Transcriptomic profiling of myeloid cells in Alzheimer's Disease brain illustrates heterogeneity of microglia endolysosomal subtypes

Katherine E. Prater*[1] (ORCID: 0000-0001-8615-207X), Kevin J. Green*[1], Wei Sun[2], Carole L. Smith[1], Kenneth L. Chiou[3], Laura Heath[4], Shannon Rose[5], C. Dirk Keene[5], Ronald Y. Kwon[6,7], Noah Snyder-Mackler (ORCID: 0000-0003-3026-6160)[3], Elizabeth E. Blue (ORCID: 0000-0002-0633-0305)[9,10], Jessica E. Young[5,6], Ali Shojaie[8], Benjamin Logsdon[4,11], Gwenn A. Garden[12], and Suman Jayadev[1,6]

1 Department of Neurology, University of Washington, Seattle, WA 98195
2 Biostatistics Program, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA
3 School of Life Sciences, Center for Evolution and Medicine, Arizona State University, Tempe, AZ
4 Sage Bionetworks, Seattle, WA
5 Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA
6 Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, WA
7 Department of Orthopaedics and Sports Medicine, University of Washington, Seattle, WA
8 Department of Biostatistics, University of Washington, Seattle, WA
9 Division of Medical Genetics, University of Washington, Seattle, WA
10 Brotman Baty Institute for Precision Medicine, Seattle, Washington, 98195, USA
11 Cajal Neuroscience, Seattle, WA
12 Department of Neurology, University of North Carolina, Chapel Hill, NC
* Equal co-first authors

Key words: AD, microglia, RNAseq, inflammation, human, APOE

Correspondence to:
Suman Jayadev, MD
sumie@uw.edu

Word Count:
Abstract: 147
Text: 4756
Methods: 1987
Figures: 7
Tables: 1
Supplemental Figures: 2 Tables: 3

Acknowledgements: Funding for these studies was provided by: RF1AG063540 and P30AG066509. Additional support was provided by the Ellison Foundation. Autopsy materials used in this study were obtained from the University of Washington Neuropathology Core, which is supported by the Alzheimer's Disease Research Center (AG05136) and the Adult Changes in Thought Study (AG006781). Conceived and designed experiments: SJ, KJG, CLS, BL, GAG. Performed the experiments: KJG, CLS, SR, CDK. Analyzed data: KEP, KJG, KLC. Biostatistics oversight: WS, AS. Wrote the manuscript: KEP, KJG, SJ. Edited the manuscript: CLS, WS, EEB, GAG. Provided code: WS, KLC, NSM, RYK, LH. Conceptualized the research and collaborations: SJ, GAG, JEY, EEB, BL.

Abstract:

Microglia-mediated neuroinflammation contributes to disease progression in Alzheimer's Disease (AD). Microglia demonstrate heterogeneous states with proposed beneficial, harmful, and disease-specific subtypes. Defining the spectrum of microglia phenotypes is crucially important to the design of neuroinflammation-modulating therapies. We performed single-nucleus RNA-seq on over 120,000 microglia nuclei isolated from dorsolateral prefrontal cortices collected from 12 AD and 10 controls. We detected several previously unrecognized microglia transcriptomic sub-populations, including three with predominant endolysosomal gene expression patterns and heterogeneous metabolic and inflammatory signatures. One endoloysomal subtype is more abundant in AD individuals and uniquely enriched for transcripts involved in nucleic acid detection and interferon signaling. AD individuals also exhibited increased expression of unfolded protein response and inflammatory genes in homeostatic microglia subpopulations. These results highlight the heterogenous nature of the microglia response to AD pathology and inform efforts to target specific subtypes of microglia in the development of novel AD therapies.

64 **Introduction:**

65 Alzheimer's Disease (AD) affects millions of individuals worldwide, with increasing
66 impact of the disease as the US/global population ages. AD is pathologically characterized by
67 the presence of extracellular amyloid-beta plaques, neuronal intracellular neurofibrillary tangles,
68 and neuroinflammation. Neuroinflammation is one hypothesized driver of AD[1–3] and microglia
69 are critical sources of neuroinflammation[4–9]. In AD brain, microglia release inflammatory factors,
70 lose protective homeostatic function, and initiate aberrant phagocytosis of synapses and
71 neurons, all of which influence the pathophysiology behind cognitive decline[5,9–11]. As such,
72 dysregulated microglial inflammatory responses have potential as future therapeutic targets. Yet
73 large gaps remain in our understanding of microglia inflammatory responses in AD brain.
74 Experimental models illustrate a spectrum of microglia states differentiated by
75 morphology, physiology, and gene expression patterns, but less is known about microglia
76 functional heterogeneity in human disease. Studies in human autopsy brain tissue are not
77 amenable to controlled hypothesis testing but can provide unique insights into human disease
78 pathogenesis. Defining the molecular signatures of specific microglial populations in human
79 brain can give context to the recognized inflammatory responses in AD as well as guide and
80 inform hypothesis testing *in vitro* and *in vivo*.

81 Prior single cell or single nucleus RNAseq studies in human and animal model brain
82 identified subpopulations of microglia with specific biological function correlates in AD[12–21].
83 Several studies have described "Disease-associated microglia" (DAMs), or alternatively
84 activated-response microglia (ARMs), a specific population of microglia observed in AD mouse
85 models and proportionately enriched in AD human cases[12,14,15]. While informative, these
86 pioneering studies involve small numbers (100's or perhaps 1000's) of microglia. Two recent
87 studies profiled 16,000–148,000 human microglia and revealed 9-12 microglia subpopulations

88 suggesting that larger numbers of sequenced cells could achieve finer resolution of biologically
89 relevant subtypes[18,20].

90 Microglia comprise 3-8% of all cells in brain tissue[22]. Therefore, many single nucleus
91 RNAseq (snRNAseq) studies profile only dozens or hundreds of microglia per individual, limiting
92 the ability to characterize the degree of transcriptional heterogeneity actually present[14,17]. We
93 posited that cellular processes upstream of the well-recognized "inflammatory" or "responsive"
94 microglia profiles would be uncovered by improved resolution of microglia phenotypes in human
95 brain. To allow better characterization of the changes occurring in microglia subpopulations in
96 AD brain, we developed a novel enrichment technique that enhances microglia representation in
97 a snRNAseq dataset and applied this approach to a well-defined cohort of AD and control
98 cases. We generated microglia transcriptional profiles from a cohort of 22 individuals with mixed
99 *APOE* genotypes. We annotated microglia subtypes with plausible biological roles and identified
100 transcriptional differences between AD and control individuals within microglial phenotypes. In
101 addition to homeostatic and inflammatory phenotypes described previously, we uncovered
102 microglial subtypes with distinct transcriptomic profiles that provide new avenues for hypotheses
103 testing in future studies of AD mechanisms.

104

105 **Results:**

106 ***Fluorescence-activated nuclei sorting for PU.1 expression enriches microglia***
107 ***nuclei 20-fold.*** We hypothesized that much larger numbers of microglia nuclei could be
108 evaluated by snRNAseq if first enriched using fluorescence-activated nuclei sorting (FANS) for
109 expression of a myeloid specific transcription factor (PU.1; Figure 1A) To confirm that PU.1
110 FANS was effective, we isolated and sequenced nuclei with and without PU.1 FANS (N=4). We
111 analyzed similar numbers of total nuclei in the unsorted (46,085; Figure 1B) and PU.1 sorted
112 (41,488; Figure 1C) datasets. The PU.1 sorted dataset contained 20X more microglia nuclei

113 defined by high expression of *C3*, *CD74*, *C1QB*, *CX3CR1*, and *SPI1* (23,310 microglia nuclei)
114 than the unsorted dataset (1,032 microglia nuclei; Figure 1B,C). The microglia nuclei observed
115 in the PU.1 sorted dataset also demonstrate further complexity as evidenced by more microglia
116 subtypes (Figure 1C). We next applied PU.1 FANS to a cohort of 22 individuals, allowing us to
117 distinguish subclusters/subpopulations with better resolution and define their biological
118 functions. After PU.1 FANS, samples retain a variety of non-myeloid cell types (Figure 2A) while
119 providing clear resolution of clusters demonstrating distinct microglia gene expression patterns
120 (63% of the nuclei; Figure 2B).

121 ***Complexity of microglia subpopulations.*** The initial dataset consisted of 205,226
122 nuclei, with 200,948 nuclei (98%) passing QC and doublet removal. Of those, 127,371 were
123 microglia, with a minimum of 1,300 and an average of 5,790 nuclei per individual. Cluster
124 analysis of nuclei with microglia-like expression identified 10 clusters (Figure 3A) characterized
125 by differentially expressed genes (DEGs) comparing the cluster to all other nuclei (Figure 3B).
126 These subpopulations also have separable biological function correlates based on pathway
127 analyses (Figure 3C). There are many definitions of pathways and no single gold standard, so
128 we performed gene set enrichment analysis (GSEA) three times, using pathways defined by
129 GO, KEGG, and Reactome. These results are in Tables S2 and S3, with Figure 3C illustrating
130 only those pathways supported by at least 2/3 GSEA analyses.

131 We identified Cluster 1, the largest cluster, as the "homeostatic" cluster based on high
132 expression of *CX3CR1* and *P2RY12*[12,14,17,20]. Genes over-expressed in Cluster 1 matched 69%
133 of genes that are over-expressed in the homeostatic cluster defined by Gerrits *et al*. (2021), the
134 largest microglia-specific snRNAseq study to-date. Cluster 1 also contained genes identified as
135 "homeostatic" from two other snRNAseq publications[12,17]. We abbreviate this "homeostatic"
136 cluster as HC1. Cluster 2, while very similar to HC1 in terms of a non-activated state, was
137 further defined by a set of genes with lower expression compared to HC1 (Figure 3B). The

genes with lower expression in Cluster 2 are known to be involved in inflammatory processes and include *SPP1*, *HIF1A*, *CD14*, and *MALAT1*[12,15,23]. Although different in their gene expression, HC1 and Cluster 2 may have similar biological function. Similar to previous publications, we also assessed DEGs for other clusters in comparison to HC1 to assess change from "homeostatic" states (Table S2).

