



# Molecular characterization of selectively vulnerable neurons in Alzheimer's disease

Kun Leng<sup>1,2,3,4,13</sup>, Emmy Li<sup>1,2,3,13</sup>, Rana Eser<sup>5</sup>, Antonia Piergies<sup>5</sup>, Rene Sit<sup>2</sup>, Michelle Tan<sup>2</sup>, Norma Neff<sup>2</sup>, Song Hua Li<sup>5</sup>, Roberta Diehl Rodriguez<sup>6</sup>, Claudia Kimie Suemoto<sup>7,8</sup>, Renata Elaine Paraizo Leite<sup>7</sup>, Alexander J. Ehrenberg<sup>5</sup>, Carlos A. Pasqualucci<sup>7</sup>, William W. Seeley<sup>5,9</sup>, Salvatore Spina<sup>5</sup>, Helmut Heinsen<sup>7,10</sup>, Lea T. Grinberg<sup>5,7,11</sup>✉ and Martin Kampmann<sup>1,2,12</sup>✉

**Alzheimer's disease (AD) is characterized by the selective vulnerability of specific neuronal populations, the molecular signatures of which are largely unknown. To identify and characterize selectively vulnerable neuronal populations, we used single-nucleus RNA sequencing to profile the caudal entorhinal cortex and the superior frontal gyrus—brain regions where neurofibrillary inclusions and neuronal loss occur early and late in AD, respectively—from postmortem brains spanning the progression of AD-type tau neurofibrillary pathology. We identified RORB as a marker of selectively vulnerable excitatory neurons in the entorhinal cortex and subsequently validated their depletion and selective susceptibility to neurofibrillary inclusions during disease progression using quantitative neuropathological methods. We also discovered an astrocyte subpopulation, likely representing reactive astrocytes, characterized by decreased expression of genes involved in homeostatic functions. Our characterization of selectively vulnerable neurons in AD paves the way for future mechanistic studies of selective vulnerability and potential therapeutic strategies for enhancing neuronal resilience.**

Selective vulnerability is a fundamental feature of neurodegenerative diseases, in which different neuronal populations show a gradient of susceptibility to degeneration. Selective vulnerability at the network level has been extensively explored in AD<sup>1–3</sup>. However, little is known about the mechanisms underlying selective vulnerability at the cellular level in AD, which could provide insight into disease mechanisms and lead to therapeutic strategies.

The entorhinal cortex (EC) is one of the first cortical brain regions to exhibit neuronal loss in AD<sup>4</sup>. Neurons in the external EC layers, especially in layer II, accumulate tau-positive neurofibrillary inclusions and die early in the course of AD<sup>5–10</sup>. However, these selectively vulnerable neurons have yet to be characterized at the molecular level. Furthermore, it is unknown whether there are differences in vulnerability among subpopulations of these neurons. Although rodent models of AD have offered important insights, the available models fail to simultaneously capture some critical disease processes, such as the accumulation of neurofibrillary inclusions and neuronal loss<sup>11</sup>, limiting the extrapolation of findings from rodent models to address selective vulnerability.

Recently, single-nucleus RNA sequencing (snRNA-seq) has enabled large-scale characterization of transcriptomic profiles of individual cells from postmortem human brain tissue<sup>12,13</sup>. However, snRNA-seq studies of AD published to date have focused on cell-type-specific differential gene expression between individuals with AD and healthy controls<sup>14,15</sup>, without explicitly addressing selective vulnerability.

Here, we performed snRNA-seq on postmortem brain tissue from a cohort of individuals spanning the progression of AD-type tau neurofibrillary pathology to characterize changes in the relative abundance of cell types and cell-type subpopulations. We discovered a selectively vulnerable subpopulation of excitatory neurons in the EC and validated the selective depletion of this subpopulation during AD progression with quantitative histopathology, using multiplex immunofluorescence in EC regions delineated by rigorous cytoarchitectonic criteria. Furthermore, we uncovered an astrocyte subpopulation likely corresponding to reactive astrocytes that showed downregulation of genes involved in homeostatic function.

## Results

**Cohort selection and cross-sample alignment.** We performed snRNA-seq on cell nuclei extracted from postmortem brain tissue (Methods) from the EC at the level of the mid-uncus area and from the superior frontal gyrus (SFG) at the level of the anterior commissure (Brodman area 8), from ten male individuals carrying the apolipoprotein E gene (*APOE*)  $\epsilon 3/\epsilon 3$  genotype representing the cortical-free, early and late stages of AD-type tau neurofibrillary pathology (Braak stages<sup>1</sup> 0, 2 and 6; Fig. 1a and Table 1). The neuropathological hallmarks of AD are amyloid plaques and neurofibrillary inclusions. Since the accumulation of neurofibrillary inclusions measured by the Braak staging system is the best correlate of clinical cognitive decline, after neuronal loss<sup>16</sup>, we reasoned that profiling

<sup>1</sup>Institute for Neurodegenerative Disease, University of California, San Francisco, San Francisco, CA, USA. <sup>2</sup>Chan Zuckerberg Biohub, San Francisco, CA, USA. <sup>3</sup>Biomedical Sciences Graduate Program, University of California, San Francisco, San Francisco, CA, USA. <sup>4</sup>Medical Scientist Training Program, University of California, San Francisco, San Francisco, CA, USA. <sup>5</sup>Memory and Aging Center, Weill Institute for Neurosciences, Department of Neurology, University of California, San Francisco, San Francisco, CA, USA. <sup>6</sup>Department of Neurology, Universidade de São Paulo, Faculdade de Medicina, São Paulo, Brazil. <sup>7</sup>Department of Pathology, Universidade de São Paulo, Faculdade de Medicina, São Paulo, Brazil. <sup>8</sup>Division of Geriatrics, Department of Clinical Medicine, Universidade de São Paulo, Faculdade de Medicina, São Paulo, Brazil. <sup>9</sup>Department of Pathology, University of California, San Francisco, San Francisco, CA, USA. <sup>10</sup>Department of Psychiatry, University of Würzburg, Würzburg, Germany. <sup>11</sup>Global Brain Health Institute, University of California, San Francisco, San Francisco, CA, USA. <sup>12</sup>Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA, USA. <sup>13</sup>These authors contributed equally: Kun Leng, Emmy Li. ✉e-mail: [lea.grinberg@ucsf.edu](mailto:lea.grinberg@ucsf.edu); [martin.kampmann@ucsf.edu](mailto:martin.kampmann@ucsf.edu)



**Table 1 | Description of postmortem cohort**

Participants used for snRNA-seq								
Participant no.	Braak stage	Sex	Age at death (years)	Postmortem interval (h)	ADNC score	CDR before death	APOE genotype	Source
1	0	M	50	13	A0, B0, C0	0	ε3/ε3	BBAS
2	0	M	60	12	A0, B0, C0	0.5	ε3/ε3	BBAS
3	0	M	71	12	A1, B0, C0	0	ε3/ε3	BBAS
4	2	M	72	15	A1, B1, C0	0	ε3/ε3	BBAS
5*	2	M	77	4.9	A2, B1, C1	0.5	ε3/ε3	UCSF
6*	2	M	87	30	A2, B1, C2	2	ε3/ε3	UCSF
7*	2	M	91	50	A1, B1, C1	0	ε3/ε3	UCSF
8*	6	M	72	6.9	A3, B3, C3	3	ε3/ε3	UCSF
9*	6	M	82	6.7	A3, B3, C3	3	ε3/ε3	UCSF
10	6	M	82	9	A3, B3, C3	3	ε3/ε3	UCSF
Participants used for immunofluorescence validation								
Participant no.	Braak stage	Sex	Age at death (years)	Postmortem interval (h)	ADNC score	CDR before death	APOE genotype	Source
5*	2	M	77	4.9	A2, B1, C1	0.5	ε3/ε3	UCSF
6*	2	M	87	30	A2, B1, C2	2	ε3/ε3	UCSF
7*	2	M	91	50	A1, B1, C1	0	ε3/ε3	UCSF
8*	6	M	72	6.9	A3, B3, C3	3	ε3/ε3	UCSF
9*	6	M	82	6.7	A3, B3, C3	3	ε3/ε3	UCSF
11	0	F	62	10.1	A1, B0, C0	0	NA	BBAS
12	0	M	64	12	A0, B0, C0	0	ε3/ε3	BBAS
13	1	M	60	19	A0, B1, C0	0	NA	BBAS
14	1	F	64	13	A1, B1, C0	0	ε3/ε3	BBAS
15	1	M	70	11	A1, B1, C0	0	ε3/ε3	BBAS
16	1	F	82	9.6	A1, B1, C0	0	NA	BBAS
17	2	F	79	18	A1, B1, C1	0	ε3/ε3	BBAS
18	2	F	81	30.3	A1, B1, C0	NA	ε3/ε3	UCSF
19	3	M	81	8.3	A2, B2, C3	1	NA	UCSF
20	3	M	84	28	A3, B2, C2	1	NA	UCSF
21	3	F	88	9.8	A3, B2, C2	0.5	ε3/ε3	UCSF
22	3	M	89	9.1	A3, B2, C2	1	ε3/ε3	UCSF
23	4	F	87	9.5	A1, B2, C3	2	ε3/ε3	UCSF
24	4	M	91	11.2	A3, B2, C2	0.5	ε3/ε3	UCSF
25	4	M	103	7.8	A1, B2, C2	NA	ε3/ε3	UCSF
26	5	M	77	8.4	A3, B3, C3	0.5	ε4/ε4	UCSF
27	5	M	85	11.2	A3, B3, C3	1	ε3/ε3	UCSF
28	5	M	86	8.6	A3, B3, C3	2	ε3/ε4	UCSF
29	5	F	87	17	A3, B3, C2	3	ε3/ε3	BBAS
30	6	F	64	7.3	A3, B3, C3	3	ε3/ε4	UCSF
31	6	F	67	9.7	A3, B3, C3	3	ε4/ε4	UCSF

Asterisks denote participants' samples used both for snRNA-seq and immunofluorescence validation. The AD neuropathological change (ADNC) score incorporates assessment of amyloid-β deposits ('A'), staging of neurofibrillary tangles ('B') and scoring of neuritic plaques ('C')<sup>47</sup>. The Clinical Dementia Rating (CDR) reflects the degree of cognitive impairment<sup>48</sup>. BBAS, Brazilian BioBank for Aging Studies; F, female; M, male; NA, not applicable; UCSF, University of California, San Francisco.

matched EC and SFG samples across different Braak stages would allow us to isolate the effect of disease progression on cell types and cell-type subpopulations.

A challenge in characterizing the impact of disease progression on different cell-type subpopulations is that these subpopulations need to be defined in a way that is independent from the effect of disease progression. Typically, cell-type subpopulations are defined

by characterizing cells into subgroups of the same cell type through cluster analysis (that is, clustering), followed by examination of marker gene expression in the resulting clusters. To remove the effect of disease progression on clustering, we performed, before clustering, cross-sample alignment of the data from each brain region using scAlign (Methods). Importantly, after identifying clusters in the alignment space, we used the original data for subsequent

analyses involving examination of gene expression, such as identifying differentially expressed genes between clusters.

**Changes in broad cell-type composition with neuropathological Alzheimer's disease progression.** After quality control (Methods), we recovered 42,528 cells from the EC and 63,608 cells from the SFG. Examination of the average number of genes and unique molecular identifiers (UMIs) detected per cell showed similar or superior transcript coverage compared to previously published AD snRNA-seq datasets<sup>14,15</sup> (Extended Data Fig. 1a,b).

After cross-sample alignment, we performed clustering and recovered clusters that demonstrated spatial grouping in *t*-stochastic neighborhood embedding (*t*-SNE) largely uncorrelated with the individual of origin (Fig. 1b,c). Furthermore, the clusters showed specific expression of cell-type markers and grouped in a manner consistent with their expression of cell-type markers in hierarchical clustering (Fig. 1d,e and Methods). For comparison, we also performed clustering without cross-sample alignment, which resulted in many clusters that were defined by individual of origin in addition to cell type (Extended Data Fig. 1c–f). Having confirmed the effectiveness of cross-sample alignment in removing the effect of technical and experimental factors on clustering, we then assigned clusters to broad cell types (that is, excitatory neurons, inhibitory neurons, astrocytes, oligodendrocytes, oligodendrocyte precursor cells, microglia and endothelial cells) based on their expression of cell-type markers (Fig. 1d,e and Methods).

Next, to assess whether the proportions of broad cell types change with disease progression, we aggregated clusters assigned to the same cell type for each participant and then computed the relative abundance of each cell type in each individual. We tested for statistical significance using beta regression and corrected for multiple testing using Holm's method (Methods). While there were not many statistically significant changes in the relative abundance of cell types, we observed a downward trend in the relative abundance of EC excitatory neurons in Braak stages 2 ( $P_{\text{unadjusted}} = 0.18$ ) and 6 ( $P_{\text{unadjusted}} = 0.02$ ) and of SFG excitatory neurons only in Braak stage 6 ( $P_{\text{unadjusted}} = 0.05$ ), consistent with early involvement of the EC and sparing of the SFG until late Braak stages, and the previously described greater vulnerability of excitatory neurons relative to inhibitory neurons in AD<sup>17,18</sup>.

### Selective vulnerability of excitatory neuron subpopulations.

Based on these observations, we next asked whether specific subpopulations of excitatory neurons show a decline in their relative abundance with disease progression, by performing subclustering of excitatory neurons in the EC and SFG after cross-sample alignment (Methods). The EC, a relatively phylogenetically conserved brain structure in mammals, is among the first cortical fields to accumulate tau-positive neurofibrillary inclusions followed by neuronal loss in AD<sup>1</sup>. The EC is a heterogeneous structure and cytoarchitectonic considerations matter when analyzing and sampling this region to avoid biased observations<sup>19</sup>. During evolution, the position of the EC changed, and the mouse medial EC (the source of our layer-specific marker genes) is generally regarded as the equivalent of the caudal EC in humans (our sampling location)<sup>20,21</sup>.

In the EC, we discerned nine excitatory neuron subpopulations (Fig. 2a–d). These subpopulations exhibited distinct expression of EC layer-specific genes identified in the mouse medial EC<sup>22</sup>. Notably, subpopulation EC:Exc.s2 showed a striking ~50% decrease in its relative abundance in Braak stage 2 compared to Braak stage 0, with no further decrease in Braak stage 6 (Fig. 2c), suggesting depletion early in disease. EC:Exc.s1 and EC:Exc.s4 similarly exhibited a ~50–60% reduction in their relative abundance in Braak stage 2. EC:Exc.s1, EC:Exc.s2 and EC:Exc.s4 expressed genes associated with mouse EC layer II (Fig. 2c), consistent with the fact that tau neurofibrillary inclusions are known to accumulate preferentially in

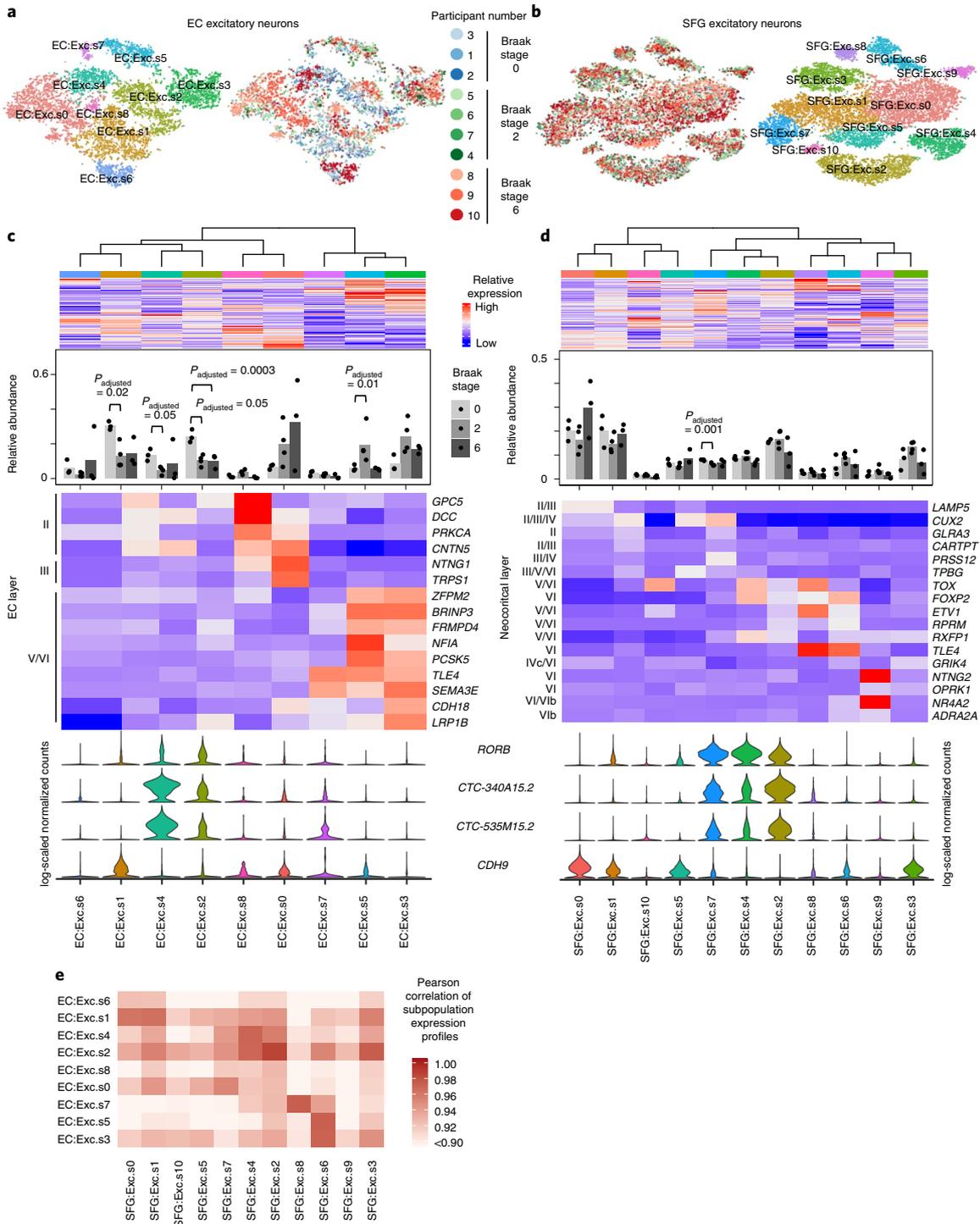
human EC layer II early in AD<sup>5–8</sup>. However, not all subpopulations expressing genes associated with mouse EC layer II showed similar levels of early vulnerability. For example, EC:Exc.s6 and EC:Exc.s8 did not demonstrate statistically significant changes in their relative abundance across disease progression. We failed to find evidence of selective vulnerability in neuronal subpopulations expressing genes associated with mouse EC layer III (EC:Exc.s0) or V/VI (EC:Exc.7, EC:Exc.s5 and EC:Exc.s3). In fact, EC:Exc.s5 exhibited a statistically significant increase in its relative abundance in Braak stage 2. Since neurons are postmitotic, this increase is likely due to the selective earlier depletion of more vulnerable excitatory neuron subpopulations, followed by later depletion of EC:Exc.s5.

