or (H) contain only mouse cells or human nuclei, respectively. cluster. Human-specific  
1708 (H) and mouse-specific (M) clusters are labeled. Arrow highlights two human nuclei that cluster  
1709 with mouse perivascular macrophages. **(E-G)** Heatmaps showing the proportion of each human  
1710 and mouse cluster (rows) that are members of each CCA cluster (columns) for GABAergic  
1711 neurons (E), glutamatergic neurons (F), and non-neuronal cells (G). Rows and columns are  
1712 hierarchically clustered, and most CCA clusters include human and mouse clusters that allows  
1713 inference of homology between these clusters.

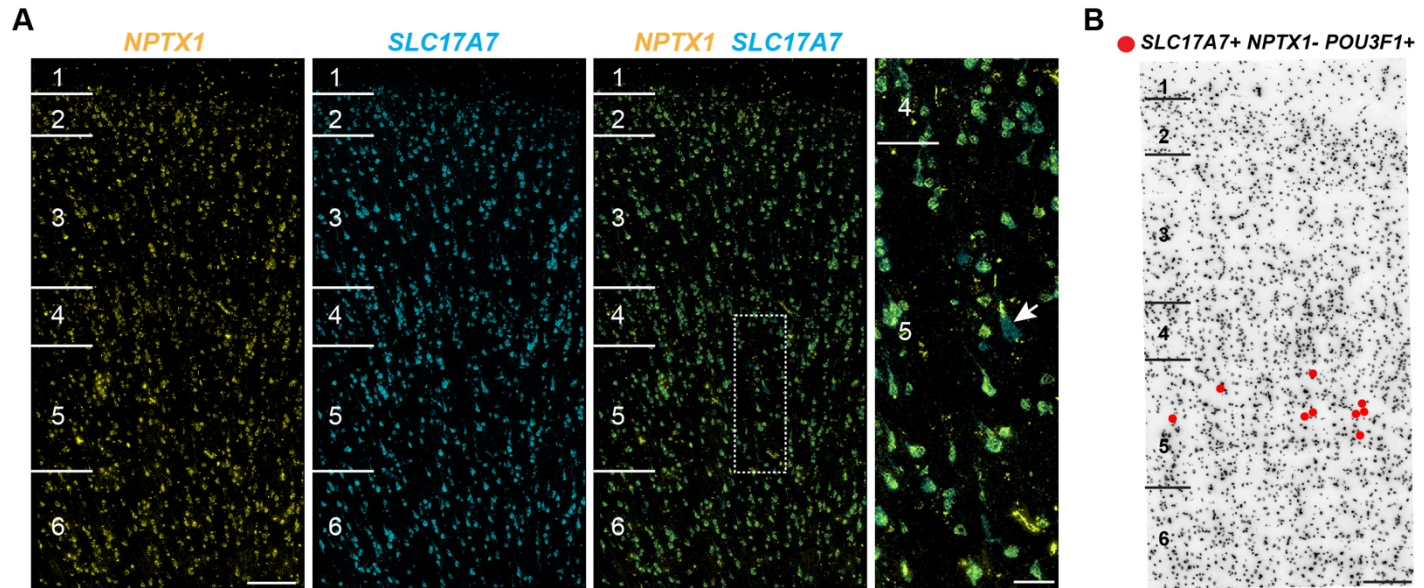


1714

1715



1716 **Extended Data Figure 10. RNAscope multiplex FISH staining for broad interneuron class**  
1717 **markers in human MTG and mouse temporal association area (TEa). (A)** A representative  
1718 inverted DAPI-stained cortical column illustrating the laminar positions of cells  
1719 combinatorially labeled with broad interneuron class markers. Human MTG is shown in  
1720 the top panels and mouse TEa is shown in the bottom panels. Left to right: red dots  
1721 mark cells that are *GAD1*<sup>+</sup>/*Gad1*<sup>+</sup>, *ADARB2*<sup>+</sup>/*Adarb2*<sup>+</sup>, and *LHX6*<sup>-</sup>/*Lhx6*<sup>-</sup> (i.e. *ADARB2*  
1722 branch interneurons); blue dots mark cells that are *GAD1*<sup>+</sup>/*Gad1*<sup>+</sup>, *ADARB2*<sup>-</sup>/*Adarb2*<sup>-</sup>,  
1723 and *LHX6*<sup>+</sup>/*Lhx6*<sup>+</sup> (i.e. *LHX6* branch interneurons); green dots mark cells that  
1724 are *GAD1*<sup>+</sup>/*Gad1*<sup>+</sup>, *ADARB2*<sup>+</sup>/*Adarb2*<sup>+</sup>, *LHX6*<sup>+</sup>/*Lhx6*<sup>+</sup> (i.e. Inh L2-6 *LAMP5 CA1* cells  
1725 [human] or *Lamp5 Lhx6* cells [mouse]); the far-right panel shows all the labeled cell  
1726 classes overlaid onto the cortical column. (B) Representative images of cells labeled  
1727 with the *GAD1*, *ADARB2*, and *LHX6* gene panel for human (top) and mouse (bottom).  
1728 Left to right: cells double positive for *GAD1* and *ADARB2*; cells double positive  
1729 for *GAD1* and *LHX6*; *GAD1*, *ADARB2*, and *LHX6* triple positive cells. Scale bars, 15μm  
1730 (human), 10μm (mouse).



1731

1732 **Extended Data Figure 11. RNAscope multiplex FISH for markers of putative pyramidal**  
1733 **tract (PT) neurons in human MTG. (A)** FISH for *NPTX1*, a marker of non-PT excitatory types  
1734 and *SLC17A7*, which is expressed in all excitatory neurons, shows that *NPTX1* labels  
1735 most *SLC17A7*+ cells across all cortical layers. The area indicated by the boxed region in the  
1736 overlaid image of *NPTX1* and *SCL17A7* staining is shown at higher magnification in the  
1737 adjacent panel to the right. One *SLC17A7*+ cell, indicated with the white arrow, is *NPTX1*-, but  
1738 the rest of the *SLC17A7*+ cells in the field of view are *NPTX1*+. Scale bars, left (200um), right  
1739 (50um). **(B)** A Representative inverted DAPI-stained cortical column overlaid with red dots that  
1740 represent *SLC17A7*+, *NPTX1*-, and *POU3F1*+ cells. *POU3F1* is a specific marker of the  
1741 putative PT type (Exc L4-5 *FEZF2 SCN4B*). Scale bar, 200µm.

## 1742 **Supplementary Data**

1743 **Semantic representation of cell type clusters.** To provide for an unambiguous, rigorous, and  
1744 informative semantic representation of the cell types defined using the single nucleus RNA  
1745 sequencing gene expression clusters, we developed a strategy for defining provisional cell  
1746 types (pCL) in an rdf representation. We previously proposed that these provisional cell types  
1747 could be defined using a combination of the tissue anatomic structure from which the assayed  
1748 specimen was derived (e.g. middle temporal gyrus - go\_sc:part\_of uberon:UBERON\_0002771),  
1749 a set of marker genes whose combination is uniquely expressed in cells of the type (e.g.  
1750 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_8620 and  
1751 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_1751), and a supertype parent class  
1752 represented in either the pCL or in the reference Cell Ontology (CL) (e.g. skos:broader :pCL78)  
1753 [Bakken 2017, Aevermann 2018]. In addition, here we have also made the distinction between  
1754 the experimental evidence supporting the existence of the provisional cell type class in the form  
1755 of a single cell expression cluster ID (mtg\_cluster\_sc:evidence "Inh L1-2 PAX6 CDH12") and  
1756 the provisional cell type class itself (e.g. pCL1). We also included additional knowledge about  
1757 the location of instances of each provisional cell type class within specific layers of the cerebral  
1758 cortex, both in terms of any layer in which the cell body of the type is found to be localized in  
1759 (mtg\_cluster\_sc:has\_soma\_location\_in) and in terms of the layer that a cell of the type is  
1760 preferentially enriched in (mtg\_cluster\_sc:enriched\_in). Finally, we included the cell cluster size  
1761 (i.e. the number of single nuclei included in the cell type cluster from this particular experiment)  
1762 as a rough estimate of the relative cell type abundance in the specimen assayed. This rdf  
1763 representation is amenable to query-based inferencing using SPARQL.