Two clusters were similar to microglia subtypes previously described in human brain. Cluster 4 was defined by expression of *FTH1* and *FTL* as well as a large number of ribosomal genes (Figure 3B). A cluster of microglia with similar gene expression phenotype is consistently identified in AD brain in both mouse and human[12–14,17,20,24]. This cluster is enriched for pathways involved in apoptosis, response to interferon-gamma (IFNγ), and many mitochondrial and respiratory functions (Figure 3C). Cluster 8 is defined by expression of classic inflammatory activation genes including *TNFAIP3, NFKB1, TRAF3, RELB* (Figure 3B)[1,25,26]. Pathway analysis revealed this cluster was enriched in NFκB signaling, IFN signaling as well as Toll-like receptor and RIG-1 mediated signaling consistent with activation of nucleic acid and DAMP recognition pathways (Figure 3C).

Next, we found three novel clusters with endolysosomal gene and pathway signatures. Cluster 3 was defined by a set of genes which are implicated in aggregate protein internalization by microglia (Figure 3B)[12,14,15,17,20,24]. DEGs in this cluster including *FCER1G*, *MERTK*, *DENND4c*, *ATG7* and *ATG16L2* suggested phagocytosis, vesicle mediated transport, and autophagy functions[26]. The pathways enriched in Cluster 3 include a variety of endosome and lysosome pathways as well as catabolism and lipid binding, but no inflammatory processes (Figure 3C). Consistent with this characterization, genes involved in glycolysis such as *HK2* and *PFKFB3* have lower expression in Cluster 3, suggesting these cells have not undergone the metabolic switch to glycolysis observed in the microglia inflammatory state[27].

162 Cluster 5 and 6 displayed an endolysosomal signature, though also appeared
163 metabolically active with distinct inflammatory characteristics. Cluster 5 was differentiated from
164 other microglia nuclei by increased expression of *HSP90AA1, HIF1A*, and *BAG3* in addition to
165 other heat shock protein (HSP) genes (Figure 3B), suggesting these cells are responding to
166 external stress. Genes driving glycolysis including *PFKFB*3 and *HK2* were more expressed in
167 Cluster 5 when compared to HC1, suggesting a switch from oxidative phosphorylation to
168 glycolysis in these cells[27]. The pathways enriched in this cluster indicate it is active in
169 endocytosis, autophagy, mitophagy (Figure 3C).

170 Cluster 6 was metabolically active, but unlike Cluster 5 appeared to have a more
171 inflammatory profile (Figure 3B). The pathways enriched within Cluster 6 demonstrate metabolic
172 activity and response to stressors similar to Cluster 5 (Fig 3C) with enrichment of glucose
173 catabolic and chaperone binding, NLR signaling and ubiquitin-like protein ligase activity. Unique
174 to Cluster 6, we found that the DNA repair genes[28] *ATM*, and *RNASEH2B* were lower in
175 expression. Cluster 6 showed increased expression of genes involved in double and single
176 strand DNA/RNA intracellular recognition (an "antiviral" gene expression pattern) including
177 *MAVS, DDX42, ADAR*, the lysosomal gene *CLC-7*, and mediators of the NLRP3 inflammasome
178 including *CASP1, TXNIP, P2RX7,* and the pattern recognition receptor *CARD9*[25,26,29–32]. The
179 downstream mediators of nucleic acid recognition and inflammasome activation include the type
180 1 interferons and we found significantly higher levels of *IRF3, IRF5*, and *IRF7*, suggesting that
181 Cluster 6 microglia are activating pathogen receptor and inflammasome pathways and inducing
182 interferon signaling molecules. While *IL1b* was increased in this cluster compared to HC1,
183 higher levels of *IL1b* and expression of other inflammatory effector molecules such as *NFkB*
184 were found in Cluster 8.

185 Cluster 7 is characterized by expression of genes involved in migration and motility
186 (Figure 3B). Pathways enriched in Cluster 7 also indicate these cells are motile, with changes to

187 cytoskeleton and membrane that indicate movement of processes or the cell itself (Figure 3C).
188 This subpopulation of microglia was also notable for expression of lipid processing genes *LPL*
189 and *PPARG*[33].

190 Finally, our dataset identified two clusters on the extreme ends of the proliferation
191 spectrum. Cluster 9 is defined from all other nuclei by genes including *TMEM2* and *HAMP*
192 (Figure 3B). *HAMP* encodes the protein hepcidin, one of the primary drivers of iron homeostasis
193 in the human body[34]. Pathways enriched in this microglia subtype involve metal ion homeostasis
194 and cytokine production while pathways that are significantly underrepresented include motility
195 (Table S2). This profile is consistent with senescent microglia[35]. In contrast, Cluster 10 is
196 defined from all other nuclei by expression of genes involved in cell cycle regulation and DNA
197 repair[28,36]. The pathways enriched in Cluster 10 confirm the relative increase of genes involved
198 in cell cycle processes and a decrease of endosome and cytokine processing genes in these
199 microglia (Figure 3C).

200 ***APOE genotype does not substantially alter clustering of microglia subtypes.*** We
201 asked whether there would be a detectable difference in the cluster profile of microglia if we
202 were to focus solely on a cohort of *APOE* ε3/ε3 genotype individuals. Despite the role of *APOE*
203 in risk and progression of AD[37], to date no studies have defined microglia subtypes in a
204 population of individuals with a specific *APOE* genotype. Since the majority (13/22) of our
205 sample were homozygous for the *APOE* ε3 allele, we generated a subset of our dataset that
206 consisted entirely of *APOE* ε3/ε3 individuals (7 controls and 6 AD pathology; 75018 microglia
207 nuclei). After re-normalizing and re-clustering, we identified 9 subpopulations of microglia
208 (Figure 4A). These clusters were defined by genes similar to those that defined the clusters in
209 the Mixed *APOE* genotype dataset (Figure 4B). Although *APOE* ε3/ε3 genotype individuals had
210 nuclei in mixed APOE cluster 9 (senescent microglia) it was not distinct after re-clustering the
211 smaller dataset. We identified clusters where the genes enriched in a cluster were very similar

212 (~60% or higher match) to the Mixed *APOE* dataset (Figure 4C). We found homeostatic,
213 neurodegenerative, inflammatory, cell cycle, and endolysosomal clusters similar to those of the
214 Mixed *APOE* cohort (Figure 4C). We again detected metabolically active clusters in the *APOE*
215 ε3/ε3 genotype dataset, suggesting that the presence of multiple distinct metabolically active
216 microglia subtypes is common in the human brain even when controlling for *APOE* genotype.

217 ***Genes implicated in AD risk and pathogenesis are differentially regulated in***
218 ***individual microglial subtypes.*** A number of studies in animals and humans have
219 characterized the "Disease Associated Microglia" phenotype and we determined whether the set
220 of genes originally defined by Keren-Shaul *et al*.[12] was enriched in a particular subtype of our
221 human microglia dataset. Utilizing GSEA, we determined whether gene sets reflecting biological
222 pathways tend to have higher expression in specific clusters. While the majority of our clusters
223 (with the exception of Clusters 9 and 10) had enrichment of approximately 35 genes in the DAM
224 gene set, Cluster 4 demonstrated the highest adjusted *p*-values (adjusted $p<0.001$) and the
225 expected direction of gene expression (Figure 5A). Cluster 6 was also enriched for multiple
226 DAM genes that differed from Cluster 4 in the expected direction (Figure 5A). This suggests that
227 in our dataset, the DAM phenotype is split across multiple clusters with differing gene
228 expression profiles and biological functions consistent with previous reports in human
229 microglia[16,17].

230 Multiple AD GWAS identified risk SNPs are associated with genes expressed in
231 microglia or myeloid cells[8]. Utilizing a list of 46 genes in SNP loci associated with altered AD
232 risk[37,38], we used GSEA to assess enrichment of this gene set within our microglia clusters. We
233 observed significant enrichment (adjusted $p<0.001$) of AD risk genes differentially expressed in
234 Cluster 6 (Figure 5B). Altered expression of genes associated with AD risk was observed in
235 both directions in Cluster 6 microglia. *PICALM* and *SORL1* are significantly lower, while others

236 including *APP*, *APOE*, and *BIN1* were significantly higher in expression (all adjusted $p<0.001$;
237 Figure 5B).

238 Having detected enrichment of AD risk genes unique to Cluster 6, we assessed whether
239 genes from gene sets representing biological processes were enriched in specific clusters. Our
240 GSEA analysis demonstrated strong enrichment (adjusted $p<0.001$) of endolysosomal genes
241 *LGMN*, *CTSL*, and *CTSB*, in Clusters 3, 6, and 7, and the pair of *CTSL* and *CSTB* in Cluster 5
242 (Figure 5C). Our heatmap demonstrates strong expression of the GO endolysosomal gene set
243 in Cluster 6, with lesser levels in Clusters 3 and 5, though not all are statistically significant
244 (Figure 5C). We also confirmed expression of inflammatory TLR signaling in Clusters 6 and 8
245 using the KEGG TLR gene set confirming the inflammatory activation phenotype of both
246 subpopulations of microglia (Figure 5D). We utilized the KEGG cell cycle gene set to confirm
247 that Cluster 10 is enriched for cell cycle genes significantly (adjusted $p<0.001$; Figure 5E).
248 These gene sets provide additional verification of the biological functions attributed to these
249 clusters utilizing genes not necessarily identified as differentially expressed in our dataset.

250 ***AD cases demonstrate increased numbers of Cluster 6 microglia.*** We determined
251 that only Cluster 6 was enriched for AD nuclei over control nuclei (adjusted $p=0.006$; Figure 6).
252 This AD-enriched cluster also has the highest differential regulation of genes related to top AD
253 risk SNPs (Figure 5B)[37,38]. This finding suggests that AD is associated with increased numbers
254 of microglia that demonstrate an endolysosomal, glycolytic, and inflammatory gene expression
255 profile. These results confirm previously reported findings that the subpopulation of microglia
256 proportionately altered in AD human brain differs from that identified in mouse models[18,39].

257 ***AD brains contain fewer microglia expressing high levels of cell cycle genes.*** In
258 addition to identifying microglia subtypes increased in AD, we addressed whether a microglia
259 subtype present in healthy aging was reduced in AD brain. We determined that Cluster 10, the
260 cluster differentially expressing cell cycle regulatory genes, is larger in control brain compared to

261 AD brain (Mixed *APOE* adjusted $p<0.001$; *APOE* ε3/ε3 adjusted $p<0.001$; Figure 6). These data
262 demonstrate that AD pathology includes a detectable reduction in a microglia subpopulation
263 enriched for cell cycle and DNA repair genes.