To identify molecular markers of selectively vulnerable excitatory neuron subpopulations in the EC (EC:Exc.s2, EC:Exc.s4 and EC:Exc.s1), we inspected transcript levels of genes differentially expressed between pairs of subpopulations and curated a set of genes that were specifically expressed by no more than four subpopulations (Extended Data Fig. 2a), which we decided was a reasonable threshold for a positive marker to be useful. We found that EC:Exc.s2 and EC:Exc.s4 specifically expressed *RORB*, *CTC-340A15.2* and *CTC-535M15.2* (Fig. 2c). *RORB* (RAR-related Orphan Receptor B) encodes a transcription factor known as a marker and developmental driver of layer IV neurons in the neocortex<sup>23–25</sup>, but is also expressed by neurons in other layers<sup>13</sup>. Little is known about the noncoding transcripts *CTC-340A15.2* and *CTC-535M15.2* in the context of neuronal identity and function. We also found that EC:Exc.s1 was marked by high expression of *CDH9* (Fig. 2c), which encodes a cadherin with neuron-specific expression. However, *CDH9* was also expressed by other excitatory neuron subpopulations in the EC, and we could not find markers that were specifically expressed only in EC:Exc.s1. Therefore, we chose to focus our analysis on EC:Exc.s2 and EC:Exc.s4.

In addition to identifying molecular markers of the selectively vulnerable EC:Exc.s2 and EC:Exc.s4 neurons, we also enumerated genes that were differentially expressed in EC:Exc.s2 and EC:Exc.s4 compared to all other excitatory neurons in the EC, controlling for differences across participants (Methods). We found that genes with higher expression in EC:Exc.s2 and EC:Exc.s4 were enriched for axon-localized proteins and voltage-gated potassium channels, whereas genes with lower expression in EC:Exc.s2 and EC:Exc.s4 were enriched for synapse- and dendrite-localized proteins and pathways involving G-protein-mediated signaling, ion transport and neurotransmitter receptor signaling (Extended Data Fig. 2b–e and Supplementary Table 1).

We also performed differential gene expression analysis across Braak stages for EC excitatory neuron subpopulations (Methods), comparing Braak stages 6 and 0, which yielded the largest number of differentially expressed genes. We found a broad decrease in expression of genes encoding pre- and postsynaptic proteins in Braak stage 6 compared to stage 0 for many EC excitatory neuron subpopulations (Extended Data Fig. 3b,d,f and Supplementary Table 2). Furthermore, we observed that the selectively vulnerable subpopulation EC:Exc.s2 had the largest number of downwardly differentially expressed genes and the strongest enrichments for pre- and postsynaptic proteins in these genes (Extended Data Fig. 3b,d). Overall, the downregulation of synapse-related genes that we have observed mirrors the findings from a recent study by Marinaro et al.<sup>26</sup>, which examined the frontal cortex in familial monogenic AD using snRNA-seq, and it is consistent with a previous study of gene expression changes in AD in the EC and other brain regions using laser capture microdissection of neurons followed by DNA microarray analysis<sup>27</sup>.

We next examined excitatory neuron subpopulations in the SFG. Similar to previous studies<sup>12,13</sup>, we found that excitatory neuron subpopulations in the SFG (11 in total) expressed distinct sets of neocortical layer-specific genes (Fig. 2b,d), recapitulating the laminar



**Fig. 2 | RORB-expressing excitatory neuron subpopulations in the EC are selectively vulnerable. a, b**, t-SNE projection of excitatory neurons from the EC (**a**) and SFG (**b**) in their respective alignment spaces, colored by individual of origin (center) or subpopulation identity (outer). **c, d**, Heat map and hierarchical clustering of subpopulations and subpopulation marker expression (top); 'high' and 'low' relative expression reflect above- and below-average expression, respectively (Methods). Relative abundance of subpopulations across Braak stages (second from top); for each brain region, statistical significance of differences in relative abundance across Braak stages (Braak 0:  $n=3$ , Braak 2:  $n=4$ , Braak 6:  $n=3$ ;  $n$ , number of participants sampled) was determined by beta regression and adjusted for multiple comparisons (Methods). Expression heat map (lower) of EC layer-specific genes identified from Ramsden et al.<sup>22</sup> (**c**). Expression heat map (lower) of neocortical layer-specific genes from Lake et al.<sup>12</sup> (**d**). Expression of selectively vulnerable subpopulation markers identified in the EC (bottom). **e**, Heat map of Pearson correlation between the gene expression profiles of EC and SFG subpopulations.

organization of the neocortex. Interestingly, SFG:Exc.s4 and SFG:Exc.s2, which were marked by the vulnerability markers that we identified in the EC (*RORB*, *CTC-340A15.2* and *CTC-535M15.2*),

trended towards decreased relative abundance only in Braak stage 6 (Fig. 2d), consistent with the late appearance of neurofibrillary inclusions in the SFG starting at Braak stage 5. Although SFG:Exc.

s7, which also expressed the EC vulnerability markers, exhibited a statistically significant decrease in relative abundance in Braak stage 2 but not 6, the magnitude of change was negligibly small.

Given that SFG:Exc.s4 and SFG:Exc.s2 expressed similar markers to EC:Exc.s4 and EC:Exc.s2, we wondered if SFG:Exc.s4 and SFG:Exc.s2 may resemble EC:Exc.s4 and EC:Exc.s2 more broadly at the transcriptome level. SFG:Exc.s4 and SFG:Exc.s2 were indeed most similar to EC:Exc.s4 and EC:Exc.s2 based on the Pearson correlation coefficient between the expression profiles of SFG and EC subpopulations (Fig. 2e). We observed the same pattern when we mapped subpopulations in the EC to those in the SFG by performing cross-sample alignment for both brain regions jointly (Extended Data Fig. 4). This similarity is consistent with the reported similarity between deep-layer neocortical excitatory neurons and EC excitatory neurons in general<sup>28</sup>. The similarity in transcriptomes of vulnerable excitatory neurons in different brain regions is intriguing and suggests similar mechanisms of selective vulnerability in different brain regions.

Although the decrease in the relative abundance of SFG:Exc.s2 and SFG:Exc.s4 in Braak stage 6 was not statistically significant after correction for multiple testing, we asked whether we could detect signs of selective vulnerability in neocortical *RORB*-expressing excitatory neurons in an independent dataset with a larger sample size. To this end, we reanalyzed data from Mathys et al.<sup>14</sup>, which profiled the prefrontal cortex from 24 participants with AD and 24 healthy controls, with our cross-sample alignment pipeline and performed subclustering of excitatory neurons. In the Mathys et al. dataset<sup>14</sup>, we discerned ten excitatory neuron subpopulations, each of which expressed distinct sets of neocortical layer-specific genes (Extended Data Fig. 5a,b), similar to Lake et al.<sup>12</sup> and our dataset. Of these ten subpopulations, Mathys:Exc.s4, Mathys:Exc.s5 and Mathys:Exc.s1 expressed *RORB* at high levels (*CTC-340A15.2* and *CTC-535M15.2* were not available in the preprocessed Mathys et al.<sup>14</sup> data). Importantly, we observed a statistically significant decrease in the relative abundance of Mathys:Exc.s4 in male participants with AD versus controls (Extended Data Fig. 5b), recapitulating the selective vulnerability observed in our dataset, which consisted only of male participants. Furthermore, gene expression correlation analysis showed that Mathys:Exc.s4 was the most similar to EC:Exc.s2 and EC:Exc.s4 (Extended Data Fig. 5c), again demonstrating similarity between selectively vulnerable excitatory neurons in the neocortex and those in the EC.

Although we did not detect any statistically significant changes in the relative abundance of *RORB*-expressing subpopulations in female participants in the Mathys et al. dataset<sup>14</sup>, Mathys:Exc.s1 trended towards decreased relative abundance in female participants with AD ( $P_{\text{unadjusted}} = 0.17$ ) and mapped to EC:Exc.s2 by gene

expression correlation (Extended Data Fig. 5b,c). Furthermore, the Marinaro et al. dataset<sup>26</sup> included both male and females with monogenic AD and also reported the selective vulnerability of two out of four *RORB*-expressing excitatory neuron subpopulations in the prefrontal cortex (ExcB1 and ExcB4)<sup>26</sup>, providing further evidence that subsets of *RORB*-expressing excitatory neurons in the neocortex are selectively vulnerable.

Considering the Mathys et al.<sup>14</sup> and the Marinaro et al.<sup>26</sup> datasets together with our dataset, it appears that while not all *RORB*-expressing excitatory neuron subpopulations in the neocortex showed signs of selective vulnerability, those that did were the most similar to *RORB*-expressing excitatory neurons in the EC, all of which showed signs of selective vulnerability.

**Validation of the selective vulnerability of *RORB*-expressing excitatory neurons.** To validate our finding that *RORB*-expressing excitatory neurons in the EC were highly vulnerable in AD, we performed multiplex immunofluorescence on postmortem brains of 26 participants with AD spanning Braak stages 0 to 6, who were devoid of non-AD neuropathological changes (Table 1). Specifically, we quantified the proportion of excitatory neurons and *RORB*-positive excitatory neurons in the EC superficial layers (that is, above layer IV; referred to as dissecans-1 (ref. <sup>19</sup>) in Fig. 3b). Given EC heterogeneity, we used rigorous cytoarchitectonic parameters to delineate the caudal EC and minimize artifactual results (Fig. 3a–c, Extended Data Fig. 6 and Methods). We used multiplex immunofluorescence<sup>29</sup> to label nuclei (DAPI), excitatory neurons (TBR1), *RORB*<sup>+</sup> neurons and phospho-tau neuronal inclusions (CP13, which detects tau phosphorylated at Ser 202). We observed a substantial reduction in the proportion of *RORB*<sup>+</sup> neurons among excitatory neurons in Braak stages 2–4 and 5–6 compared to Braak stages 0–1 (Fig. 3d,e). Furthermore, by analyzing a subset of individuals with AD, we detected phospho-tau (CP13) preferentially in *RORB*<sup>+</sup> compared to *RORB*<sup>-</sup> excitatory neurons (Fig. 3f,g). Thus, our results substantiate that *RORB*-expressing excitatory neurons are highly vulnerable in AD and support a model in which their depletion is a consequence of accumulating tau neurofibrillary changes.

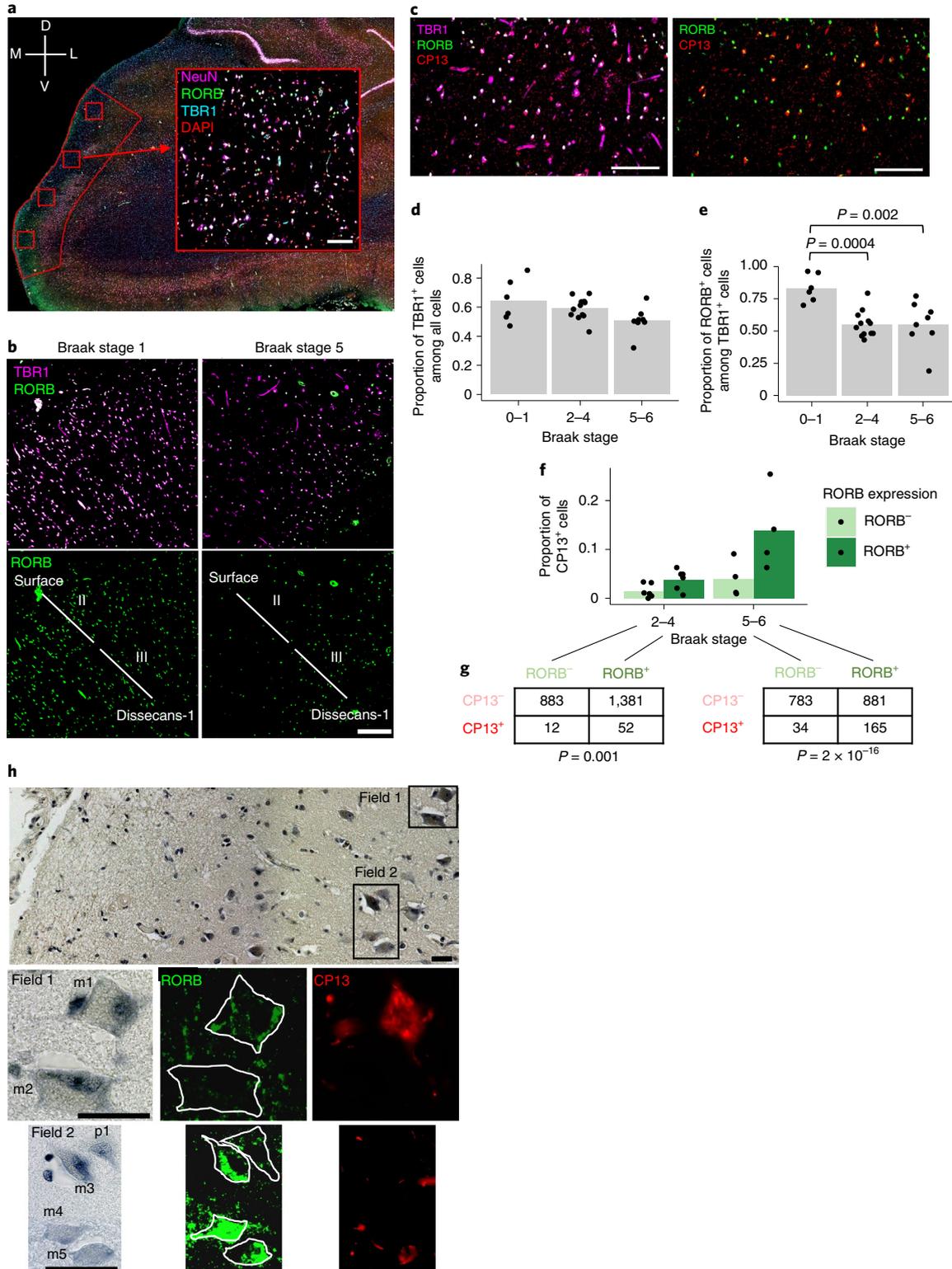
Given that large multipolar neurons of ‘stellate’ morphology in EC layer II are particularly vulnerable in AD<sup>5–8</sup>, we examined the morphological features of *RORB*<sup>+</sup> excitatory neurons in layer II by overlaying immunofluorescence with Nissl staining. We found that *RORB*<sup>+</sup> excitatory neurons adopted various shapes, including pyramidal and multipolar morphologies (Fig. 3h). Conversely, some large multipolar neurons were *RORB*<sup>-</sup> (Fig. 3h). Our results are consistent with the known vulnerability of large multipolar EC layer II neurons and demonstrate that molecular characterization of vulnerable neurons refines the results of morphological studies.