1764

1765 Aevermann BD, Novotny M, Bakken T, Miller JA, Diehl AD, Osumi-Sutherland D, Lasken RS,  
1766 Lein ES, Scheuermann RH. Cell type discovery using single-cell transcriptomics: implications  
1767 for ontological representation. Hum Mol Genet. 2018 May 1;27(R1):R40-R47. doi:  
1768 10.1093/hmg/ddy100. PubMed PMID: 29590361; PubMed Central PMCID: PMC5946857.

1769

1770 Bakken T, Cowell L, Aevermann BD, Novotny M, Hodge R, Miller JA, Lee A, Chang I,  
1771 McCarrison J, Pulendran B, Qian Y, Schork NJ, Lasken RS, Lein ES, Scheuermann RH. Cell  
1772 type discovery and representation in the era of high-content single cell phenotyping. BMC  
1773 Bioinformatics. 2017 Dec 21;18(Suppl 17):559. doi: 10.1186/s12859-017-1977-1. PubMed  
1774 PMID: 29322913; PubMed Central PMCID: PMC5763450.

1775

1776 # \*\*\*\*\* Required\_Resources \*\*\*\*\*

1777

1778 @prefix rdf: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>> .

1779 @prefix rdfs: <<http://www.w3.org/2000/01/rdf-schema#>> .

1780 @prefix xsd: <<http://www.w3.org/2001/XMLSchema#>> .

1781 @prefix owl: <<http://www.w3.org/2002/07/owl#>> .

1782 @prefix : <[http://www.jcvi.org/cl\\_ext/mtg\\_cluster#](http://www.jcvi.org/cl_ext/mtg_cluster#)> .  
1783 @prefix mtg\_cluster\_sc: <[http://www.jcvi.org/cl\\_ext/mtg\\_cluster\\_schema#](http://www.jcvi.org/cl_ext/mtg_cluster_schema#)> .  
1784 @prefix go\_sc: <[http://www.jcvi.org/cl\\_ext/go\\_schema#](http://www.jcvi.org/cl_ext/go_schema#)> .  
1785 @prefix cl: <<http://purl.obolibrary.org/obo/cl#>> .  
1786 @prefix hugo: <<http://ncicb.nci.nih.gov/xml/owl/evs/hugo#>> .  
1787 @prefix uberon: <<http://purl.obolibrary.org/obo/uberon#>> .  
1788 @prefix skos: <<http://www.w3.org/2004/02/skos/core#>> .  
1789  
1790 <[http://www.jcvi.org/cl\\_ext/mtg\\_cluster](http://www.jcvi.org/cl_ext/mtg_cluster)>  
1791 rdf:type owl:Ontology ;  
1792 owl:versionInfo "\$Id: mtg\_cluster\_v4.ttl, Ver 1.0, May 23, 2018, Mohamed Keshk. \$" .  
1793  
1794  
1795 # \*\*\*\*\* MTG\_Cluster\_Annotation Concepts \*\*\*\*\*  
1796  
1797 :pCL1  
1798 a skos:Concept ;  
1799 mtg\_cluster\_sc:id "pCL1" ;  
1800 skos:broader :pCL78 ;  
1801 rdfs:label "PAX6|CDH12-expressing cerebral cortex MTG GABAergic interneuron" ;  
1802 mtg\_cluster\_sc:evidence "Inh L1-2 PAX6 CDH12" ;  
1803 go\_sc:part\_of uberon:UBERON\_0002771 ;  
1804 mtg\_cluster\_sc:enriched\_in "cortical\_layer1" ;  
1805 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1806 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
1807 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_8620 ;  
1808 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_1751 ;  
1809 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1810 mtg\_cluster\_sc:cluster\_size "90"^^xsd:int ;

1811 .  
1812 :pCL2  
1813 a skos:Concept ;  
1814 mtg\_cluster\_sc:id "pCL2" ;  
1815 skos:broader :pCL78 ;  
1816 rdfs:label "PAX6|TNFAIP8L3-expressing cerebral cortex MTG GABAergic interneuron" ;  
1817 mtg\_cluster\_sc:evidence "Inh L1-2 PAX6 TNFAIP8L3" ;  
1818 go\_sc:part\_of uberon:UBERON\_0002771 ;  
1819 mtg\_cluster\_sc:enriched\_in "cortical\_layer2" ;  
1820 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1821 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
1822 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_8620 ;  
1823 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_20620 ;  
1824 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1825 mtg\_cluster\_sc:cluster\_size "16"^^xsd:int ;  
1826 .  
1827 :pCL3  
1828 a skos:Concept ;  
1829 mtg\_cluster\_sc:id "pCL3" ;  
1830 skos:broader :pCL78 ;  
1831 rdfs:label "SST|NMBR-expressing cerebral cortex MTG GABAergic interneuron" ;  
1832 mtg\_cluster\_sc:evidence "Inh L1 SST NMBR" ;  
1833 go\_sc:part\_of uberon:UBERON\_0002771 ;  
1834 mtg\_cluster\_sc:enriched\_in "cortical\_layer1" ;  
1835 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1836 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_11329 ;  
1837 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_7843 ;  
1838 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1839 mtg\_cluster\_sc:cluster\_size "283"^^xsd:int ;



1840 .  
1841 :pCL4  
1842 a skos:Concept ;  
1843 mtg\_cluster\_sc:id "pCL4" ;  
1844 skos:broader :pCL78 ;  
1845 rdfs:label "LAMP5|LCP2-expressing cerebral cortex MTG GABAergic interneuron" ;  
1846 mtg\_cluster\_sc:evidence "Inh L1-4 LAMP5 LCP2" ;  
1847 go\_sc:part\_of uberon:UBERON\_0002771 ;  
1848 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
1849 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1850 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
1851 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
1852 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
1853 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_16097 ;  
1854 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_6529 ;  
1855 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1856 mtg\_cluster\_sc:cluster\_size "356"^^xsd:int ;  
1857 .  
1858 :pCL5  
1859 a skos:Concept ;  
1860 mtg\_cluster\_sc:id "pCL5" ;  
1861 skos:broader :pCL78 ;  
1862 rdfs:label "LAMP5|DBP-expressing cerebral cortex MTG GABAergic interneuron" ;  
1863 mtg\_cluster\_sc:evidence "Inh L1-2 LAMP5 DBP" ;  
1864 go\_sc:part\_of uberon:UBERON\_0002771 ;  
1865 mtg\_cluster\_sc:enriched\_in "cortical\_layer1" ;  
1866 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1867 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
1868 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_16097 ;

1869 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_2697 ;  
1870 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1871 mtg\_cluster\_sc:cluster\_size "21"^^xsd:int ;  
1872 .  
1873 :pCL6  
1874 a skos:Concept ;  
1875 mtg\_cluster\_sc:id "pCL6" ;  
1876 skos:broader :pCL78 ;  
1877 rdfs:label "LAMP5|CA1-expressing cerebral cortex MTG GABAergic interneuron" ;  
1878 mtg\_cluster\_sc:evidence "Inh L2-6 LAMP5 CA1" ;  
1879 go\_sc:part\_of uberon:UBERON\_0002771 ;  
1880 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
1881 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
1882 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
1883 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
1884 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
1885 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
1886 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_16097 ;  
1887 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_1368 ;  
1888 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1889 mtg\_cluster\_sc:cluster\_size "256"^^xsd:int ;  
1890 .  
1891 :pCL7  
1892 a skos:Concept ;  
1893 mtg\_cluster\_sc:id "pCL7" ;  
1894 skos:broader :pCL78 ;  
1895 rdfs:label "SST|CHRNA4-expressing cerebral cortex MTG GABAergic interneuron" ;  
1896 mtg\_cluster\_sc:evidence "Inh L1 SST CHRNA4" ;  
1897 go\_sc:part\_of uberon:UBERON\_0002771 ;

