

1 Single nucleus sequencing fails to detect microglial activation in human tissue

2 *N. Thrupp*^{1,2}, *C. Sala Frigerio*^{1,2,3}, *L. Wolfs*^{1,2}, *N. G. Skene*⁴, *S. Poovathingal*^{1,2}, *Y. Fourné*^{1,2}, *P. M. Matthews*⁴, *T.*
3 *Theys*⁵, *R. Mancuso*^{1,2}, *B. de Strooper*^{1,2,3,6*}, *M. Fiers*^{1,2,3*}

4

5 ¹ *Centre for Brain and Disease Research, Flanders Institute for Biotechnology (VIB), Leuven, Belgium*

6 ² *Department of Neurosciences and Leuven Brain Institute, KU Leuven, Leuven, Belgium*

7 ³ *UK Dementia Research Institute at UCL, University College London, London, UK*

8 ⁴ *UK Dementia Research Institute at Imperial College London and Department of Brain Science, Imperial College London, London, UK*

9 ⁵ *Department of Neurosciences, Research group experimental neurosurgery and neuroanatomy, KU Leuven, Leuven, Belgium*

10 * *corresponding authors*

11 ⁶ *lead contact*

12

13 **Abstract**

14 Single nucleus RNA-Seq (snRNA-Seq) methods are used as an alternative to single cell
15 RNA-Seq methods, as they allow transcriptomic profiling of frozen tissue. However, it is
16 unclear whether snRNA-Seq is able to detect cellular state in human tissue. Indeed, snRNA-
17 Seq analyses of human brain samples have failed to detect a consistent microglial activation
18 signature in Alzheimer's Disease. A comparison of microglia from single cells and single
19 nuclei of four human subjects reveals that ~1% of genes is depleted in nuclei compared to
20 whole cells. This small population contains 18% of genes previously implicated in microglial
21 activation, including *APOE*, *CST3*, *FTL*, *SPPI*, and *CD74*. We confirm our findings across
22 multiple previous single nucleus and single cell studies. Given the low sensitivity of snRNA-
23 Seq to this population of activation genes, we conclude that snRNA-Seq is not suited to
24 detecting cellular activation in microglia in human disease.

25

26 **Keywords**

27 Microglia ; activation ; Alzheimer's Disease ; single nucleus RNA-Seq

28

29

30 **Introduction**

31

32 Single cell approaches allow us to study cell-to-cell heterogeneity (Habib et al., 2017), in
33 brain material however, it is difficult to dissociate individual cells (Habib et al., 2017; Lake et
34 al., 2016). This is further complicated if one is interested in studying the human brain, where
35 often only frozen material is available. One alternative to study cellular transcriptional
36 heterogeneity in brain tissue is single nucleus transcriptomics. Single nucleus RNA-Seq
37 (snRNA-Seq) studies have shown concordance between single cell and single nucleus
38 transcriptome profiles in mice (Bakken et al., 2018; Habib et al., 2017; Lake et al., 2017), but
39 have limited the comparison to the identification of major cell types. It is unclear whether a
40 snRNA-Seq approach is equally effective in identifying dynamic cellular substates such as
41 microglial activation in human tissue.

42

43 A recent breakthrough in the field of Alzheimer's Disease (AD) using single cell RNA-Seq
44 (scRNA-Seq) demonstrated clearly that microglia become activated in response to amyloid
45 plaques in mouse models (Keren-Shaul et al., 2017). This response comprises a
46 transcriptional switch to a state called Activation Response Microglia (ARM) (Sala Frigerio
47 et al., 2019), or Disease-Associated Microglia (DAM, MGnD) (Keren-Shaul et al., 2017;
48 Krasemann et al., 2017). Ample evidence suggests that this microglial response is also
49 relevant in human AD: microglia are believed to play a role in amyloid clearance (Efthymiou
50 and Goate, 2017) and complement-mediated synapse loss (Fonseca et al., 2017), and
51 histological studies have demonstrated considerable microgliosis around plaques in humans
52 (McGeer et al., 1987). In addition, there is significant overlap between those genes involved
53 in the microglial response, and genes within loci carrying AD genetic risk, as identified in
54 Genome-Wide Association Studies (GWAS) (Efthymiou and Goate, 2017; Jansen et al.,
55 2019; Kunkle et al., 2019; Lambert et al., 2013; Marioni et al., 2018), for example, *APOE*,
56 *TREM2*, *APOC1*, *CD33* (Sala Frigerio et al., 2019). Most recently, the engrafting of human
57 microglia into AD mouse models, followed by single cell RNA-sequencing, identified 66
58 DAM genes relevant to human activation¹⁵, and a bulk RNA-Seq study of AD patients
59 identified 64 DAM genes¹⁶. In stark contrast, a number of high-profile snRNA-Seq studies of
60 microglia in human AD (Del-Aguila et al., 2019; Grubman et al., 2019; Mathys et al., 2019;
61 Zhou et al., 2020) have not recovered a consistent microglial activation signature. A recent

62 cluster analysis by Mathys *et al.* of 48 AD patients and controls reported only 28 of 257
63 orthologous activation genes in common with the DAM signature (Mathys *et al.*, 2019).
64 Differential expression analysis between AD and control patients revealed 22 genes
65 upregulated in AD patients (5 overlapping with the DAM signature). Of these AD genes, only
66 8 were also upregulated in another snRNA-Seq study of human AD (Grubman *et al.*, 2019),
67 and only 4 were also upregulated in another snRNA-Seq study of AD TREM2 variants (Zhou
68 *et al.*, 2020). The AD TREM2 variant study also only identified 11 DAM genes enriched in
69 AD patients compared with controls. Del Aguila *et al.*, analysing single nucleus
70 transcriptomics from 3 AD patients, were unable to recapitulate an activation signature (Del-
71 Aguila *et al.*, 2019). This has led to speculation that there is no such DAM signature in
72 humans.

73

74 Here we compared the performance of snRNA-Seq to scRNA-Seq for the analysis of
75 microglia from human cortical biopsies, and demonstrated that technical limitations inherent
76 to snRNA-Seq provide a more likely explanation for this lack of consistency in snRNA-Seq
77 studies of AD. We confirmed our results using publicly-available data.

78

79 **Results**

80

81 **snRNA-Seq recovers major cell types from human tissue, but not microglial state**

82

83 scRNA-Seq of FACS-sorted microglia was performed on temporal cortices of four human
84 subjects who had undergone neocortical resection (see Supplementary Table 1 for subject
85 data)(Mancuso *et al.*, 2019). We generated snRNA-Seq libraries from these same subjects.
86 Following quality filtering, PCA analysis and clustering of 37,060 nuclei, we identified 7
87 major cell types (Supplementary Fig. 1a, b): oligodendrocytes (ODC, 34.0%), excitatory
88 neurons (27.0%), interneurons (11.2%), oligodendrocyte precursors (OPC 9.4%), microglia
89 (11.3%), astrocytes (6.0%), and endothelial cells (1.1%). We focus here on the microglial
90 population, which was extracted from the main dataset.

91

92 We first checked whether clustering analysis of single nuclei could recover subpopulations of
93 microglia comparable to the single cell approach. A comparison of single nucleus and single
94 cell clustering suggested that we could only partially recover similar microglial subcluster
95 structure using both methods (see Supplementary Text and Supplementary Fig. 1c-e).

96

97 **Gene expression profiling of human nuclei and cells**

98

99 To compare gene abundance in single microglial cells (14,823 cells) and nuclei (3,940
100 nuclei), we performed a differential abundance analysis between cells and nuclei from the 4
101 subjects (Fig. 1a). As demonstrated in previous studies (Bakken et al., 2018; Gerrits et al.,
102 2019; Habib et al., 2017; Lake et al., 2017), the majority of genes showed similar normalized
103 abundance levels in cells and nuclei, with 98.6% of genes falling along the diagonal in Fig.
104 1a (Pearson's correlation coefficient = 0.92 , $p < 2.2e-16$). However, we identified a group of
105 246 genes (1.1% of detected genes) that was less abundant in nuclei (fold change < -2 , $p_{adj} <$
106 0.05 , blue in Fig. 1a). A second population of 68 genes (0.3%) was found to be more
107 abundant in nuclei (fold change > 2 , $p_{adj} < 0.05$, red in Fig. 1a). Additionally, 3,248 genes
108 were exclusively detected in cells, and 5,068 genes exclusively detected in nuclei.

109

110 The observed differences in abundance between cells and nuclei were consistent across all
111 four subjects (Fig. 1b, Supplementary Fig. 2a). Downsampling of cellular reads indicated
112 that differences in abundance were not the result of different sequencing depths
113 (Supplementary Fig. 2b,c). The full differential abundance results can be found in
114 Supplementary Table 2.

115

116 To assess the robustness of this finding, we used our nuclei-abundant genes and cell-
117 abundant genes to compare enrichment across all pairs of 8 publicly-available single cell or
118 single nucleus datasets (Supplementary Table 3, Fig. 1c). We consistently found our nuclei-
119 depleted (cell-abundant) genes to be depleted in other single nucleus microglia compared to
120 single cell microglia (mean microglial Z-score of cell-abundant genes was 7.95 when
121 comparing cells to nuclei, whereas cell-to-cell comparisons yielded a mean of 0.01, and
122 nuclei-to-nuclei comparisons yielded a mean of 0.81, for Z-scores with $p_{adj} < 0.05$). We also

123 found our nuclei-abundant genes to be consistently enriched in other microglial nuclei
124 compared with microglial cells (mean microglial Z-score -2.99 compared to -2.33 in nuclei
125 against nuclei, no significant enrichment was found in cell-to-cell comparisons with $p_{\text{adj}} <$
126 0.05).

127

128 To assess functional enrichment among genes found to be more abundant in cells or nuclei,
129 we ranked all genes according to log fold change (genes with a low abundance in nuclei had a
130 negative log fold change) and performed a Gene Set Enrichment Analysis (GSEA,
131 Subramanian et al., 2005) against gene markers from previous studies (Fig. 2a). For these
132 analyses, a positive Normalised Enrichment Score (NES) represented nuclear enrichment,
133 and a negative NES represented nuclear depletion. As expected, cytoplasmic RNA (defined
134 by Bahar Halpern et al., 2015) was clearly enriched among genes found to be more abundant
135 in cells (NES = -1.98, $p_{\text{adj}} = 3.6\text{e-}05$), as was mitochondrial mRNA (NES = -1.71, $p_{\text{adj}} = 1.6\text{e-}$
136 04, gene set extracted from Ensembl's BioMart (Zerbino et al., 2018)). mRNA found to be
137 more abundant in the nucleus by (Bahar Halpern et al., 2015) tended towards enrichment in
138 nuclei but was not significant (NES = 0.87, $p_{\text{adj}} = 8.2\text{e-}01$), which is to be expected as
139 scRNA-Seq captures both nuclear and cytoplasmic RNA. RNA of genes coding for ribosomal
140 proteins were also depleted in nuclei (NES = -2.28, $p_{\text{adj}} = 3.6\text{e-}05$), as previously described¹.
141 Genes with shorter coding sequences (CDS) were depleted in nuclei (NES = -1.38, $p_{\text{adj}} =$
142 2.5e-02), while longer CDS were enriched (NES = 2.07, $p_{\text{adj}} = 2.1\text{e-}05$), as already observed
143 in earlier snRNA-Seq studies³. Finally, the genes defined by Gerrits (Gerrits et al., 2019) as
144 cellular-enriched in a differential analysis of microglial cells versus (fresh) nuclei in humans
145 were also enriched in cells in our data, showing a NES score of -2.15 ($p_{\text{adj}} = 3.6\text{e-}05$).