264 ***Microglia transcriptomic progression takes multiple paths.*** Microglia are known to
265 enter various activated and quiescent states. We utilized trajectory and lineage analyses to
266 determine which subtypes of microglia may be end state versus transition state as a hypothesis-
267 generating exercise. Both trajectory and lineage analyses utilize gene expression changes to
268 predict which clusters ("cellular states") may proceed from one to another. Using Monocle3[40],
269 we applied trajectory reconstruction and pseudotime analysis on the Mixed *APOE* dataset of 22
270 individuals (Supplemental Figure 1A). The pseudotime analysis suggests that Cluster 4 and
271 Cluster 6 are later pseudotime potential endpoints (Supplemental Figure 1A). We additionally
272 applied Slingshot[41] to identify lineages between subpopulations of microglia clusters. When we
273 required Slingshot to use HC1 as the starting end of its lineage algorithm, it produced three
274 lineages with endpoints of Clusters 6, 4, and 8 and all three lineages transitioned through
275 Cluster 10 (Supplemental Figure 1B). Both Monocle3 and Slingshot identified Cluster 4 and
276 Cluster 6 as endpoints or later pseudotime phenotypes, suggesting consistency for these two
277 results across multiple bioinformatic approaches. While these findings require replication, they
278 suggest that microglia may pass through multiple states to reach one of a set of endpoints that
279 represent neurodegenerative, endolysosomal, and inflammatory processes.

280 ***Homeostatic microglia differ between AD and controls.*** One of the challenges to
281 identifying microglia single nuclei transcriptomic change in AD subjects is to assay sufficient
282 numbers of microglia for gene expression analysis. Previous manuscripts have not attempted to
283 define gene expression differences between AD and Control brains within specific microglia
284 subtypes. By increasing the yield of microglia nuclei per participant, we had adequate power to

285 test for gene expression changes between AD and controls as two cohorts each contributing 12
286 or 10 people, respectively.

287 Comparison of HC1 from the AD cohort to controls revealed altered expression of
288 immune genes associated with endoplasmic reticulum stress and the unfolded protein
289 response. *HSPA5* (FC = 3.11, adjusted *p*=0.056), and *EIF2AK3* (FC = 0.70, adjusted *p*=0.083;
290 Figure 7A)[42–44] were highly expressed in AD brain compared to control. *EIF2AK3*, also known
291 as *PERK*, is activated by viral infection[42]. These genes have also been identified as influenced
292 by or influencing the production of amyloid-beta[42,43]. *CD14* (FC = 1.67, adjusted *p*=0.055; Figure
293 7A) was also more highly expressed in AD brain in HC1, and is known to interact with TLR4 to
294 enhance phagocytosis of amyloid-beta in addition to providing multiple activation paths including
295 interferon stimulation[23]. Together, the increased expression of these genes in AD brain within
296 HC1 cluster microglia suggest a more inflammatory gene expression pattern, even within this
297 "homeostatic" microglia subpopulation.

298 Cluster 3, which we identified as consistent with a phagocytic but not inflammatory state
299 also demonstrated enhanced expression of genes associated with inflammation including
300 *NFKB1* (FC=0.62, adjusted *p*=0.052) and *ZNF143* (FC=0.80, adjusted *p*=0.084) and altered
301 metabolism including *ATP13A3* (FC=1.12, adjusted *p*=0.078, Figure 7B) in AD brain compared
302 to controls. *ZNF143* is a transcription factor induced by DNA damage[45]. *ATP13A3* is localized
303 primarily to recycling endosomes[46] and may play a role in heavy metal and polyamine
304 trafficking. The differential expression of these genes in AD Cluster 3 suggests an enhancement
305 of inflammatory and altered metabolic processing in AD brain in a subpopulation of microglia
306 that did not possess an inflammatory phenotype.

307 In contrast to HC1 and Cluster 3, we did not identify any genes altered in Cluster 6. The
308 majority of the genes altered in AD within the inflammatory Cluster 8 were not associated with
309 inflammation or stress pathways. Although the genes associated with inflammation and ER

310 stress were more highly expressed in AD brain in HC1 and Cluster 3, they were unchanged in
311 Clusters 8 (Figure S2) or 10 (Table S2). These results confirm that the changes we observed in
312 the less activated subpopulations of microglia were driven by AD gene expression changes only
313 within specific microglia subpopulations.

314

315 **Discussion:**

316 This study utilized a novel sorting technique to enrich our snRNAseq dataset for
317 microglia, yielding better resolution of the diversity of microglia subpopulations and their
318 alterations in AD brain. We identified ten unique microglia subpopulations including previously
319 described "homeostatic", and "inflammatory" phenotypes, as well as novel subtypes with a
320 range of activation phenotypes and distinct features. Furthermore, we found diversity of
321 microglia subtypes with endolysosomal phenotypes and that one endolysosomal subtype in
322 particular appears to link endolysosomal function with interferon responses pathways. This
323 cluster is both enriched in AD individuals and expression of AD risk genes.

324 Immune cells are metabolically dynamic, modulating the switch from oxidative
325 phosphorylation to glycolysis in response to cell stress and activation of pattern recognition
326 receptors[27]. This metabolic reprogramming is also implicated in AD pathophysiologic
327 change[47,48]. In our study, at least two Clusters (5 and 6), seemed to have made the metabolic
328 switch to glycolysis. Cluster 5 was enriched in mitophagy and autophagy which, if functioning
329 appropriately, is a homeostatic mechanism during activation and aging[49]. In line with this
330 interpretation, the Cluster 5 phenotype did not appear to be initiating a simultaneous
331 inflammatory response in contrast to Cluster 6. A recent proteomic analysis of mouse microglia
332 has demonstrated a similar microglia autophagic lysosomal subtype present in both an AD
333 model and healthy brain[50].

The Cluster 6 subtype was distinct as proportionately larger in AD individuals, suggesting the microenvironment and stimuli contributing to this phenotype was more likely to be found in AD brain. In addition to lysosomal and vesicular function, Cluster 6 was enriched for pathways characterized by unfolded protein response and aggrephagy. We speculate this profile implies that Cluster 6 cells are initiating pathways to process accumulated protein aggregations. Cluster 6 also demonstrated expression of genes involved in the detection of cytosolic nucleic acids which can lead to activation[29,31,51]. Roy *et al*. recently reported that amyloid fibrils containing nucleic acids can induce a type I interferon response and subsequent synapse loss in an animal AD model[52]. The gene expression and pathway enrichment of Cluster 6 suggests a similar model of interferon signaling induction. Cross talk between interferon signaling and the inflammasome contributes to AD pathogenesis, and the NLRP3 inflammasome which is activated by lysosome rupture is a key component to AD neuroinflammation[3]. This study in human AD tissue suggests that the detection and response to RNA/DNA is occurring in specific microglia subtypes also enriched for endolysosomal genes and pathways.

Together, Clusters 3, 5, and 6 represent a diversity of endocytosis, vesicle trafficking, endolysosomal and autophagosome pathways defined within the dataset. Seminal work since the 1990s has implicated a key role of endolysosomal dysfunction in AD[53]. While innovative *in vitro* and *in vivo* studies in immune cells have mapped the path from lysosomal destabilization to inflammasome activation, to date there has not yet been clear evidence from human patient tissue that these processes may occur in particular subpopulations of microglia. As the first to report that a specific cluster of microglia with endolysosomal function is both enriched for AD nuclei, AD risk gene expression, and interferon activation, our results require further replication in other AD cohorts, and suggest the need for *in vitro* studies where specific transitions between different endolysosomal phenotypes can be initiated and detected experimentally.

359 Multiple studies have reproducibly identified differential expression of at least 20 AD risk
360 genes in brain from both mouse and human[15,17,18,20,24,54,55]. In mouse studies, AD risk genes
361 were found to be enriched in homeostatic microglia, activated response microglia[15], and
362 modules enriched for viral response and autophagy pathways[24]. In human brain, several studies
363 have identified AD risk genes across multiple phenotypes of microglia including both
364 homeostatic and activated subpopulations[17,18]. Many of the top International Genomics of
365 Alzheimer's Project (IGAP) genes are involved in the endolysosomal system[37], and we found
366 AD endolysosomal risk genes are selectively differentially regulated in Cluster 6. The
367 endolysosomal network is critical to maintaining neural cell homeostasis and has long been
368 implicated in AD pathogenesis. Impaired microglial endolysosomal function is often discussed in
369 the context of insufficient amyloid clearance[53,56]. Yet the endolysosomal system in microglia
370 maintains a critical role in identifying and processing foreign microbes including initiation of TLR
371 and interferon signaling[25,49]. In this study, we resolved lysosomal active microglia into a non-
372 inflammatory autophagic phenotype (Cluster 5) and a dysregulated inflammatory phenotype
373 (Cluster 6), finding the endolysosomal AD risk genes to be more strongly regulated in the latter.

374 One premise behind subtyping myeloid cells in brain tissue is that the disease process
375 may result in heterogeneous populations of microglia, some of which may be more relevant to
376 target for therapy. We found that a microglial subtype characterized by genes related to cell
377 cycle and DNA repair was depleted in AD compared to control. Microglia population dynamics
378 are known to play important roles in neurodegeneration and neuroinflammation[57]. Further
379 studies are needed both to confirm our findings of a decrease of cell-cycle associated microglia
380 in human AD and whether the decrease is due to a transition toward senescent phenotypes in
381 AD brain, which may be regionally dependent and related to neuropathological features.

382 The influence of aging is unavoidable in a study with post-mortem AD and age-matched
383 brain. Aging signatures have been described, unique to human, distinct from the activation

384 profile from murine AD models, yet sharing similarities with human AD signatures[55]. Genes
385 suggestive of an "aging signature" are expressed in all clusters in this study[58] consistent with
386 our older age cohort. Inflammaging, the immune changes associated with age and correlated
387 with chronic low-grade inflammation[59], may not only confound interpretations of gene
388 expression profiles attributable to AD, but may also contribute to the disease mechanisms
389 hypothesized to drive AD. For instance, lysosomal function and ability to maintain homeostasis
390 decreases with age[60], thus the interaction between common genetic risk, aging, and experience
391 may in aggregate raise the risk of AD over time. Additional studies exploring differences
392 between younger controls and early-onset AD may help to explore the aging, inflammaging, and
393 AD specific signatures.