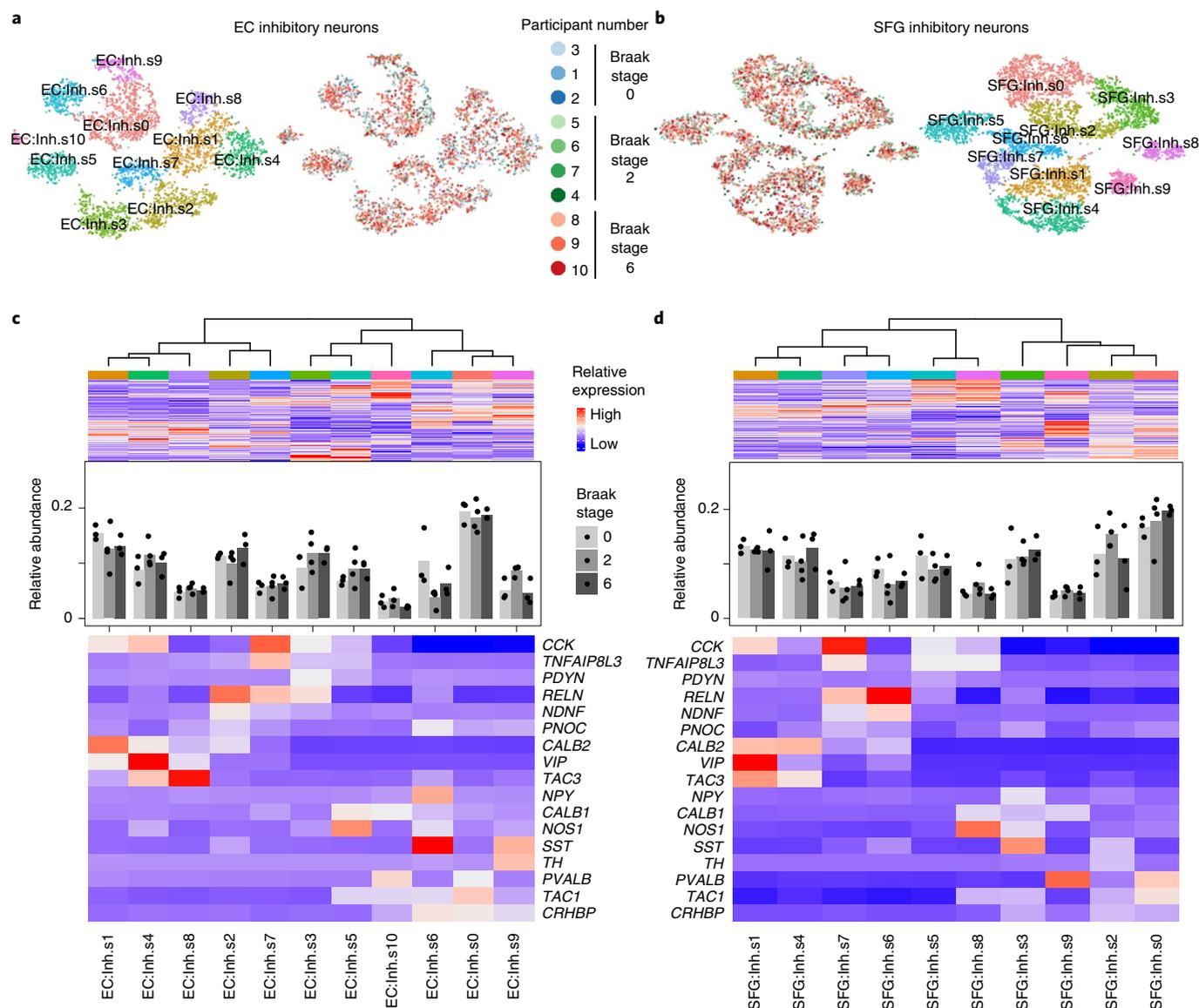
**Fig. 3 | Immunofluorescence of the EC validates selective vulnerability of *RORB*-expressing excitatory neurons.** **a**, The method for extracting regions of interest (ROIs) is illustrated with a representative brain slice used for immunofluorescence (pseudo-colored: DAPI in orange, *RORB* in green, TBR1 in blue and NeuN in pink) with the EC delineated in red. Four ROIs (red squares) were randomly distributed along the superficial layers of the EC and extracted for quantification after masking neurons (Methods). The insert shows a representative ROI image (with different pseudo-coloring). The anatomical orientation of the slice is provided in the top left corner (D, dorsal; V, ventral; M, medial; L, lateral). **b**, Representative *RORB* staining in a Braak stage 1 sample (left) versus a Braak stage 5 sample (right), shown with (top) and without (bottom) excitatory neurons marked by TBR1 staining. The EC layers captured in the image are demarcated in the lower images (Methods and Extended Data Fig. 6). **c**, Representative CP13 staining in a Braak stage 6 sample, shown together with TBR1 and *RORB* staining (left) or only with *RORB* staining (right). **d,e**, Proportion of TBR1<sup>+</sup> cells among all cells (**d**) or proportion of *RORB*<sup>+</sup> cells among TBR1<sup>+</sup> cells (**e**) averaged across ROIs for each individual across groups of Braak stages; statistical significance of differences in the above proportions across groups of Braak stages (Braak 0–1:  $n = 6$ , Braak 2–4:  $n = 12$ , Braak 5–6:  $n = 8$ ;  $n$ , number of participants sampled) was determined by beta regression without adjustment for multiple comparisons. **f**, Proportion of CP13<sup>+</sup> cells in *RORB*<sup>-</sup> or *RORB*<sup>+</sup> excitatory neurons (that is, TBR1<sup>+</sup> cells) averaged across ROIs for each individual across groups of Braak stages. **g**, Contingency tables of raw counts of TBR1<sup>+</sup> cells based on their *RORB* or CP13 staining status summed across ROIs and individuals for each group of Braak stages (Braak 2–4:  $n = 6$ , Braak 5–6:  $n = 4$  participants sampled); Fisher’s exact test  $P$  value (two-sided). **h**, Representative image of EC layer II neurons stained with gallocyanin (top) with the corresponding *RORB* and CP13 immunofluorescence signal shown in selected fields (field 1: middle; field 2: bottom). *RORB*<sup>+</sup> neurons include both large multipolar neurons (m1, m3, m4 and m5) and pyramidal neurons (p1). One large multipolar neuron (m2) is *RORB*<sup>-</sup>. The neuronal somas are outlined manually in white in the *RORB* immunofluorescence images. Scale bars: 100  $\mu\text{m}$  (**a–c**); 15  $\mu\text{m}$  (**h**). For all data, the experiment was performed once.

**Lack of differences in vulnerability of inhibitory neuron subpopulations.** We next examined inhibitory neurons, which are more resistant to tau pathology than excitatory neurons in AD<sup>17,18</sup>. In both brain regions, inhibitory neuron subpopulations expressed distinct sets of inhibitory neuron subtype markers (Fig. 4a–d), consistent with previous studies<sup>12,13</sup>. We did not detect statistically significant changes in the relative abundance of inhibitory neurons subpopulations in the EC or SFG (Fig. 4c,d), or in the prefrontal cortex in the Mathys et al.

dataset<sup>14</sup> (Extended Data Fig. 7). Although Marinaro et al. reported broad depletion of inhibitory neuron subpopulations in familial monogenic AD, there was no strong evidence of selective vulnerability in particular inhibitory neuron subpopulations relative to other inhibitory neuron subpopulations in work by Marinaro et al.<sup>26</sup>

**Analysis of glial subpopulations.** Glial cells have emerged as important players in AD. We found a trend towards increased





**Fig. 4 | Inhibitory neuron subpopulations do not consistently show differences in resilience or vulnerability to AD progression. a, b,** *t*-SNE projection of inhibitory neurons from the EC (**a**) and SFG (**b**) in their respective alignment spaces, colored by individual of origin (center) or subpopulation identity (outer). **c, d,** Heat map and hierarchical clustering of subpopulations and subpopulation marker expression (top); 'high' and 'low' relative expression reflect above- and below-average expression, respectively (Methods). Relative abundance of subpopulations across Braak stages (Braak 0:  $n=3$ , Braak 2:  $n=4$ , Braak 6:  $n=3$ ;  $n$ , number of participants sampled) was determined by beta regression and adjusted for multiple comparisons (Methods). Expression heat map of inhibitory neuron molecular subtype markers from Lake et al.<sup>12</sup> (bottom).

relative abundance of microglia in the EC in with AD progression (Fig. 1f), consistent with microgliosis. Next, we asked whether a specific transcriptional state of microglia was associated with AD in our dataset. Recent single-cell profiling of microglia from mouse models of AD identified disease-associated microglia<sup>30</sup> (DAM), the transcriptional signature of which overlaps only partially with that of human microglia found in AD<sup>31</sup>. Considering the possibility that DAM may cluster separately from homeostatic microglia after cross-sample alignment, we performed subclustering of microglia in our dataset, discerning four subpopulations in the EC and five subpopulations in the SFG (Extended Data Fig. 8a,b). However, similar to Thrupp et al.<sup>32</sup>, we were unable to detect the expression of the majority of homeostatic microglia markers and DAM markers in our dataset or in that of Mathys et al.<sup>14</sup> (Extended Data Fig. 8d–f), which may be due to the relatively low number

of genes captured in microglia compared to other cell types (Fig. 1h,i) and the depletion of many DAM markers in nuclei compared to whole cells<sup>32</sup>.

We next examined oligodendrocytes, which were shown by Mathys et al.<sup>14</sup> to exhibit a strong transcriptional response in AD. Subclustering of oligodendrocytes in the EC and SFG revealed subpopulations (EC:Oligo.s0 and EC:Oligo.s4; SFG:Oligo.s1 and SFG:Oligo.s2) that exhibited higher expression of AD-associated oligodendrocyte genes from the Mathys et al. dataset<sup>14</sup>, that is, genes with higher expression in the AD-associated subpopulation Oli0 in work by Mathys et al.<sup>14</sup> (Extended Data Fig. 9d,e). Although the function of these genes in the context of AD is largely unknown, a spatial transcriptomics study of AD<sup>33</sup> has recently implicated a subset of these genes in the response of oligodendrocytes to amyloid plaques (for example, *CRYAB* and *QDPR*).

Finally, we turned our attention to astrocytes. While reactive astrocytes are ubiquitously associated with AD pathology<sup>34</sup>, only few studies to date have directly profiled reactive astrocytes due to the difficulty of specifically isolating reactive astrocytes<sup>35,36</sup>. Similarly to our interrogation of microglia, we asked if reactive astrocytes would cluster separately from nonreactive astrocytes following cross-sample alignment. After subclustering of astrocytes in our dataset, we discerned four subpopulations in the EC and six subpopulations in the SFG (Fig. 5a–d). In each brain region, at least one subpopulation (EC: Astro.3, SFG: Astro.s4 and SFG: Astro.s5) expressed dramatically higher levels of *GFAP*, which we refer to as *GFAP*<sub>high</sub> astrocytes (Fig. 5c,d). In the EC, *GFAP*<sub>high</sub> astrocytes also expressed *CD44* and *HSPB1*, markers of pan-reactive astrocytes<sup>37</sup>; *TNC*, which is upregulated in stab-wound reactive astrocytes<sup>38</sup>; and *HSP90AA1*, which is upregulated in reactive astrocytes associated with middle cerebral artery occlusion<sup>39</sup> (Fig. 5c,d). Interestingly, in the SFG, *GFAP*<sub>high</sub> astrocytes consisted of two subpopulations, one marked by higher expression of *CD44* and *TNC*, both of which are involved in interactions with the extracellular matrix, and the other marked by higher expression of *HSPB1* and *HSP90AA1*, both of which are chaperones involved in proteostasis. For downregulated genes, *GFAP*<sub>high</sub> astrocytes consistently expressed lower levels of genes associated with glutamate/GABA homeostasis (*SLC1A2*, *SLC1A3*, *GLUL* and *SLC6A11*) and synaptic adhesion/maintenance (*NRXN1*, *CADM2*, *PTN* and *GPC5*), suggesting a loss of homeostatic function.

Examination of all differentially expressed genes in *GFAP*<sub>high</sub> astrocytes showed significant overlap with differentially expressed genes from reactive astrocytes in a mouse model of spinal cord injury<sup>40</sup> (Fig. 5e). Overlapping downregulated genes included the previously noted genes associated with glutamate homeostasis and synaptic adhesion/maintenance and also genes related to lipid metabolism, cytoskeleton and extracellular matrix, and transporters (Fig. 5f,g).

Finally, to confirm the presence of *GFAP*<sub>high</sub> astrocytes in an independent dataset, we performed subclustering of astrocytes from the Mathys et al. dataset<sup>14</sup> after cross-sample alignment, which yielded three subpopulations (Extended Data Fig. 10a,b). Indeed, we found that Mathys: Astro.s2 behaved identically compared to *GFAP*<sub>high</sub> astrocytes in our dataset in upregulating reactive astrocyte markers and downregulating genes associated with glutamate/GABA homeostasis and synaptic adhesion (Extended Data Fig. 10b). Furthermore, the differentially expressed genes in Mathys: Astro.s3 overlapped highly with those in *GFAP*<sub>high</sub> astrocytes in our dataset (Extended Data Fig. 10c).

## Discussion

Selective vulnerability is a fundamental feature of neurodegenerative diseases, including AD. Past studies have characterized the most vulnerable neurons in AD based on topography and morphology. For instance, EC layer II neurons are more vulnerable than EC

layer III pyramidal neurons<sup>8–10</sup>. However, the molecular signature of selectively vulnerable neurons in AD is largely unknown.

Using a combination of snRNA-seq and quantitative neuropathology in postmortem human brains, we discovered that in the caudal EC, specific excitatory neuron subpopulations defined by snRNA-seq were selectively vulnerable in AD. These neurons expressed genes associated with layer II of the mouse medial EC, consistent with the known vulnerability of neurons in superficial layers of the human EC in AD<sup>5–8</sup>. We identified and validated *RORB* as a marker of these selectively vulnerable excitatory neuron subpopulations. Selectively vulnerable *RORB*<sup>+</sup> excitatory neurons included both large multipolar neurons and pyramidal neurons. Our findings demonstrate that morphology alone is insufficient to determine selective vulnerability.

We also found that tau neuronal inclusions, a chief neuropathological hallmark in AD, preferentially accumulated in *RORB*<sup>+</sup> excitatory neurons in the EC. To uncover potential cell biological mechanisms underlying the vulnerability of EC *RORB*<sup>+</sup> excitatory neurons, we compared the gene expression profiles of EC *RORB*-expressing excitatory neurons against all other EC excitatory neurons, which revealed differences in the expression of genes encoding synapse- versus axon-localized proteins, potassium channel subunits, G-protein signaling molecules and neurotransmitter receptor signaling molecules. Future studies utilizing in vitro and animal models of AD together with techniques for manipulating gene expression such as CRISPR inhibition and activation<sup>41–43</sup> will make it possible to address these potential mechanistic connections among *RORB* expression, phospho-tau accumulation, and vulnerability.

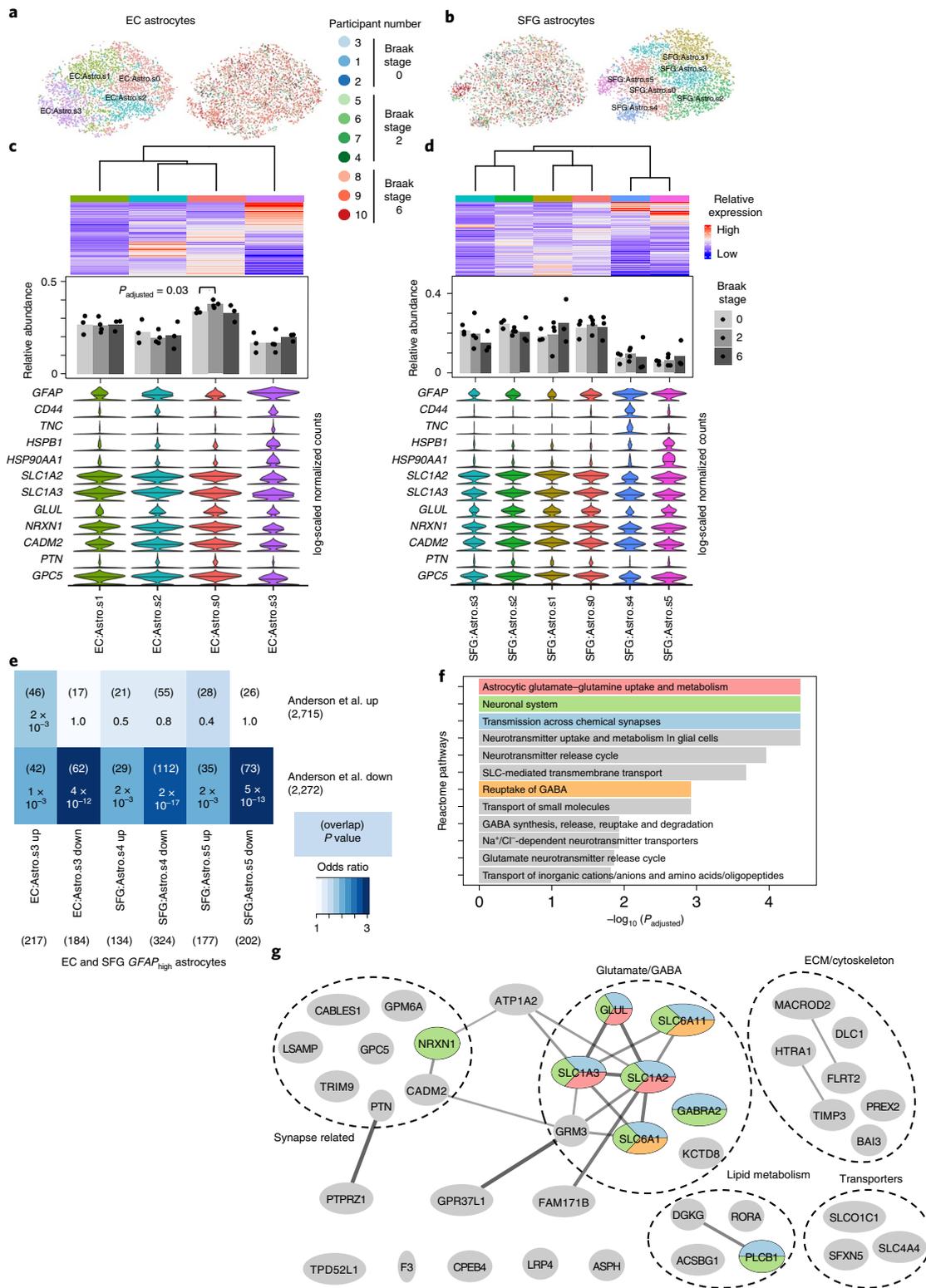
In neocortical areas, layers III and V are the first to accumulate tau neurofibrillary inclusions in AD<sup>1,44,45</sup>. Our dataset, together with our reanalysis of datasets from Mathys et al.<sup>14</sup> and Marinaro et al.<sup>26</sup>, suggests that in the neocortex, vulnerable excitatory neuron subpopulations express *RORB* and have a similar transcriptional profile as selectively vulnerable neurons in the EC, although not all neocortical *RORB*-expressing neurons are vulnerable. Given that *RORB* is known to function as a developmental driver of neuronal subtype identity in the neocortex<sup>23–25</sup>, we hypothesize that the vulnerability of *RORB*-expressing excitatory neuron subpopulations in different brain regions may be caused by gene expression programs driven by *RORB* and potentially other subtype-determining transcription factors. Further mechanistic studies involving the perturbation of *RORB* expression in cell-based or animal models of AD are necessary to test this hypothesis.