1898 mtg\_cluster\_sc:enriched\_in "cortical\_layer1" ;  
1899 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1900 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_11329 ;  
1901 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_1958 ;  
1902 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1903 mtg\_cluster\_sc:cluster\_size "52"^^xsd:int ;  
1904 .  
1905 :pCL8  
1906 a skos:Concept ;  
1907 mtg\_cluster\_sc:id "pCL8" ;  
1908 skos:broader :pCL78 ;  
1909 rdfs:label "GAD1|MC4R-expressing cerebral cortex MTG GABAergic interneuron" ;  
1910 mtg\_cluster\_sc:evidence "Inh L1-2 GAD1 MC4R" ;  
1911 go\_sc:part\_of uberon:UBERON\_0002771 ;  
1912 mtg\_cluster\_sc:enriched\_in "cortical\_layer1" ;  
1913 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1914 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
1915 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_4092 ;  
1916 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_6932 ;  
1917 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1918 mtg\_cluster\_sc:cluster\_size "107"^^xsd:int ;  
1919 .  
1920 :pCL9  
1921 a skos:Concept ;  
1922 mtg\_cluster\_sc:id "pCL9" ;  
1923 skos:broader :pCL78 ;  
1924 rdfs:label "SST|BAGE2-expressing cerebral cortex MTG GABAergic interneuron" ;  
1925 mtg\_cluster\_sc:evidence "Inh L1-2 SST BAGE2" ;  
1926 go\_sc:part\_of uberon:UBERON\_0002771 ;

1927 mtg\_cluster\_sc:enriched\_in "cortical\_layer1" ;  
1928 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1929 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
1930 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_11329 ;  
1931 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_15723 ;  
1932 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1933 mtg\_cluster\_sc:cluster\_size "108"^^xsd:int ;  
1934 .  
1935 :pCL10  
1936 a skos:Concept ;  
1937 mtg\_cluster\_sc:id "pCL10" ;  
1938 skos:broader :pCL80 ;  
1939 rdfs:label "PAX6|SYT6-expressing cerebral cortex MTG GABAergic interneuron" ;  
1940 mtg\_cluster\_sc:evidence "Inh L1-3 PAX6 SYT6" ;  
1941 go\_sc:part\_of uberon:UBERON\_0002771 ;  
1942 mtg\_cluster\_sc:enriched\_in "cortical\_layer2" ;  
1943 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1944 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
1945 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
1946 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_8620 ;  
1947 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_18638 ;  
1948 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1949 mtg\_cluster\_sc:cluster\_size "29"^^xsd:int ;  
1950 .  
1951 :pCL11  
1952 a skos:Concept ;  
1953 mtg\_cluster\_sc:id "pCL11" ;  
1954 skos:broader :pCL80 ;  
1955 rdfs:label "TSPAN12-expressing cerebral cortex MTG GABAergic interneuron" ;

1956 mtg\_cluster\_sc:evidence "Inh L1-2 VIP TSPAN12" ;  
1957 go\_sc:part\_of uberon:UBERON\_0002771 ;  
1958 mtg\_cluster\_sc:enriched\_in "cortical\_layer1" ;  
1959 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1960 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
1961 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_21641 ;  
1962 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1963 mtg\_cluster\_sc:cluster\_size "42"^^xsd:int ;  
1964 .  
1965 :pCL12  
1966 a skos:Concept ;  
1967 mtg\_cluster\_sc:id "pCL12" ;  
1968 skos:broader :pCL80 ;  
1969 rdfs:label "CHRNA6-expressing cerebral cortex MTG GABAergic interneuron" ;  
1970 mtg\_cluster\_sc:evidence "Inh L1-4 VIP CHRNA6" ;  
1971 go\_sc:part\_of uberon:UBERON\_0002771 ;  
1972 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
1973 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1974 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
1975 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
1976 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
1977 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_15963 ;  
1978 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1979 mtg\_cluster\_sc:cluster\_size "25"^^xsd:int ;  
1980 .  
1981 :pCL13  
1982 a skos:Concept ;  
1983 mtg\_cluster\_sc:id "pCL13" ;  
1984 skos:broader :pCL80 ;



1985 rdfs:label "ADAMTSL1-expressing cerebral cortex MTG GABAergic interneuron" ;  
1986 mtg\_cluster\_sc:evidence "Inh L1-3 VIP ADAMTSL1" ;  
1987 go\_sc:part\_of uberon:UBERON\_0002771 ;  
1988 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
1989 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
1990 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
1991 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
1992 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_14632 ;  
1993 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
1994 mtg\_cluster\_sc:cluster\_size "72"^^xsd:int ;  
1995 .  
1996 :pCL14  
1997 a skos:Concept ;  
1998 mtg\_cluster\_sc:id "pCL14" ;  
1999 skos:broader :pCL80 ;  
2000 rdfs:label "PENK-expressing cerebral cortex MTG GABAergic interneuron" ;  
2001 mtg\_cluster\_sc:evidence "Inh L1-4 VIP PENK" ;  
2002 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2003 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2004 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2005 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2006 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2007 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2008 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_8831 ;  
2009 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2010 mtg\_cluster\_sc:cluster\_size "17"^^xsd:int ;  
2011 .  
2012 :pCL15  
2013 a skos:Concept ;

2014 mtg\_cluster\_sc:id "pCL15" ;  
2015 skos:broader :pCL80 ;  
2016 rdfs:label "QPCT-expressing cerebral cortex MTG GABAergic interneuron" ;  
2017 mtg\_cluster\_sc:evidence "Inh L2-6 VIP QPCT" ;  
2018 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2019 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2020 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2021 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2022 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2023 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2024 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2025 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_9753 ;  
2026 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2027 mtg\_cluster\_sc:cluster\_size "37"^^xsd:int ;  
2028 .  
2029 :pCL16  
2030 a skos:Concept ;  
2031 mtg\_cluster\_sc:id "pCL16" ;  
2032 skos:broader :pCL80 ;  
2033 rdfs:label "HS3ST3A1-expressing cerebral cortex MTG GABAergic interneuron" ;  
2034 mtg\_cluster\_sc:evidence "Inh L3-6 VIP HS3ST3A1" ;  
2035 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2036 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2037 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2038 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2039 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2040 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2041 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_5196 ;  
2042 mtg\_cluster\_sc:neuron\_type "GABAergic" ;

2043 mtg\_cluster\_sc:cluster\_size "80"^^xsd:int ;  
2044 .  
2045 :pCL17  
2046 a skos:Concept ;  
2047 mtg\_cluster\_sc:id "pCL17" ;  
2048 skos:broader :pCL80 ;  
2049 rdfs:label "PCDH20-expressing cerebral cortex MTG GABAergic interneuron" ;  
2050 mtg\_cluster\_sc:evidence "Inh L1-2 VIP PCDH20" ;  
2051 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2052 mtg\_cluster\_sc:enriched\_in "cortical\_layer2" ;  
2053 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2054 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2055 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_14257 ;  
2056 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2057 mtg\_cluster\_sc:cluster\_size "61"^^xsd:int ;  
2058 .  
2059 :pCL18  
2060 a skos:Concept ;  
2061 mtg\_cluster\_sc:id "pCL18" ;  
2062 skos:broader :pCL80 ;  
2063 rdfs:label "SERPINF1-expressing cerebral cortex MTG GABAergic interneuron" ;  
2064 mtg\_cluster\_sc:evidence "Inh L2-5 VIP SERPINF1" ;  
2065 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2066 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2067 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2068 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2069 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2070 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2071 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_8824 ;