394 As expected, given its frequency and strong association with AD, we observe more
395 copies of the *APOE* ε4 allele in AD cases than controls within our sample reflecting the
396 availability of tissues from our brain bank. The degree of similarity between microglia subtypes
397 when comparing the mixed *APOE* group and the *APOE* ε3/ε3 cohort was notable. However, the
398 limited sample size remains a significant caveat to drawing strong conclusions. The profile of
399 gene expression changes observed in AD cases within non-inflammatory microglia
400 subpopulations altered depending on whether the analysis was performed in an *APOE* ε3/ε3 or
401 a mixed genotype. This finding may not be surprising given the established role of the *APOE*
402 allele in AD risk as well as work suggesting mechanistic roles for APOE in AD immune
403 pathophysiology. Nevertheless, the results underline the importance of balancing and more
404 preferably, selecting samples based on *APOE* genotype and the need for larger cohorts to allow
405 for appropriate comparison.

406 Previous bulk RNAseq studies have identified upregulation of genes associated with
407 inflammation in AD brain[24,39]. While studies in mouse models have begun to differentiate
408 different forms of inflammatory phenotypes or the timing of inflammation in AD, these have yet

409 to be defined in human[24,39,54]. The significant advantage of our study is that we identified AD-
410 specific gene expression changes within the subpopulations of microglia defined by the
411 snRNAseq dataset. Cluster 1 in both Mixed *APOE* and *APOE* ε3/ε3 cohorts was considered
412 "homeostatic"[12,17,20]. In our dataset, homeostatic nuclei from AD subjects differed from their
413 control homeostatic counterparts. While both AD and control nuclei were equally represented in
414 this group, AD subjects revealed a transcriptomic signature consistent with induction of
415 inflammatory programs. This finding, that cells annotated as "homeostatic" are associated with
416 differing transcriptional profiles between control and AD, may be a relevant consideration in
417 future, larger studies to identify the AD associated cellular states.

418 Like all snRNAseq studies, there are limitations to our study. Although we analyzed a
419 subset of all *APOE* ε3/ε3 genotype individuals and identified similar subpopulations of microglia,
420 we cannot directly distinguish the impact of *APOE* ε4 in this cohort. Future studies in cohorts
421 balanced to study the interaction of sex and *APOE* genotype will be necessary to tailor
422 interventions effectively. While our PU.1 sorting technique provides a unique way to enrich our
423 dataset for microglia nuclei and allows for greater depth of resolution at an individual level, it
424 potentially selects for a specific population of microglia. There may be other subpopulations of
425 microglia that are underrepresented or not found in our dataset because of varying levels of
426 PU.1 expression. Although gene expression is a useful molecular tool for cellular subtyping, it
427 does not always directly describe protein expression, localization, or activity[61]. Future studies to
428 assess the correlation between gene and protein expression at a spatial level can provide
429 further context to our understanding of the phenotypic heterogeneity captured here.
430 Nevertheless, this novel work demonstrates the improved resolution that can be achieved in
431 autopsy brain tissue revealing phenotypes recognized *in vitro* and in other inflammatory models,
432 but previously unidentified in human AD brain.

433 Studies in animal models have described a spectrum of microglia phenotypes, though
434 confirming that diversity is present in human AD is complicated by the technical challenges of
435 capturing sufficient microglia from brain. Our novel enrichment technique allowed us to identify
436 and assess alterations within specific subpopulations of microglia isolated from human post-
437 mortem brain. Our identification of multiple internalization and trafficking clusters with varying
438 metabolic and inflammatory states provides important information for further studies to replicate
439 and investigate. The phenotype-specific alterations in both composition and gene expression in
440 AD brain provide additional evidence that specific targeting of detrimental biological function
441 rather than blanket approaches to therapeutic targeting of microglia will be critical moving
442 forward.

443

444 **Methods:**

445 Human Brain Tissue: Dorsolateral prefrontal cortex (dlPFC) tissue from human brains
446 was obtained from the Neuropathology Core of the Alzheimer's Disease Research Center
447 (ADRC) at the University Washington (UW) following informed consent approved by the UW
448 Institutional Review Board (IRB). Patients (n=12) were confirmed post-mortem to have AD
449 according to their ADNC score of 2-3 (Table 1). Control individuals (n=10) did not meet AD
450 diagnosis criteria post-mortem and had ADNC scores of 0-1 indicating low or no neuropathology
451 (Table 1).

452 Brain samples were obtained during rapid autopsy, flash-frozen in either liquid nitrogen
453 or isopentane and stored at -80ºC. Sample criteria for inclusion in this study included post-
454 mortem interval (PMI) ≤ 10 hrs, low comorbid pathology such as Lewy Bodies and hippocampal
455 sclerosis, and a brain pH at autopsy ≥ six (Table S1 provides additional demographic and
456 clinical characteristics).

457 Isolation of Nuclei from Flash Frozen Tissue for Unsorted snRNA-seq: Nuclei from
458 frozen rapid autopsy samples were isolated using protocols adapted from 10x Genomics
459 Demonstrated Protocols and De Groot *et al*.[62]. Samples were processed on ice using pre-chilled
460 solutions and RNAse-free instruments and equipment unless otherwise stated. Briefly, four
461 2mm punches of dlPFC gray matter were collected using a biopsy punch (Fisher Scientific,
462 Waltham, MA) into a 1.5mL microcentrifuge tube on dry ice. Brain punches were homogenized
463 on ice in 50µl of Nuclei Lysis Buffer (NLB: 10mM Tris-HCl pH 7.4, 10mM NaCl, 3mM $MgCl_2$,
464 0.1% NP-40 Alternative (Calbiochem, La Jolla, CA) 0.5% Protector RNase Inhibitor (Sigma
465 Aldrich, St Louis, MO) and 100µM Aurintricarboxylic acid (ATA) in nuclease-free water) using a
466 disposable pestle (USA Scientific, Ocala, FL). The pestle was rinsed into the sample tube with
467 900µL of NLB and the homogenate was thoroughly mixed using a regular-bore P1000 tip. The
468 homogenate was incubated at 4ºC under gentle agitation for 10 min, pelleted at 500 x g for 7
469 min at 4ºC and resuspended in 500µL Nuclei Suspension Solution (NSS: phosphate buffered
470 saline (PBS) supplemented with 100µM ATA, 1% bovine serum albumin (BSA), and
471 0.5% Protector RNase Inhibitor). The nuclei suspension was carefully layered onto 900µL of
472 Percoll/Myelin Gradient Buffer (PMGB) consisting of 51% myelin gradient buffer (MGB)[62], 27%
473 Percoll (GE Healthcare, Uppsala, Sweden), 3% 10X Hanks Balanced Salt Solution without
474 calcium, magnesium (HBSS; Fisher Scientific), 2.5% 1.5M NaCl, 0.1% 10mM ATA in MGB, and
475 0.5% Protector RNase Inhibitor. The gradient was centrifuged at 950 x g for 20 min at 4°C with
476 slow acceleration and no brake. Myelin and supernatant were aspirated and the nuclei pellet
477 was resuspended in Resuspension Buffer (RB: PBS containing 1% BSA and 0.5% Protector
478 RNase Inhibitor) at a concentration of 1000 nuclei/µL and proceeded immediately to single-
479 nuclei RNA sequencing (snRNA-seq).

480 Isolation of Nuclei from Flash Frozen Tissue for Fluorescence-activated Nuclei Sorting
481 (FANS): Briefly, 400-500mg of dlPFC was collected into a 1.5 ml microcentrifuge tube on dry

482 ice. Brain tissue was homogenized on ice in 100µl of NLB supplemented with the following
483 protease and phosphatase inhibitors: 1X cOmplete Mini Protease Inhibitor Cocktail (Sigma-
484 Aldrich), 0.5% Phosphatase Inhibitor Cocktail 2 (Sigma-Aldrich) and 1mM phenylmethylsulfonyl
485 fluoride (PMSF; Tocris Bioscience, Ellisville, MO) using a disposable pestle (USA Scientific,
486 Ocala, FL). The pestle was rinsed with 800 µl of NLB supplemented with protease and
487 phosphatase inhibitors into the sample tube and the homogenate was thoroughly mixed using a
488 regular-bore P1000 tip. The homogenate was incubated at 4ºC under gentle agitation for 10min,
489 pelleted at 500 x g for 7 min at 4ºC and resuspended in 900µL PMGB supplemented with
490 protease and phosphatase inhibitors. The nuclei suspension was gently overlaid with 300µL
491 NSS supplemented with protease and phosphatase inhibitors. The gradient was centrifuged at
492 950 x g for 20 min at 4ºC with slow acceleration and no brake. The myelin and supernatant
493 were aspirated and the nuclei pellet proceeded immediately to FANS.

494 Fluorescence Activated Nuclei Sorting (FANS): Nuclei were washed with cold FANS
495 media (10% fetal bovine serum (FBS), 10mM HEPES, 100µM ATA, 10% 10X HBSS, 0.5%
496 Protector RNase Inhibitor, protease and phosphatase inhibitors, and 1% saponin in nuclease-
497 free water) and resuspended in 1mL of FANS media at a concentration of 2-2.5x10^6 nuclei/mL.
498 Nuclei were blocked with 1% Human BD Fc Block (clone Fc1.3216, BD Biosciences, San Jose,
499 CA) and incubated on ice for 10 min. Nuclei were labeled with either anti-PU.1-PE (clone 9G7,
500 1:50, Cell Signaling Technology, Danvers, MA) or IgG-PE isotype control (clone DA1E, 1:50,
501 Cell Signaling Technology) for 4 hours on ice. Nuclei were washed three times with cold FANS
502 media and resuspended in 250 (isotype control) or 500µL (PU.1) FANS media supplemented
503 with 10µg/mL DAPI (Sigma-Aldrich). Nuclei were sorted using a FACSAria III Cell Sorter (BD
504 Biosciences) equipped with a 70µM nozzle until 30,000 PU.1-positive nuclei were collected in
505 RB. Sorted nuclei were centrifuged at 1,000 x g for 10 min at 4ºC. The nuclei pellet was

506 resuspended in RB at a concentration of 1000 nuclei/μL and proceeded immediately to snRNA-
507 seq.