A previous study suggested changes in the number of neurons expressing calbindin and parvalbumin, which tend to mark inhibitory neurons, in EC layer II in AD<sup>46</sup>. Here, we found no evidence of selective vulnerability in inhibitory neurons subpopulations in EC layer II or any other layer. Inhibitory neurons in EC superficial layers show a gradient of abundance in the various EC regions<sup>20</sup>, which could confound the results. However, given that we used

**Fig. 5 | *GFAP*<sub>high</sub> astrocytes show signs of dysfunction in glutamate homeostasis and synaptic support. a,b**, t-SNE projection of astrocytes from the EC (a) and SFG (b) in their respective alignment spaces, colored by individual of origin (center) or subpopulation identity (outer). **c,d**, Heat map and hierarchical clustering of subpopulations and subpopulation marker expression (top; 'high' and 'low' relative expression reflect above- and below-average expression, respectively; Methods). Relative abundance of subpopulations across Braak stages (middle); for each brain region, statistical significance of differences in relative abundance across Braak stages (Braak 0:  $n = 3$ , Braak 2:  $n = 4$ , Braak 6:  $n = 3$ ;  $n$ , number of participants sampled) was determined by beta regression and adjusted for multiple comparisons (Methods). Expression of genes associated with reactive astrocytes, with the median expression level marked by a solid line (bottom). **e**, Enrichment analysis of overlap between differentially expressed genes in *GFAP*<sub>high</sub> astrocytes versus differentially expressed genes in reactive astrocytes from Anderson et al.<sup>40</sup> The number of genes in each gene set and the number of overlapping genes are shown in parentheses; the hypergeometric test  $P$  values (one-sided, corrected for multiple testing using the Benjamini–Hochberg procedure) are shown without parentheses. **f**, Enrichment of Reactome pathways in downregulated genes in *GFAP*<sub>high</sub> astrocytes, with selected terms highlighted in color. **g**, Functional association network (Methods) of downregulated genes shared between EC and SFG *GFAP*<sub>high</sub> astrocytes that overlap with those from the Anderson et al. data<sup>40</sup>. Genes with stronger associations are connected by thicker lines. Genes that belong to selected gene sets in **f** are highlighted in color. ECM, extracellular matrix.

strict cytoarchitectonic criteria to sample the EC, it is unlikely that our results reflect comparisons of different EC areas across the participants with AD. Evidence suggests that these inhibitory neurons undergo changes in morphology and function rather than loss in sporadic AD<sup>46</sup>. Thus, our results do not preclude the possibility that inhibitory neuron subpopulations may be differentially affected by AD progression at the morphological and functional level, even if neuronal loss is not apparent.

Accumulating evidence highlights the importance of glial changes in AD. We discovered an astrocyte subpopulation expressing high levels of *GFAP*, which we termed *GFAP*<sub>high</sub> astrocytes, in both the EC and SFG, and in the prefrontal cortex from work by Mathys et al.<sup>14</sup> We found that *GFAP*<sub>high</sub> astrocytes expressed higher levels of other genes associated with reactive astrocytes and lower levels of genes involved in astrocyte homeostatic functions. Furthermore, we found a high degree of overlap between genes



differentially expressed in *GFAP*<sub>high</sub> astrocytes and genes differentially expressed in reactive astrocytes from a mouse model of spinal injury<sup>40</sup>. Thus, we believe that *GFAP*<sub>high</sub> astrocytes correspond to reactive astrocytes in AD, which may have compromised homeostatic function.

Our study has several methodological strengths. First, the post-mortem cohort used for snRNA-seq and histopathological validation consists of well-characterized individuals with AD, devoid of non-AD pathology. To minimize confounders in the snRNA-seq results, we selected only male participants with an *APOE* ε3/ε3 genotype. Second, we sequenced a large number of nuclei from each individual (~10,000 nuclei per individual, compared to ~1,700 nuclei per individual in the Mathys et al. dataset<sup>14</sup>) from two brain regions per individual (~4,000 nuclei from the EC and ~6,000 nuclei from the SFG per individual). Third, we used strict cytoarchitectonic criteria to sample brain regions for snRNA-seq and histopathological validation, instead of broadly defined sampling areas used by previous studies. Fourth, our focus was on defining cell-type subpopulations that showed changes in relative abundance between disease stages, which can reflect important disease processes such as neuronal loss, and to define the genes characteristic of these subpopulations. By defining cell-type subpopulations independently of disease progression, we could compare gene expression between subpopulations within participants while controlling for differences among participants; this is more robust than comparing gene expression in a given subpopulation across groups of individuals, which can be influenced by differences in confounding factors between the groups. Lastly, by validating our findings using a new multiplex immunofluorescence approach<sup>29</sup>, we could quantify the relative abundance of excitatory neurons and RORB<sup>+</sup> neurons and also demonstrate that RORB<sup>+</sup> excitatory neurons were preferentially affected by neurofibrillary inclusions.

A limitation of our study is that we only included male participants carrying the *APOE* ε3/ε3 genotype in the snRNA-seq analysis. We included females and individuals carrying the *APOE* ε4 allele associated with AD risk in our histopathological validation, but caution should be taken before generalizing our results to these groups. Future studies will provide a systematic analysis of the impact of sex and *APOE* status on selective vulnerability in AD.

In conclusion, our study contributes a pioneering characterization of selectively vulnerable neuronal populations in AD using snRNA-seq profiling of paired brain regions from the same participants, which were all carefully curated participants with AD and controls. These results will inform future studies of the mechanistic basis of selective vulnerability in both animal and in vitro models, such as human induced pluripotent stem cells-derived neurons, in which the deployment of CRISPR inhibition and activation technology enables elucidation of the functional consequences of transcriptional changes<sup>41,43</sup>.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41593-020-00764-7>.

Received: 4 April 2020; Accepted: 20 November 2020;  
Published online: 11 January 2021

### References

- Braak, H. & Braak, E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* **82**, 239–259 (1991).
- Scholl, M. et al. PET imaging of tau deposition in the aging human brain. *Neuron* **89**, 971–982 (2016).
- Seeley, W. W., Crawford, R. K., Zhou, J., Miller, B. L. & Greicius, M. D. Neurodegenerative diseases target large-scale human brain networks. *Neuron* **62**, 42–52 (2009).
- Braak, H. & Braak, E. Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiol. Aging* **16**, 271–278 (1995).
- Price, J. L. et al. Neuron number in the entorhinal cortex and CA1 in preclinical Alzheimer disease. *Arch. Neurol.* **58**, 1395–1402 (2001).
- Stranahan, A. M. & Mattson, M. P. Selective vulnerability of neurons in layer II of the entorhinal cortex during aging and Alzheimer's disease. *Neural Plast.* **2010**, 108190 (2010).
- Van Hoesen, G. W., Hyman, B. T. & Damasio, A. R. Entorhinal cortex pathology in Alzheimer's disease. *Hippocampus* **1**, 1–8 (1991).
- Gomez-Isla, T. et al. Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. *J. Neurosci.* **16**, 4491–4500 (1996).
- Braak, H. & Braak, E. The human entorhinal cortex: normal morphology and lamina-specific pathology in various diseases. *Neurosci. Res.* **15**, 6–31 (1992).
- Kordower, J. H. et al. Loss and atrophy of layer II entorhinal cortex neurons in elderly people with mild cognitive impairment. *Ann. Neurol.* **49**, 202–213 (2001).
- Drummond, E. & Wisniewski, T. Alzheimer's disease: experimental models and reality. *Acta Neuropathol.* **133**, 155–175 (2017).
- Lake, B. B. et al. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586–1590 (2016).
- Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
- Mathys, H. et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019).
- Grubman, A. et al. A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. *Nat. Neurosci.* **22**, 2087–2097 (2019).
- Nelson, P. T. et al. Correlation of Alzheimer disease neuropathologic changes with cognitive status: a review of the literature. *J. Neuropathol. Exp. Neurol.* **71**, 362–381 (2012).
- Hof, P. R. et al. Parvalbumin-immunoreactive neurons in the neocortex are resistant to degeneration in Alzheimer's disease. *J. Neuropathol. Exp. Neurol.* **50**, 451–462 (1991).
- Fu, H. et al. A tau homeostasis signature is linked with the cellular and regional vulnerability of excitatory neurons to tau pathology. *Nat. Neurosci.* **22**, 47–56 (2019).
- Heinsen, H. et al. Quantitative investigations on the human entorhinal area: left-right asymmetry and age-related changes. *Anat. Embryol. (Berl.)* **190**, 181–194 (1994).
- Kobro-Flatmoen, A. & Witter, M. P. Neuronal chemo-architecture of the entorhinal cortex: a comparative review. *Eur. J. Neurosci.* **50**, 3627–3662 (2019).
- Naumann, R. K. et al. Conserved size and periodicity of pyramidal patches in layer 2 of medial/caudal entorhinal cortex. *J. Comp. Neurol.* **524**, 783–806 (2016).
- Ramsden, H. L., Surmeli, G., McDonagh, S. G. & Nolan, M. F. Laminal and dorsoventral molecular organization of the medial entorhinal cortex revealed by large-scale anatomical analysis of gene expression. *PLoS Comput. Biol.* **11**, e1004032 (2015).
- Jabaudon, D., Shnider, S. J., Tischfield, D. J., Galazo, M. J. & Macklis, J. D. RORBeta induces barrel-like neuronal clusters in the developing neocortex. *Cereb. Cortex* **22**, 996–1006 (2012).
- Oishi, K., Aramaki, M. & Nakajima, K. Mutually repressive interaction between Brn1/2 and Rorb contributes to the establishment of neocortical layer 2/3 and layer 4. *Proc. Natl Acad. Sci. USA* **113**, 3371–3376 (2016).
- Nakagawa, Y. & O'Leary, D. D. Dynamic patterned expression of orphan nuclear receptor genes RORalpha and RORbeta in developing mouse forebrain. *Dev. Neurosci.* **25**, 234–244 (2003).
- Marinero, F. et al. Molecular and cellular pathology of monogenic Alzheimer's disease at single cell resolution. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.07.14.202317> (2020).
- Liang, W. S. et al. Altered neuronal gene expression in brain regions differentially affected by Alzheimer's disease: a reference data set. *Physiol. Genomics* **33**, 240–256 (2008).
- Franjic, D. et al. Molecular diversity among adult human hippocampal and entorhinal cells. Preprint at *bioRxiv* <https://doi.org/10.1101/2019.12.31.889139> (2019).
- Ehrenberg, A. J. et al. A manual multiplex immunofluorescence method for investigating neurodegenerative diseases. *J. Neurosci. Meth.* <https://doi.org/10.1016/j.jneumeth.2020.108708> (2020).
- Keren-Shaul, H. et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell* **169**, 1276–1290 e1217 (2017).
- Srinivasan, K. et al. Alzheimer's patient brain myeloid cells exhibit enhanced aging and unique transcriptional activation. Preprint at *bioRxiv* <https://doi.org/10.1101/610345> (2019).

32. Thrupp, N., et al. Single nucleus sequencing fails to detect microglial activation in human tissue. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.04.13.035386> (2020).
  33. Chen, W. T. et al. Spatial transcriptomics and in situ sequencing to study Alzheimer's disease. *Cell* **182**, 976–991 (2020).
  34. Perez-Nievas, B. G. & Serrano-Pozo, A. Deciphering the Astrocyte reaction in Alzheimer's disease. *Front Aging Neurosci.* **10**, 114 (2018).
  35. Simpson, J. E. et al. Microarray analysis of the astrocyte transcriptome in the aging brain: relationship to Alzheimer's pathology and APOE genotype. *Neurobiol. Aging* **32**, 1795–1807 (2011).
  36. Sekar, S. et al. Alzheimer's disease is associated with altered expression of genes involved in immune response and mitochondrial processes in astrocytes. *Neurobiol. Aging* **36**, 583–591 (2015).
  37. Liddelow, S. A. et al. Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* **541**, 481–487 (2017).
  38. Laywell, E. D. et al. Enhanced expression of the developmentally regulated extracellular matrix molecule tenascin following adult brain injury. *Proc. Natl Acad. Sci. USA* **89**, 2634–2638 (1992).
  39. Zamanian, J. L. et al. Genomic analysis of reactive astrogliosis. *J. Neurosci.* **32**, 6391–6410 (2012).
  40. Anderson, M. A. et al. Astrocyte scar formation aids central nervous system axon regeneration. *Nature* **532**, 195–200 (2016).
  41. Kampmann, M. A CRISPR approach to neurodegenerative diseases. *Trends Mol. Med.* **23**, 483–485 (2017).
  42. Kampmann, M. CRISPR-based functional genomics for neurological disease. *Nat. Rev. Neurol.* **16**, 465–480 (2020).
  43. Tian, R. et al. CRISPR interference-based platform for multimodal genetic screens in human iPSC-Derived neurons. *Neuron* **104**, 239–255.e212 (2019).
  44. Hof, P. R. & Morrison, J. H. Neocortical neuronal subpopulations labeled by a monoclonal antibody to calbindin exhibit differential vulnerability in Alzheimer's disease. *Exp. Neurol.* **111**, 293–301 (1991).
  45. Hof, P. R., Cox, K. & Morrison, J. H. Quantitative analysis of a vulnerable subset of pyramidal neurons in Alzheimer's disease: I. Superior frontal and inferior temporal cortex. *J. Comp. Neurol.* **301**, 44–54 (1990).
  46. Mikkonen, M., Alafuzoff, I., Tapiola, T., Soininen, H. & Miettinen, R. Subfield- and layer-specific changes in parvalbumin, calretinin and calbindin-D28K immunoreactivity in the entorhinal cortex in Alzheimer's disease. *Neuroscience* **92**, 515–532 (1999).
  47. Montine, T. J. et al. National Institute on Aging-Alzheimer's association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach. *Acta Neuropathol.* **123**, 1–11 (2012).
  48. Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A. & Martin, R. L. A new clinical scale for the staging of dementia. *Br. J. Psychiatry* **140**, 566–572 (1982).
- Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
- © The Author(s), under exclusive licence to Springer Nature America, Inc. 2021

## Methods

**Postmortem cohort.** This study was approved by the University of Sao Paulo institutional review board and deemed nonhuman participant research by UCSF. De-identified human postmortem brain tissue was supplied by the Neurodegenerative Disease Brain Bank (NDBB) at UCSF and the BBAS from the University of Sao Paulo<sup>49</sup>. The NDBB receives brain donations from participants enrolled in the UCSF Memory and Aging Center research programs. The BBAS is population-based and houses a high percentage of pathologically and clinically normal control participants who are not available in the NDBB. Neuropathological assessments were performed using standardized protocols and followed internationally accepted criteria for neurodegenerative diseases<sup>50–52</sup>. The brain samples used in this study contained a broad burden of AD-type pathology and were selected to be free from non-AD pathology, including Lewy body disease, TDP-43 proteinopathies, primary tauopathies and cerebrovascular changes. Argyrophilic grain disease was not an exclusion criterion based on its high prevalence and lack of correlation with major clinical symptoms<sup>53–55</sup>. In total, the cohort included 10 participants with AD who underwent snRNA-seq, representing Braak stages 0, 2 and 6, all *APOE*  $\epsilon$ 3/ $\epsilon$ 3 genotypes and 26 participants with AD who underwent neuroanatomical analysis, representing Braak stages 0–6<sup>1,56</sup>, ranging from 2–5 individuals for each Braak stage. Table 1 depicts the characteristics of the 31 participants with AD.

**Isolation of nuclei from frozen postmortem human brain tissue.** Isolation of nuclei was performed similarly as previously described<sup>57</sup>. Briefly, frozen brain tissue was dounce homogenized in 5 ml of lysis buffer (0.25 M sucrose, 25 mM KCl, 5 mM MgCl<sub>2</sub>, 20 mM Tricine-KOH (pH 7.8), 1 mM dithiothreitol, 0.15 mM spermine, 0.5 mM spermidine, 1× protease inhibitor (Sigma, 4693159001) and RNase inhibitor (Promega, N2615)). Following initial dounce homogenization, IGEPAL-630 was added to a final concentration of 0.3%, and the sample was homogenized with five more strokes. The solution was then filtered through a 40- $\mu$ m cell filter and mixed with Optiprep (Sigma, D1556-250ML) to create a 25% Optiprep solution. This solution was then layered onto a 30%/40% Optiprep gradient and centrifuged at 10,000g for 18 min using the SW41-Ti rotor. The nuclei were collected at the 30%/40% Optiprep interface.

**Droplet-based single-nucleus RNA sequencing.** Droplet-based snRNA-seq was performed using the Chromium Single Cell 3' Reagent Kits v2 from 10X Genomics. Nuclei were resuspended to a concentration of 1,000 nuclei per  $\mu$ l in 30% Optiprep solution before loading according to the manufacturer's protocol, with 10,000 nuclei recovered per sample as the target. cDNA fragment analysis was performed using the Agilent 4200 TapeStation System. Sequencing and quality control were performed as described by The Tabula Muris Consortium<sup>58</sup>.

**Preprocessing of snRNA-seq data.** Sequencing data generated from snRNA-seq libraries were demultiplexed using Cell Ranger (version 2.1.0) with cellranger mkfastq. To align reads, we first generated our own pre-mRNA GRCh38 reference genome using cellranger mkref to account for introns that may be eliminated using the default GRCh38 reference genome. Alignment and gene expression quantification was then performed using cellranger count with default settings.

**Exploratory analysis of EC and SFG data.** For each sample, the raw gene–barcode matrix outputted by Cell Ranger was converted into a SingleCellExperiment (SCE) object in R (version 3.5.1) using the read10xCounts function from the DropletUtils package<sup>59</sup> (version 1.2.2). Droplets containing nuclei were then distinguished from empty droplets using DropletUtils::emptyDrops with the parameter false discovery rate (FDR) = 0.01, and then nuclei (hereafter referred to as 'cells') with less than 200 UMIs were discarded. Afterward, SCE objects corresponding to each sample were merged into a single SCE object for downstream processing and analyses.

For normalization of raw counts, to avoid artifacts caused by data sparsity, the approach of Lun et al.<sup>60</sup> was adopted: for each sample, cells were first clustered using a graph-based method followed by pooling counts across cells in each cluster to obtain pool-based size factors, which were then deconvoluted to yield cell-based size factors. Clustering was performed using the quickCluster function from the scanr package<sup>61</sup> (version 1.10.2) with the parameters method = igrph, min.mean = 0.1, irlba.args = list (maxit = 1,000) and the block parameter set to a character vector containing the sample identity of each cell. Size factors were computed using scanr::computeSumFactors with the parameter min.mean = 0.1 and the cluster parameter set to a character vector containing the cluster identity of each cell; cells with negative size factors were removed. Normalization followed by log transformation was then performed using the normalize function from the scater package<sup>62</sup> (version 1.10.1).