2072 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2073 mtg\_cluster\_sc:cluster\_size "55"^^xsd:int ;  
2074 .  
2075 :pCL19  
2076 a skos:Concept ;  
2077 mtg\_cluster\_sc:id "pCL19" ;  
2078 skos:broader :pCL80 ;  
2079 rdfs:label "TYR-expressing cerebral cortex MTG GABAergic interneuron" ;  
2080 mtg\_cluster\_sc:evidence "Inh L2-5 VIP TYR" ;  
2081 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2082 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2083 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2084 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2085 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2086 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2087 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_12442 ;  
2088 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2089 mtg\_cluster\_sc:cluster\_size "62"^^xsd:int ;  
2090 .  
2091 :pCL20  
2092 a skos:Concept ;  
2093 mtg\_cluster\_sc:id "pCL20" ;  
2094 skos:broader :pCL80 ;  
2095 rdfs:label "CHRM2-expressing cerebral cortex MTG GABAergic interneuron" ;  
2096 mtg\_cluster\_sc:evidence "Inh L1-3 VIP CHRM2" ;  
2097 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2098 mtg\_cluster\_sc:enriched\_in "cortical\_layer2" ;  
2099 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2100 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;

2101 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2102 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_1951 ;  
2103 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2104 mtg\_cluster\_sc:cluster\_size "175"^^xsd:int ;  
2105 .  
2106 :pCL21  
2107 a skos:Concept ;  
2108 mtg\_cluster\_sc:id "pCL21" ;  
2109 skos:broader :pCL80 ;  
2110 rdfs:label "CBLN1-expressing cerebral cortex MTG GABAergic interneuron" ;  
2111 mtg\_cluster\_sc:evidence "Inh L2-4 VIP CBLN1" ;  
2112 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2113 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2114 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2115 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2116 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2117 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_1543 ;  
2118 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2119 mtg\_cluster\_sc:cluster\_size "67"^^xsd:int ;  
2120 .  
2121 :pCL22  
2122 a skos:Concept ;  
2123 mtg\_cluster\_sc:id "pCL22" ;  
2124 skos:broader :pCL80 ;  
2125 rdfs:label "CCDC184-expressing cerebral cortex MTG GABAergic interneuron" ;  
2126 mtg\_cluster\_sc:evidence "Inh L1-3 VIP CCDC184" ;  
2127 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2128 mtg\_cluster\_sc:enriched\_in "cortical\_layer2" ;  
2129 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;



2130 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2131 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2132 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_33749 ;  
2133 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2134 mtg\_cluster\_sc:cluster\_size "64"^^xsd:int ;  
2135 .  
2136 :pCL23  
2137 a skos:Concept ;  
2138 mtg\_cluster\_sc:id "pCL23" ;  
2139 skos:broader :pCL80 ;  
2140 rdfs:label "GGH-expressing cerebral cortex MTG GABAergic interneuron" ;  
2141 mtg\_cluster\_sc:evidence "Inh L1-3 VIP GGH" ;  
2142 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2143 mtg\_cluster\_sc:enriched\_in "cortical\_layer2" ;  
2144 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2145 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2146 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2147 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_4248 ;  
2148 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2149 mtg\_cluster\_sc:cluster\_size "68"^^xsd:int ;  
2150 .  
2151 :pCL24  
2152 a skos:Concept ;  
2153 mtg\_cluster\_sc:id "pCL24" ;  
2154 skos:broader :pCL80 ;  
2155 rdfs:label "LBH-expressing cerebral cortex MTG GABAergic interneuron" ;  
2156 mtg\_cluster\_sc:evidence "Inh L1-2 VIP LBH" ;  
2157 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2158 mtg\_cluster\_sc:enriched\_in "cortical\_layer2" ;

2159 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2160 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2161 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_29532 ;  
2162 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2163 mtg\_cluster\_sc:cluster\_size "47"^^xsd:int ;  
2164 .  
2165 :pCL25  
2166 a skos:Concept ;  
2167 mtg\_cluster\_sc:id "pCL25" ;  
2168 skos:broader :pCL80 ;  
2169 rdfs:label "CASC6-expressing cerebral cortex MTG GABAergic interneuron" ;  
2170 mtg\_cluster\_sc:evidence "Inh L2-3 VIP CASC6" ;  
2171 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2172 mtg\_cluster\_sc:enriched\_in "cortical\_layer2" ;  
2173 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2174 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2175 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_49076 ;  
2176 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2177 mtg\_cluster\_sc:cluster\_size "45"^^xsd:int ;  
2178 .  
2179 :pCL26  
2180 a skos:Concept ;  
2181 mtg\_cluster\_sc:id "pCL26" ;  
2182 skos:broader :pCL80 ;  
2183 rdfs:label "SPAG17-expressing cerebral cortex MTG GABAergic interneuron" ;  
2184 mtg\_cluster\_sc:evidence "Inh L2-4 VIP SPAG17" ;  
2185 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2186 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2187 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;

2188 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2189 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2190 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_26620 ;  
2191 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2192 mtg\_cluster\_sc:cluster\_size "33"^^xsd:int ;  
2193 .  
2194 :pCL27  
2195 a skos:Concept ;  
2196 mtg\_cluster\_sc:id "pCL27" ;  
2197 skos:broader :pCL80 ;  
2198 rdfs:label "OPRM1-expressing cerebral cortex MTG GABAergic interneuron" ;  
2199 mtg\_cluster\_sc:evidence "Inh L1-4 VIP OPRM1" ;  
2200 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2201 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2202 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2203 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2204 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2205 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2206 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_8156 ;  
2207 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2208 mtg\_cluster\_sc:cluster\_size "52"^^xsd:int ;  
2209 .  
2210 :pCL28  
2211 a skos:Concept ;  
2212 mtg\_cluster\_sc:id "pCL28" ;  
2213 skos:broader :pCL81 ;  
2214 rdfs:label "NPY-expressing cerebral cortex MTG GABAergic interneuron" ;  
2215 mtg\_cluster\_sc:evidence "Inh L3-6 SST NPY" ;  
2216 go\_sc:part\_of uberon:UBERON\_0002771 ;

2217 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2218 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2219 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2220 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2221 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2222 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_7955 ;  
2223 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2224 mtg\_cluster\_sc:cluster\_size "15"^^xsd:int ;  
2225 .  
2226 :pCL29  
2227 a skos:Concept ;  
2228 mtg\_cluster\_sc:id "pCL29" ;  
2229 skos:broader :pCL81 ;  
2230 rdfs:label "HPGD-expressing cerebral cortex MTG GABAergic interneuron" ;  
2231 mtg\_cluster\_sc:evidence "Inh L3-6 SST HPGD" ;  
2232 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2233 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2234 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2235 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2236 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2237 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2238 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_5154 ;  
2239 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2240 mtg\_cluster\_sc:cluster\_size "60"^^xsd:int ;  
2241 .  
2242 :pCL30  
2243 a skos:Concept ;  
2244 mtg\_cluster\_sc:id "pCL30" ;  
2245 skos:broader :pCL81 ;

2246 rdfs:label "B3GAT2-expressing cerebral cortex MTG GABAergic interneuron" ;  
2247 mtg\_cluster\_sc:evidence "Inh L4-6 SST B3GAT2" ;  
2248 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2249 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2250 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2251 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2252 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2253 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_922 ;  
2254 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2255 mtg\_cluster\_sc:cluster\_size "182"^^xsd:int ;  
2256 .  
2257 :pCL31  
2258 a skos:Concept ;  
2259 mtg\_cluster\_sc:id "pCL31" ;  
2260 skos:broader :pCL81 ;  
2261 rdfs:label "KLHDC8A-expressing cerebral cortex MTG GABAergic interneuron" ;  
2262 mtg\_cluster\_sc:evidence "Inh L5-6 SST KLHDC8A" ;  
2263 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2264 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2265 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2266 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2267 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_25573 ;  
2268 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2269 mtg\_cluster\_sc:cluster\_size "63"^^xsd:int ;  
2270 .  
2271 :pCL32  
2272 a skos:Concept ;  
2273 mtg\_cluster\_sc:id "pCL32" ;  
2274 skos:broader :pCL81 ;