508 Single Nuclei RNA-Sequencing (snRNA-seq): Single nuclei libraries were generated
509 using the Chromium Next GEM Single Cell 3' GEM, Library and Gel Bead Kit v3 (10x Genomics,
510 Pleasanton, CA) according to the manufacturer's protocol. Briefly, a target capture of 10,000
511 nuclei were loaded onto a channel of a Chromium Next GEM Chip B (10x Genomics) to
512 generate Gel Bead-in-Emulsions (GEMs). GEMs underwent reverse transcription to barcode
513 RNA followed by amplification, fragmentation and 5' adaptor and sample index attachment.
514 Gene expression libraries were sequenced on the NovaSeq 6000 platform (Illumina, San Diego,
515 CA) using a S2 100 cycle flow cell.

516 Alignment and Quality Control: Gene counts were obtained by aligning reads to the hg38
517 genome (GRCh38-1.2.0) using CellRanger 3.0.2 software (10x Genomics). Reads mapping to
518 precursor mRNA were included to account for unspliced nuclear transcripts. The majority of our
519 analysis was performed in R[63]. Droplets from 22 PU.1 sorted samples were combined using
520 Seurat v3.3[64]. Unsorted and PU.1 sorted droplets isolated from the same 4 subjects were
521 combined using Seurat and analyzed in the same manner. Droplets containing less than 350
522 UMIs, less than 350 genes, or greater than 1% mitochondrial genes were excluded from
523 analysis. The Seurat object was split by individual sample and ambient RNA was removed from
524 the remaining droplets using SoupX[65]. The level of ambient RNA contamination was estimated
525 using expression of mitochondrial genes, with a maximum contamination estimate set to 20%,
526 and removed. The majority of samples were estimated to have 5-12% ambient RNA
527 contamination, within the expected range for 10X RNAseq studies. Droplets containing multiple
528 nuclei were scored using Scrublet[66] with an estimated 10% doublet rate, as determined by 10x
529 Genomics. Doublet thresholds were determined manually at a local minimum in doublet score

530 and doublets were removed. 200,948 nuclei with an average of 1,787 genes per nucleus
531 remained in the dataset for further analysis.

532 Normalization and Cell Clustering: Normalization and clustering of the nuclei were
533 performed using Seurat v3.3[64]. Data were normalized for read depth using Pearson residuals
534 and mitochondrial gene content was regressed out using Seurat's `SCTransform` function[67].
535 Individual sample variability was removed using Seurat's `Anchoring` and `Integration`
536 functions[64]. Integration features were identified from the top 5000 genes across the samples
537 using `SelectIntegrationFeatures`. The data was prepared using `PrepSCTIntegration`.
538 The anchors were found using `FindIntegrationAnchors` with no reference sample, cca
539 reduction, and 30 dimensions. The top 5000 variable genes were kept and used during principle
540 component (PC) analysis using the `RunPCA` function in Seurat. 15 PCs were used to create a
541 shared nearest neighbors graph with k=20 using the `FindNeighbors` function in Seurat. The
542 modularity function was optimized using a resolution of 0.2 to determine clusters using the
543 Louvain algorithm with multilevel refinement to determine broad cell-types. Dimensionality
544 reduction of the clusters was performed using the `RunUMAP` function in Seurat. Each cluster
545 met a minimum threshold of 30 defining DEGs and was comprised of nuclei from >10% of the
546 cohort (more than two individuals).

547 Clusters were annotated for cell-type using manual evaluation for a set of known genetic
548 markers[68]. A new Seurat object was made containing only the microglia nuclei (N = 127,371).
549 Normalization, individual variability removal, integration, and shared nearest neighbors graph
550 creation were repeated as above on the microglia nuclei. 20 PCs were chosen to account for a
551 significant amount of the variance based on the elbow plot of the microglia dataset. Clusters
552 were determined using the Leiden algorithm with `method=igraph` and `weights=true`. The
553 appropriate clustering resolution was determined using 0.1 steps over a range of 0.1-0.8. The
554 number of clusters were chosen using the maximum number of clusters for which no cluster

555 contained a majority of nuclei from one sample and where the clusters appeared to contain
556 higher densities of nuclei. Clusters were highly conserved across analysis by Louvain, Louvain
557 with multilevel refinement, and Leiden algorithms. Distribution of nuclei within each cluster was
558 calculated using the '`chisq.test`' function in R[63] to compare the actual percentage of nuclei
559 from either the AD or control group within the cluster to the expected proportion of nuclei that
560 would be contributed based on our dataset composition. P-values from the chi-squared tests
561 were adjusted using FDR and considered significant if adjusted $p<0.05$.

562 snRNA-seq Differential Gene Expression and Gene Set Enrichment Analyses:
563 Differential gene expression analysis of the clusters was performed on the RNA assay of the
564 Seurat object using the `FindAllMarkers` function in Seurat, with the MAST algorithm. Genes
565 tested had expression in at least 25% of the nuclei in the cluster. Differentially expressed genes
566 (DEGs) had an FDR adjusted p-value<0.05 and a log fold change>1.25. Cluster 1 was
567 determined to be inactivated microglia nuclei. Differential gene expression analysis was
568 repeated as above comparing each other cluster to Cluster 1. GO, KEGG and Biocarta pathway
569 sets, version 7.2, were downloaded from http://www.gsea-msigdb.org/gsea/downloads.jsp.
570 Gene set enrichment analysis (GSEA) was performed using the `GSEA` function in
571 ClusterProfiler[69] modified to use a set seed for reproducibility, using the above pathway sets.
572 Genes entered into the GSEA analysis were ordered according to the log fold change so that up
573 and downregulated genes were listed together. Enriched pathways had an FDR adjusted *p*-
574 value<0.05. We considered pathways to be representative if significant results included similar
575 genes and biological functions in at least two of the three major databases (GO, KEGG,
576 Reactome).

577 Group Differential Gene Expression and Gene Set Enrichment Analyses: For each
578 sample and cluster, gene expression counts from the RNA assay of the Seurat object were
579 combined from each cell to generate a "pseudobulk" dataset. Genes expressed in fewer than

580 the number of nuclei in the smallest cluster were removed. Differential gene expression analysis
581 was performed using DESeq2. We initially included age, sex, and *APOE* ε4 allele status as
582 interaction terms in the model, but the interaction with sex did not contribute to the variance in
583 gene expression observed, so our final model included only age and *APOE* ε4 allele status as
584 interaction terms. For the subset of *APOE* ε3/ε3 genotype individuals, we only included age as
585 an interaction term in the final model after confirmation that sex was still not a significant
586 contributor. We confirmed that there were no major artifacts in DEGs determined by our
587 analysis by replicating the results utilizing an additional robust normalization method for zero-
588 inflated datasets[70]. Due to the large number of comparisons, DEGs had an FDR adjusted *p*-
589 value<0.1 and a fold change>1.25. GSEA was performed as above with the genes ordered by
590 the Wald statistic. Enriched pathways had an FDR adjusted *p*-value<0.05. As with the DEGs
591 described above, pathways were considered representative if they appeared in at least two of
592 the three major databases (GO, KEGG, Reactome).

593 Trajectory and Lineage Analysis: Trajectory analysis was performed using Monocle3[40].
594 The data were transferred to a cell dataset (cds) object and Monocle3 "`learn_graph`" was run
595 on the whole dataset. Lineage analysis was performed using Slingshot[41]. PCA and UMAP
596 embeddings were extracted from the Seurat object. We applied the algorithm both with and
597 without a defined starting point to the mixed *APOE* genotype dataset.

598 Visualizations: Visualizations were generated using ggplot2, pheatmap, and Seurat
599 native functionality[64].

600

601 **References:**

602 1. Calsolaro, V. & Edison, P. Neuroinflammation in Alzheimer's disease: Current evidence and future
603 directions. *Alzheimer's & Dementia* **12**, 719–732 (2016).
604 2. Webers, A., Heneka, M. T. & Gleeson, P. A. The role of innate immune responses and
605 neuroinflammation in amyloid accumulation and progression of Alzheimer's disease. *Immunology*
606 *& Cell Biology* **98**, 28–41 (2020).
607 3. Milner, M. T., Maddugoda, M., Götz, J., Burgener, S. S. & Schroder, K. The NLRP3 inflammasome
608 triggers sterile neuroinflammation and Alzheimer's disease. *Curr Opin Immunol* **68**, 116–124
609 (2021).
610 4. Streit, W. J., Xue, Q.-S., Tischer, J. & Bechmann, I. Microglial pathology. *Acta Neuropathol Commun*
611 **2**, (2014).
612 5. Colonna, M. & Butovsky, O. Microglia Function in the Central Nervous System During Health and
613 Neurodegeneration. *Annu Rev Immunol* **35**, 441–468 (2017).
614 6. Wolf, S. A., Boddeke, H. W. G. M. & Kettenmann, H. Microglia in Physiology and Disease. *Annu Rev*
615 *Physiol* **79**, 619–643 (2017).
616 7. Prinz, M., Jung, S. & Priller, J. Microglia Biology: One Century of Evolving Concepts. *Cell* **179**, 292–
617 311 (2019).
618 8. McQuade, A. & Blurton-Jones, M. Microglia in Alzheimer's Disease: Exploring How Genetics and
619 Phenotype Influence Risk. *Journal of Molecular Biology* **431**, 1805–1817 (2019).
620 9. Salter, M. W. & Stevens, B. Microglia emerge as central players in brain disease. *Nat. Med.* **23**,
621 1018–1027 (2017).
622 10. Hansen, D. V., Hanson, J. E. & Sheng, M. Microglia in Alzheimer's disease. *J Cell Biol* **217**, 459–472
623 (2018).