Before dimensionality reduction, highly variable genes were identified for each sample separately using the approach of Lun et al.<sup>61</sup>, that is, the variance of each gene was decomposed into a technical and biological component. Technical variance was assumed as Poisson and modeled using scanr::makeTechTrend. The mean-variance trend across genes was fitted using scanr::trendVar with parameters use.spikes = FALSE and loess.args = list (span = 0.05), and the trend slot of the resulting fit object was then set to the output of scanr::makeTechTrend. Biological variance was extracted from the total variance using scanr::decomposeVar with

the above fit object as the input. Finally, highly variable genes that were preserved across samples were identified by combining the variance decompositions with scanr::combineVar, using Stouffer's z-score method for meta-analysis (method = z), which assigns more weight to samples with more cells.

For initial data exploration, genes with a combined biological variance of greater than 0 were used as the feature set for dimensionality reduction by principal-component analysis using scanr::parallelPCA, which uses Horn's parallel analysis to decide how many principal components to retain, with the parameter approx = TRUE. Clustering was then performed on the retained principal components using the FindClusters function from the Seurat package<sup>63</sup> (version 2.3.4) with the parameter resolution = 0.8, which required conversion of SCE objects to Seurat objects using Seurat::Convert. To visualize the clusters, t-SNE was performed on the retained principal components using scater::runTSNE with parameters perplexity = 30 and rand\_seed = 100.

**Cross-sample alignment of SFG and EC data.** Initial data exploration revealed that clustering was driven by individual of origin in addition to cell-type identity, which makes it difficult to analyze changes in the relative abundance or gene expression of a given cell type across disease progression or brain regions. To recover clusters defined mainly by cell-type identity, data were aligned across samples from each brain region using scAlign<sup>64</sup> (version 1.0.0), which leverages a neural network to learn a low-dimensional alignment space in which cells from different datasets group by biological function independent of technical and experimental factors. As noted by Johansen & Quon<sup>64</sup>, scAlign converges faster with little loss of performance when the input data is represented by principal components or canonical correlation vectors. Therefore, before running scAlign, the top 2,000 genes with the highest combined biological variance were used as the feature set for canonical correlation analysis (CCA), which was implemented using Seurat::RunMultiCCA with parameter num.cc = 15. The number of canonical coordinates to use for scAlign was determined by the elbow method using Seurat::MetageneBicorPlot. scAlign was then run on the cell loadings along the top ten canonical correlation vectors with the following parameters: options = scAlignOptions (steps = 10,000, log.every = 5,000, architecture = large, num.dim = 64), encoder.data = cca, supervised = none, run.encoder = TRUE, run.decoder = FALSE, log.results = TRUE and device = CPU. Clustering was then performed on the full dimensionality of the output from scAlign using Seurat::FindClusters with parameters resolution = 0.8 for the SFG and resolution = 0.6 for the EC. Clusters were visualized with t-SNE using Seurat::RunTSNE on the full dimensionality of the output from scAlign using do.fast = TRUE. Alignment using scAlign followed by clustering was also performed for all samples from both brain regions jointly.

To assign clusters identified in the aligned subspace generated by scAlign to major brain cell types, the following marker genes were used: *SLC1A7* and *CAMK2A* for excitatory neurons, *GAD1* and *GAD2* for inhibitory neurons, *SLC1A2* and *AQP4* for astrocytes, *MBP* and *MOG* for oligodendrocytes, *PDGFRA* and *SOX10* for oligodendrocyte precursor cells, *CD74* and *CX3CR1* for microglia/myeloid cells and *CLDN5* and *FLT1* for endothelial cells. Clusters expressing markers for more than one cell type, most likely reflecting doublets, were removed from downstream analyses.

**Cell-type-specific subclustering (subpopulation) analysis.** To identify cell-type subpopulations, cells from all samples belonging to a given major cell type were extracted for sample-level recomputation of size factors and highly variable genes. CCA was then performed using the top 1,000 genes with the highest combined biological variance as the feature set, followed by alignment of the first 10 to 12 canonical coordinates with scAlign using the parameter steps = 2,500. The full dimensionality of the output from scAlign was used for subclustering (using resolution = 0.4) and t-SNE. Analyzing cells from each brain region separately, marker genes for subpopulations were identified using scanr::findMarkers with parameters direction = up, pval.type = any, lfc = 0.58 and the block parameter set to a character vector corresponding to the sample identity for each cell. Subpopulations that expressed markers for more than one cell type were removed from downstream analyses.

**Identification of differentially expressed genes in cell-type subpopulations.** To identify genes differentially expressed by a cell-type subpopulation compared to all other subpopulations in a way that accounts for true biological replication (that is, at the level of individuals), UMI counts of cells from the same individual belonging to the subpopulation of interest or all other subpopulations were summed to obtain 'pseudo-bulk' samples, which were then analyzed using edgeR<sup>65</sup> (version 3.24.3) following the approach recommended by Amezcua et al.<sup>65</sup>. An FDR cutoff of 0.1 was used.

**Heat map of relative gene expression across cell types or cell-type subpopulations.** For heat maps of relative gene expression across cell types or cell-type subpopulations, log-scaled normalized counts of each gene were z-score transformed across all cells and then averaged across cells in each cluster to enhance visualization of differences among clusters. Thus, genes with high relative expression had above-average expression (positive z-scores) and genes with low relative expression had below-average expression (negative z-scores).

**Functional association network analysis and pathway enrichment analysis of differentially expressed genes.** Differentially expressed genes were visualized as a functional association network using STRING-db<sup>67</sup> (version 11), a protein–protein association network based on known physical interactions, functional associations, coexpression and other metrics, and Cytoscape<sup>68</sup> (version 3.7.2), a network visualization software. When generating the networks, the STRING-db association confidence score cutoff was set to 0.5, and the network layout was optimized for visualization using yFiles Organic Layout. For pathway enrichment analysis, enrichments for Gene Ontology terms and Reactome Pathways were also obtained through STRING-db, using an FDR cutoff of 0.05.

**Entorhinal cortex layer-specific genes.** Due to the lack of published data on layer-specific genes for the human EC, layer-specific genes in the mouse medial entorhinal cortex (MEC) were obtained from Ramsden et al.<sup>22</sup>; the MEC is the most phylogenetically similar to the human caudal EC<sup>20,21</sup> used in this study. Specifically, genes with expression specific for layers II, III and V/VI of the mouse MEC according to the S4 Dataset excel spreadsheet in the supplementary information of Ramsden et al.<sup>22</sup> were mapped to human genes and cross-referenced against genes differentially expressed across EC excitatory neuron subclusters (obtained using `scanr::findMarkers` without setting `direction = up`).

**Reanalysis of the Mathys et al. dataset.** To reanalyze the data from Mathys et al.<sup>14</sup> using our cross-sample alignment approach, the filtered matrix of UMI counts ('Data/Gene Expression (RNA-seq)/Processed/filtered\_count\_matrix.mtx') and associated row and column metadata ('filtered\_gene\_row\_names.txt' and 'filtered\_column\_metadata.txt') were downloaded from the AMP-AD Knowledge Portal (syn18485175). The experimental and clinical metadata files were downloaded from 'Data/Metadata/'. The filtered UMI counts matrix and the associated row and column metadata were then converted to a SCE object for analysis, and the relevant experimental and clinical metadata (for example, 'pathologic diagnosis of AD') were merged with the SCE object. The cell-type assignments from Mathys et al.<sup>14</sup> provided in the column metadata were used for subclustering.

**Functional annotation of differentially expressed genes in *GFAP*<sub>high</sub> astrocytes.** We obtained the functional annotation for differentially expressed genes from GeneCards<sup>69</sup> and verified the primary literature references for glutamate/GABA-related genes<sup>70–75</sup> and synaptic adhesion/maintenance-related genes<sup>76–79</sup>.

**Quantitative histopathological assessment using multiplex immunofluorescence.** *Delineation of the caudal EC.* We used archival paraffin blocks from the UCSF/NDBB and BBAS (Table 1). First, we collected blocks sampling the hippocampal formation anterior to the lateral geniculate body from the ten participants with AD used for the snRNA-seq and another 30 participants with AD spanning all Braak stages<sup>1</sup>. To determine if the caudal EC region was present, 8- $\mu$ m thick sections of each block underwent hematoxylin and eosin staining (Extended Data Fig. 8a). We took digital images of the stained sections and aligned each one the most approximate section from a large collection of 400- $\mu$ m thick serial coronal sections of whole-brain hemispheres stained for galocyanin (provided by H.H.)<sup>19,80</sup> (Extended Data Fig. 8b). We eliminated blocks from five participants with AD used for snRNA-seq and four of the extra individuals with AD for lack of caudal EC. Next, with the aid of the paired galocyanin sections, we delineated the borders of the caudal EC in each case (Extended Data Fig. 8a).

The EC is considered a peri- or allocortex, depending on the author<sup>8</sup>. EC parcellation and cytoarchitectonic definitions have been a matter of debate, and here, we are adopting the cytoarchitectonic definitions proposed by Heinsen and colleagues<sup>19</sup>, which is based on the examination of thick histological preparations and considered the definitions proposed by Insausti & Amaral (six layers)<sup>81</sup> and Braak & Braak (three layers)<sup>9</sup>. In thick histological sections, the caudal entorhinal region features well-delineated clusters of stellate or principal cells in layer II (pre-alpha clusters) and three lamina dissecans<sup>19</sup>. The external dissecans (dissecans-ext) divides layers II and III and is particularly prominent in the caudal EC. Dissecans-1 (diss-1) corresponds to layer IV of work by Insausti<sup>82</sup> and the lamina dissecans of data from Braak & Braak<sup>9</sup> and Rose<sup>83</sup>. The most internal dissecans (dissecans-2 or diss-2) is hardly appreciated in thin sections but easy to visualize in thick sections. It roughly corresponds to layer Vc of the caudal subregions of work by Insausti<sup>82</sup>.

**Multiplex immunofluorescence.** Next, for each tissue sample from participants with AD, an 8- $\mu$ m thick, formalin-fixed and paraffin-embedded coronal section underwent immunofluorescence against TBR1, RORB and phospho-tau (CP13) as described below. TBR1 (T-box, brain 1) is a transcription factor that has a role in differentiation of glutamatergic neurons and is a marker for excitatory neurons, including EC excitatory neurons<sup>84,85</sup>. In summary, sections were deparaffinized and incubated in 3.0% hydrogen peroxide (Fisher, H325-500) in methanol to inactivate endogenous peroxidase. Antigen retrieval was performed in 1× Tris-EDTA heat-induced epitope retrieval solution (TES500; PBS with 0.05% Tween 20 (PBS-T)) at pH9 in an autoclave at 121 °C for five min. To reduce nonspecific background staining, sections were blocked with 5% Milk/PBS-T. To avoid cross-reactions between primary antibodies that were raised against the same species, an antibody stripping step using 0.80%  $\beta$ -mercaptoethanol/10% sodium

dodecyl sulfate in 12.5% Tris-HCL was performed after the tyramide-signal amplification development for RORB.

Sections were first incubated overnight in primary antibody against RORB (1:400 dilution; rabbit, HPA008393, Millipore Sigma), which was later developed in goat anti-rabbit HRP (1:400 dilution; R-05072-500, Advanta) with Alexa Fluor 488 tyramide-signal amplification (1:100 dilution; B40953, Thermo Fisher). Next, sections were stripped of RORB primary antibody and then incubated overnight in a cocktail of primary antibodies against TBR1 (1:100 dilution; Rabbit, ab31940, Abcam) and CP13 (1:800 dilution; mouse, phospho-tau serine 202; a gift from P. Davies), all of which were later developed with secondary antibodies and fluorophores: for TBR1, Alexa Fluor 546 conjugated anti-rabbit secondary (1:200 dilution; A-11010, Thermo Fisher) was used; and for CP13, biotinylated anti-mouse (1:400 dilution; BA-2000, Vector Laboratory) with streptavidin Alexa Fluor 790 (1:250 dilution; S11378, Thermo Fisher) was used. Sections were then counterstained with DAPI diluted in PBS (1:5,000 dilution; D1306, Invitrogen). Finally, sections were incubated in Sudan Black B (199664-25g, Sigma) to reduce autofluorescence and coverslipped using Prolong antifade mounting medium (P36980, Invitrogen). A quality-control slide was used to verify the efficacy of the antibody stripping process. A detailed description of the method is provided by Ehrenberg et al.<sup>29</sup>. Sections were scanned using a Zeiss AxioScan slide scanner.

For generating the images shown in Fig. 3h, a section from participant no. 6 (Braak stage 2; Table 1) was stained with galocyanin-chrome alum following standard methods<sup>19</sup>. The section was placed on a coverslip and scanned using a Zeiss AxioScan slide scanner. Next, the section was removed from the coverslip and underwent immunofluorescence for RORB and CP13 as described above. Then, the section was placed on a coverslip and scanned once more.

**Neuronal quantification.** The caudal EC delineations carried out in the hematoxylin and eosin-stained slides were then transferred to the immunostained images. Within these borders, we randomly placed four 500 × 500- $\mu$ m ROIs overlaying the EC external layers (I to III), which we identified as being external to dissecans-1. We then extracted the ROIs for quantification in ImageJ (Fig. 3). The number of excitatory neurons was quantified by segmenting the TBR1 signal, using a threshold to create a mask and the segmentation editor plugin to manually remove all nonneuronal artifacts and vessels. The number of RORB<sup>+</sup> excitatory neurons was then counted using the mask of excitatory (TBR1<sup>+</sup>) neurons in the segmentation editor and manually removing all neurons not expressing RORB. All segmentations were manually verified for quality control. Quantification was performed blinded to the neuropathological diagnosis. We quantified phospho-tau (CP13) staining in two ROIs in a subset of the participants with AD, using the same FIJI protocol.

**Statistics. Beta regression.** For each brain region, the relative abundance of a given cell cluster or cell type, ranging from 0 to 1, was computed for each sample, treated as an independent measurement and assumed to follow a beta distribution (although this was not formally tested). To determine the statistical significance of changes in the relative abundance of a given cluster or cell type across Braak stages, beta regression<sup>86</sup> was performed using the `betareg` package (version 3.1-1), with the formula `relative.abundance ~ braak.stage` for both the mean and precision models, and the bias-corrected maximum-likelihood estimator (`type = BC`). The statistical significance of changes in the proportion of TBR1<sup>+</sup> cells and RORB<sup>+</sup> cells among TBR1<sup>+</sup> cells obtained from immunofluorescence validation was assessed similarly as above using beta regression. To correct for multiple hypothesis testing for each family of tests (for example, testing all cell-type subpopulations for a brain region), Holm's method was used to adjust *P* values obtained from beta regression to control the family-wise type I error rate at 0.05.

**Fisher's exact test.** For Fig. 3g, two-sided Fisher's exact test was used to calculate the statistical significance of the observed enrichment of CP13 staining in RORB<sup>+</sup> excitatory neurons. The test was performed in R using `fisher.test` with `alternative = 'two-sided'`.

**Hypergeometric test.** For Fig. 5e and Extended Data Fig. 10c, a one-sided hypergeometric test (implemented in R with the package `GeneOverlap`, version 1.18.0) was used to calculate the statistical significance of the observed gene overlaps. The *P* values were adjusted for multiple testing using the Benjamini–Hochberg method.

**Randomization.** Data collection for the snRNA-seq or immunostaining validation was not randomized or blocked.

**Sample sizes.** No statistical methods were used to predetermine sample sizes, but our sample sizes are comparable to those reported previously<sup>14,15</sup>.

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

## Data availability

The raw snRNA-seq sequencing data and unfiltered UMI count matrices are available on the Gene Expression Omnibus (GEO) under accession code

GSE147528. Single-cell data after quality control is available for download at Synapse.org under the Synapse ID [syn21788402](https://syn21788402). Post-quality-control data can also be explored interactively through the CellXGene platform at <https://kammannlab.ucsf.edu/ad-brain>. Data from Mathys *et al.*<sup>14</sup> was downloaded from Synapse under Synapse ID [syn18485175](https://syn18485175).

### Code availability

We provide the full bioinformatics pipeline for the analysis of snRNA-seq data in this paper at <https://kammannlab.ucsf.edu/ad-brain-analysis>.