2275 rdfs:label "NPM1P10-expressing cerebral cortex MTG GABAergic interneuron" ;  
2276 mtg\_cluster\_sc:evidence "Inh L5-6 SST NPM1P10" ;  
2277 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2278 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2279 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2280 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2281 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_7912 ;  
2282 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2283 mtg\_cluster\_sc:cluster\_size "79"^^xsd:int ;  
2284 .  
2285 :pCL33  
2286 a skos:Concept ;  
2287 mtg\_cluster\_sc:id "pCL33" ;  
2288 skos:broader :pCL81 ;  
2289 rdfs:label "GXYLT2-expressing cerebral cortex MTG GABAergic interneuron" ;  
2290 mtg\_cluster\_sc:evidence "Inh L4-6 SST GXYLT2" ;  
2291 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2292 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2293 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2294 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2295 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2296 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_33383 ;  
2297 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2298 mtg\_cluster\_sc:cluster\_size "41"^^xsd:int ;  
2299 .  
2300 :pCL34  
2301 a skos:Concept ;  
2302 mtg\_cluster\_sc:id "pCL34" ;  
2303 skos:broader :pCL81 ;

2304 rdfs:label "STK32A-expressing cerebral cortex MTG GABAergic interneuron" ;  
2305 mtg\_cluster\_sc:evidence "Inh L4-5 SST STK32A" ;  
2306 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2307 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2308 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2309 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2310 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_28317 ;  
2311 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2312 mtg\_cluster\_sc:cluster\_size "93"^^xsd:int ;  
2313 .  
2314 :pCL35  
2315 a skos:Concept ;  
2316 mtg\_cluster\_sc:id "pCL35" ;  
2317 skos:broader :pCL81 ;  
2318 rdfs:label "CALB1-expressing cerebral cortex MTG GABAergic interneuron" ;  
2319 mtg\_cluster\_sc:evidence "Inh L1-3 SST CALB1" ;  
2320 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2321 mtg\_cluster\_sc:enriched\_in "cortical\_layer2" ;  
2322 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2323 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2324 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2325 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_1434 ;  
2326 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2327 mtg\_cluster\_sc:cluster\_size "279"^^xsd:int ;  
2328 .  
2329 :pCL36  
2330 a skos:Concept ;  
2331 mtg\_cluster\_sc:id "pCL36" ;  
2332 skos:broader :pCL81 ;

2333 rdfs:label "ADGRG6-expressing cerebral cortex MTG GABAergic interneuron" ;  
2334 mtg\_cluster\_sc:evidence "Inh L3-5 SST ADGRG6" ;  
2335 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2336 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2337 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2338 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2339 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2340 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_13841 ;  
2341 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2342 mtg\_cluster\_sc:cluster\_size "132"^^xsd:int ;  
2343 .  
2344 :pCL37  
2345 a skos:Concept ;  
2346 mtg\_cluster\_sc:id "pCL37" ;  
2347 skos:broader :pCL81 ;  
2348 rdfs:label "FRZB-expressing cerebral cortex MTG GABAergic interneuron" ;  
2349 mtg\_cluster\_sc:evidence "Inh L2-4 SST FRZB" ;  
2350 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2351 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2352 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2353 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2354 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2355 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_3959 ;  
2356 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2357 mtg\_cluster\_sc:cluster\_size "64"^^xsd:int ;  
2358 .  
2359 :pCL38  
2360 a skos:Concept ;  
2361 mtg\_cluster\_sc:id "pCL38" ;

2362 skos:broader :pCL81 ;  
2363 rdfs:label "TH-expressing-expressing cerebral cortex MTG GABAergic interneuron" ;  
2364 mtg\_cluster\_sc:evidence "Inh L5-6 SST TH" ;  
2365 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2366 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2367 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2368 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2369 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_11782 ;  
2370 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2371 mtg\_cluster\_sc:cluster\_size "27"^^xsd:int ;  
2372 .  
2373 :pCL39  
2374 a skos:Concept ;  
2375 mtg\_cluster\_sc:id "pCL39" ;  
2376 skos:broader :pCL82 ;  
2377 rdfs:label "GLP1R-expressing cerebral cortex MTG GABAergic interneuron" ;  
2378 mtg\_cluster\_sc:evidence "Inh L5-6 GAD1 GLP1R" ;  
2379 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2380 mtg\_cluster\_sc:enriched\_in "cortical\_layer6" ;  
2381 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2382 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2383 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_4324 ;  
2384 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2385 mtg\_cluster\_sc:cluster\_size "27"^^xsd:int ;  
2386 .  
2387 :pCL40  
2388 a skos:Concept ;  
2389 mtg\_cluster\_sc:id "pCL40" ;  
2390 skos:broader :pCL82 ;

2391 rdfs:label "LGR5-expressing cerebral cortex MTG GABAergic interneuron" ;  
2392 mtg\_cluster\_sc:evidence "Inh L5-6 PVALB LGR5" ;  
2393 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2394 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2395 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2396 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2397 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_4504 ;  
2398 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2399 mtg\_cluster\_sc:cluster\_size "52"^^xsd:int ;  
2400 .  
2401 :pCL41  
2402 a skos:Concept ;  
2403 mtg\_cluster\_sc:id "pCL41" ;  
2404 skos:broader :pCL82 ;  
2405 rdfs:label "MEPE-expressing cerebral cortex MTG GABAergic interneuron" ;  
2406 mtg\_cluster\_sc:evidence "Inh L4-5 PVALB MEPE" ;  
2407 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2408 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2409 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2410 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2411 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_13361 ;  
2412 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2413 mtg\_cluster\_sc:cluster\_size "64"^^xsd:int ;  
2414 .  
2415 :pCL42  
2416 a skos:Concept ;  
2417 mtg\_cluster\_sc:id "pCL42" ;  
2418 skos:broader :pCL82 ;  
2419 rdfs:label "WFDC2-expressing cerebral cortex MTG GABAergic interneuron" ;



2420 mtg\_cluster\_sc:evidence "Inh L2-4 PVALB WFDC2" ;  
2421 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2422 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2423 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2424 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2425 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2426 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_15939 ;  
2427 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2428 mtg\_cluster\_sc:cluster\_size "387"^^xsd:int ;  
2429 .  
2430 :pCL43  
2431 a skos:Concept ;  
2432 mtg\_cluster\_sc:id "pCL43" ;  
2433 skos:broader :pCL82 ;  
2434 rdfs:label "SULF1-expressing cerebral cortex MTG GABAergic interneuron" ;  
2435 mtg\_cluster\_sc:evidence "Inh L4-6 PVALB SULF1" ;  
2436 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2437 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2438 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2439 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2440 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2441 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_20391 ;  
2442 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2443 mtg\_cluster\_sc:cluster\_size "167"^^xsd:int ;  
2444 .  
2445 :pCL44  
2446 a skos:Concept ;  
2447 mtg\_cluster\_sc:id "pCL44" ;  
2448 skos:broader :pCL82 ;

2449 rdfs:label "SST|MIR548F2-expressing cerebral cortex MTG GABAergic interneuron" ;  
2450 mtg\_cluster\_sc:evidence "Inh L5-6 SST MIR548F2" ;  
2451 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2452 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2453 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2454 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2455 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_11329 ;  
2456 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_35306 ;  
2457 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2458 mtg\_cluster\_sc:cluster\_size "80"^^xsd:int ;  
2459 .  
2460 :pCL45  
2461 a skos:Concept ;  
2462 mtg\_cluster\_sc:id "pCL45" ;  
2463 skos:broader :pCL82 ;  
2464 rdfs:label "SCUBE3-expressing cerebral cortex MTG GABAergic interneuron" ;  
2465 mtg\_cluster\_sc:evidence "Inh L2-5 PVALB SCUBE3" ;  
2466 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2467 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2468 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2469 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2470 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2471 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2472 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_13655 ;  
2473 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2474 mtg\_cluster\_sc:cluster\_size "32"^^xsd:int ;  
2475 .  
2476 :pCL46  
2477 a skos:Concept ;