624 11. Streit, W. J., Khoshbouei, H. & Bechmann, I. Dystrophic microglia in late-onset Alzheimer's disease.
625 *Glia* **68**, 845–854 (2020).
626 12. Keren-Shaul, H. *et al.* A Unique Microglia Type Associated with Restricting Development of
627 Alzheimer's Disease. *Cell* **169**, 1276-1290.e17 (2017).
628 13. Mathys, H. *et al.* Temporal Tracking of Microglia Activation in Neurodegeneration at Single-Cell
629 Resolution. *Cell Rep* **21**, 366–380 (2017).
630 14. Mathys, H. *et al.* Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337
631 (2019).
632 15. Sala Frigerio, C. *et al.* The Major Risk Factors for Alzheimer's Disease: Age, Sex, and Genes
633 Modulate the Microglia Response to Aβ Plaques. *Cell Rep* **27**, 1293-1306.e6 (2019).
634 16. Alsema, A. M. *et al.* Profiling Microglia From Alzheimer's Disease Donors and Non-demented
635 Elderly in Acute Human Postmortem Cortical Tissue. *Front Mol Neurosci* **13**, 134 (2020).
636 17. Nguyen, A. T. *et al.* APOE and TREM2 regulate amyloid-responsive microglia in Alzheimer's disease.
637 *Acta Neuropathol* **140**, 477–493 (2020).
638 18. Olah, M. *et al.* Single cell RNA sequencing of human microglia uncovers a subset associated with
639 Alzheimer's disease. *Nat Commun* **11**, (2020).
640 19. Zhou, Y. *et al.* Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and -
641 independent cellular responses in Alzheimer's disease. *Nat Med* **26**, 131–142 (2020).
642 20. Gerrits, E. *et al.* Distinct amyloid-β and tau-associated microglia profiles in Alzheimer's disease.
643 *Acta Neuropathol* (2021) doi:10.1007/s00401-021-02263-w.
644 21. Shahidehpour, R. K. *et al.* Dystrophic microglia are associated with neurodegenerative disease and
645 not healthy aging in the human brain. *Neurobiology of Aging* **99**, 19–27 (2021).
646 22. Lawson, L. J., Perry, V. H. & Gordon, S. Turnover of resident microglia in the normal adult mouse
647 brain. *Neuroscience* **48**, 405–415 (1992).

648 23. Becher, B., Fedorowicz, V. & Antel, J. P. Regulation of CD14 expression on human adult central
649 nervous system-derived microglia. *J Neurosci Res* **45**, 375–381 (1996).
650 24. Rexach, J. E. *et al.* Tau Pathology Drives Dementia Risk-Associated Gene Networks toward Chronic
651 Inflammatory States and Immunosuppression. *Cell Reports* **33**, 108398 (2020).
652 25. Schoggins, J. W. *et al.* A diverse array of gene products are effectors of the type I interferon
653 antiviral response. *Nature* **472**, 481–485 (2011).
654 26. Wang, D. *et al.* The role of NLRP3-CASP1 in inflammasome-mediated neuroinflammation and
655 autophagy dysfunction in manganese-induced, hippocampal-dependent impairment of learning
656 and memory ability. *Autophagy* **13**, 914–927 (2017).
657 27. Lauro, C. & Limatola, C. Metabolic Reprograming of Microglia in the Regulation of the Innate
658 Inflammatory Response. *Front Immunol* **11**, 493 (2020).
659 28. Reijns, M. A. M. *et al.* Enzymatic removal of ribonucleotides from DNA is essential for mammalian
660 genome integrity and development. *Cell* **149**, 1008–1022 (2012).
661 29. Subramanian, N., Natarajan, K., Clatworthy, M. R., Wang, Z. & Germain, R. N. The adapter MAVS
662 promotes NLRP3 mitochondrial localization and inflammasome activation. *Cell* **153**, 348–361
663 (2013).
664 30. Di Virgilio, F., Dal Ben, D., Sarti, A. C., Giuliani, A. L. & Falzoni, S. The P2X7 Receptor in Infection and
665 Inflammation. *Immunity* **47**, 15–31 (2017).
666 31. Drummond, R. A. *et al.* CARD9+ microglia promote antifungal immunity via IL-1β- and CXCL1-
667 mediated neutrophil recruitment. *Nat Immunol* **20**, 559–570 (2019).
668 32. Li, L. *et al.* Thioredoxin-Interacting Protein (TXNIP) Associated NLRP3 Inflammasome Activation in
669 Human Alzheimer's Disease Brain. *Journal of Alzheimer's Disease* **68**, 255–265 (2019).
670 33. Bozina, T. *et al.* Effects of lipoprotein lipase and peroxisome proliferator-activated receptor-
671 gamma gene variants on metabolic syndrome traits. *Coll Antropol* **37**, 801–808 (2013).

672 34. Ganz, T. & Nemeth, E. HEPCIDIN AND IRON HOMEOSTASIS. *Biochim Biophys Acta* **1823**, 1434–1443
673 (2012).
674 35. Angelova, D. M. & Brown, D. R. Microglia and the aging brain: are senescent microglia the key to
675 neurodegeneration? *Journal of Neurochemistry* **151**, 676–688 (2019).
676 36. Mjelle, R. *et al.* Cell cycle regulation of human DNA repair and chromatin remodeling genes. *DNA*
677 *Repair* **30**, 53–67 (2015).
678 37. Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci
679 and implicates Aβ, tau, immunity and lipid processing. *Nat Genet* **51**, 414–430 (2019).
680 38. Lambert, J.-C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for
681 Alzheimer's disease. *Nat Genet* **45**, 1452–1458 (2013).
682 39. Friedman, B. A. *et al.* Diverse Brain Myeloid Expression Profiles Reveal Distinct Microglial
683 Activation States and Aspects of Alzheimer's Disease Not Evident in Mouse Models. *Cell Reports*
684 **22**, 832–847 (2018).
685 40. Cao, J. *et al.* The single cell transcriptional landscape of mammalian organogenesis. *Nature* **566**,
686 496–502 (2019).
687 41. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics.
688 *BMC Genomics* **19**, 477 (2018).
689 42. O'Connor, T. *et al.* Phosphorylation of the Translation Initiation Factor eIF2α Increases BACE1
690 Levels and Promotes Amyloidogenesis. *Neuron* **60**, 988–1009 (2008).
691 43. Correani, V. *et al.* Plasma Membrane Protein Profiling in Beta-Amyloid-Treated Microglia Cell Line.
692 *PROTEOMICS* **17**, 1600439 (2017).
693 44. Ghemrawi, R. & Khair, M. Endoplasmic Reticulum Stress and Unfolded Protein Response in
694 Neurodegenerative Diseases. *Int J Mol Sci* **21**, 6127 (2020).

695 45. Ishiguchi, H. *et al.* ZNF143 activates gene expression in response to DNA damage and binds to
696 cisplatin-modified DNA. *Int J Cancer* **111**, 900–909 (2004).
697 46. Sørensen, D. M. *et al.* Parkinson disease related ATP13A2 evolved early in animal evolution. *PLoS*
698 *One* **13**, e0193228 (2018).
699 47. Baik, S. H. *et al.* A Breakdown in Metabolic Reprogramming Causes Microglia Dysfunction in
700 Alzheimer's Disease. *Cell Metab* **30**, 493-507.e6 (2019).
701 48. Johnson, E. C. B. *et al.* Large-scale proteomic analysis of Alzheimer's disease brain and
702 cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and
703 astrocyte activation. *Nat Med* **26**, 769–780 (2020).
704 49. Van Acker, Z. P., Bretou, M. & Annaert, W. Endo-lysosomal dysregulations and late-onset
705 Alzheimer's disease: impact of genetic risk factors. *Mol Neurodegener* **14**, 20 (2019).
706 50. Burns, J. C. *et al.* Differential accumulation of storage bodies with aging defines discrete subsets of
707 microglia in the healthy brain. *Elife* **9**, e57495 (2020).
708 51. Song, X. *et al.* DNA repair inhibition leads to active export of repetitive sequences to the cytoplasm
709 triggering an inflammatory response. *J Neurosci* JN-RM-0845-21 (2021)
710 doi:10.1523/JNEUROSCI.0845-21.2021.
711 52. Roy, E. R. *et al.* Type I interferon response drives neuroinflammation and synapse loss in Alzheimer
712 disease. *J Clin Invest* **130**, 1912–1930 (2020).
713 53. Nixon, R. A. Amyloid precursor protein and endosomal-lysosomal dysfunction in Alzheimer's
714 disease: inseparable partners in a multifactorial disease. *FASEB J* **31**, 2729–2743 (2017).
715 54. Rangaraju, S. *et al.* Identification and therapeutic modulation of a pro-inflammatory subset of
716 disease-associated-microglia in Alzheimer's disease. *Mol Neurodegener* **13**, 24 (2018).
717 55. Srinivasan, K. *et al.* Alzheimer's Patient Microglia Exhibit Enhanced Aging and Unique
718 Transcriptional Activation. *Cell Reports* **31**, 107843 (2020).

719 56. Podleśny-Drabiniok, A., Marcora, E. & Goate, A. M. Microglial Phagocytosis: A Disease-Associated
720 Process Emerging from Alzheimer's Disease Genetics. *Trends in Neurosciences* **43**, 965–979 (2020).
721 57. Martin-Estebane, M. & Gomez-Nicola, D. Targeting Microglial Population Dynamics in Alzheimer's
722 Disease: Are We Ready for a Potential Impact on Immune Function? *Front Cell Neurosci* **14**, 149
723 (2020).
724 58. Olah, M. *et al.* A transcriptomic atlas of aged human microglia. *Nat Commun* **9**, 539 (2018).
725 59. Franceschi, C., Garagnani, P., Parini, P., Giuliani, C. & Santoro, A. Inflammaging: a new immune–
726 metabolic viewpoint for age-related diseases. *Nat Rev Endocrinol* **14**, 576–590 (2018).
727 60. Cuervo, A. M. & Dice, J. F. When lysosomes get old☆. *Experimental Gerontology* **35**, 119–131
728 (2000).
729 61. Koussounadis, A., Langdon, S. P., Um, I. H., Harrison, D. J. & Smith, V. A. Relationship between
730 differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. *Sci*
731 *Rep* **5**, 10775 (2015).
732 62. De Groot, C. j. a., Hulshof, S., Hoozemans, J. j. m. & Veerhuis, R. Establishment of microglial cell
733 cultures derived from postmortem human adult brain tissue: Immunophenotypical and functional
734 characterization. *Microscopy Research and Technique* **54**, 34–39 (2001).
735 63. R Core Team. *R: A language and environment for statistical computing*. (R Foundation for
736 Statistical Computing, 2020).
737 64. Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902.e21 (2019).
738 65. Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet-based single-
739 cell RNA sequencing data. *Gigascience* **9**, giaa151 (2020).
740 66. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell Doublets in
741 Single-Cell Transcriptomic Data. *Cell Syst* **8**, 281-291.e9 (2019).