146

147 To further characterise genes with higher or lower abundance in nuclei compared with cells,
148 we performed GSEA, using Gene Ontology (GO) terms extracted from MSigDb (Liberzon et
149 al., 2011) against the ranked log fold change. We selected the 100 terms with the highest
150 NES, and the 100 terms with the lowest NES ($p_{\text{adj}} < 0.05$). Given the high overlap in terms,
151 we clustered ontology terms based on the number of shared genes, in order to define “super”
152 GO clusters (Supplementary Fig. 2e,f). We repeated the GSEA analysis using these *super-*
153 GO clusters (Fig. 2b, Supplementary Tables 4 and 5) and observed an enrichment of neuronal
154 and synaptic terms in nuclei-abundant genes (also shown in the red population in Fig.1a). We

155 suspect a synaptosome contamination during centrifugation. This is supported by the
156 enrichment of synaptosome genes (NES = 1.82, $p_{\text{adj}} = 3.6\text{e-}05$, Fig. 2a; Hafner et al., 2019)
157 and ambient RNA – mRNA originating not from cells/nuclei but from free-floating
158 transcripts in the solution (Macosko et al., 2015) – (NES = 1.71, $p_{\text{adj}} = 9.2\text{e-}05$, Fig. 2a,
159 Supplementary Fig. 2d) within the nucleus-abundant genes. The two gene sets share a strong
160 overlap (Supplementary Table 6). These genes, although enriched in nuclei compared with
161 cellular levels, still show low abundance (most of these genes show a normalised abundance
162 of no more than 2 – Fig. 1a).

163

164 **Activation genes identified in mouse models of AD are depleted in human nuclei**

165

166 More interesting was the depletion of immune-related genes in nuclei (Fig. 2b). We therefore
167 tested whether microglial activation genes^{7,8,17} were also depleted in nuclei (Fig. 2c,
168 Supplementary Tables 5 and 6). Remarkably, we found a strong depletion of genes associated
169 with mouse microglial activation: 45 of 257 orthologous DAM genes (Keren-Shaul et al.,
170 2017), (NES = -2.16, $p_{\text{adj}} = 3.6\text{e-}05$, Fig. 2c,d), and 28 of 200 orthologous ARM genes (Sala
171 Frigerio et al., 2019) (NES = -2.01, $p_{\text{adj}} = 3.6\text{e-}05$, Fig. 2c,e), confirming that mouse
172 microglial activation genes were less abundant in nuclei. Genes upregulated by LPS
173 stimulation in mice (Gerrits et al., 2019) also showed depletion in nuclei (NES = -1.86, $p_{\text{adj}} =$
174 $3.6\text{e-}05$, Fig. 2c, Supplementary Figure 2g).

175

176 **Activation genes identified in mouse studies of AD are depleted in human nuclei**

177

178 We next examined genes that were identified as markers of the human microglial response to
179 AD in the recent snRNA-Seq study by (Mathys et al., 2019) (Fig. 2c, f, g). Markers of this
180 response (referred to by Mathys *et al.* as “Mic1”) had a NES score of -2.14 ($p_{\text{adj}} = 3.6\text{e-}05$),
181 indicating that they were depleted in nuclei (Fig. 2c). The study identified 28 DAM genes as
182 marker genes of the Mic1 response cluster (shown in orange in Fig. 2f); however the majority
183 of DAM genes were not recovered using their snRNA-Seq protocol (purple in Fig. 2f). Fig.
184 2g shows in green all the markers of the human activation cluster Mic1. Clearly, DAM genes
185 and other Mic1 markers showed higher abundance in cells relative to nuclei (confirming the

186 NES score in Fig. 2c). Further, it seems likely that the recovered DAM genes (orange in Fig.
187 2f) and *Mic1* markers in general (green in Fig. 2g), were detected in the original snRNA-Seq
188 experiment owing to their higher nuclear abundance compared with the nuclear abundance of
189 other genes, including those DAM genes that were not recovered (purple in Fig. 2f).

190

191 **Discussion**

192

193 In summary, in our comparison of snRNA-Seq and scRNA-Seq of human microglia, we
194 identified a set of genes (1.1% of the gene population) with at least 2-fold lower abundance in
195 nuclei compared to their cellular levels (Fig. 1a-b). This small set is strongly enriched for
196 genes previously associated with microglial activation in mouse models of AD, for example
197 *APOE*, *CST3*, *FTL*, *SPP1*, and *CD74* (Fig. 2b-e). Thus, while our work agrees with previous
198 experiments demonstrating that snRNA-Seq can determine cell type (Supplementary Fig.
199 1a,b), we argue that there are important limitations when studying cellular state in humans.
200 This limitation is likely responsible for the difficulty in identifying consistent DAM- or
201 ARM-like gene populations in the human brain in snRNA-Seq-based studies. We identified
202 similar patterns of depletion in other single nucleus microglia (Fig. 2c).

203

204 Examination of data from the Mathys *et al.* study of human nuclei in AD (Mathys *et al.*,
205 2019) shows that only genes with higher nuclear abundance levels were detected (Fig. 2f, g).
206 This suggests that the discordance between human and mouse microglial activation is at least
207 in part a consequence of limitations in the technology, rather than biological differences
208 between the species as current snRNA-Seq suggest. Deeper sequencing (or increased sample
209 sizes) may compensate for this lack of sensitivity. However, the sparse nature of snRNA-Seq
210 and the high level of heterogeneity in human samples, combined with the fact that many
211 relevant genes have a more than two-fold lower abundance in nuclei (*e.g.* *APOE* fold change
212 = 2.57, *CST3* fold change = 3.44, *FTL* fold change = 6.53), strongly suggests that this will
213 remain a problem.

214

215 While our data is (at least partially) in agreement (Fig. 2a, c) with Gerrits *et al.* (Gerrits *et al.*,
216 2019) which also compares nuclei with cells in human microglia, they did not report a

217 nuclear depletion of activation genes. We suspect the reason for this is (a) the low human
218 sample number (n=2); (b) for the cluster analysis, Gerrits *et al.* scaled cell and nuclei
219 expression to mitochondrial and ribosomal reads, essentially masking differences between
220 nuclei and cells, and (c) for the differential expression analysis, Gerrits *et al.* compared fresh
221 cells to fresh nuclei, as opposed to frozen nuclei.

222

223 Alternative approaches may be more suitable for generating a brain atlas of human disease
224 such as AD, particularly where we are limited to frozen material. *In situ* spatial
225 transcriptomics (ST) negates issues related to tissue dissociation and cell or nucleus isolation
226 (Ståhl *et al.*, 2016), while at the same time retaining spatial information. This approach has
227 recently been applied to examine transcriptomic changes and identify genes that are co-
228 expressed across multiple cell types in the amyloid plaque niche of the mouse brain (Chen *et*
229 *al.*, 2019). In humans, a similar methodology was recently applied to identify pathway
230 dysregulation and regional differences in cellular states of the postmortem spinal tissue of
231 Amyotrophic Lateral Sclerosis (ALS) patients states (Maniatis *et al.*, 2019). Its application to
232 AD patients may shed light on transcriptomic changes occurring in microglia which localize
233 near plaques, and may also provide insights into the crosstalk occurring between
234 neighbouring cells.

235

236 In conclusion, while snRNA-Seq offers a viable alternative to scRNA-Seq for identification
237 of cell types in tissue where cell dissociation is problematic, its utility for detecting cellular
238 states in disease is limited.

239

240 **Data Availability**

241

242 Sequencing data from single microglial cells is available on GEO (accession number
243 GSE137444). Sequencing data from single nuclei will be made available on GEO.

244

245 **Code Availability**

246

247 Analysis of previous datasets was performed using the EWCE package (Skene and Grant,
248 2016) for R and the MicroglialDepletion package
249 (<https://github.com/NathanSkene/MicroglialDepletion>).

250

251 **Acknowledgements**

252

253 Work in the De Strooper laboratory was supported by the European Union (ERC- 834682 -
254 CELLPHASE_AD), the Fonds voor Wetenschappelijk Onderzoek (FWO), KU Leuven,
255 VIB, UK-DRI (Medical Research council, ARUK and Alzheimer Society), a Methusalem
256 grant from KU Leuven and the Flemish Government, the “Geneeskundige Stichting Koningin
257 Elisabeth”, Opening the Future campaign of the Leuven Universitair Fonds (LUF), the
258 Belgian Alzheimer Research Foundation (SAO-FRA) and the Alzheimer’s Association USA.
259 Bart De Strooper is holder of the Bax-Vanluffelen Chair for Alzheimer’s Disease. Cell
260 sorting was performed at the KU Leuven FACS core facility, and sequencing was carried out
261 by the VIB Nucleomics Core. Renzo Mancuso is recipient of a postdoctoral fellowship from
262 the Alzheimer’s Association, USA. Paul M. Matthews acknowledges generous support from
263 the Edmond J. Safra Foundation and Lily Safra, the NIHR Biomedical Research Centre at
264 Imperial College and the UK Dementia Research Institute.

265

266 **Author contributions**

267

268 Conceptualization: B.D.S., M.F., C.S.F., R.M. ; surgery and extraction of patient tissue
269 samples: T.T. ; Investigation: R.M., L.W., S.P. ; formal analysis: N.T., Y.F., M.F., N.G.S.,
270 P.M.M. ; Writing – original draft: N.T., C.S.F., M.F. ; writing – review & editing: B.D.S.,
271 R.M., N.G.S., P.M.M. ; supervision: M.F., B.D.S.

272

273 **Competing interests**

274

275 The authors declare that there are no competing interests.

276

277 **Materials and Methods**

278

279 **Isolation of human primary microglial cells**

280 Human primary microglial cells from the Mancuso *et. al* study (Mancuso et al., 2019) were
281 used. Briefly, microglia were FAC-sorted from brain tissue samples resected from the lateral
282 temporal neocortex of 4 epilepsy patients during neurosurgery. The full protocol is described
283 in the original study. All procedures followed protocols approved by the local Ethical
284 Committee (protocol number S61186). Sequencing was performed as described for the nuclei.
285

286 **Isolation of nuclei from human subjects**

287 Nuclei from frozen biopsy tissue were isolated as follows: brain tissue was sliced on dry ice,
288 then homogenized manually (15 gentle strokes) with 1mL ice-cold Homogenisation Buffer
289 (HB) with 5 μ L Rnasin Plus. The homogenate was strained with a 70 μ m strainer and washed
290 with 1.65mL to a final volume of 2.65mL. 2.65mL of Gradient medium was added ($V_f =$
291 5.3mL). To isolate the nuclei, the sample was added to a 4mL 29% cushion using a P1000,
292 and the weight adjusted with HB. The sample was centrifuged in a SW41Ti rotor at 7,700
293 rpm for 30 minutes at 4°C. The supernatant was removed with a plastic Pasteur pipette,
294 followed by removal of the lower supernatant with P200. Nuclei were resuspended in 200 μ L
295 of resuspension buffer, transferred to a new tube, washed again with 100-200 μ L resuspension
296 buffer, and pooled with the previous solution. Clumps were disrupted by pipetting with P200,
297 then filtered through a Falcon tube with 0.35 μ m strainer. 9 μ L of sample was mixed with 1 μ L
298 of propidium iodide (PI) stain, loaded onto a LUNA-FL slide and allowed to settle for 30
299 seconds. We viewed nuclei with the LUNA-FL Automated cell counter to check numbers and
300 shape.

301

302 **Single nucleus sequencing**

303 RNA sequencing was performed using the 10X Genomics Single Cell 3' Reagent Kit (v2)
304 according to manufacturer protocols. cDNA libraries from fresh-frozen nuclei were
305 sequenced on an Illumina HiSeq platform 4000. Supplementary Table 1 provides sequencing
306 information per sample (for cells and nuclei).

307

308 **Single nucleus analysis**

309 *Alignment.* Cellranger v2.1.1 was used to demultiplex and align sequencing output to a
310 human reference genome (assembly hg38 build 95). We used a “pre-mRNA” database to
311 align single nuclei to exons as well as introns (10x Genomics, n.d.). Following alignment,
312 nuclei from one patient sample (RM101.1) were removed due to poor quality (low read and
313 gene count). See Supplementary Table 1 for sample information. Unfiltered count matrices
314 were used for downstream analysis.