2478 mtg\_cluster\_sc:id "pCL46" ;  
2479 skos:broader :pCL77 ;  
2480 rdfs:label "LAMP5|LTK-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2481 mtg\_cluster\_sc:evidence "Exc L2 LAMP5 LTK" ;  
2482 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2483 mtg\_cluster\_sc:enriched\_in "cortical\_layer2" ;  
2484 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2485 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_16097 ;  
2486 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_6721 ;  
2487 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2488 mtg\_cluster\_sc:cluster\_size "812"^^xsd:int ;  
2489 .  
2490 :pCL47  
2491 a skos:Concept ;  
2492 mtg\_cluster\_sc:id "pCL47" ;  
2493 skos:broader :pCL77 ;  
2494 rdfs:label "LINC00507|GLP2R-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2495 mtg\_cluster\_sc:evidence "Exc L2-4 LINC00507 GLP2R" ;  
2496 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2497 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2498 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2499 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2500 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2501 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_43558 ;  
2502 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_4325 ;  
2503 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2504 mtg\_cluster\_sc:cluster\_size "170"^^xsd:int ;  
2505 .  
2506 :pCL48

2507 a skos:Concept ;  
2508 mtg\_cluster\_sc:id "pCL48" ;  
2509 skos:broader :pCL77 ;  
2510 rdfs:label "LINC00507|FREM3-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2511 mtg\_cluster\_sc:evidence "Exc L2-3 LINC00507 FREM3" ;  
2512 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2513 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2514 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2515 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2516 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_43558 ;  
2517 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_25172 ;  
2518 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2519 mtg\_cluster\_sc:cluster\_size "2284"^^xsd:int ;  
2520 .  
2521 :pCL49  
2522 a skos:Concept ;  
2523 mtg\_cluster\_sc:id "pCL49" ;  
2524 skos:broader :pCL77 ;  
2525 rdfs:label "THEMIS|C1QL3-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2526 mtg\_cluster\_sc:evidence "Exc L5-6 THEMIS C1QL3" ;  
2527 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2528 mtg\_cluster\_sc:enriched\_in "cortical\_layer6" ;  
2529 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2530 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2531 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_21569 ;  
2532 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_19359 ;  
2533 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2534 mtg\_cluster\_sc:cluster\_size "1537"^^xsd:int ;  
2535 .

2536 :pCL50  
2537 a skos:Concept ;  
2538 mtg\_cluster\_sc:id "pCL50" ;  
2539 skos:broader :pCL77 ;  
2540 rdfs:label "RORB|CARM1P1-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2541 mtg\_cluster\_sc:evidence "Exc L3-4 RORB CARM1P1" ;  
2542 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2543 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2544 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2545 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2546 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10259 ;  
2547 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_23392 ;  
2548 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2549 mtg\_cluster\_sc:cluster\_size "280"^^xsd:int ;  
2550 .  
2551 :pCL51  
2552 a skos:Concept ;  
2553 mtg\_cluster\_sc:id "pCL51" ;  
2554 skos:broader :pCL77 ;  
2555 rdfs:label "RORB|ESR1-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2556 mtg\_cluster\_sc:evidence "Exc L3-5 RORB ESR1" ;  
2557 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2558 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2559 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2560 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2561 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2562 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10259 ;  
2563 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_3467 ;  
2564 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;



2565 mtg\_cluster\_sc:cluster\_size "1428"^^xsd:int ;  
2566 .  
2567 :pCL52  
2568 a skos:Concept ;  
2569 mtg\_cluster\_sc:id "pCL52" ;  
2570 skos:broader :pCL77 ;  
2571 rdfs:label "RORB|COL22A1-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2572 mtg\_cluster\_sc:evidence "Exc L3-5 RORB COL22A1" ;  
2573 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2574 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2575 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2576 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2577 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2578 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10259 ;  
2579 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_22989 ;  
2580 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2581 mtg\_cluster\_sc:cluster\_size "160"^^xsd:int ;  
2582 .  
2583 :pCL53  
2584 a skos:Concept ;  
2585 mtg\_cluster\_sc:id "pCL53" ;  
2586 skos:broader :pCL77 ;  
2587 rdfs:label "RORB|FILIP1L-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2588 mtg\_cluster\_sc:evidence "Exc L3-5 RORB FILIP1L" ;  
2589 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2590 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2591 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2592 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2593 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;

2594 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10259 ;  
2595 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_24589 ;  
2596 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2597 mtg\_cluster\_sc:cluster\_size "153"^^xsd:int ;  
2598 .  
2599 :pCL54  
2600 a skos:Concept ;  
2601 mtg\_cluster\_sc:id "pCL54" ;  
2602 skos:broader :pCL77 ;  
2603 rdfs:label "RORB|TWIST2-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2604 mtg\_cluster\_sc:evidence "Exc L3-5 RORB TWIST2" ;  
2605 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2606 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2607 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2608 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2609 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2610 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10259 ;  
2611 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_20670 ;  
2612 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2613 mtg\_cluster\_sc:cluster\_size "93"^^xsd:int ;  
2614 .  
2615 :pCL55  
2616 a skos:Concept ;  
2617 mtg\_cluster\_sc:id "pCL55" ;  
2618 skos:broader :pCL77 ;  
2619 rdfs:label "RORB|FOLH1B-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2620 mtg\_cluster\_sc:evidence "Exc L4-5 RORB FOLH1B" ;  
2621 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2622 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;

2623 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2624 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2625 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10259 ;  
2626 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_13636 ;  
2627 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2628 mtg\_cluster\_sc:cluster\_size "870"^^xsd:int ;  
2629 .  
2630 :pCL56  
2631 a skos:Concept ;  
2632 mtg\_cluster\_sc:id "pCL56" ;  
2633 skos:broader :pCL77 ;  
2634 rdfs:label "RORB|SEMA3E-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2635 mtg\_cluster\_sc:evidence "Exc L4-6 RORB SEMA3E" ;  
2636 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2637 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2638 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2639 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2640 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2641 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10259 ;  
2642 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10727 ;  
2643 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2644 mtg\_cluster\_sc:cluster\_size "777"^^xsd:int ;  
2645 .  
2646 :pCL57  
2647 a skos:Concept ;  
2648 mtg\_cluster\_sc:id "pCL57" ;  
2649 skos:broader :pCL77 ;  
2650 rdfs:label "RORB|DAPK2-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2651 mtg\_cluster\_sc:evidence "Exc L4-5 RORB DAPK2" ;

2652 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2653 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2654 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2655 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2656 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10259 ;  
2657 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_2675 ;  
2658 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2659 mtg\_cluster\_sc:cluster\_size "173"^^xsd:int ;  
2660 .  
2661 :pCL58  
2662 a skos:Concept ;  
2663 mtg\_cluster\_sc:id "pCL58" ;  
2664 skos:broader :pCL77 ;  
2665 rdfs:label "RORB|TTC12-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2666 mtg\_cluster\_sc:evidence "Exc L5-6 RORB TTC12" ;  
2667 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2668 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2669 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2670 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2671 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10259 ;  
2672 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_23700 ;  
2673 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2674 mtg\_cluster\_sc:cluster\_size "167"^^xsd:int ;  
2675 .  
2676 :pCL59  
2677 a skos:Concept ;  
2678 mtg\_cluster\_sc:id "pCL59" ;  
2679 skos:broader :pCL77 ;  
2680 rdfs:label "RORB|C1R-expressing cerebral cortex MTG Glutamatergic neuron" ;