742 67. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data
743 using regularized negative binomial regression. *Genome Biol* **20**, 296 (2019).
744 68. Hodge, R. D. *et al.* Conserved cell types with divergent features in human versus mouse cortex.
745 *Nature* **573**, 61–68 (2019).
746 69. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation*
747 **2**, (2021).
748 70. Chen, L. *et al.* GMPR: A robust normalization method for zero-inflated count data with application
749 to microbiome sequencing data. *PeerJ* **6**, e4600 (2018).
750

751

752 **Table 1. Post-mortem brain sample demographics.**

Group	Males	Females	Avg. Age	Avg. PMI	ADNC	E2/3	E3/3	E3/4	E4/4
Ctrl	4	6	85.9	5.53	0-1	2	7	1	
AD	3	9	86.5	4.67	2-3		6	5	1
Total	7	15	86.23	5.42		2	13	6	1

753

754 Ctrl = Control, AD = Alzheimer's Disease pathology, Avg. Age = years of age at death, PMI = post-
755 mortem interval in hours, ADNC = Alzheimer's Disease Neuropathic Change, *APOE* Genotypes: *APOE*
756 alleles 2/3 (E2/3), *APOE* alleles 3/3 (E3/3), *APOE* alleles 3/4 (E3/4), *APOE* alleles 4/4 (E4/4)

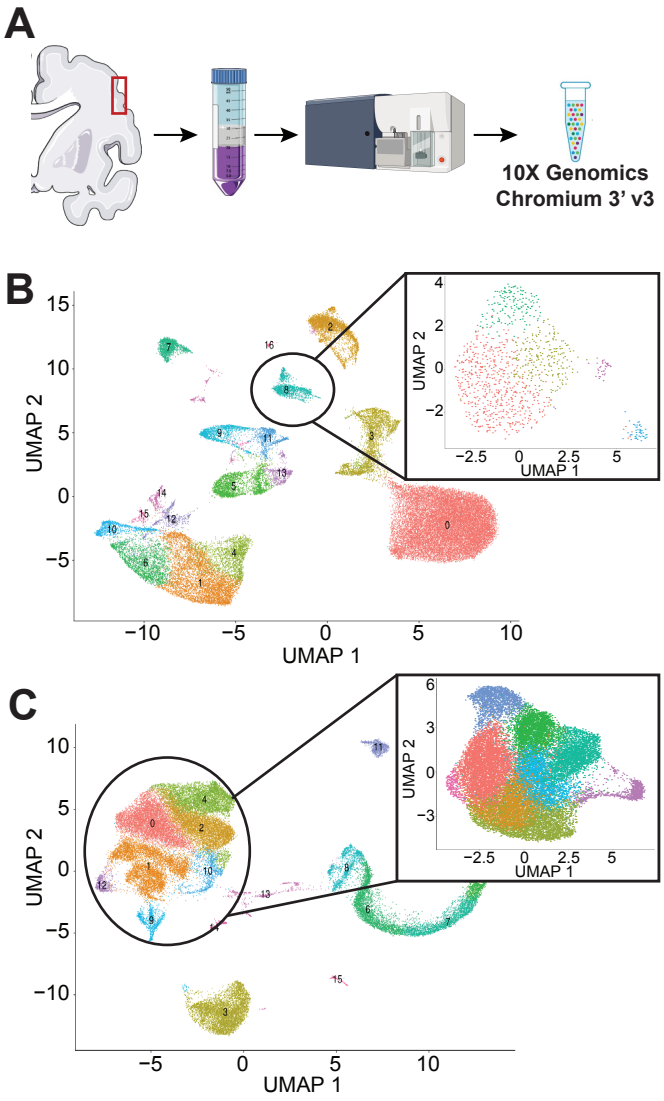


Figure 1. PU.1 enrichment increases the number of microglia nuclei and enhances microglia subpopulation resolution in snRNAseq studies. A) Nuclei isolated from dorsolateral prefrontal cortex grey matter of postmortem human brain tissue sorted by fluorescence-activated nuclei sorting using the myeloid marker PU.1 and then sequenced on the 10X Genomics Chromium 3' v3 platform. (Created in part with BioRender.com) **B)** Unsorted snRNAseq data from four samples demonstrates multiple brain cell types and a small population of microglia (N = 1032 cells) that can be further subdivided into five subpopulations. **C)** After PU.1 enrichment, a snRNAseq dataset from the same four individuals contains a large number of microglia (N = 23,310 cells), and these microglia can be further subclustered into nine subpopulations.

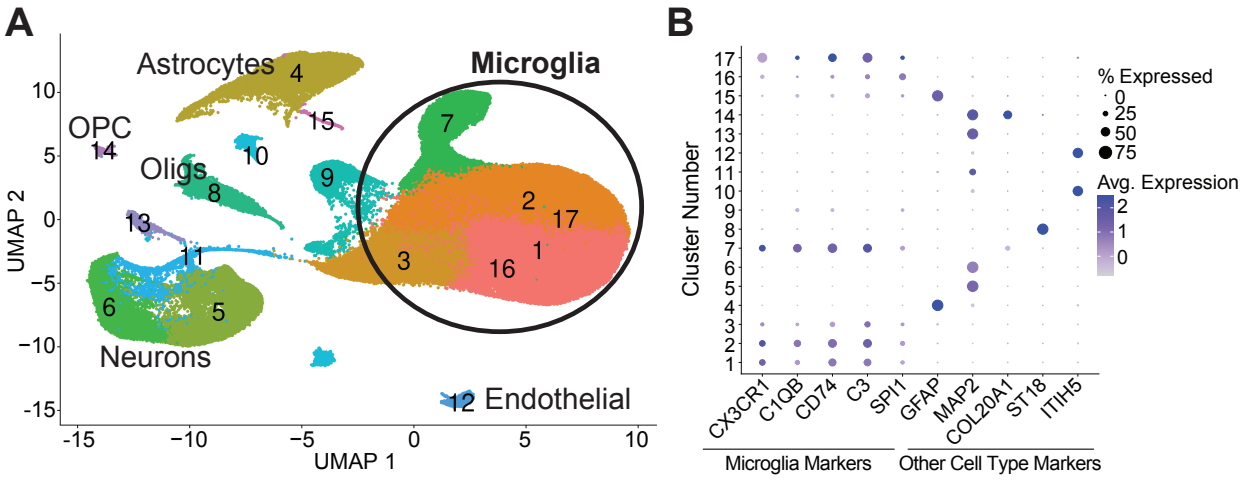


Figure 2. PU.1 enrichment provides increased numbers of microglia nuclei. A) UMAP of all nuclei in the 22 subject dataset demonstrates that while other cell types including neurons, astrocytes, oligodendrocytes (Oligs) and their progenitors (OPC) as well as endothelial cells are present, six clusters including the three largest are composed of microglia nuclei. **B)** Representative cell type marker genes (x-axis) with the percent of nuclei that express a gene (size of dot) in each cluster (distributed along Y axis) and the average expression level (color intensity) are shown for microglia (*CX3CR1, C1QB, CD74*, and *C3*), astrocytes (*GFAP*), neurons (*MAP2*), OPCs (*COL20A1*), oligodendrocytes (*ST18*), and endothelial cells (*ITIH5*) for each cluster.