315 *Extraction of microglial nuclei.* Data was processed using the Seurat v3.0.2 package (Butler
316 et al., 2018; Stuart et al., 2019) in R v3.6.1. For each patient, the count matrix was filtered to
317 exclude nuclei with fewer than 100 genes. Counts were depth-normalised, scaled by 10,000
318 and log-transformed. *FindVariableFeatures* was run using a variance-stabilising
319 transformation (“vst”) to identify the 2,000 most variable genes in each sample. Data from
320 the 4 patients was then integrated using Seurat’s *FindIntegrationAnchors* with default
321 parameters, and *IntegrateData* using 40 principal components (PCs). The dataset consisted of
322 37,060 nuclei, with a mean read depth of 4,305 counts per nucleus, and 1,791 genes per
323 nucleus. Integrated data was scaled (default Seurat parameters). We ran a Principal
324 Components Analysis (PCA), then calculated Uniform Manifold Approximation and
325 Projection (UMAP) embeddings using 40 PCs. We identified clusters using Seurat’s
326 *FindNeighbours* and *FindClusters* functions, again using 40 PCs. Based on abundance of
327 known celltype markers, we assigned each cluster to a cell type. Microglial clusters were
328 identified using known markers including *P2RY12*, *CSF1R*, *CX3CR1*, and extracted for
329 downstream analysis.

330 *Pre-processing of microglial nuclei per patient.* Microglia from each patient sample were
331 analysed individually as described for all cell types above, with the following modifications:
332 raw counts were filtered to remove genes and counts that were ± 3 standard deviations away
333 from the median value. After normalization, doublets were identified using DoubletFinder
334 v2.0.2 (McGinnis et al., 2019) using 40 PCs, assuming a 7.5% doublet rate. Following
335 removal of doublets, filtering and Seurat normalization were performed again. Data from
336 patients was then integrated and clusters were identified as above. We discarded small
337 clusters than contained markers for microglia as well as other cell types. After pre-
338 processing, 3,927 nuclei remained, with a mean count depth of 1,295 and a mean gene count
339 of 879 genes per nucleus.

340

341 **Single cell analysis**

342 Full details of single cell processing are available in Mancuso *et al.* (Mancuso et al., 2019).

343 Only cells from the four patients included in the single nucleus study were used here.

344

345 **Comparisons of single cells and single nuclei**

346 *Cluster analysis (see supplementary text).* In order to identify microglial cell states in the
347 nuclei data we calculated gene markers for each cluster using Seurat's *FindMarkers* function,
348 selecting only markers with a positive fold change. Gene markers for cell clusters were
349 extracted from the original Mancuso *et al.* study. Markers for nuclei and cells are available in
350 Supplementary Table 7. For the analysis, we kept the top 40 significant markers ($p_{\text{adj}} < 0.05$)
351 based on log fold change for the nuclear clusters and cellular clusters. For each nucleus, we
352 calculated the mean abundance levels of each cell cluster marker set against the aggregated
353 abundance of random control gene sets, using Seurat's *AddModuleScore* function. This gave
354 us the MS40 score for each marker set. We performed two-sided Fisher's Exact tests with
355 Benjamini Hochberg corrections to determine the overlap of cell cluster markers with nuclear
356 cluster markers (selecting the top 40 markers for each set), using the union of all genes in the
357 cell and nuclei datasets as a background ($p_{\text{adj}} < 0.05$ was considered significant).

358 *Differential Abundance.* We discarded all non-microglial clusters (brain macrophages,
359 neutrophils), leaving 3,721 nuclei and 14,435 cells. Differential abundance analysis was
360 performed with the Seurat package, using a two-sided Wilcoxon rank sum test, with a
361 Bonferroni correction for multiple testing. Genes with $p_{\text{adj}} < 0.05$ and fold change $> |2|$ were
362 considered significant. As Seurat applies a pseudocount of +1 to data before calculating log
363 fold changes, a fold change of 2 corresponds to a log fold change of 0.63. Log fold changes
364 calculated by Seurat were used for further analysis in gene set enrichment analysis.

365 *Scatter plots.* We calculated the mean of the normalized abundance levels for cells and for
366 nuclei, and log-transformed these values.

367 *Assessment of nuclear-enriched or cell-enriched gene sets in public scRNA-Seq and snRNA-*
368 *Seq datasets.* We followed the methodology described in (Skene et al., 2018): genes that were
369 significantly more abundant in nuclei or more abundant in cells (see "Differential abundance"
370 methodology above) were used, creating two gene sets. 8 public datasets were reduced to
371 contain six major cell types: pyramidal neurons, interneurons, astrocytes, interneurons,
372 microglia and oligodendrocyte precursors. Within each dataset, for each gene in our gene sets,

373 we calculated a celltype specificity score using the EWCE R package (Github version
374 committed July 29, 2019; Skene and Grant, 2016). For each pair of datasets, X and Y , we
375 subtracted the mean microglial specificity score of Y from X . We then calculated the same
376 scores for 10,000 random gene sets: the probability and z-score for the difference in specificity
377 for the dendritic genes is calculated using these. Finally, the depletion z-score for each gene set
378 was equal to: (mean subtracted microglial specificity score – bootstrapped mean) /
379 (bootstrapped standard deviation). A large positive z-score thus indicated that the gene set was
380 depleted in microglia of dataset Y relative to dataset X . Benjamini-Hochberg multiple testing
381 corrections were applied.

382 *Public datasets.* For the Karolinska Institutet (KI) dataset (Skene et al., 2018), we used S1
383 pyramidal neurons. For the Zeisel 2018 dataset (Zeisel et al., 2018) we used all ACTE* cells
384 as astrocytes, TEGLU* as pyramidal neurons, TEINH* as interneurons, OPC as
385 oligodendrocyte precursors and MGL* as microglia. For the Saunders dataset (Saunders et
386 al., 2018), we used all Neuron.Slc17a7 celltypes from the frontal cortex (FC), hippocampus
387 (HC) or posterior cortex (PC) as pyramidal neurons; all Neuron.Gad1Gad2 cell types from
388 FC, HC or PC as interneurons; Polydendrocyte as OPCs; Astrocyte as astrocytes, and
389 Microglia as microglia. The Lake datasets both came from a single publication (Lake et al.,
390 2018) which had data from frontal cortex, visual cortex and cerebellum. The cerebellum data
391 was not used here. Data from frontal and visual cortices were analyzed separately. All other
392 datasets - Dronc Human (Habib et al., 2017), Dronc Mouse (Habib et al., 2017), Allen
393 Institute for Brain Science (AIBS) (Hodge et al., 2019), Tasic (Tasic et al., 2016) and Habib
394 (Habib et al., 2016) – were used as described previously (Skene et al., 2018). Supplementary
395 Table 3 lists all datasets. An R package is available for the analysis at
396 <https://github.com/NathanSkene/MicroglialDepletion>.

397 *Functional analysis.* We performed Gene Set Enrichment Analysis (GSEA) using the R
398 package fgsea v1.8.0 (Sergushichev, 2016), using default parameters. Gene sets were mapped
399 against a list of genes ranked according to fold change between cellular abundance and
400 nuclear abundance. Gene ontology (GO) sets were obtained from MSigDB (Liberzon et al.,
401 2011; Subramanian et al., 2005). Other gene sets were obtained from previous studies (see
402 Supplementary Table 6). $p_{adj} < 0.05$ (Benjamini-Hochberg correction) was considered
403 significant.

404 *Clustering of gene ontology terms.* GSEA of GO terms resulted in many functional categories
405 with overlapping genes. In order to reduce this redundancy, the top and bottom 100 GO terms

406 according to normalized enrichment score (with $p_{adj} < 0.05$) were clustered as follows: a
407 Jaccard index (the size of the intersection of the two datasets, divided by the size of the union
408 of the two datasets, multiplied by 100) of the overlapping genes was calculated between each
409 significant GO set. The resulting similarity matrix was converted to a dissimilarity matrix,
410 and hierarchical clustering was performed on the matrix. We selected a k value of 16 to group
411 the GO terms based on the hierarchical clustering (see Supplementary Table 2). Gene sets
412 were merged, and each new “super” GO was assigned an annotation manually. GSEA
413 analysis was performed on these *super*-GO gene sets as described above.

414 *Gene sets from previous studies.* We extracted gene sets from previous studies for this
415 analysis. A full list of gene sets is available in Supplementary Table 6. Where data was
416 selected from mouse datasets, we converted the mouse gene to its human ortholog using R’s
417 BioMaRt package v2.40.5 (Durinck et al., 2009), selecting only orthologs that displayed 1-to-
418 1 orthology. For the ARM gene set we selected the top 200 ARM genes based on log fold
419 change (Sala Frigerio et al., 2019). For the Gerrits human gene set, we took the union of all
420 genes that showed significant differential abundance between cells and nuclei (microglia)
421 from donor 1 and donor 2 (Gerrits et al., 2019). For the LPS gene set, we took the union of all
422 genes significantly upregulated in LPS in cells and in nuclei (microglia) from the Gerrits
423 study (Gerrits et al., 2019).

424 *Downsampling of cell counts.* To match cell sequencing depth to nucleus sequencing depth
425 (see Supplementary Fig. 2b,c), we sampled without replacement the number of reads in the
426 cells by a proportion of 0.32, using the *downsampleMatrix* function of the DropletUtils R
427 package v1.4.3 (Griffiths et al., 2018; Lun et al., 2019). This resulted in a read depth of 1,304
428 compared with the original read depth of 3,979 reads per cell.

429 *Definition of ambient RNA profile in nuclei.* We extracted nuclei with less than 700 counts
430 from the original unfiltered raw count matrix of all cell types (resulting in 2,414 nuclei with a
431 mean read depth of 590), and summed the gene counts, under the assumption that these were
432 empty drops rather than nuclei. We took the top 150 genes to represent the ambient RNA
433 profile. The mean read depth of these genes in the empty drops was 121 reads per cell.