2681 mtg\_cluster\_sc:evidence "Exc L4-6 RORB C1R" ;  
2682 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2683 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2684 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2685 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2686 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2687 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10259 ;  
2688 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_1246 ;  
2689 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2690 mtg\_cluster\_sc:cluster\_size "160"^^xsd:int ;  
2691 .  
2692 :pCL60  
2693 a skos:Concept ;  
2694 mtg\_cluster\_sc:id "pCL60" ;  
2695 skos:broader :pCL77 ;  
2696 rdfs:label "FEZF2|SCN4B-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2697 mtg\_cluster\_sc:evidence "Exc L4-5 FEZF2 SCN4B" ;  
2698 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2699 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2700 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2701 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2702 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_13506 ;  
2703 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10592 ;  
2704 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2705 mtg\_cluster\_sc:cluster\_size "25"^^xsd:int ;  
2706 .  
2707 :pCL61  
2708 a skos:Concept ;  
2709 mtg\_cluster\_sc:id "pCL61" ;

2710 skos:broader :pCL77 ;  
2711 rdfs:label "THEMIS|DCSTAMP-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2712 mtg\_cluster\_sc:evidence "Exc L5-6 THEMIS DCSTAMP" ;  
2713 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2714 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2715 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2716 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2717 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_21569 ;  
2718 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_18549 ;  
2719 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2720 mtg\_cluster\_sc:cluster\_size "53"^^xsd:int ;  
2721 .  
2722 :pCL62  
2723 a skos:Concept ;  
2724 mtg\_cluster\_sc:id "pCL62" ;  
2725 skos:broader :pCL77 ;  
2726 rdfs:label "THEMIS|CRABP1-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2727 mtg\_cluster\_sc:evidence "Exc L5-6 THEMIS CRABP1" ;  
2728 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2729 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2730 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2731 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2732 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_21569 ;  
2733 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_2338 ;  
2734 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2735 mtg\_cluster\_sc:cluster\_size "147"^^xsd:int ;  
2736 .  
2737 :pCL63  
2738 a skos:Concept ;



2739 mtg\_cluster\_sc:id "pCL63" ;  
2740 skos:broader :pCL77 ;  
2741 rdfs:label "THEMIS|FGF10-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2742 mtg\_cluster\_sc:evidence "Exc L5-6 THEMIS FGF10" ;  
2743 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2744 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2745 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2746 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2747 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_21569 ;  
2748 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_3666 ;  
2749 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2750 mtg\_cluster\_sc:cluster\_size "78"^^xsd:int ;  
2751 .  
2752 :pCL64  
2753 a skos:Concept ;  
2754 mtg\_cluster\_sc:id "pCL64" ;  
2755 skos:broader :pCL77 ;  
2756 rdfs:label "FEZF2|IL26-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2757 mtg\_cluster\_sc:evidence "Exc L4-6 FEZF2 IL26" ;  
2758 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2759 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2760 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2761 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2762 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2763 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_13506 ;  
2764 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_17119 ;  
2765 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2766 mtg\_cluster\_sc:cluster\_size "344"^^xsd:int ;  
2767 .

2768 :pCL65  
2769 a skos:Concept ;  
2770 mtg\_cluster\_sc:id "pCL65" ;  
2771 skos:broader :pCL77 ;  
2772 rdfs:label "FEZF2|ABO-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2773 mtg\_cluster\_sc:evidence "Exc L5-6 FEZF2 ABO" ;  
2774 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2775 mtg\_cluster\_sc:enriched\_in "cortical\_layer6" ;  
2776 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2777 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2778 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_13506 ;  
2779 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_79 ;  
2780 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2781 mtg\_cluster\_sc:cluster\_size "373"^^xsd:int ;  
2782 .  
2783 :pCL66  
2784 a skos:Concept ;  
2785 mtg\_cluster\_sc:id "pCL66" ;  
2786 skos:broader :pCL77 ;  
2787 rdfs:label "FEZF2|SCUBE1-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2788 mtg\_cluster\_sc:evidence "Exc L6 FEZF2 SCUBE1" ;  
2789 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2790 mtg\_cluster\_sc:enriched\_in "cortical\_layer6" ;  
2791 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2792 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_13506 ;  
2793 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_13441 ;  
2794 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2795 mtg\_cluster\_sc:cluster\_size "52"^^xsd:int ;  
2796 .

2797 :pCL67  
2798 a skos:Concept ;  
2799 mtg\_cluster\_sc:id "pCL67" ;  
2800 skos:broader :pCL77 ;  
2801 rdfs:label "IL15-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2802 mtg\_cluster\_sc:evidence "Exc L5-6 SLC17A7 IL15" ;  
2803 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2804 mtg\_cluster\_sc:enriched\_in "cortical\_layer6" ;  
2805 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2806 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2807 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_5977 ;  
2808 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2809 mtg\_cluster\_sc:cluster\_size "56"^^xsd:int ;  
2810 .  
2811 :pCL68  
2812 a skos:Concept ;  
2813 mtg\_cluster\_sc:id "pCL68" ;  
2814 skos:broader :pCL77 ;  
2815 rdfs:label "FEZF2|OR2T8-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2816 mtg\_cluster\_sc:evidence "Exc L6 FEZF2 OR2T8" ;  
2817 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2818 mtg\_cluster\_sc:enriched\_in "cortical\_layer6" ;  
2819 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2820 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_13506 ;  
2821 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_15020 ;  
2822 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2823 mtg\_cluster\_sc:cluster\_size "19"^^xsd:int ;  
2824 .  
2825 :pCL69

2826 a skos:Concept ;  
2827 mtg\_cluster\_sc:id "pCL69" ;  
2828 skos:broader :pCL77 ;  
2829 rdfs:label "FEZF2|EFTUD1P1-expressing cerebral cortex MTG Glutamatergic neuron" ;  
2830 mtg\_cluster\_sc:evidence "Exc L5-6 FEZF2 EFTUD1P1" ;  
2831 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2832 mtg\_cluster\_sc:enriched\_in "cortical\_layer6" ;  
2833 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2834 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2835 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_13506 ;  
2836 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_31739 ;  
2837 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2838 mtg\_cluster\_sc:cluster\_size "314"^^xsd:int ;  
2839 .  
2840 :pCL70  
2841 a skos:Concept ;  
2842 mtg\_cluster\_sc:id "pCL70" ;  
2843 skos:broader :pCL83 ;  
2844 rdfs:label "PDGFRA-expressing MTG Oligodendrocyte precursor cell" ;  
2845 mtg\_cluster\_sc:evidence "OPC L1-6 PDGFRA" ;  
2846 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2847 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2848 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2849 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2850 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2851 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2852 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2853 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2854 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_8803 ;

2855 mtg\_cluster\_sc:cluster\_size "238"^^xsd:int ;  
2856 .  
2857 :pCL71  
2858 a skos:Concept ;  
2859 mtg\_cluster\_sc:id "pCL71" ;  
2860 skos:broader :pCL89 ;  
2861 rdfs:label "SLC14A1-expressing MTG astrocyte" ;  
2862 mtg\_cluster\_sc:evidence "Astro L1-6 FGFR3 SLC14A1" ;  
2863 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2864 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2865 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2866 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2867 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2868 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2869 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2870 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2871 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_10918 ;  
2872 mtg\_cluster\_sc:cluster\_size "230"^^xsd:int ;  
2873 .  
2874 :pCL72  
2875 a skos:Concept ;  
2876 mtg\_cluster\_sc:id "pCL72" ;  
2877 skos:broader :pCL89 ;  
2878 rdfs:label "GFAP-expressing MTG astrocyte" ;  
2879 mtg\_cluster\_sc:evidence "Astro L1-2 FGFR3 GFAP" ;  
2880 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2881 mtg\_cluster\_sc:enriched\_in "cortical\_layer2" ;  
2882 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2883 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;