### References

49. Grinberg, L. T. *et al.* Brain bank of the Brazilian aging brain study group - a milestone reached and more than 1,600 collected brains. *Cell Tissue Bank* **8**, 151–162 (2007).
50. Hyman, B. T. *et al.* National institute on Aging-Alzheimer's association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement.* **8**, 1–13 (2012).
51. Suemoto, C. K. *et al.* Neuropathological diagnoses and clinical correlates in older adults in Brazil: A cross-sectional study. *PLoS Med.* **14**, e1002267 (2017).
52. Cairns, N. J. *et al.* Neuropathologic diagnostic and nosologic criteria for frontotemporal lobar degeneration: consensus of the consortium for frontotemporal lobar degeneration. *Acta Neuropathol.* **114**, 5–22 (2007).
53. Ferrer, L., Santpere, G. & van Leeuwen, F. W. Argyrophilic grain disease. *Brain* **131**, 1416–1432 (2008).
54. Rodriguez, R. D. & Grinberg, L. T. Argyrophilic grain disease: An underestimated tauopathy. *Dement. Neuropsychol.* **9**, 2–8 (2015).
55. Rodriguez, R. D. *et al.* Argyrophilic grain disease: Demographics, clinical, and neuropathological features from a large autopsy study. *J. Neuropathol. Exp. Neurol.* **75**, 628–635 (2016).
56. Braak, H., Thal, D. R., Ghebremedhin, E. & Del Tredici, K. Stages of the pathologic process in Alzheimer disease: age categories from 1 to 100 years. *J. Neuropathol. Exp. Neurol.* **70**, 960–969 (2011).
57. Mo, A. *et al.* Epigenomic signatures of neuronal diversity in the mammalian brain. *Neuron* **86**, 1369–1384 (2015).
58. Tabula Muris, C. *et al.* Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* **562**, 367–372 (2018).
59. Lun, A. T. L. *et al.* EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* **20**, 63 (2019).
60. Lun, A. T., Bach, K. & Marioni, J. C. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biol.* **17**, 75 (2016).
61. Lun, A. T., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res* **5**, 2122 (2016).
62. McCarthy, D. J., Campbell, K. R., Lun, A. T. & Wills, Q. F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179–1186 (2017).
63. Stuart, T. *et al.* Comprehensive Integration of Single-Cell data. *Cell* **177**, 1888–1902 e1821 (2019).
64. Johansen, N. & Quon, G. scAlign: a tool for alignment, integration, and rare cell identification from scRNA-seq data. *Genome Biol.* **20**, 166 (2019).
65. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
66. Amezquita, R. A. *et al.* Orchestrating single-cell analysis with bioconductor. *Nat. Methods* **17**, 137–145 (2020).
67. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
68. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498–2504 (2003).
69. Stelzer, G. *et al.* The GeneCards Suite: From gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinforma.* **54**, 13031–131033 (2016).
70. Arriza, J. L. *et al.* Functional comparisons of three glutamate transporter subtypes cloned from human motor cortex. *J. Neurosci.* **14**, 5559–5569 (1994).
71. Borden, L. A. *et al.* Cloning of the human homologue of the GABA transporter GAT-3 and identification of a novel inhibitor with selectivity for this site. *Receptors Channels* **2**, 207–213 (1994).
72. Gendreau, S. *et al.* A trimeric quaternary structure is conserved in bacterial and human glutamate transporters. *J. Biol. Chem.* **279**, 39505–39512 (2004).
73. Häberle, J. *et al.* Congenital glutamine deficiency with glutamine synthetase mutations. *N. Engl. J. Med.* **353**, 1926–1933 (2005).
74. Kawakami, H., Tanaka, K., Nakayama, T., Inoue, K. & Nakamura, S. Cloning and expression of a human glutamate transporter. *Biochem. Biophys. Res. Commun.* **199**, 171–176 (1994).
75. Melzer, N., Biela, A. & Fahlke, C. Glutamate modifies ion conduction and voltage-dependent gating of excitatory amino acid transporter-associated anion channels. *J. Biol. Chem.* **278**, 50112–50119 (2003).
76. Südhof, T. C. Synaptic neurexin complexes: A molecular code for the logic of neural circuits. *Cell* **171**, 745–769 (2017).
77. Pellissier, F., Gerber, A., Bauer, C., Ballivet, M. & Ossipow, V. The adhesion molecule Necl-3/SynCAM-2 localizes to myelinated axons, binds to oligodendrocytes and promotes cell adhesion. *BMC Neurosci.* **8**, 90 (2007).
78. González-Castillo, C., Ortuño-Sahagún, D., Guzmán-Brambila, C., Pallàs, M. & Rojas-Mayorquín, A. E. Pleiotrophin as a central nervous system neuromodulator, evidences from the hippocampus. *Front. Cell Neurosci.* **8**, 443 (2014).
79. Siddiqui, T. J. *et al.* An LRRTM4-HSPG complex mediates excitatory synapse development on dentate gyrus granule cells. *Neuron* **79**, 680–695 (2013).
80. Heinsen, H., Arzberger, T. & Schmitz, C. Celloidin mounting (embedding without infiltration) - a new, simple and reliable method for producing serial sections of high thickness through complete human brains and its application to stereological and immunohistochemical investigations. *J. Chem. Neuroanat.* **20**, 49–59 (2000).
81. Insausti, R. & Amaral, D. G. Entorhinal cortex of the monkey: IV. Topographical and laminar organization of cortical afferents. *J. Comp. Neurol.* **509**, 608–641 (2008).
82. Insausti, R., Muñoz-Lopez, M., Insausti, A. M. & Artacho-Perula, E. The human periallocortex: layer pattern in presubiculum, parasubiculum and entorhinal cortex. A Review. *Front. Neuroanat.* **11**, 84 (2017).
83. Rose, S. Vergleichende messungen im allocortex bei tier und mensch. *J. Psychol. Neurol.* **34**, 250–255 (1927).
84. Fu, H. *et al.* Tau pathology induces excitatory neuron loss, grid cell dysfunction, and spatial memory deficits reminiscent of early Alzheimer's disease. *Neuron* **93**, 533–541.e535 (2017).
85. Hevner, R. F. *et al.* Tbr1 regulates differentiation of the preplate and layer 6. *Neuron* **29**, 353–366 (2001).
86. Ferrari, S. L. P. & Cribari-Neto, F. Beta regression for modelling rates and proportions. *J. Appl. Stat.* **31**, 799–815 (2004).

### Acknowledgements

We thank A. Pisco, A. Maynard, S. Darmanis and the MACA team at the Chan Zuckerberg Biohub for advice on analysis 10X library preparation and reagents. We thank members of the Kampmann laboratory (A. Samelson, X. Guo, R. Tian and B. Rooney) for feedback on the manuscript. This work was supported by National Institutes of Health (NIH) awards F30 AG066418 (to K.L.), K08 AG052648 (to S.S.), R56 AG057528 (to M.K. and L.T.G.), K24 AG053435 (to L.T.G.), U54 NS100717 (to L.T.G. and M.

### Author contributions

K.L., E.L., L.T.G. and M.K. conceptualized and led the overall project. K.L., L.T.G. and M.K. wrote the manuscript, with input from all co-authors. K.L. analyzed snRNA-Seq data and visualized results. E.L. generated snRNA-Seq data, with support from R.S., M.T., and N.N. R.D.R., C.K.S., R.E.P.L., A.E., C.A.P. W.W.S., and S.S. contributed to neuropathological data generation and analysis, R.E., A.P. and H.H. contributed to neuropathological data analysis, and S.H.L. contributed to neuropathological method development.

### Ethics Declarations

This project was approved the ethical committee of the University of Sao Paulo (for tissue transfer) and deemed non-human subject research by UCSF.

### Competing interests

The authors declare no competing interests.

### Additional information

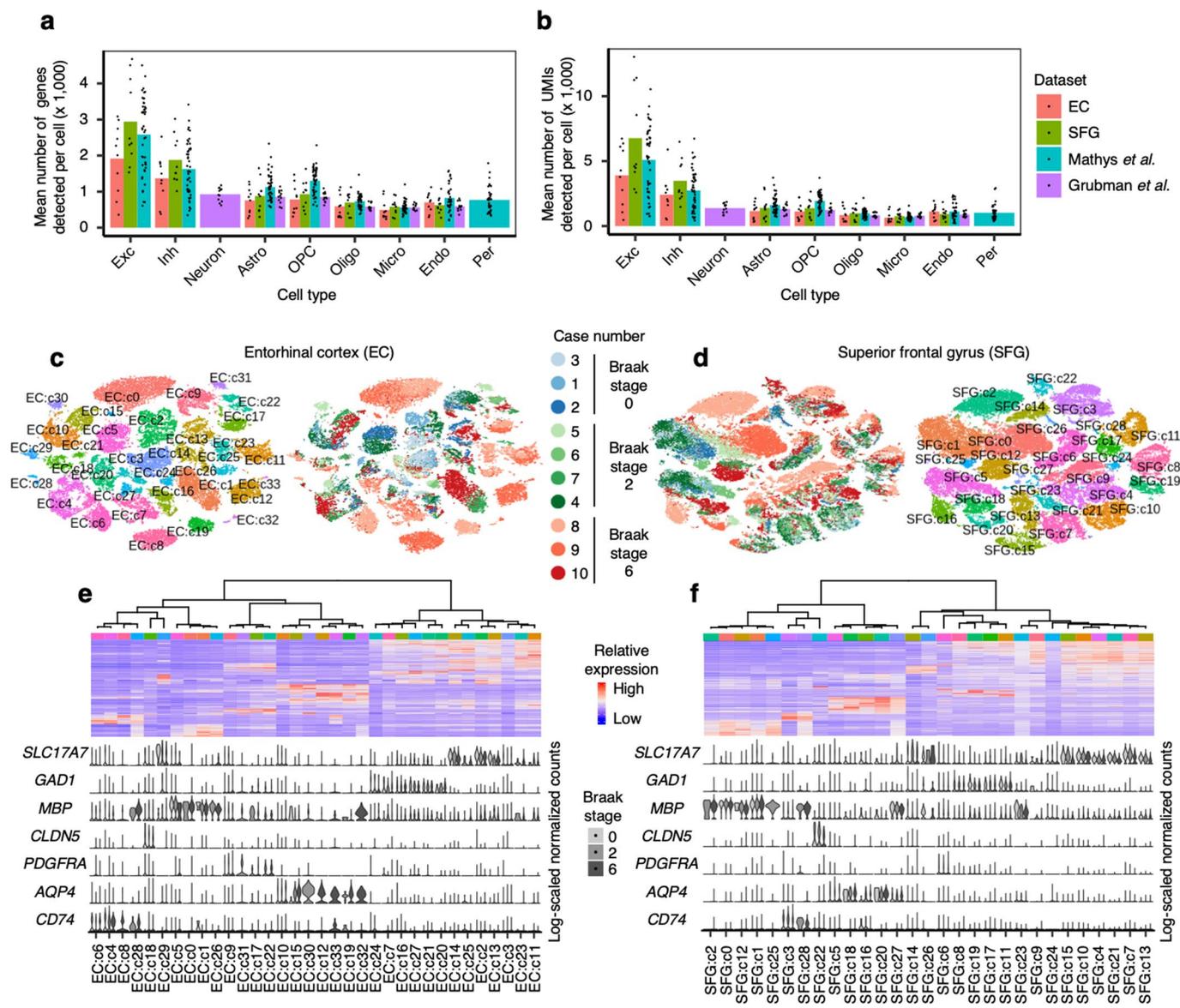
**Extended data** is available for this paper at <https://doi.org/10.1038/s41593-020-00764-7>.

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41593-020-00764-7>.

**Correspondence and requests for materials** should be addressed to L.T.G. or M.K.

**Peer review information** *Nature Neuroscience* thanks Asgeir Kobro-Flatmoen, Menno Witter, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

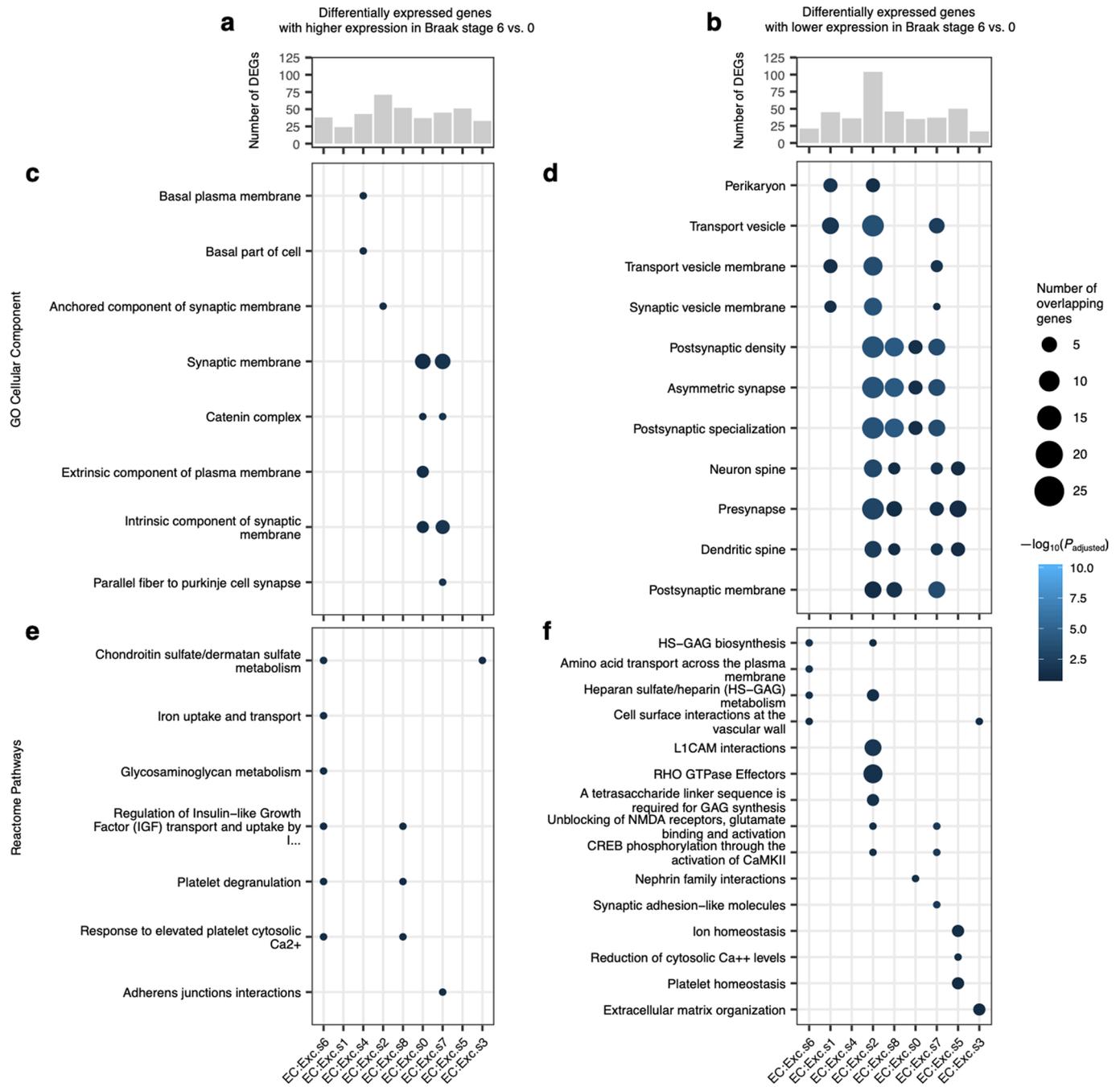
**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).



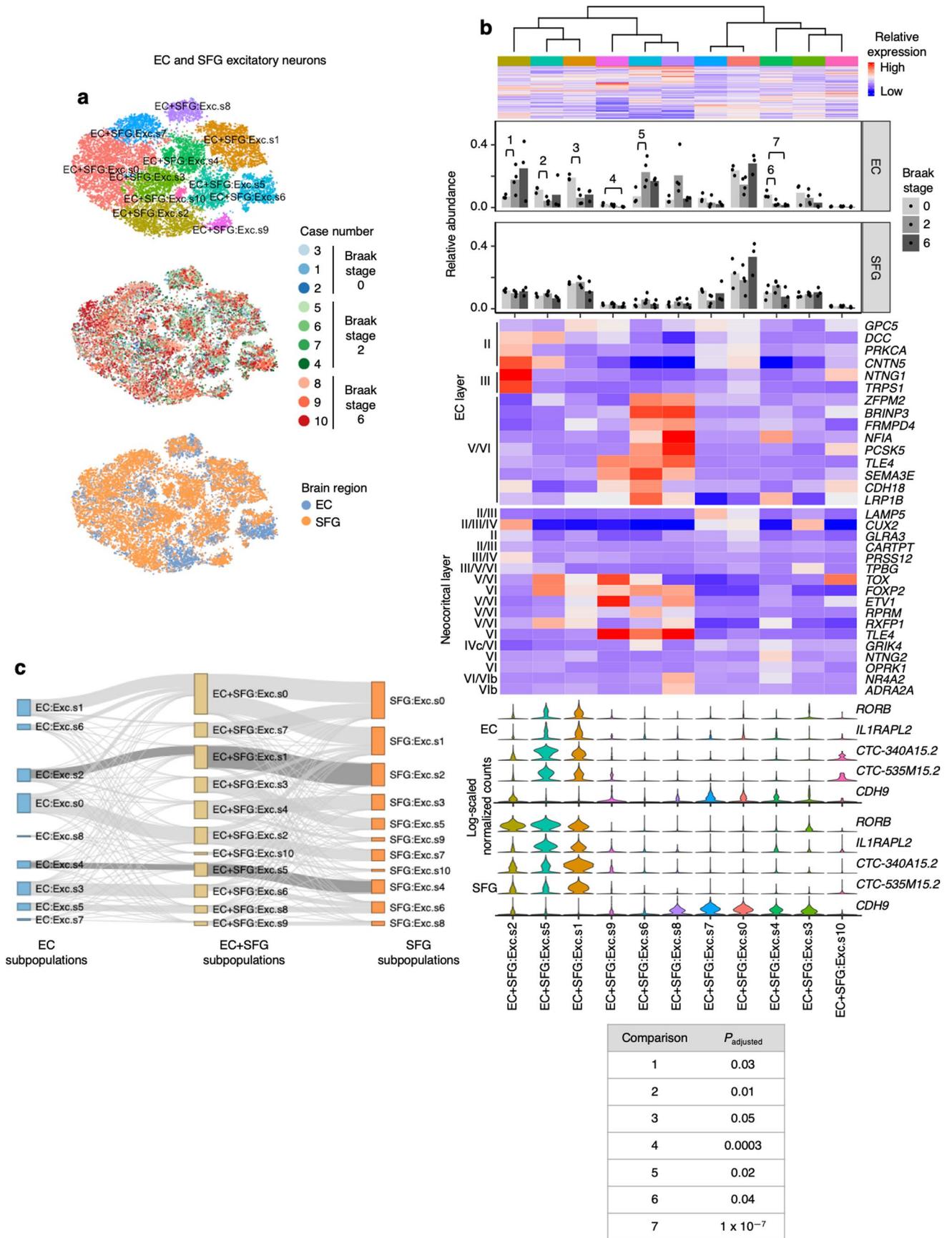
**Extended Data Fig. 1 | Data quality and initial clustering without cross-sample alignment.** **a-b**, Mean number of genes (**a**) or UMIs (**b**) detected per cell across individual samples for major cell types identified in each dataset. Grubman *et al.*<sup>15</sup> did not resolve excitatory neurons from inhibitory neurons. Pericytes were identified only in Mathys *et al.*<sup>14</sup> Cell type abbreviations: Exc – excitatory neurons, Oligo – oligodendrocytes, Astro – astrocytes, Inh – inhibitory neurons, OPC – oligodendrocyte precursor cells, Micro – microglia, Endo – endothelial cells, Per – pericytes. **c-d**, tSNE projection of cells from the EC (**c**) and SFG (**d**) clustered without first performing cross-sample alignment, colored by individual of origin (center) or cluster assignment (outer). **e-f**, Heatmap and hierarchical clustering of clusters and cluster marker expression (top subpanels); “High” and “Low” relative expression reflect above- and below-average expression, respectively (see Methods). Expression of cell type markers (bottom subpanels).