434

435

436 **References**

- 437 10x Genomics, n.d. Creating a Reference Package with cellranger mkref -Software -Single Cell Gene Expression -Official
438 10x Genomics Support [WWW Document]. URL [https://support.10xgenomics.com/single-cell-gene-](https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/advanced/references#premrna)
439 [expression/software/pipelines/latest/advanced/references#premrna](https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/advanced/references#premrna) (accessed 11.13.19).
- 440 Bahar Halpern, K., Caspi, I., Lemze, D., Levy, M., Landen, S., Elinav, E., Ulitsky, I., Itzkovitz, S., 2015. Nuclear Retention
441 of mRNA in Mammalian Tissues. *Cell Rep.* 13, 2653–2662. <https://doi.org/10.1016/j.celrep.2015.11.036>
- 442 Bakken, T.E., Hodge, R.D., Miller, J.A., Yao, Z., Nguyen, T.N., Aebermann, B., Barkan, E., Bertagnolli, D., Casper, T.,
443 Dee, N., Garren, E., Goldy, J., Graybuck, L.T., Kroll, M., Lasken, R.S., Lathia, K., Parry, S., Rimorin, C.,
444 Scheuermann, R.H., Schork, N.J., Shehata, S.I., Tieu, M., Phillips, J.W., Bernard, A., Smith, K.A., Zeng, H., Lein,
445 E.S., Tasic, B., 2018. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLOS*
446 *ONE* 13, e0209648. <https://doi.org/10.1371/journal.pone.0209648>
- 447 Butler, A., Hoffman, P., Smibert, P., Papalexi, E., Satija, R., 2018. Integrating single-cell transcriptomic data across different
448 conditions, technologies, and species. *Nat. Biotechnol.* 36, 411–420. <https://doi.org/10.1038/nbt.4096>
- 449 Chen, W.-T., Lu, A., Craessaerts, K., Pavie, B., Sala Frigerio, C., Mancuso, R., Qian, X., Lalakova, J., Kühnemund, M.,
450 Voytyuk, I., Wolfs, L., Snellinx, A., Munck, S., Jurek, A., Fernandez Navarro, J., Saido, T.C., Lundeberg, J., Fiers,
451 M., De Strooper, B., 2019. Spatial and temporal transcriptomics reveal microglia-astroglia crosstalk in the
452 amyloid- β plaque cell niche of Alzheimer’s disease (preprint). *Neuroscience*. <https://doi.org/10.1101/719930>
- 453 Del-Aguila, J.L., Li, Z., Dube, U., Mihindikulasuriya, K.A., Budde, J.P., Fernandez, M.V., Ibanez, L., Bradley, J., Wang, F.,
454 Bergmann, K., Davenport, R., Morris, J.C., Holtzman, D.M., Perrin, R.J., Benitez, B.A., Dougherty, J., Cruchaga,
455 C., Harari, O., 2019. A single-nuclei RNA sequencing study of Mendelian and sporadic AD in the human brain.
456 *Alzheimers Res. Ther.* 11, 71. <https://doi.org/10.1186/s13195-019-0524-x>
- 457 Durinck, S., Spellman, P.T., Birney, E., Huber, W., 2009. Mapping Identifiers for the Integration of Genomic Datasets with
458 the R/Bioconductor package biomaRt. *Nat. Protoc.* 4, 1184–1191. <https://doi.org/10.1038/nprot.2009.97>
- 459 Efthymiou, A.G., Goate, A.M., 2017. Late onset Alzheimer’s disease genetics implicates microglial pathways in disease risk.
460 *Mol. Neurodegener.* 12, 43. <https://doi.org/10.1186/s13024-017-0184-x>
- 461 Fonseca, M.I., Chu, S.-H., Hernandez, M.X., Fang, M.J., Modarresi, L., Selvan, P., MacGregor, G.R., Tenner, A.J., 2017.
462 Cell-specific deletion of C1qa identifies microglia as the dominant source of C1q in mouse brain. *J.*
463 *Neuroinflammation* 14, 48. <https://doi.org/10.1186/s12974-017-0814-9>
- 464 Friedman, B.A., Srinivasan, K., Ayalon, G., Meilandt, W.J., Lin, H., Huntley, M.A., Cao, Y., Lee, S.-H., Haddick, P.C.G.,
465 Ngu, H., Modrusan, Z., Larson, J.L., Kaminker, J.S., Brug, M.P. van der, Hansen, D.V., 2018. Diverse Brain
466 Myeloid Expression Profiles Reveal Distinct Microglial Activation States and Aspects of Alzheimer’s Disease Not
467 Evident in Mouse Models. *Cell Rep.* 22, 832–847. <https://doi.org/10.1016/j.celrep.2017.12.066>
- 468 Gerrits, E., Heng, Y., Boddeke, E.W.G.M., Eggen, B.J.L., 2019. Transcriptional profiling of microglia; current state of the
469 art and future perspectives. *Glia* n/a. <https://doi.org/10.1002/glia.23767>
- 470 Griffiths, J.A., Richard, A.C., Bach, K., Lun, A.T.L., Marioni, J.C., 2018. Detection and removal of barcode swapping in
471 single-cell RNA-seq data. *Nat. Commun.* 9, 1–6. <https://doi.org/10.1038/s41467-018-05083-x>
- 472 Grubman, A., Chew, G., Ouyang, J.F., Sun, G., Choo, X.Y., McLean, C., Simmons, R.K., Buckberry, S., Vargas-Landin,
473 D.B., Poppe, D., Pflueger, J., Lister, R., Rackham, O.J.L., Petretto, E., Polo, J.M., 2019. A single-cell atlas of
474 entorhinal cortex from individuals with Alzheimer’s disease reveals cell-type-specific gene expression regulation.
475 *Nat. Neurosci.* 1–11. <https://doi.org/10.1038/s41593-019-0539-4>
- 476 Habib, N., Avraham-Davidi, I., Basu, A., Burks, T., Shekhar, K., Hofree, M., Choudhury, S.R., Aguet, F., Gelfand, E.,
477 Ardlie, K., Weitz, D.A., Rozenblatt-Rosen, O., Zhang, F., Regev, A., 2017. Massively parallel single-nucleus
478 RNA-seq with DroNc-seq. *Nat. Methods* 14, 955–958. <https://doi.org/10.1038/nmeth.4407>
- 479 Habib, N., Li, Y., Heidenreich, M., Swiech, L., Avraham-Davidi, I., Trombetta, J.J., Hession, C., Zhang, F., Regev, A.,
480 2016. Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science* 353, 925–928.
481 <https://doi.org/10.1126/science.aad7038>

- 482 Hafner, A.-S., Donlin-Asp, P.G., Leitch, B., Herzog, E., Schuman, E.M., 2019. Local protein synthesis is a ubiquitous
483 feature of neuronal pre- and postsynaptic compartments. *Science* 364. <https://doi.org/10.1126/science.aau3644>
- 484 Hasselmann, J., Coburn, M.A., England, W., Figueroa Velez, D.X., Kiani Shabestari, S., Tu, C.H., McQuade, A.,
485 Kolahdouzan, M., Echeverria, K., Claes, C., Nakayama, T., Azevedo, R., Coufal, N.G., Han, C.Z., Cummings,
486 B.J., Davtyan, H., Glass, C.K., Healy, L.M., Gandhi, S.P., Spitale, R.C., Blurton-Jones, M., 2019. Development of
487 a Chimeric Model to Study and Manipulate Human Microglia In Vivo. *Neuron*.
488 <https://doi.org/10.1016/j.neuron.2019.07.002>
- 489 Hodge, R.D., Bakken, T.E., Miller, J.A., Smith, K.A., Barkan, E.R., Graybuck, L.T., Close, J.L., Long, B., Johansen, N.,
490 Penn, O., Yao, Z., Eggermont, J., Höllt, T., Levi, B.P., Shehata, S.I., Aevermann, B., Beller, A., Bertagnolli, D.,
491 Brouner, K., Casper, T., Cobbs, C., Dalley, R., Dee, N., Ding, S.-L., Ellenbogen, R.G., Fong, O., Garren, E.,
492 Goldy, J., Gwinn, R.P., Hirschstein, D., Keene, C.D., Keshk, M., Ko, A.L., Lathia, K., Mahfouz, A., Maltzer, Z.,
493 McGraw, M., Nguyen, T.N., Nyhus, J., Ojemann, J.G., Oldre, A., Parry, S., Reynolds, S., Rimorin, C.,
494 Shapovalova, N.V., Somasundaram, S., Szafer, A., Thomsen, E.R., Tieu, M., Quon, G., Scheuermann, R.H.,
495 Yuste, R., Sunkin, S.M., Lelieveldt, B., Feng, D., Ng, L., Bernard, A., Hawrylycz, M., Phillips, J.W., Tasic, B.,
496 Zeng, H., Jones, A.R., Koch, C., Lein, E.S., 2019. Conserved cell types with divergent features in human versus
497 mouse cortex. *Nature* 1–8. <https://doi.org/10.1038/s41586-019-1506-7>
- 498 Jansen, I.E., Savage, J.E., Watanabe, K., Bryois, J., Williams, D.M., Steinberg, S., Sealock, J., Karlsson, I.K., Hägg, S.,
499 Athanasiu, L., Voyle, N., Proitsi, P., Witoelar, A., Stringer, S., Aarsland, D., Almdahl, I.S., Andersen, F., Bergh,
500 S., Bettella, F., Bjornsson, S., Brækhus, A., Bråthen, G., de Leeuw, C., Desikan, R.S., Djurovic, S., Dumitrescu,
501 L., Fladby, T., Hohman, T.J., Jonsson, P.V., Kiddle, S.J., Rongve, A., Saltvedt, I., Sando, S.B., Selbæk, G., Shoai,
502 M., Skene, N.G., Snaedal, J., Stordal, E., Ulstein, I.D., Wang, Y., White, L.R., Hardy, J., Hjerling-Leffler, J.,
503 Sullivan, P.F., van der Flier, W.M., Dobson, R., Davis, L.K., Stefansson, H., Stefansson, K., Pedersen, N.L.,
504 Ripke, S., Andreassen, O.A., Posthuma, D., 2019. Genome-wide meta-analysis identifies new loci and functional
505 pathways influencing Alzheimer’s disease risk. *Nat. Genet.* <https://doi.org/10.1038/s41588-018-0311-9>
- 506 Keren-Shaul, H., Spinrad, A., Weiner, A., Matcovitch-Natan, O., Dvir-Szternfeld, R., Ulland, T.K., David, E., Baruch, K.,
507 Lara-Astaiso, D., Toth, B., Itzkovitz, S., Colonna, M., Schwartz, M., Amit, I., 2017. A Unique Microglia Type
508 Associated with Restricting Development of Alzheimer’s Disease. *Cell* 169, 1276-1290.e17.
509 <https://doi.org/10.1016/j.cell.2017.05.018>
- 510 Krasemann, S., Madore, C., Cialic, R., Baufeld, C., Calcagno, N., Fatimy, R.E., Beckers, L., O’Loughlin, E., Xu, Y., Fanek,
511 Z., Greco, D.J., Smith, S.T., Tweet, G., Humulock, Z., Zrzavy, T., Conde-Sanroman, P., Gacias, M., Weng, Z.,
512 Chen, H., Tjon, E., Mazaheri, F., Hartmann, K., Madi, A., Ulrich, J.D., Glatzel, M., Worthmann, A., Heeren, J.,
513 Budnik, B., Lemere, C., Ikezu, T., Heppner, F.L., Litvak, V., Holtzman, D.M., Lassmann, H., Weiner, H.L.,
514 Ochando, J., Haass, C., Butovsky, O., 2017. The TREM2-APOE Pathway Drives the Transcriptional Phenotype of
515 Dysfunctional Microglia in Neurodegenerative Diseases. *Immunity* 47, 566-581.e9.
516 <https://doi.org/10.1016/j.immuni.2017.08.008>
- 517 Kunkle, B.W., Grenier-Boley, B., Sims, R., Bis, J.C., Damotte, V., Naj, A.C., Boland, A., Vronskaya, M., Lee, S.J. van der,
518 Amlie-Wolf, A., Bellenguez, C., Frizatti, A., Chouraki, V., Martin, E.R., Sleegers, K., Badarinarayan, N.,
519 Jakobsdottir, J., Hamilton-Nelson, K.L., Moreno-Grau, S., Olaso, R., Raybould, R., Chen, Y., Kuzma, A.B.,
520 Hiltunen, M., Morgan, T., Ahmad, S., Vardarajan, B.N., Epelbaum, J., Hoffmann, P., Boada, M., Beecham, G.W.,
521 Garnier, J.-G., Harold, D., Fitzpatrick, A.L., Valladares, O., Moutet, M.-L., Gerrish, A., Smith, A.V., Qu, L., Bacq,
522 D., Denning, N., Jian, X., Zhao, Y., Zompo, M.D., Fox, N.C., Choi, S.-H., Mateo, I., Hughes, J.T., Adams, H.H.,
523 Malamon, J., Sanchez-Garcia, F., Patel, Y., Brody, J.A., Dombroski, B.A., Naranjo, M.C.D., Daniilidou, M.,
524 Eiriksdottir, G., Mukherjee, S., Wallon, D., Uphill, J., Aspelund, T., Cantwell, L.B., Garzia, F., Galimberti, D.,
525 Hofer, E., Butkiewicz, M., Fin, B., Scarpini, E., Sarnowski, C., Bush, W.S., Meslage, S., Kornhuber, J., White,
526 C.C., Song, Y., Barber, R.C., Engelborghs, S., Sordon, S., Voijnovic, D., Adams, P.M., Vandenbergh, R.,