2884 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_4235 ;  
2885 mtg\_cluster\_sc:cluster\_size "61"^^xsd:int ;  
2886 .  
2887 :pCL73  
2888 a skos:Concept ;  
2889 mtg\_cluster\_sc:id "pCL73" ;  
2890 skos:broader :pCL86 ;  
2891 rdfs:label "OPALIN-expressing MTG Oligodendrocyte" ;  
2892 mtg\_cluster\_sc:evidence "Oligo L1-6 OPALIN" ;  
2893 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2894 mtg\_cluster\_sc:enriched\_in "cortical\_layer5" ;  
2895 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2896 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2897 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2898 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2899 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2900 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2901 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_20707 ;  
2902 mtg\_cluster\_sc:cluster\_size "313"^^xsd:int ;  
2903 .  
2904 :pCL74  
2905 a skos:Concept ;  
2906 mtg\_cluster\_sc:id "pCL74" ;  
2907 skos:broader :pCL87 ;  
2908 rdfs:label "NOSTRIN-expressing cerebral cortex MTG endothelial cell" ;  
2909 mtg\_cluster\_sc:evidence "Endo L2-6 NOSTRIN" ;  
2910 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2911 mtg\_cluster\_sc:enriched\_in "cortical\_layer4" ;  
2912 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;



2913 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2914 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer4" ;  
2915 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer5" ;  
2916 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer6" ;  
2917 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_20203 ;  
2918 mtg\_cluster\_sc:cluster\_size "9"^^xsd:int ;  
2919 .  
2920 :pCL75  
2921 a skos:Concept ;  
2922 mtg\_cluster\_sc:id "pCL75" ;  
2923 skos:broader :pCL88 ;  
2924 rdfs:label "TYROBP-expressing MTG Microglial cell" ;  
2925 mtg\_cluster\_sc:evidence "Micro L1-3 TYROBP" ;  
2926 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2927 mtg\_cluster\_sc:enriched\_in "cortical\_layer3" ;  
2928 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer1" ;  
2929 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer2" ;  
2930 mtg\_cluster\_sc:has\_soma\_location\_in "cortical\_layer3" ;  
2931 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_12449 ;  
2932 mtg\_cluster\_sc:cluster\_size "63"^^xsd:int ;  
2933 .  
2934 :pCL76  
2935 a skos:Concept ;  
2936 mtg\_cluster\_sc:id "pCL76" ;  
2937 skos:broader :pCL90 ;  
2938 rdfs:label "GAD1-expressing cerebral cortex MTG GABAergic interneuron" ;  
2939 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2940 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_4092 ;  
2941 mtg\_cluster\_sc:neuron\_type "GABAergic" ;

2942 .  
2943 :pCL77  
2944 a skos:Concept ;  
2945 mtg\_cluster\_sc:id "pCL77" ;  
2946 skos:broader :pCL91 ;  
2947 rdfs:label "SLC17A7-expressing MTG Glutamatergic neuron" ;  
2948 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2949 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_16704 ;  
2950 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;  
2951 .  
2952 :pCL78  
2953 a skos:Concept ;  
2954 mtg\_cluster\_sc:id "pCL78" ;  
2955 skos:broader :pCL76 ;  
2956 rdfs:label "ADARB2-expressing cerebral cortex MTG GABAergic interneuron" ;  
2957 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2958 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_227 ;  
2959 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2960 .  
2961 :pCL79  
2962 a skos:Concept ;  
2963 mtg\_cluster\_sc:id "pCL79" ;  
2964 skos:broader :pCL76 ;  
2965 rdfs:label "LHX6-expressing cerebral cortex MTG GABAergic interneuron" ;  
2966 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2967 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_21735 ;  
2968 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2969 .  
2970 :pCL80

2971 a skos:Concept ;  
2972 mtg\_cluster\_sc:id "pCL80" ;  
2973 skos:broader :pCL78 ;  
2974 rdfs:label "VIP-expressing cerebral cortex MTG GABAergic interneuron" ;  
2975 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2976 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_12693 ;  
2977 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2978 .  
2979 :pCL81  
2980 a skos:Concept ;  
2981 mtg\_cluster\_sc:id "pCL81" ;  
2982 skos:broader :pCL79 ;  
2983 rdfs:label "SST-expressing cerebral cortex MTG GABAergic interneuron" ;  
2984 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2985 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_11329 ;  
2986 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2987 .  
2988 :pCL82  
2989 a skos:Concept ;  
2990 mtg\_cluster\_sc:id "pCL82" ;  
2991 skos:broader :pCL79 ;  
2992 rdfs:label "PVALB-expressing cerebral cortex MTG GABAergic interneuron" ;  
2993 go\_sc:part\_of uberon:UBERON\_0002771 ;  
2994 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_9704 ;  
2995 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
2996 .  
2997 :pCL83  
2998 a skos:Concept ;  
2999 mtg\_cluster\_sc:id "pCL83" ;

3000 skos:broader cl:CL\_0002453 ;  
3001 rdfs:label "MTG Oligodendrocyte precursor cell" ;  
3002 go\_sc:part\_of uberon:UBERON\_0002771 ;  
3003 .  
3004 :pCL84  
3005 a skos:Concept ;  
3006 mtg\_cluster\_sc:id "pCL84" ;  
3007 skos:broader cl:CL\_0002605 ;  
3008 rdfs:label "MTG Astrocyte of the cerebral cortex" ;  
3009 go\_sc:part\_of uberon:UBERON\_0002771 ;  
3010 .  
3011 :pCL86  
3012 a skos:Concept ;  
3013 mtg\_cluster\_sc:id "pCL86" ;  
3014 skos:broader cl:CL\_0000128 ;  
3015 rdfs:label "MTG Oligodendrocyte" ;  
3016 go\_sc:part\_of uberon:UBERON\_0002771 ;  
3017 .  
3018 :pCL87  
3019 a skos:Concept ;  
3020 mtg\_cluster\_sc:id "pCL87" ;  
3021 skos:broader cl:CL\_1001602 ;  
3022 rdfs:label "cerebral cortex MTG endothelial cell" ;  
3023 go\_sc:part\_of uberon:UBERON\_0002771 ;  
3024 .  
3025 :pCL88  
3026 a skos:Concept ;  
3027 mtg\_cluster\_sc:id "pCL88" ;  
3028 skos:broader cl:CL\_0000129 ;

3029 rdfs:label "MTG Microglial cell" ;  
3030 go\_sc:part\_of uberon:UBERON\_0002771 ;  
3031 .  
3032 :pCL89  
3033 a skos:Concept ;  
3034 mtg\_cluster\_sc:id "pCL89" ;  
3035 skos:broader :pCL84 ;  
3036 rdfs:label "FGFR3-expressing MTG astrocyte" ;  
3037 go\_sc:part\_of uberon:UBERON\_0002771 ;  
3038 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_3690 ;  
3039 .  
3040 :pCL90  
3041 a skos:Concept ;  
3042 mtg\_cluster\_sc:id "pCL90" ;  
3043 skos:broader cl:CL\_0010011 ;  
3044 rdfs:label "cerebral cortex MTG GABAergic interneuron" ;  
3045 go\_sc:part\_of uberon:UBERON\_0002771 ;  
3046 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_4092 ;  
3047 mtg\_cluster\_sc:neuron\_type "GABAergic" ;  
3048 .  
3049 :pCL91  
3050 a skos:Concept ;  
3051 mtg\_cluster\_sc:id "pCL91" ;  
3052 skos:broader cl:CL\_0000679 ;  
3053 rdfs:label "MTG Glutamatergic neuron" ;  
3054 go\_sc:part\_of uberon:UBERON\_0002771 ;  
3055 mtg\_cluster\_sc:selectively\_expresses hugo:HGNC\_16704 ;  
3056 mtg\_cluster\_sc:neuron\_type "Glutamatergic" ;