**Extended Data Fig. 2 | Expression of selected EC excitatory neuron subpopulation markers and pathway enrichment analysis of differentially expressed genes in selectively vulnerable EC excitatory neuron subpopulations.** **a**, Expression heatmap of genes that are specifically expressed by four or fewer EC excitatory neuron subpopulations; “High” and “Low” relative expression reflect above- and below-average expression, respectively (see Methods). **b-d**, Enrichment analysis against Gene Ontology Cellular Component terms or Reactome Pathways (**b,d**) and functional association network analysis (**c,e**; see Methods) of genes with higher (**b-c**) or lower expression (**d-e**) in RORB+ vulnerable EC excitatory neurons, with selected terms highlighted by color. In panels **c** and **e**, genes with stronger associations are connected by thicker lines, and genes without known associations are not shown.

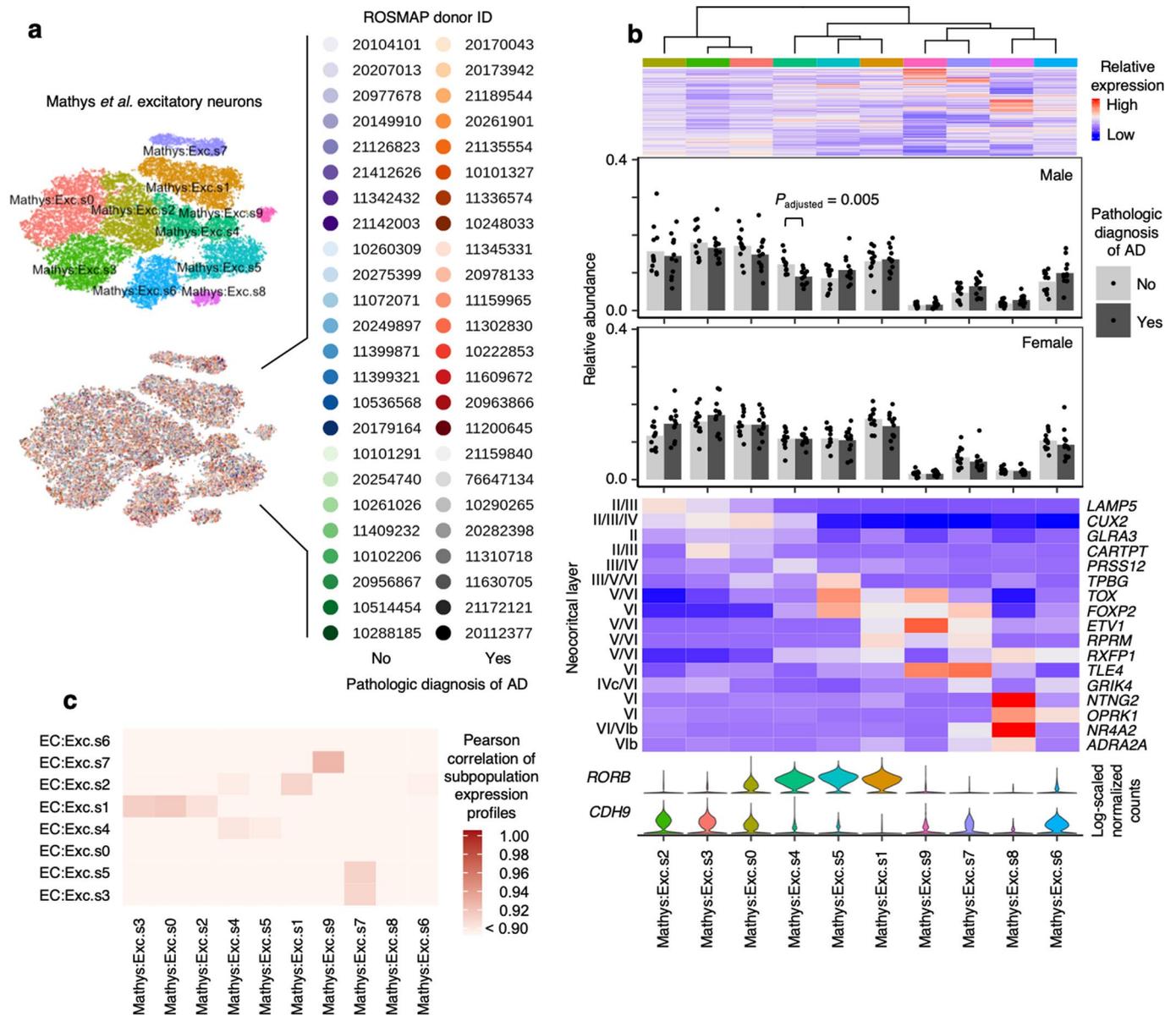


**Extended Data Fig. 3 | Differential expression analysis across Braak stages for EC excitatory neuron subpopulations. a-b**, Number of differentially expressed genes in EC excitatory neuron subpopulations with higher (a) or lower (b) expression in Braak stage 6 vs. Braak stage 0. **c-f**, Enrichment analysis against Gene Ontology Cellular Component terms (c-d) or Reactome Pathways (e-f) of differentially expressed genes in EC excitatory neuron subpopulations with higher (c,e) or lower (d,f) expression in Braak stage 6 vs. Braak stage 0.

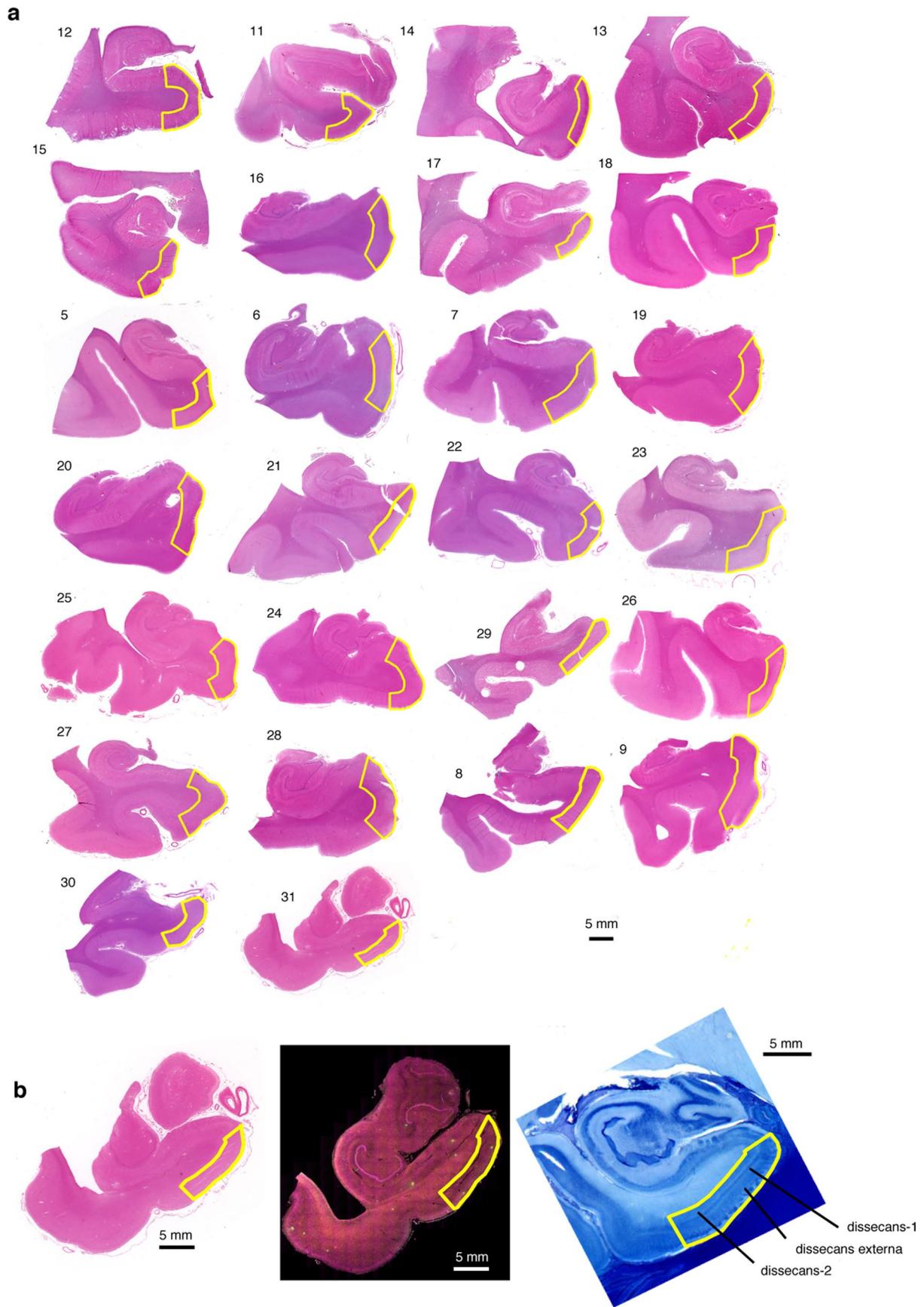


Extended Data Fig. 4 | See next page for caption.

**Extended Data Fig. 4 | Alignment of EC and SFG maps homologous excitatory neuron subpopulations.** **a**, tSNE projection of excitatory neurons from the EC and SFG in the joint alignment space, colored by subpopulation identity (top), individual of origin (middle), or brain region (bottom). **b**, Heatmap and hierarchical clustering of subpopulations and subpopulation marker expression (top subpanel); “High” and “Low” relative expression reflect above- and below-average expression, respectively (see Methods). Relative abundance of subpopulations across Braak stages (second and third subpanels); for each brain region, statistical significance of differences in relative abundance across Braak stages (Braak 0  $n=3$ , Braak 2  $n=4$ , Braak 6  $n=3$ , where  $n$  is the number of individuals sampled) was determined by beta regression and adjusted for multiple comparisons (see Methods). Expression heatmap of EC layer-specific genes identified from Ramsden *et al.*<sup>22</sup> (fourth subpanel). Expression heatmap of neocortical layer-specific genes from Lake *et al.*<sup>12</sup> (fifth subpanel). Expression of selectively vulnerable EC excitatory neuron subpopulation markers by excitatory neurons in the EC (sixth subpanel) or SFG (bottom subpanel). Significant beta regression  $P$  values (adjusted for multiple testing) are shown in a table at the bottom of the panel. **c**, Sankey diagram connecting subpopulation identity of excitatory neurons in the EC alignment space and the SFG alignment space to subpopulation identity in the EC+SFG alignment space. The links connecting EC:Exc.s2 and EC:Exc.s4 to SFG:Exc.s2 and SFG:Exc.s4, respectively, are highlighted.

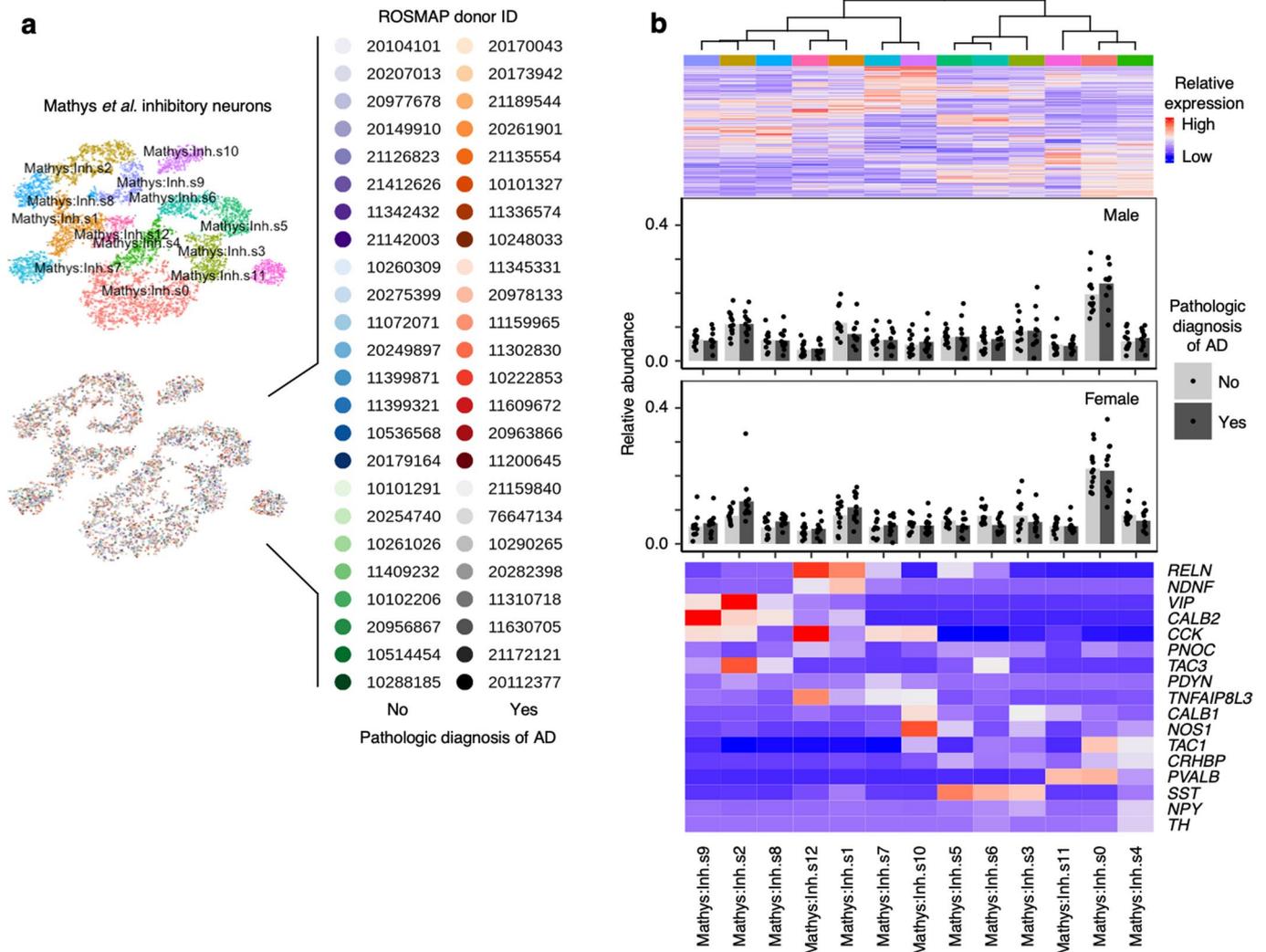


**Extended Data Fig. 5 | Cross-sample alignment of excitatory neurons from Mathys *et al.* recapitulates selective vulnerability in a RORB-expressing subpopulation.** **a**, tSNE projection of excitatory neurons from Mathys *et al.*<sup>14</sup> in the alignment space, colored by subpopulation identity (top) or individual of origin (bottom). **b**, Heatmap and hierarchical clustering of subpopulations and subpopulation marker expression (top subpanel); “High” and “Low” relative expression reflect above- and below-average expression, respectively (see Methods). Relative abundance of subpopulations in in AD cases vs. controls, separated by sex (second and third subpanels); for each sex, statistical significance of differences in relative abundance between AD cases vs. controls (cases  $n=12$ , controls  $n=12$ , where  $n$  is the number of individuals sampled) was determined by beta regression and adjusted for multiple comparisons (see Methods). Expression heatmap of neocortical layer-specific genes from Lake *et al.*<sup>12</sup> (fourth subpanel). Expression of selectively vulnerable EC excitatory neuron subpopulation markers (bottom subpanel). **c**, Heatmap of Pearson correlation between the gene expression profiles of excitatory neuron subpopulations from the EC vs. those from the prefrontal cortex in Mathys *et al.*<sup>14</sup>.

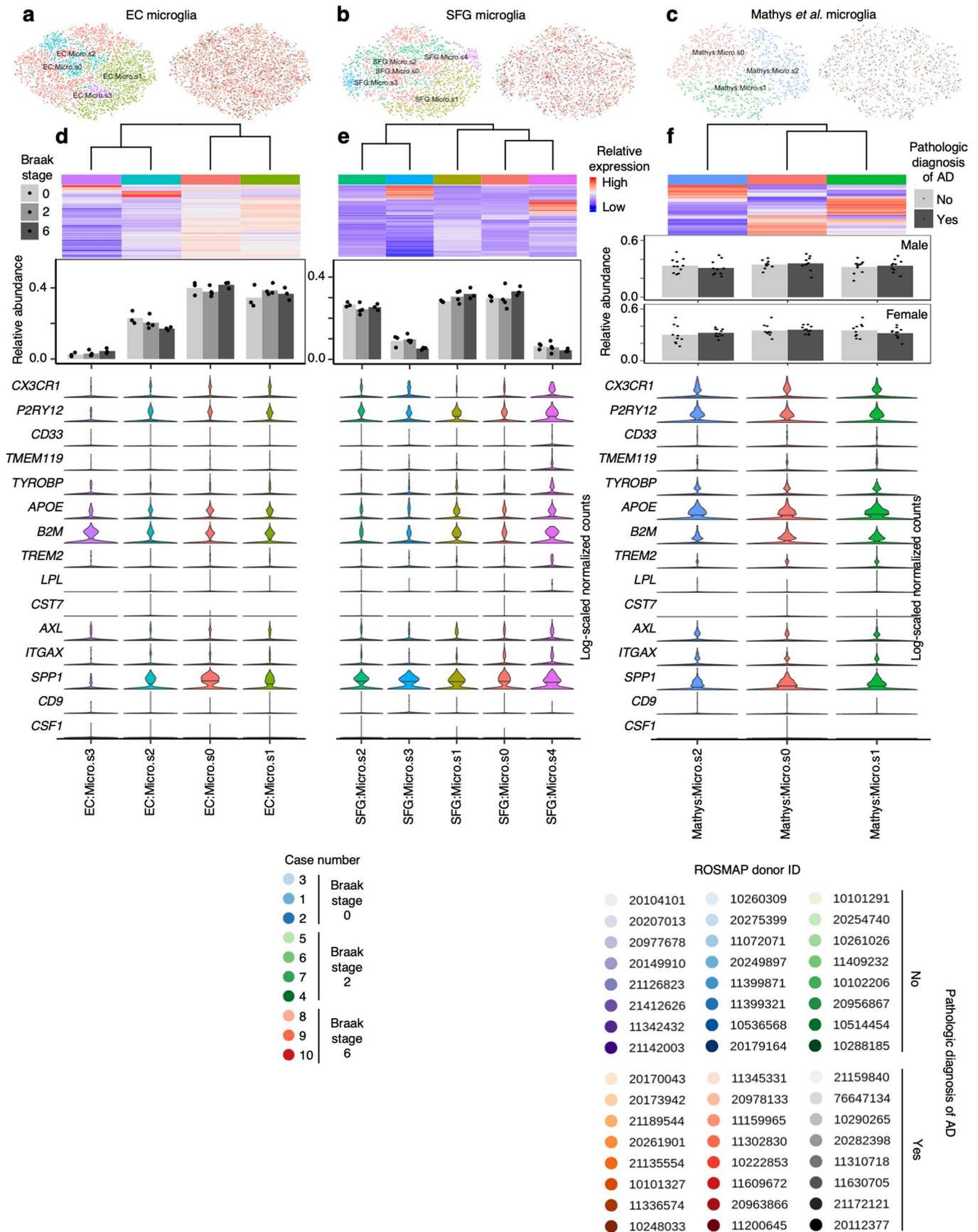


Extended Data Fig. 6 | See next page for caption.