527 Mayhaus, M., Cupples, L.A., Albert, M.S., Deyn, P.P.D., Gu, W., Himali, J.J., Beekly, D., Squassina, A.,
528 Hartmann, A.M., Orellana, A., Blacker, D., Rodriguez-Rodriguez, E., Lovestone, S., Garcia, M.E., Doody, R.S.,
529 Munoz-Fernandez, C., Sussams, R., Lin, H., Fairchild, T.J., Benito, Y.A., Holmes, C., Karamujic-Comić, H.,
530 Frosch, M.P., Thonberg, H., Maier, W., Roschupkin, G., Ghatti, B., Giedraitis, V., Kawalia, A., Li, S., Huebinger,
531 R.M., Kilander, L., Moebus, S., Hernández, I., Kamboh, M.I., Brundin, R., Turton, J., Yang, Q., Katz, M.J.,
532 Concari, L., Lord, J., Beiser, A.S., Keene, C.D., Helisalmi, S., Kloszewska, I., Kukull, W.A., Koivisto, A.M.,
533 Lynch, A., Tarraga, L., Larson, E.B., Haapasalo, A., Lawlor, B., Mosley, T.H., Lipton, R.B., Solfrizzi, V., Gill,
534 M., Longstreth, W.T., Montine, T.J., Frisardi, V., Diez-Fairen, M., Rivadeneira, F., Petersen, R.C., Deramecourt,
535 V., Alvarez, I., Salani, F., Ciaramella, A., Boerwinkle, E., Reiman, E.M., Fievet, N., Rotter, J.I., Reisch, J.S.,
536 Hanon, O., Cupidi, C., Uitterlinden, A.G.A., Royall, D.R., Dufouil, C., Maletta, R.G., Rojas, I. de, Sano, M.,
537 Brice, A., Cecchetti, R., George-Hyslop, P.S., Ritchie, K., Tsolaki, M., Tsuang, D.W., Dubois, B., Craig, D., Wu,
538 C.-K., Soininen, H., Avramidou, D., Albin, R.L., Fratiglioni, L., Germanou, A., Apostolova, L.G., Keller, L.,
539 Koutroumani, M., Arnold, S.E., Panza, F., Gkatzima, O., Asthana, S., Hannequin, D., Whitehead, P., Atwood,
540 C.S., Caffarra, P., Hampel, H., Quintela, I., Carracedo, Á., Lannfelt, L., Rubinsztein, D.C., Barnes, L.L., Pasquier,
541 F., Frölich, L., Barral, S., McGuinness, B., Beach, T.G., Johnston, J.A., Becker, J.T., Passmore, P., Bigio, E.H.,
542 Schott, J.M., Bird, T.D., Warren, J.D., Boeve, B.F., Lupton, M.K., Bowen, J.D., Proitsi, P., Boxer, A., Powell, J.F.,
543 Burke, J.R., Kauwe, J.S.K., Burns, J.M., Mancuso, M., Buxbaum, J.D., Bonuccelli, U., Cairns, N.J., McQuillin, A.,
544 Cao, C., Livingston, G., Carlson, C.S., Bass, N.J., Carlsson, C.M., Hardy, J., Carney, R.M., Bras, J., Carrasquillo,
545 M.M., Guerreiro, R., Allen, M., Chui, H.C., Fisher, E., Masullo, C., Crocco, E.A., DeCarli, C., Bisceglia, G., Dick,
546 M., Ma, L., Duara, R., Graff-Radford, N.R., Evans, D.A., Hodges, A., Faber, K.M., Scherer, M., Fallon, K.B.,
547 Riemenschneider, M., Fardo, D.W., Heun, R., Farlow, M.R., Kölsch, H., Ferris, S., Leber, M., Foroud, T.M.,
548 Heuser, I., Galasko, D.R., Giegling, I., Gearing, M., Hüll, M., Geschwind, D.H., Gilbert, J.R., Morris, J., Green,
549 R.C., Mayo, K., Growdon, J.H., Feulner, T., Hamilton, R.L., Harrell, L.E., Dricchel, D., Honig, L.S., Cushion, T.D.,
550 Huentelman, M.J., Hollingworth, P., Hulette, C.M., Hyman, B.T., Marshall, R., Jarvik, G.P., Meggy, A., Abner,
551 E., Menzies, G.E., Jin, L.-W., Leonenko, G., Real, L.M., Jun, G.R., Baldwin, C.T., Grozeva, D., Karydas, A.,
552 Russo, G., Kaye, J.A., Kim, R., Jessen, F., Kowall, N.W., Vellas, B., Kramer, J.H., Vardy, E., LaFerla, F.M.,
553 Jöckel, K.-H., Lah, J.J., Dichgans, M., Leverenz, J.B., Mann, D., Levey, A.I., Pickering-Brown, S., Lieberman,
554 A.P., Klopp, N., Lunetta, K.L., Wichmann, H.-E., Lyketsos, C.G., Morgan, K., Marson, D.C., Brown, K.,
555 Martiniuk, F., Medway, C., Mash, D.C., Nöthen, M.M., Masliah, E., Hooper, N.M., McCormick, W.C., Daniele,
556 A., McCurry, S.M., Bayer, A., McDavid, A.N., Gallacher, J., McKee, A.C., Bussche, H. van den, Mesulam, M.,
557 Brayne, C., Miller, B.L., Riedel-Heller, S., Miller, C.A., Miller, J.W., Al-Chalabi, A., Morris, J.C., Shaw, C.E.,
558 Myers, A.J., Wiltfang, J., O'Bryant, S., Olichney, J.M., Alvarez, V., Parisi, J.E., Singleton, A.B., Paulson, H.L.,
559 Collinge, J., Perry, W.R., Mead, S., Peskind, E., Cribbs, D.H., Rossor, M., Pierce, A., Ryan, N.S., Poon, W.W.,
560 Nacmias, B., Potter, H., Sorbi, S., Quinn, J.F., Sacchinelli, E., Raj, A., Spalletta, G., Raskind, M., Caltagirone, C.,
561 Bossù, P., Orfei, M.D., Reisberg, B., Clarke, R., Reitz, C., Smith, A.D., Ringman, J.M., Warden, D., Roberson,
562 E.D., Wilcock, G., Rogaeva, E., Bruni, A.C., Rosen, H.J., Gallo, M., Rosenberg, R.N., Ben-Shlomo, Y., Sager,
563 M.A., Mecocci, P., Saykin, A.J., Pastor, P., Cuccaro, M.L., Vance, J.M., Schneider, J.A., Schneider, L.S., Slifer,
564 S., Seeley, W.W., Smith, A.G., Sonnen, J.A., Spina, S., Stern, R.A., Swerdlow, R.H., Tang, M., Tanzi, R.E.,
565 Trojanowski, J.Q., Troncoso, J.C., Deerlin, V.M.V., Eldik, L.J.V., Vinters, H.V., Vonsattel, J.P., Weintraub, S.,
566 Welsh-Bohmer, K.A., Wilhelmsen, K.C., Williamson, J., Wingo, T.S., Woltjer, R.L., Wright, C.B., Yu, C.-E., Yu,
567 L., Saba, Y., Pilotto, A., Bullido, M.J., Peters, O., Crane, P.K., Bennett, D., Bosco, P., Coto, E., Boccardi, V.,
568 Jager, P.L.D., Lleo, A., Warner, N., Lopez, O.L., Ingelsson, M., Deloukas, P., Cruchaga, C., Graff, C., Gwilliam,
569 R., Fornage, M., Goate, A.M., Sanchez-Juan, P., Kehoe, P.G., Amin, N., Ertekin-Taner, N., Berr, C., DeBette, S.,
570 Love, S., Launer, L.J., Younkin, S.G., Dartigues, J.-F., Corcoran, C., Ikram, M.A., Dickson, D.W., Nicolas, G.,
571 Champion, D., Tschanz, J., Schmidt, H., Hakonarson, H., Clarimon, J., Munger, R., Schmidt, R., Farrer, L.A.,

572 Broeckhoven, C.V., O'Donovan, M.C., DeStefano, A.L., Jones, L., Haines, J.L., Deleuze, J.-F., Owen, M.J.,
573 Gudnason, V., Mayeux, R., Escott-Price, V., Psaty, B.M., Ramirez, A., Wang, L.-S., Ruiz, A., Duijn, C.M. van,
574 Holmans, P.A., Seshadri, S., Williams, J., Amouyel, P., Schellenberg, G.D., Lambert, J.-C., Pericak-Vance, M.A.,
575 2019. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau,
576 immunity and lipid processing. *Nat. Genet.* 51, 414–430. <https://doi.org/10.1038/s41588-019-0358-2>
577 Lake, B.B., Ai, R., Kaeser, G.E., Salathia, N.S., Yung, Y.C., Liu, R., Wildberg, A., Gao, D., Fung, H.-L., Chen, S.,
578 Vijayaraghavan, R., Wong, J., Chen, A., Sheng, X., Kaper, F., Shen, R., Ronaghi, M., Fan, J.-B., Wang, W., Chun,
579 J., Zhang, K., 2016. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human
580 brain. *Science* 352, 1586–1590. <https://doi.org/10.1126/science.aaf1204>
581 Lake, B.B., Chen, S., Sos, B.C., Fan, J., Kaeser, G.E., Yung, Y.C., Duong, T.E., Gao, D., Chun, J., Kharchenko, P.V.,
582 Zhang, K., 2018. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain.
583 *Nat. Biotechnol.* 36, 70–80. <https://doi.org/10.1038/nbt.4038>
584 Lake, B.B., Codeluppi, S., Yung, Y.C., Gao, D., Chun, J., Kharchenko, P.V., Linnarsson, S., Zhang, K., 2017. A
585 comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type
586 expression from nuclear RNA. *Sci. Rep.* 7, 6031. <https://doi.org/10.1038/s41598-017-04426-w>
587 Lambert, J.C., Ibrahim-Verbaas, C.A., Harold, D., Naj, A.C., Sims, R., Bellenguez, C., DeStefano, A.L., Bis, J.C., Beecham,
588 G.W., Grenier-Boley, B., Russo, G., Thornton-Wells, T.A., Jones, N., Smith, A.V., Chouraki, V., Thomas, C.,
589 Ikram, M.A., Zelenika, D., Vardarajan, B.N., Kamatani, Y., Lin, C.F., Gerrish, A., Schmidt, H., Kunkle, B.,
590 Dunstan, M.L., Ruiz, A., Bioreau, M.T., Choi, S.H., Reitz, C., Pasquier, F., Cruchaga, C., Craig, D., Amin, N.,
591 Berr, C., Lopez, O.L., De Jager, P.L., Deramecourt, V., Johnston, J.A., Evans, D., Lovestone, S., Letenneur, L.,
592 Moron, F.J., Rubinsztein, D.C., Eiriksdottir, G., Sleegers, K., Goate, A.M., Fievet, N., Huentelman, M.W., Gill,
593 M., Brown, K., Kamboh, M.I., Keller, L., Barberger-Gateau, P., McGuinness, B., Larson, E.B., Green, R., Myers,
594 A.J., Dufouil, C., Todd, S., Wallon, D., Love, S., Rogaeva, E., Gallacher, J., St George-Hyslop, P., Clarimon, J.,
595 Lleo, A., Bayer, A., Tsuang, D.W., Yu, L., Tsolaki, M., Bossu, P., Spalletta, G., Proitsi, P., Collinge, J., Sorbi, S.,
596 Sanchez-Garcia, F., Fox, N.C., Hardy, J., Deniz Naranjo, M.C., Bosco, P., Clarke, R., Brayne, C., Galimberti, D.,
597 Mancuso, M., Matthews, F., Moebus, S., Mecocci, P., Del Zompo, M., Maier, W., Hampel, H., Pilotto, A.,
598 Bullido, M., Panza, F., Caffarra, P., Nacmias, B., Gilbert, J.R., Mayhaus, M., Lannefelt, L., Hakonarson, H.,
599 Pichler, S., Carrasquillo, M.M., Ingelsson, M., Beekly, D., Alvarez, V., Zou, F., Valladares, O., Younkin, S.G.,
600 Coto, E., Hamilton-Nelson, K.L., Gu, W., Razquin, C., Pastor, P., Mateo, I., Owen, M.J., Faber, K.M., Jonsson,
601 P.V., Combarros, O., O'Donovan, M.C., Cantwell, L.B., Soininen, H., Blacker, D., Mead, S., Mosley, T.H.J.,
602 Bennett, D.A., Harris, T.B., Fratiglioni, L., Holmes, C., de Bruijn, R.F., Passmore, P., Montine, T.J., Bettens, K.,
603 Rotter, J.I., Brice, A., Morgan, K., Foroud, T.M., Kukull, W.A., Hannequin, D., Powell, J.F., Nalls, M.A., Ritchie,
604 K., Lunetta, K.L., Kauwe, J.S., Boerwinkle, E., Riemenschneider, M., Boada, M., Hiltunen, M., Martin, E.R.,
605 Schmidt, R., Rujescu, D., Wang, L.S., Dartigues, J.F., Mayeux, R., Tzourio, C., Hofman, A., Nothen, M.M., Graff,
606 C., Psaty, B.M., Jones, L., Haines, J.L., Holmans, P.A., Lathrop, M., Pericak-Vance, M.A., Launer, L.J., Farrer,
607 L.A., van Duijn, C.M., Van Broeckhoven, C., Moskvina, V., Seshadri, S., Williams, J., Schellenberg, G.D.,
608 Amouyel, P., 2013. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's
609 disease. *Nat. Genet.* 45, 1452–1458. <https://doi.org/10.1038/ng.2802>
610 Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., Mesirov, J.P., 2011. Molecular signatures
611 database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740. <https://doi.org/10.1093/bioinformatics/btr260>
612 Lun, A.T.L., Riesenfeld, S., Andrews, T., Dao, T.P., Gomes, T., Marioni, J.C., participants in the 1st Human Cell Atlas
613 Jamboree, 2019. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA
614 sequencing data. *Genome Biol.* 20, 63. <https://doi.org/10.1186/s13059-019-1662-y>
615 Macosko, E.Z., Basu, A., Satija, R., Nemes, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N.,
616 Martersteck, E.M., Trombetta, J.J., Weitz, D.A., Sanes, J.R., Shalek, A.K., Regev, A., McCarroll, S.A., 2015.

- 617 Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 161, 1202–
618 1214. <https://doi.org/10.1016/j.cell.2015.05.002>
- 619 Mancuso, R., Daele, J.V.D., Fattorelli, N., Wolfs, L., Balusu, S., Burton, O., Liston, A., Sierksma, A., Fourne, Y.,
620 Poovathingal, S., Arranz-Mendiguren, A., Frigerio, C.S., Claes, C., Serneels, L., Theys, T., Perry, V.H., Verfaillie,
621 C., Fiers, M., Strooper, B.D., 2019. Stem-cell-derived human microglia transplanted in mouse brain to study
622 human disease. *Nat. Neurosci.* 1–6. <https://doi.org/10.1038/s41593-019-0525-x>
- 623 Maniatis, S., Äijö, T., Vickovic, S., Braine, C., Kang, K., Mollbrink, A., Fagegaltier, D., Saiz-Castro, G., Cuevas, M.,
624 Watters, A., Lundeberg, J., Bonneau, R., Phatnani, H., 2019. Spatiotemporal dynamics of molecular pathology in
625 amyotrophic lateral sclerosis 6.
- 626 Marioni, R.E., Harris, S.E., Zhang, Q., McRae, A.F., Hagenaars, S.P., Hill, W.D., Davies, G., Ritchie, C.W., Gale, C.R.,
627 Starr, J.M., Goate, A.M., Porteous, D.J., Yang, J., Evans, K.L., Deary, I.J., Wray, N.R., Visscher, P.M., 2018.
628 GWAS on family history of Alzheimer’s disease. *Transl. Psychiatry* 8, 99. [https://doi.org/10.1038/s41398-018-
0150-6](https://doi.org/10.1038/s41398-018-
629 0150-6)
- 630 Mathys, H., Davila-Velderrain, J., Peng, Z., Gao, F., Mohammadi, S., Young, J.Z., Menon, M., He, L., Abdurrob, F., Jiang,
631 X., Martorell, A.J., Ransohoff, R.M., Hafler, B.P., Bennett, D.A., Kellis, M., Tsai, L.-H., 2019. Single-cell
632 transcriptomic analysis of Alzheimer’s disease. *Nature* 1. <https://doi.org/10.1038/s41586-019-1195-2>
- 633 McGeer, P.L., Itagaki, S., Tago, H., McGeer, E.G., 1987. Reactive microglia in patients with senile dementia of the
634 Alzheimer type are positive for the histocompatibility glycoprotein HLA-DR. *Neurosci. Lett.* 79, 195–200.
635 [https://doi.org/10.1016/0304-3940\(87\)90696-3](https://doi.org/10.1016/0304-3940(87)90696-3)
- 636 McGinnis, C.S., Murrow, L.M., Gartner, Z.J., 2019. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data
637 Using Artificial Nearest Neighbors. *Cell Syst.* 8, 329-337.e4. <https://doi.org/10.1016/j.cels.2019.03.003>
- 638 Sala Frigerio, C., Wolfs, L., Fattorelli, N., Thrupp, N., Voytyuk, I., Schmidt, I., Mancuso, R., Chen, W.-T., Woodbury, M.E.,
639 Srivastava, G., Möller, T., Hudry, E., Das, S., Saido, T., Karran, E., Hyman, B., Perry, V.H., Fiers, M., Strooper,
640 B.D., 2019. The Major Risk Factors for Alzheimer’s Disease: Age, Sex, and Genes Modulate the Microglia
641 Response to A β Plaques. *Cell Rep.* 27, 1293-1306.e6. <https://doi.org/10.1016/j.celrep.2019.03.099>
- 642 Saunders, A., Macosko, E.Z., Wysoker, A., Goldman, M., Krienen, F.M., de Rivera, H., Bien, E., Baum, M., Bortolin, L.,
643 Wang, S., Goeva, A., Nemes, J., Kamitaki, N., Brumbaugh, S., Kulp, D., McCarroll, S.A., 2018. Molecular
644 Diversity and Specializations among the Cells of the Adult Mouse Brain. *Cell* 174, 1015-1030.e16.
645 <https://doi.org/10.1016/j.cell.2018.07.028>
- 646 Sergushichev, A.A., 2016. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic
647 calculation. *bioRxiv* 060012. <https://doi.org/10.1101/060012>
- 648 Skene, N.G., Bryois, J., Bakken, T.E., Breen, G., Crowley, J.J., Gaspar, H.A., Giusti-Rodriguez, P., Hodge, R.D., Miller,
649 J.A., Muñoz-Manchado, A.B., O’Donovan, M.C., Owen, M.J., Pardiñas, A.F., Ryge, J., Walters, J.T.R.,
650 Linnarsson, S., Lein, E.S., Sullivan, P.F., Hjerling-Leffler, J., 2018. Genetic identification of brain cell types
651 underlying schizophrenia. *Nat. Genet.* 50, 825–833. <https://doi.org/10.1038/s41588-018-0129-5>
- 652 Skene, N.G., Grant, S.G.N., 2016. Identification of Vulnerable Cell Types in Major Brain Disorders Using Single Cell
653 Transcriptomes and Expression Weighted Cell Type Enrichment. *Front. Neurosci.* 10.
654 <https://doi.org/10.3389/fnins.2016.00016>
- 655 Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O.,
656 Huss, M., Mollbrink, A., Linnarsson, S., Codeluppi, S., Borg, Å., Pontén, F., Costea, P.I., Sahlén, P., Mulder, J.,
657 Bergmann, O., Lundeberg, J., Frisén, J., 2016. Visualization and analysis of gene expression in tissue sections by
658 spatial transcriptomics. *Science* 353, 78–82. <https://doi.org/10.1126/science.aaf2403>
- 659 Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., Satija,
660 R., 2019. Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888-1902.e21.
661 <https://doi.org/10.1016/j.cell.2019.05.031>

662 Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L.,
663 Golub, T.R., Lander, E.S., Mesirov, J.P., 2005. Gene set enrichment analysis: A knowledge-based approach for
664 interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550.
665 <https://doi.org/10.1073/pnas.0506580102>

666 Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T.,
667 Bertagnolli, D., Goldy, J., Shapovalova, N., Parry, S., Lee, C., Smith, K., Bernard, A., Madisen, L., Sunkin, S.M.,
668 Hawrylycz, M., Koch, C., Zeng, H., 2016. Adult mouse cortical cell taxonomy revealed by single cell
669 transcriptomics. *Nat. Neurosci.* 19, 335–346. <https://doi.org/10.1038/nn.4216>

670 Zeisel, A., Hochgerner, H., Lönnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Häring, M., Braun, E., Borm, L.E.,
671 La Manno, G., Codeluppi, S., Furlan, A., Lee, K., Skene, N., Harris, K.D., Hjerling-Leffler, J., Arenas, E., Ernfors,
672 P., Marklund, U., Linnarsson, S., 2018. Molecular Architecture of the Mouse Nervous System. *Cell* 174, 999-
673 1014.e22. <https://doi.org/10.1016/j.cell.2018.06.021>

674 Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C.G.,
675 Gil, L., Gordon, L., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O.G., Janacek, S.H., Juettemann, T., To, J.K.,
676 Laird, M.R., Lavidas, I., Liu, Z., Loveland, J.E., Maurel, T., McLaren, W., Moore, B., Mudge, J., Murphy, D.N.,
677 Newman, V., Nuhn, M., Ogeh, D., Ong, C.K., Parker, A., Patricio, M., Riat, H.S., Schuilenburg, H., Sheppard, D.,
678 Sparrow, H., Taylor, K., Thormann, A., Vullo, A., Walts, B., Zadissa, A., Frankish, A., Hunt, S.E., Kostadima, M.,
679 Langridge, N., Martin, F.J., Muffato, M., Perry, E., Ruffier, M., Staines, D.M., Trevanion, S.J., Aken, B.L.,
680 Cunningham, F., Yates, A., Flicek, P., 2018. Ensembl 2018. *Nucleic Acids Res.* 46, D754–D761.
681 <https://doi.org/10.1093/nar/gkx1098>

682 Zhou, Y., Song, W.M., Andhey, P.S., Swain, A., Levy, T., Miller, K.R., Poliani, P.L., Cominelli, M., Grover, S., Gilfillan,
683 S., Cella, M., Ulland, T.K., Zaitsev, K., Miyashita, A., Ikeuchi, T., Sainouchi, M., Kakita, A., Bennett, D.A.,
684 Schneider, J.A., Nichols, M.R., Beausoleil, S.A., Ulrich, J.D., Holtzman, D.M., Artyomov, M.N., Colonna, M.,
685 2020. Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent
686 cellular responses in Alzheimer’s disease. *Nat. Med.* 26, 131–142. <https://doi.org/10.1038/s41591-019-0695-9>