**Extended Data Fig. 6 | Delineation of the EC for each case used in immunofluorescence validation. a,** The borders of the caudal EC delineated on sections stained with hematoxylin and eosin (H&E) for all 26 cases used in immunofluorescence validation (Table 1). **b,** Borders of the EC were determined with the aid of 400  $\mu\text{m}$  thick serial coronal sections of whole-brain hemispheres stained with gallocyanin (see Methods). Each H&E section (left) along with its corresponding immunofluorescence image (middle) was aligned to the most approximate gallocyanin section (right), in which the the dissecans layers (diss-1, diss-2, and diss-ext) characteristic of the caudal EC were easier to visualize. This was then used to guide delineation of the EC on the H&E and immunofluorescence sections. For more details on the cytoarchitectonic definitions used to define the caudal EC, please consult Heinsen *et al.*<sup>19</sup>.

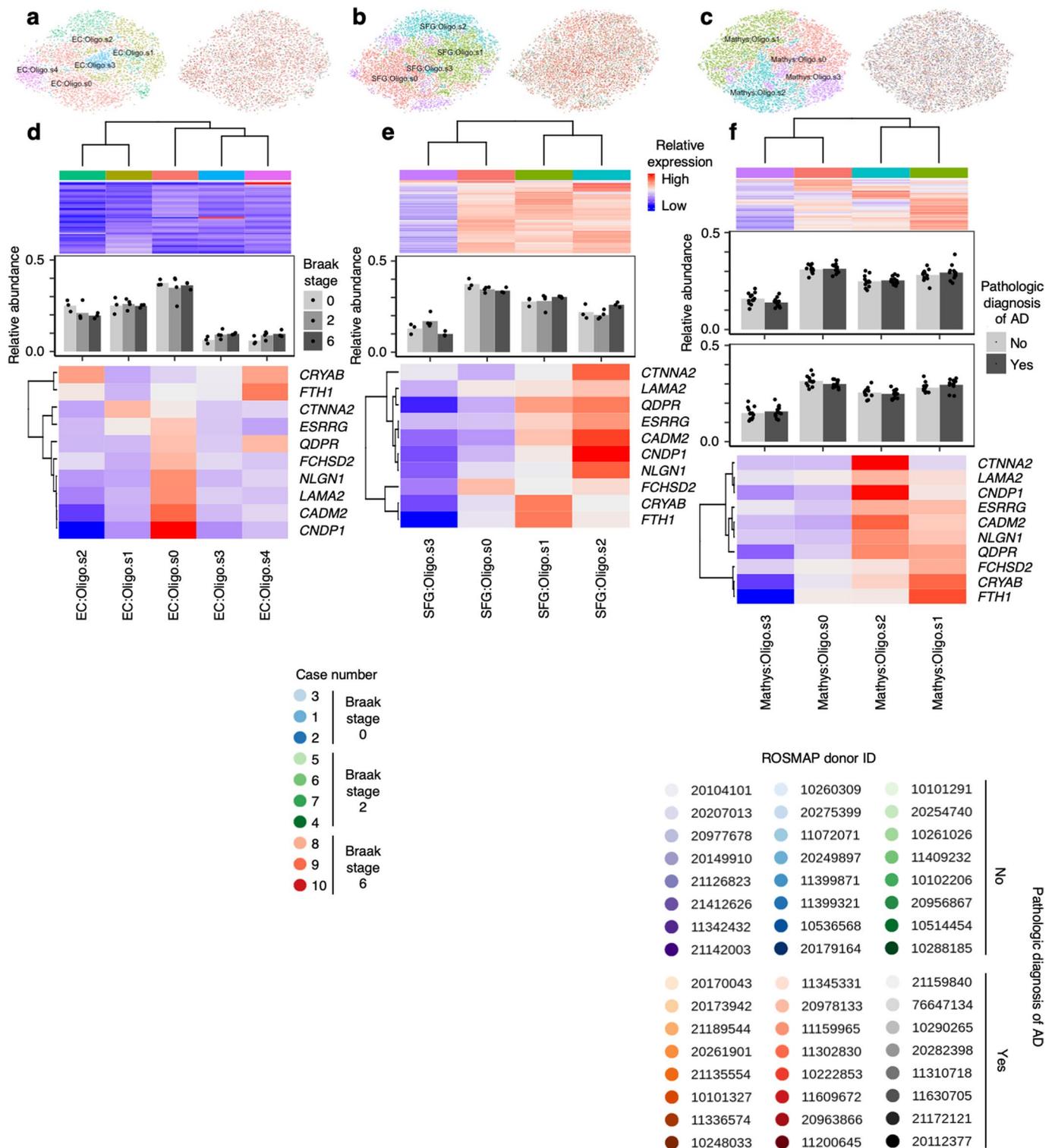


**Extended Data Fig. 7 | Inhibitory neurons from Mathys *et al.* also do not show differences in resilience or vulnerability to AD. a**, tSNE projection of inhibitory neurons from Mathys *et al.*<sup>14</sup> in the alignment space, colored by subpopulation identity (top) or individual of origin (bottom). **b**, Heatmap and hierarchical clustering of subpopulations and subpopulation markers (top subpanel); “High” and “Low” relative expression reflect above- and below-average expression, respectively (see Methods). Relative abundance of subpopulations in AD cases vs. controls, separated by sex (second and third subpanels); for each sex, statistical significance of differences in relative abundance between AD cases vs. controls (cases n=12, controls n=12, where n is the number of individuals sampled) was determined by beta regression and adjusted for multiple comparisons (see Methods). Expression heatmap of inhibitory neuron subtype markers from Lake *et al.*<sup>12</sup> (bottom subpanel).

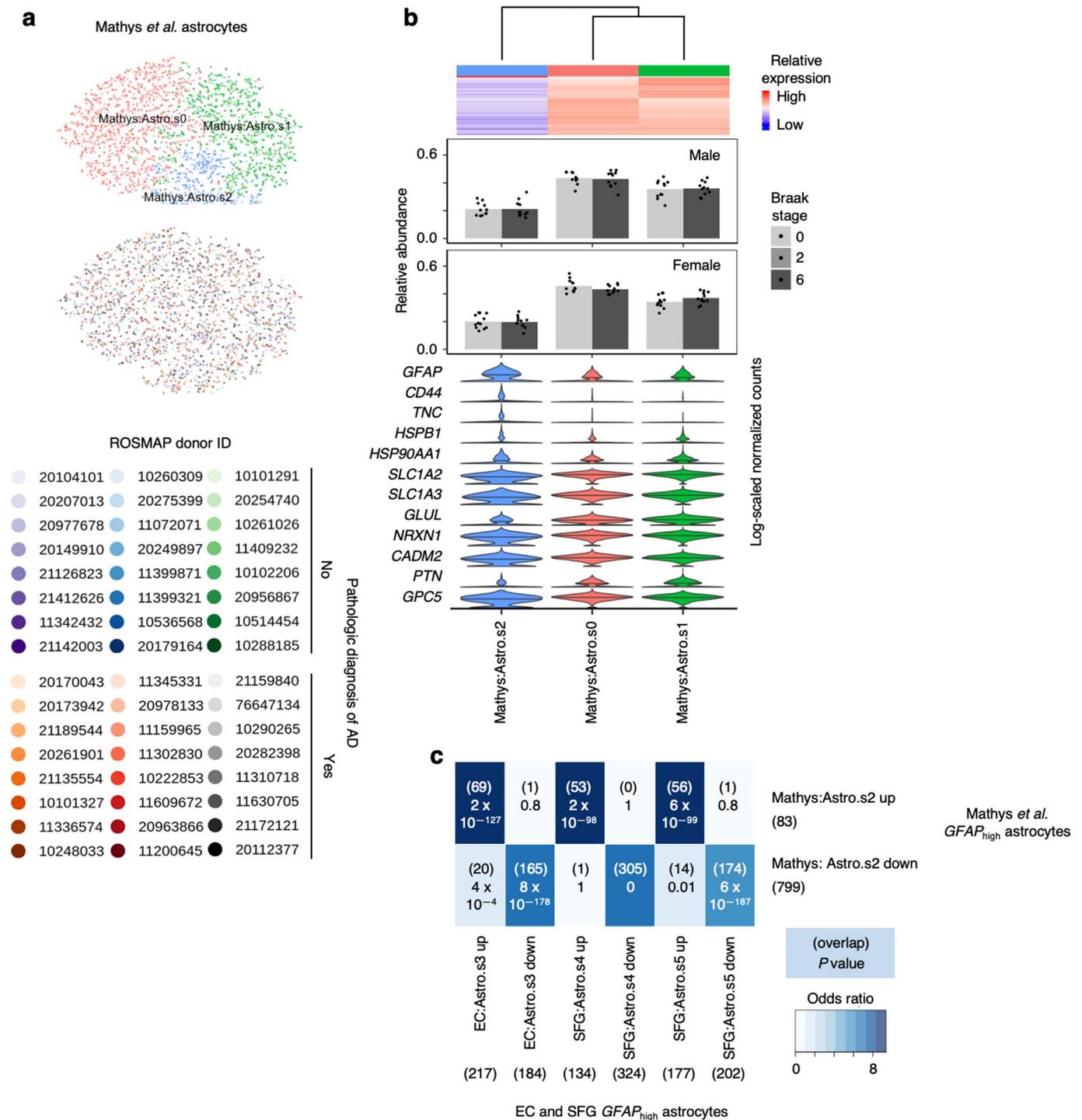


Extended Data Fig. 8 | See next page for caption.

**Extended Data Fig. 8 | Subclustering of microglia does not sufficiently resolve disease associated microglia signature.** **a-c,** tSNE projection of astrocytes from the EC (**a**), SFG (**b**), and Mathys *et al.*<sup>14</sup> (**c**) in their respective alignment spaces, colored by subpopulation identity (left) or individual of origin (right). **d-f,** Heatmap and hierarchical clustering of subpopulations and subpopulation marker expression (top subpanels); “High” and “Low” relative expression reflect above- and below-average expression, respectively (see Methods). Relative abundance of subpopulations (middle subpanels) across Braak stages in the EC and SFG (for each brain region, Braak 0 n=3, Braak 2 n=4, Braak 6 n=3, where n is the number of individuals sampled) or between AD cases vs. controls in Mathys *et al.*<sup>14</sup> (for each sex, cases n =12, controls n =12, where n is the number of individuals sampled); statistical significance of differences in relative abundance was determined by beta regression and adjusted for multiple comparisons (see Methods). Expression of disease associated microglia markers, with median expression level marked by line (bottom subpanels).



**Extended Data Fig. 9 | Subclustering of oligodendrocytes identifies subpopulations with higher expression of AD-associated oligodendrocyte markers from Mathys et al.** **a-c**, tSNE projection of oligodendrocytes from the EC (**a**), SFG (**b**), and Mathys et al.<sup>14</sup> (**c**) in their respective alignment spaces, colored by subpopulation identity (left) or individual of origin (right). **d-f**, Heatmap and hierarchical clustering of subpopulations and subpopulation marker expression (top subpanels); “High” and “Low” relative expression reflect above- and below-average expression, respectively (see Methods). Relative abundance of subpopulations (middle subpanels) across Braak stages in the EC and SFG (for each brain region, Braak 0 n=3, Braak 2 n=4, Braak 6 n=3, where n is the number of individuals sampled) or between AD cases vs. controls in Mathys et al.<sup>14</sup> (for each sex, cases n=12, controls n=12, where n is the number of individuals sampled); statistical significance of differences in relative abundance was determined by beta regression and adjusted for multiple comparisons (see Methods). Relative expression of AD-associated oligodendrocyte subpopulation markers from Mathys et al.<sup>14</sup> (bottom subpanels).



**Extended Data Fig. 10 | Astrocyte subpopulations with high GFAP expression from Mathys *et al.* are highly similar to those from the EC and SFG.**

**a**, tSNE projection of astrocytes from Mathys *et al.*<sup>14</sup> in the alignment subspace, colored by subpopulation identity (top) or individual of origin (bottom).

**b**, Heatmap and hierarchical clustering of subpopulations and subpopulation marker expression (top subpanel); “High” and “Low” relative expression reflect above- and below-average expression, respectively (see Methods). Relative abundance of subpopulations in AD cases vs. controls, separated by sex (middle subpanels); for each sex, statistical significance of differences in relative abundance between AD cases vs. controls (cases n=12, controls n=12, where n is the number of individuals sampled) was determined by beta regression and adjusted for multiple comparisons (see Methods). Expression of genes associated with reactive astrocytes, with median expression level marked by line (bottom subpanel).

**c**, Enrichment analysis of overlap between differentially expressed genes in astrocytes with high GFAP expression from Mathys *et al.*<sup>14</sup> vs. differentially expressed genes in astrocytes with high GFAP expression from the EC and SFG; the number of genes in each gene set and the number of overlapping genes are shown in parentheses, and the hypergeometric test p-values are shown without parentheses.

## Reporting Summary

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

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

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

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

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

Data collection

Quantification of cell numbers from immunofluorescence validation was performed with ImageJ (version 1.8.0).

Data analysis

Alignment of raw sequencing data from snRNA-seq was performed using Cellranger (version 2.1.0). Analysis of snRNA-seq data was performed in R (version 3.5.1) using the following packages: DropletUtils (version 1.2.2.), scran (version 1.10.2), scater (version 1.10.1), Seurat (version 2.3.4), scAlign (version 1.0.0), edgeR (version 3.24.3), and betareg (version 3.1-1). Network visualization and pathway enrichment analysis was performed using String-db (v11) and Cytoscape (version 3.7.2). We provide the full bioinformatics pipeline for the analysis of snRNA-Seq data in this paper at <https://kampmannlab.ucsf.edu/ad-brain-analysis>.

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

### Data

Policy information about [availability of data](#)

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

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

The raw snRNA-seq sequencing data and unfiltered UMI count matrices are available on the Gene Expression Omnibus (GEO) under the accession GSE147528. Single-cell data after quality control is available for download in synapse.org at under the Synapse ID syn21788402 (<https://www.synapse.org/#!Synapse:syn21788402>). Post- quality control data can also be explored interactively through the CellXGene platform at <https://kampmannlab.ucsf.edu/ad-brain>. Data from Mathys et al.<sup>21</sup> was downloaded from Synapse under Synapse ID syn18485175 (<https://www.synapse.org/#!Synapse:syn18485175>, "Data/Gene Expression (RNA-seq)/Processed/filtered\_count\_matrix.mtx"). Cytoscape was downloaded from <https://cytoscape.org/> and String-db was accessed at <https://string-db.org/>.

## Field-specific reporting

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

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

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

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

Sample size	The brain samples used in this study contained a broad burden of AD-type pathology and were selected to be free from non-AD pathology including Lewy body disease, TDP-43 proteinopathies, primary tauopathies, and cerebrovascular changes. Argyrophilic grain disease (AGD) was not an exclusion criterion based on its high prevalence and lack of correlation with significant clinical symptoms. In total, the cohort included 10 cases who underwent snRNA-seq, representing Braak stages 0, 2 and 6, all ApoE 3/3, and 26 cases who underwent neuroanatomical analysis, representing Braak stages 0-6, ranging from 2-5 individuals per Braak stage. We did not perform statistical analysis to per-determine sample size given that the snRNA-seq study was meant to be exploratory and that we expected to focus on large effects. The sample sizes that we chose are comparable to previous publications (Mathys et al. and Marinaro et al.).
Data exclusions	None of the samples described above were excluded from the relevant analyses.
Replication	The decrease in the relative abundance of RORB+ excitatory neurons in the entorhinal cortex observed in the snRNA-seq data was confirmed by immunofluorescence in a larger cohort of cases. The validation experiment was performed once.
Randomization	Since our study employed a case-control design, randomization does not apply. We matched cases and controls for sex and APOE genotype, and for age as much as possible. Exact matching for age was not possible given the extremely low availability of Braak 0 cases with ages matching those of Braak 2 and Braak 6 cases.
Blinding	The experimenter was blinded to the Braak stage for quantification of cell numbers for immunofluorescence validation.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

### Methods

n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

## Antibodies

Antibodies used	Primary antibodies: RORB (1:400, rabbit, HPA008393, Millipore Sigma), TBR1 (1:100, rabbit, ab31940, Abcam), CP13 (1:800, mouse, phospho-tau serine 202, gift of Peter Davies, NY; the original source of the antibody is a non-commercial hybridoma). Secondary antibodies: goat anti-rabbit HRP (1:400, R-05072-500, Advansta), AlexaFluor 546 conjugated goat anti-rabbit IgG (1:200, A-11010, Thermo Fisher), biotinylated horse anti-mouse IgG (1:400, BA-2000, Vector Laboratory).
Validation	Each primary antibody was validated for immunofluorescence on human brain slices by testing for the expected staining pattern based on which cell types are known to express the target. The RORB and TBR1 antibodies has been validated for immunofluorescence and immunohistochemistry respectively on their manufacturer's websites.