687

688

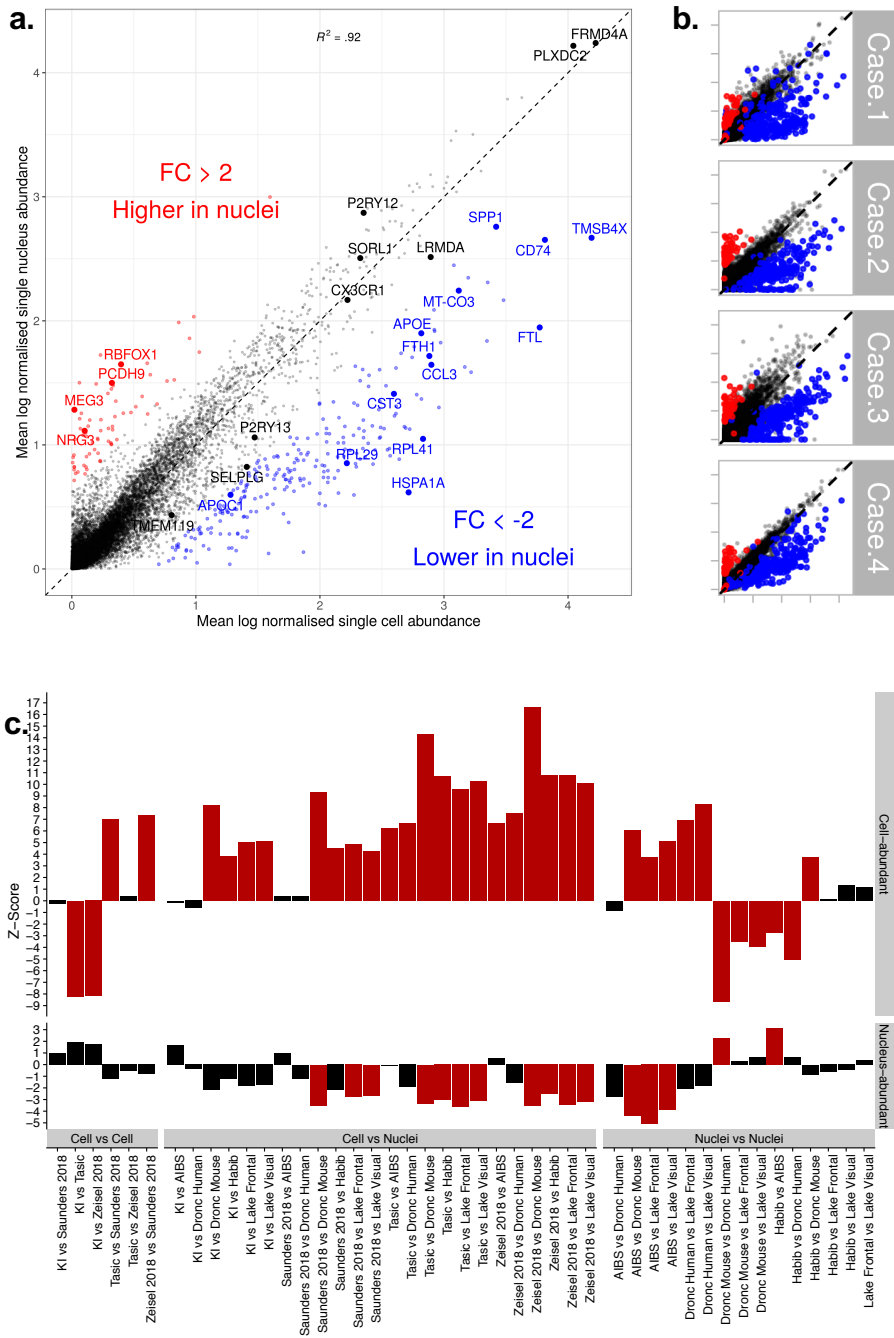
689

690 **Figures**

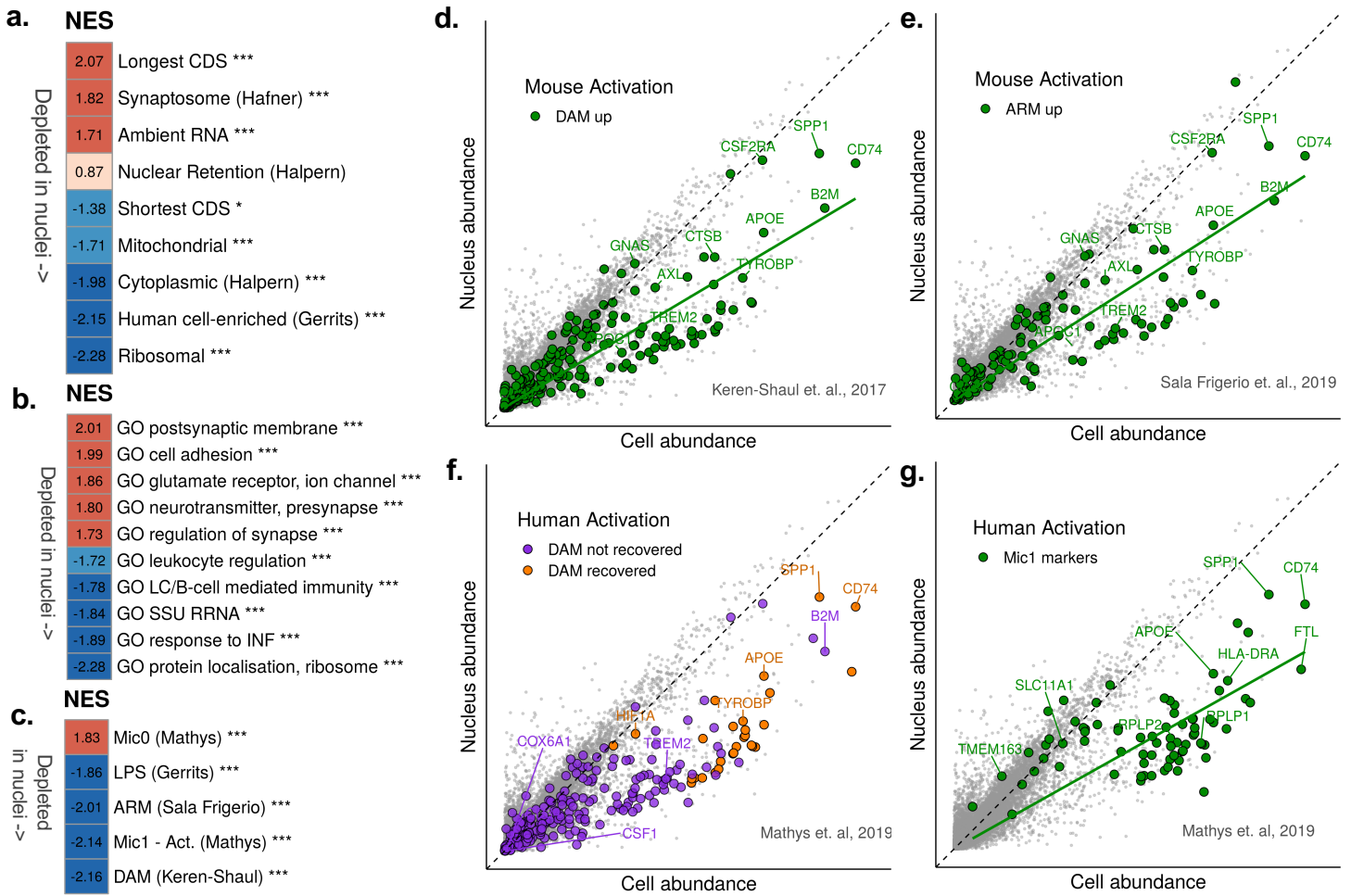
691

692 **Fig. 1: Gene abundance in single microglial cells versus single microglial nuclei of**
693 **human cortical tissue. a.** Mean normalised gene abundance in cells (x axis) and nuclei. (y
694 axis). Red: genes with significantly higher abundance in nuclei ($p_{\text{adj}} < 0.05$, fold change > 2).
695 Blue: genes that are significantly less abundance in nuclei ($p_{\text{adj}} < 0.05$, fold change < -2).
696 Genes were normalized to read depth (per cell), scaled by 10,000 and log-transformed.
697 *MALATI* (which had normalized abundance levels of 6.0 and 6.9 respectively in cells and
698 nuclei) has been removed for visualisation purposes. FC = fold change. Full results are
699 available in Supplementary Table 2. **b.** Scatterplot as in a), per patient (with the same genes
700 highlighted). Supplementary Table 1 contains patient data. **c.** Each bar represents a
701 comparison between two datasets (X versus Y), with the bootstrapped z -scores representing
702 the extent to which cell-enriched genes (top panel) and nuclear-enriched genes (bottom panel)
703 have lower specificity for microglia in dataset Y relative to that in dataset X . Larger z -scores
704 indicate greater depletion of genes, and red bars indicate a statistically significant depletion
705 ($p_{\text{adj}} < 0.05$, by bootstrapping). KI = Karolinska Institutet ; AIBS = Allen Institute for Brain
706 Science.

707



708 **Fig. 2: Functional analysis of genes that are enriched or depleted in nuclei. a.** Gene Set
709 Enrichment Analysis (GSEA) of gene sets related to cellular location and gene coding
710 sequence length (CDS). Background genes were ranked according to log fold change of
711 nuclei versus cells. Red: higher Normalised Enrichment Score (NES), i.e. more genes
712 associated with nuclear enrichment ; blue: negative NES scores (depletion in nuclei). ***
713 represents significance ($p_{adj} < 0.0005$). CDS = coding sequence. **b.** GSEA of *super*-Gene
714 Ontology gene sets against ranked nucleus-cell log fold changes. Only top and bottom
715 categories (according to NES) are shown. Colours as in a). GO = Gene Ontology ; SSU
716 RRNA = small subunit ribosomal RNA ; INF = Interferon ; LC = leukocyte. **c.** GSEA of
717 selected gene sets from previous studies of microglial activation, against fold change as in a).
718 *** represents significance ($p_{adj} < 0.0005$). Mic0 = markers of microglial cluster 0 in human
719 brain tissue ; Mic1 = markers of microglial cluster 1 (response to plaques) defined by
720 (Mathys et al., 2019) in human brain tissue. ARM = Activation Response Microglia (Sala
721 Frigerio *et al.*⁸). DAM = Disease-Associated Microglia (Keren-Shaul *et al.*⁷) **d.** Scatterplot as
722 in Fig. 1a), highlighting in green the DAM genes. A regression line for the highlighted genes
723 is shown in green (slope = 0.60). **e.** Scatterplot as in d), highlighting in green the ARM genes.
724 A regression line for the highlighted genes is shown in green (slope = 0.64). **f.** Scatterplot as
725 in d), highlighting the DAM genes recovered in the study of human activation in AD (Mathys
726 et al., 2019). Purple: DAM genes not recovered in their study ; orange: DAM genes recovered
727 in their study. **g.** Scatterplot as in d), Green: human activation marker genes defined by
728 (Mathys et al., 2019). A regression line for the highlighted genes is shown in green (slope =
729 0.56). Gene sets are available in Supplementary Table 6.



Supplementary Data

Figures

Fig. S1: Clustering of single nuclei from human cortical tissue.

Fig. S2: Gene abundance in single microglial cells versus single microglial nuclei of human cortical tissue.

Text

Supplementary Text: Clustering of microglial cells and nuclei in human cortical tissue.

Tables

Supplementary Table 1: Patient Metadata

Supplementary Table 2: Differential abundance nuclei versus cells

Supplementary Table 3: Public single cell / single nucleus datasets

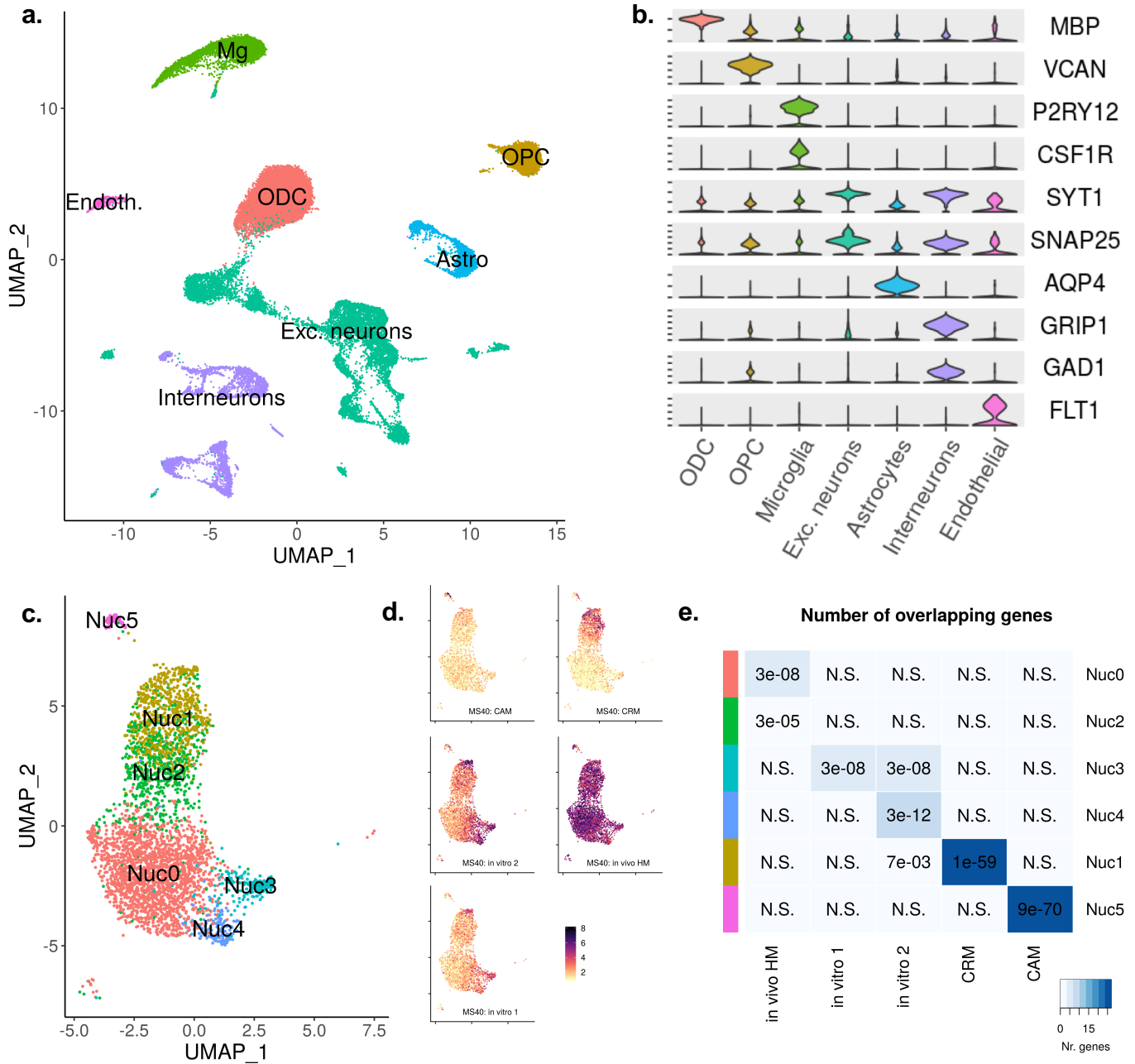
Supplementary Table 4: Clustering of Gene Ontology (GO) terms

Supplementary Table 5: Results of GSEA analysis

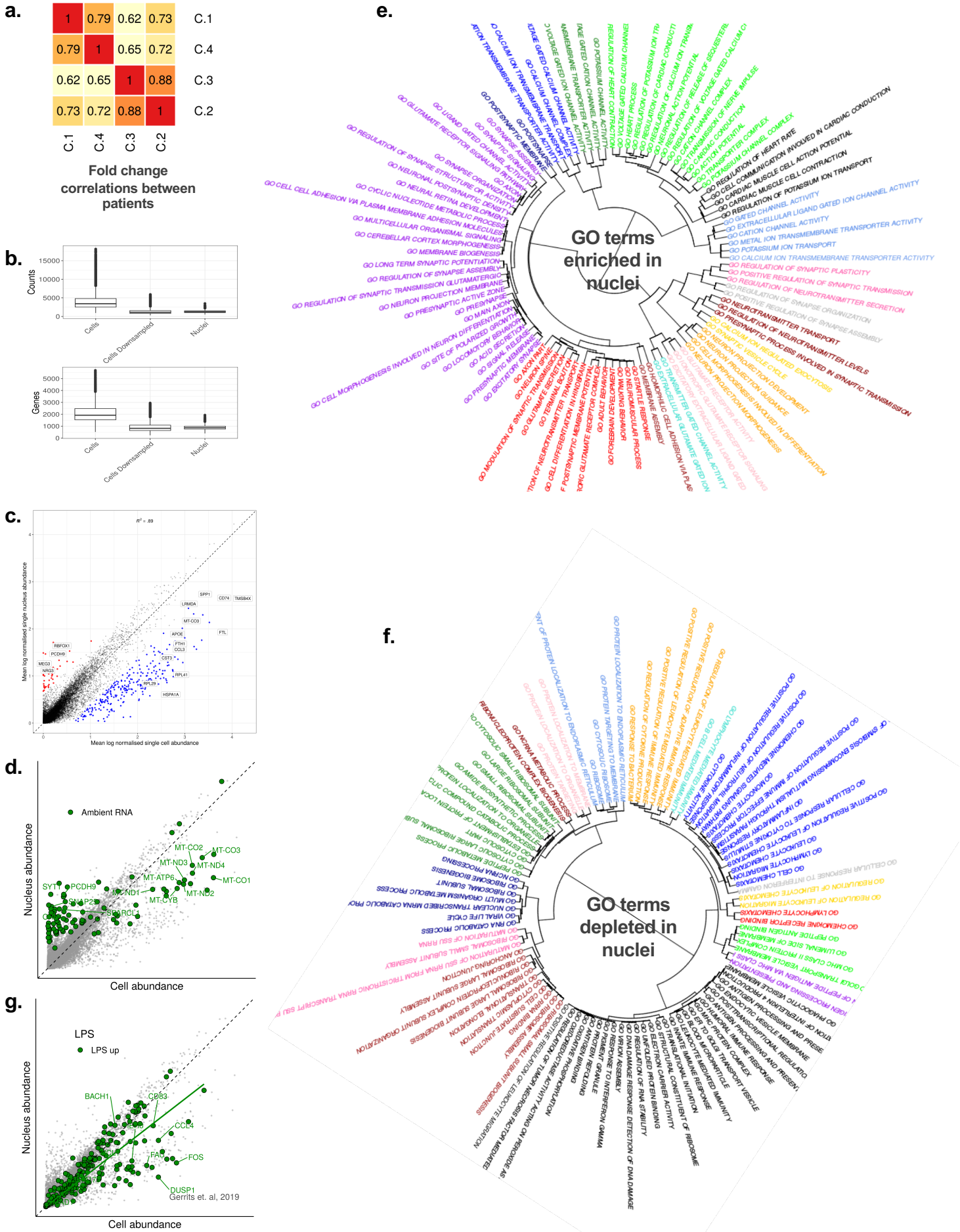
Supplementary Table 6: Lists of all gene sets

Supplementary Table 7: Cluster markers for single nucleus cluster analysis

730 **Fig. S1: Clustering of single nuclei from human tissue. a.** UMAP plot of 37,060 nuclei
731 from cortical tissue of 4 neurosurgical patients, coloured according to cell type. Mg =
732 Microglia ; OPC = oligodendrocyte precursor cells ; ODC = oligodendrocytes ; Astro =
733 Astrocytes ; Endoth = Endothelial cells ; Exc. neurons = excitatory neurons. **b.** Violin plots
734 show selected markers of the different cell types (data is normalised for count depth and log-
735 transformed). **c.** UMAP plot of 3,721 microglial nuclei from cortical tissue of 4 neurosurgical
736 patients, coloured according to cluster number, after *in silico* extraction of microglia (based
737 on markers such as *P2RY12*) and reclustering. **d.** Module scores for gene sets extracted from
738 the original Mancuso *et al.* single cell microglia study (Mancuso *et al.*, 2019). The top 40
739 genes according to log fold change were selected for each gene set. **e.** Overlap of top 40
740 marker genes from cellular clusters on the horizontal axis (Mancuso *et al.*) and nuclear
741 clusters on the vertical axis. The blue scale represents the number of genes in common,
742 numbers represent p_{adj} values. Vertical coloured bars correspond to the clusters shown in c).
743 N.S. = not significant ($p_{adj} > 0.05$). MS40 = Module Score of top 40 gene markers ; CAM =
744 macrophages ; CRM = cytokine response ; *in vitro* 1 = activation-like module (similar to *in*
745 *vitro* macrophages) ; *in vitro* 2 = activation-like module (similar to *in vitro* monocytes) ; *in*
746 *vivo* HM = homeostatic. Nuc = Nuclear clusters. Cluster markers are provided in
747 Supplementary Table 7.
748
749



750 **Fig. S2: Gene abundance in single microglial cells versus single microglial nuclei of**
751 **human cortical tissue. a.** Correlation matrix of gene abundance fold changes (single cell vs
752 single nucleus abundance) between patients. **b.** Downsampling of reads: boxplots for
753 numbers of reads (top) and numbers of genes (bottom) for single cells before downsampling,
754 single cells after downsampling, and single nuclei. Boxplots show median, with 25% and
755 75% quantiles. **c.** Scatterplot of mean gene abundance in cells against mean gene abundance
756 in nuclei (as in Fig. 1a) after downsampling of reads in cells. Data is normalised to count
757 depth and log-transformed. Points in red represent genes with significantly higher abundance
758 in nuclei, while those in blue are significantly less abundant in nuclei ($p_{\text{adj}} < 0.05$, fold change
759 $> |2|$). **d.** Scatter plot, as in Fig. 1a) showing the ambient mRNA in green (the same dataset
760 was used in Fig. 2d). Ambient RNA is defined as the 150 most abundant genes in the 700
761 nuclei with the lowest total read counts. **e, f.** Dendrograms of **e.** top 100 Gene Ontology (GO)
762 terms enriched in nuclei, and **f.** top 100 GO terms depleted in nuclei. GO terms were
763 clustered based on overlap between their gene sets. The colours show how GO terms were
764 clustered. These clusters are described in Supplementary Table 4. **g.** Scatterplot as in Fig. 1a),
765 highlighting in green genes that are upregulated during LPS stimulation in mice (Gerrits et
766 al., 2019). A regression line for the highlighted genes is shown in green (slope = 0.78).
767
768



769 **Supplementary Text: Clustering of microglial cells and nuclei in human cortical tissue**

770

771 We sequenced nuclei from cortical tissue of 4 neurosurgical patients. Single cell sequencing
772 of FAC-sorted microglia was performed on cortical tissue of the same patients in a previous
773 study (Mancuso et al., 2019). Subject data is available in Supplementary Table 1. Following
774 quality filtering, data integration, PCA analysis and clustering of nuclei, we identified 7 main
775 cell types in 37,060 nuclei: oligodendrocytes (ODC, 34.0%), excitatory neurons (27.0%),
776 interneurons (11.3%), oligodendrocyte precursors (OPC 9.4%), microglia (11.3%), astrocytes
777 (6.0%), and endothelial cells (1.1%). Supplementary Fig. 1a and Supplementary Fig. 1b show
778 UMAP embeddings for all nuclei, coloured by cell type, and selected markers for each cell
779 type, respectively.

780

781 Microglial nuclei were isolated and reclustered. We identified 3,721 microglia (expressing
782 *MEF2A*, *P2RY12*, *CX3CR1*, *CSF1R*), a macrophage cluster (enriched for *CD163* and *MRC1*,
783 67 nuclei), a neutrophil cluster (72 nuclei), and a cluster containing microglial as well as
784 astrocytic markers (marked by *GFAP*, 68 nuclei). The neutrophil and ambiguous clusters
785 were discarded, leaving only microglia and brain macrophages for downstream analysis
786 (Supplementary Fig. 1c). Cluster markers are provided in Supplementary Table 6.

787

788 In order to determine if nuclei could recover microglial clusters identified in cells, we
789 selected the top 40 markers defined by Mancuso *et al.* (Mancuso et al., 2019) for each of the
790 clusters they identified in the original analysis of microglial cells. For each nucleus, we
791 scored each set of markers based on the abundance of those markers in the nucleus, using
792 Seurat's *AddModuleScore* function. These scores, referred to as MS40 scores, are highlighted
793 in Supplementary Fig. 1d. Our nuclei were able to recover a cytokine response cluster
794 (CRM), marked by *CCL3*, *CCL4*, and an activation-like cluster, equivalent to the “*in vitro*
795 microglia” identified in the original study (original markers included *APOC1*, *GPNMB*,
796 *SPPI*, *APOE*). Homeostatic markers appeared ubiquitously through-out the nuclei dataset,
797 and we were not able to distinguish a reduction of these markers in the activation-like
798 response cluster, as we would expect from transcriptomic profiling of microglia in mice
799 (Keren-Shaul et al., 2017; Sala Frigerio et al., 2019). Finally, the CAM (macrophage) cluster

800 (*CDI63*, *MRC1*), separated out from the bulk of the microglia, and was easily-recognisable
801 based on its MS40 score. Cluster markers are provided in Supplementary Table 6.

802

803 In order to quantify the differences between cells and nuclei in more detail, we examined the
804 overlap of the top 40 markers between nuclei clusters and cell clusters (Supplementary Fig.
805 1e). The cell macrophage (CAM) and cell cytokine (CRM) clusters showed the largest
806 overlaps with Nuc1 and Nuc7 (27 and 24 of 40 markers, respectively). Other clusters only
807 showed overlaps of between 1 and 5 genes. Cluster Nuc3 showed similar overlaps between
808 “*in vitro* 1” and “*in vitro* 2” (5 genes). Cluster Nuc0 showed an overlap of 5 genes with “*in*
809 *vivo* HM”, and cluster Nuc2 showed an overlap of 2 genes with “*in vivo* HM”. Cluster Nuc4
810 showed similarities with the “*in vitro* 2” cluster, suggesting it could be a cluster of activation,
811 however all 5 overlapping genes were mitochondrial genes. Cluster Nuc3 markers *RPS12*,
812 *TPT1*, *FTL*, *RPS18* and *EEF1A1* also appeared as markers of “*in vitro* 2”.

813

814 We performed similar analyses using more markers, however we found that introducing more
815 markers resulted in nuclei markers overlapping with more than one cell cluster. We also
816 noticed that introducing more markers resulted in overlaps between markers of the cellular
817 clusters with each other. Selecting 40 markers allowed us to align cellular and nuclear
818 clusters in an almost one-to-one fashion (see Supplementary Fig. 1e).

819

820 Overall, cytokine clusters and macrophage clusters were recovered well using single nucleus
821 methods, however, differences between other microglial subpopulations were not
822 convincingly recovered.

823