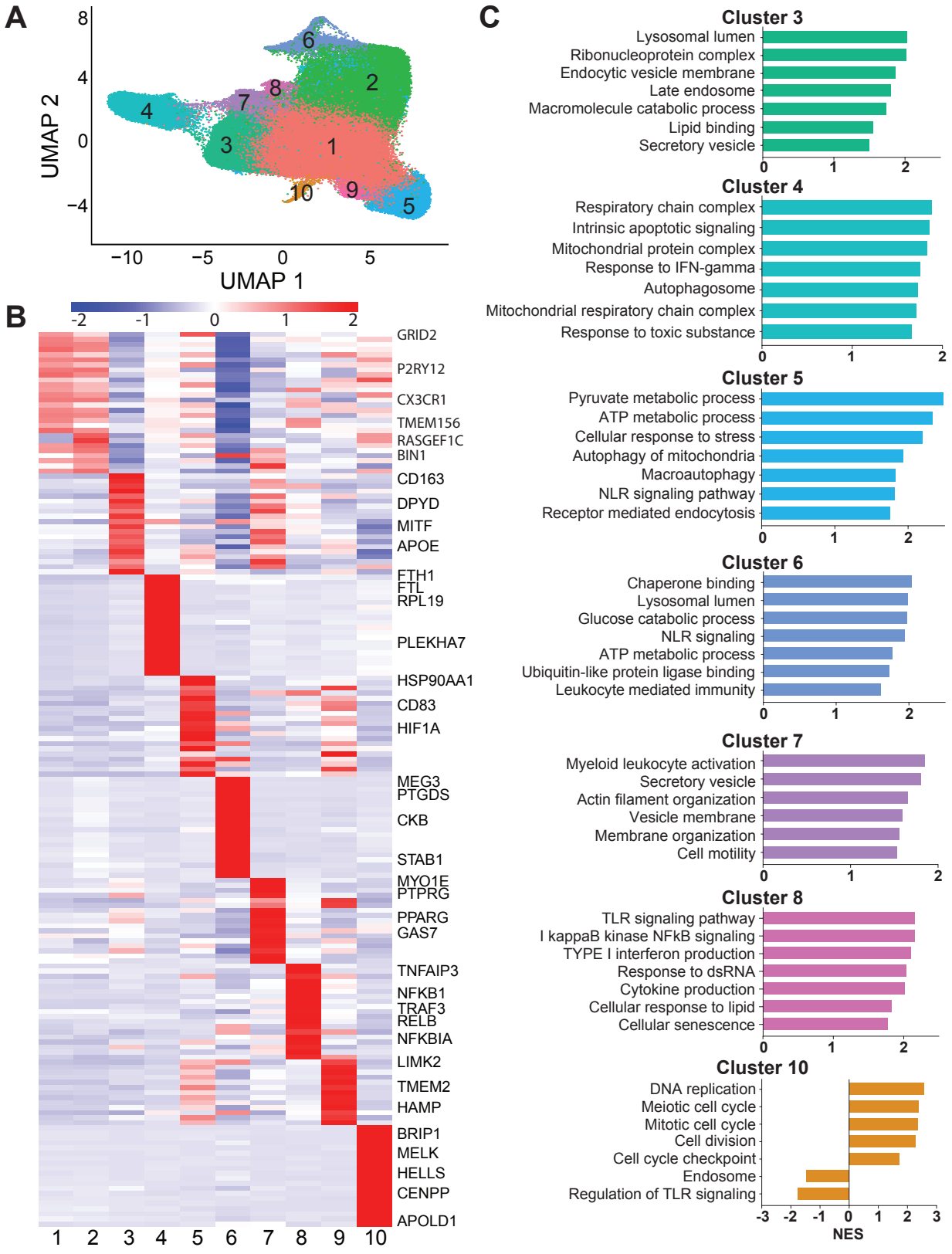
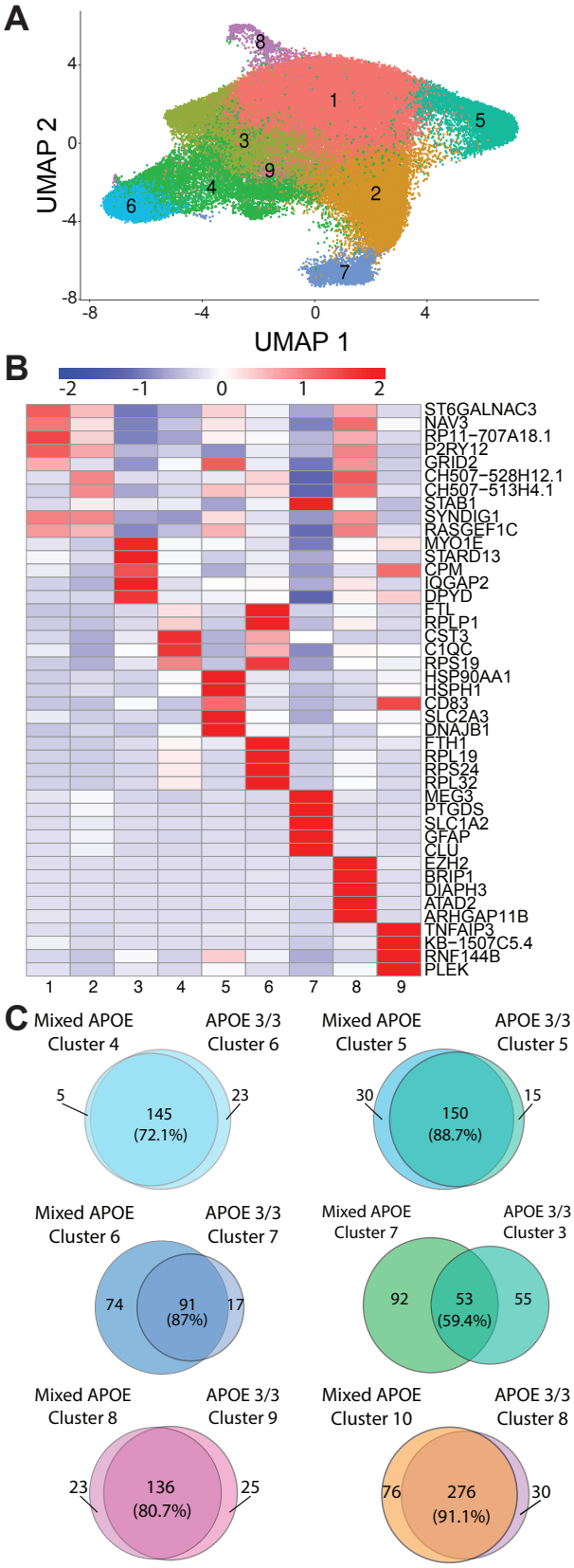


Figure 3. Microglia subpopulations reveal a wide diversity of gene expression and biological pathway correlates. A) UMAP of unbiased clustering on the mixed APOE genotype 22 sample dataset

779 contains 10 microglia subpopulations. **B)** Differential expression analysis comparing each cluster to all
780 other clusters demonstrates distinct gene expression profiles for each. The top 25 genes from each cluster
781 are displayed in the heatmap, with gene names of particular biological interest annotated on the right.
782 Cluster 1 is high in canonical microglia genes (*CX3CR1* and *P2RY12*) thought to be highly expressed in
783 an unactivated state, so we label this cluster "homeostatic". **C)** GSEA analysis of genes that differentiate
784 each cluster from cluster 1 ("homeostatic") suggest distinct biological pathways and function correlates
785 for each cluster. Pathways displayed are GO, but are representative of pathways that were significantly
786 altered in two of KEGG, Reactome, and GO databases.

Figure 4. APOE 3/3 genotype does not substantially alter microglial clustering in human autopsy brain. A) UMAP of unbiased clustering on 13 samples of only APOE 3/3 individuals shows 9 clusters. **B)** Similar to the clusters identified in the mixed *APOE* genotype dataset, the clusters identified in our *APOE* ε3/ε3 genotype dataset are highly distinct by gene expression. The top 5 genes are displayed for each cluster. **C)** Venn diagrams demonstrating overlap between clusters from the Mixed *APOE* and *APOE* ε3/ε3 cohorts demonstrating significant similarity in gene expression profiles.



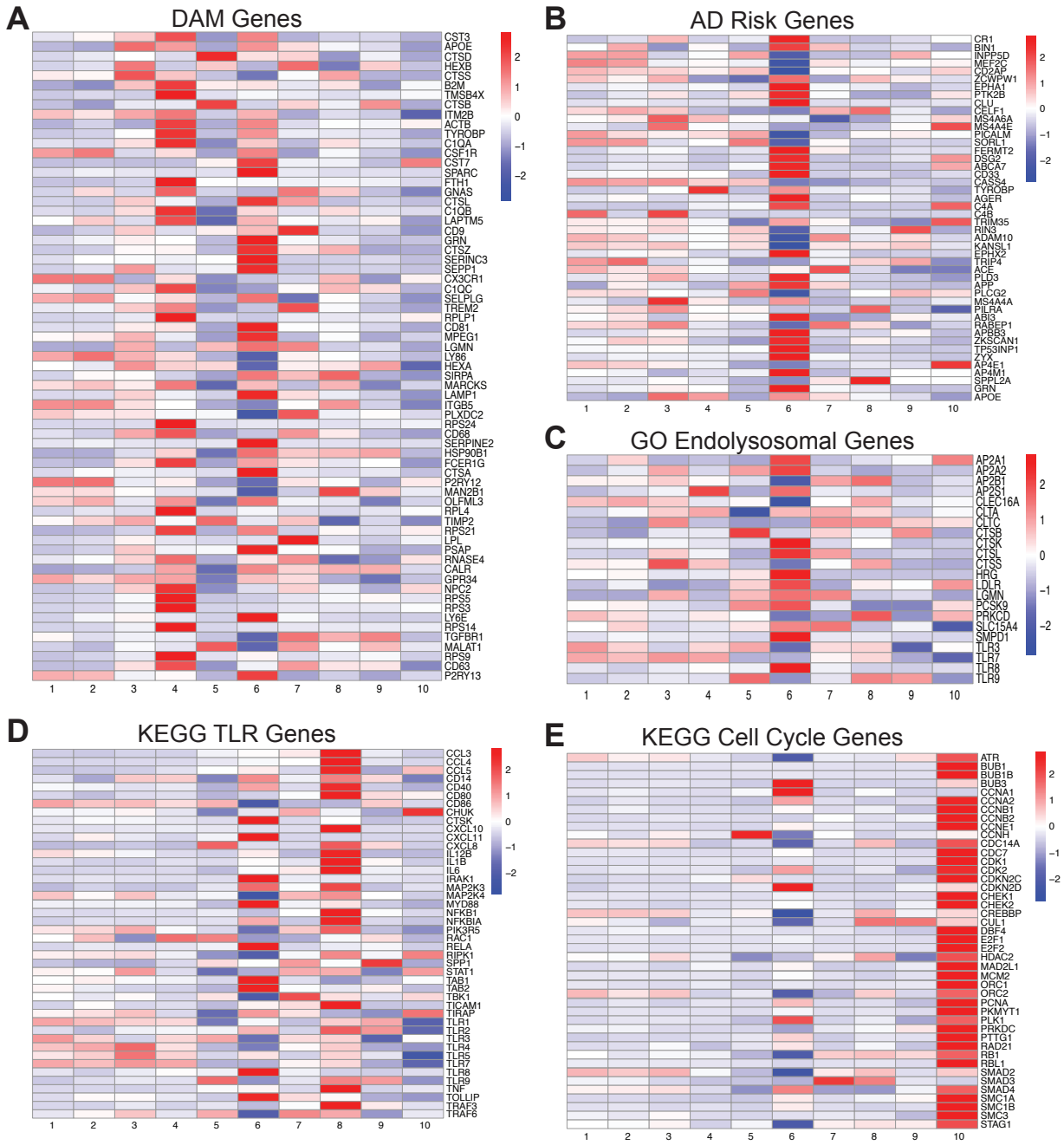


Figure 5. Previously defined gene lists demonstrate unique expression in microglia subpopulations, including AD risk genes specifically enriched in Cluster 6. A) Diseas-assocated microglia (DAM) genes as defined by Keren-Shaul et al. (2017) were statistically enriched in Cluster 4 using GSEA analysis, and the heatmap demonstrates upregulation of many of these genes. Cluster 6 also demonstrates upregulation of some DAM genes, distinct from those in Cluster 4. **B)** Genes associated with AD risk SNPs are highly enriched in Cluster 6. **C)** Heatmap of genes from the GO endolysosomal pathway gene set demonstrating significant enrichment in Cluster 6, and Clusters 3, 5, and 7. **D)** Heatmap of representative genes from the KEGG toll-like receptor pathway gene set illustrating significant enrichment in Cluster 8, of many genes (Supplemental data) as would be expected given the inflammatory

phenotype. **E)** Gene expression of representative genes from the KEGG cell cycle pathway gene set demonstrates significant enrichment in Cluster 10.

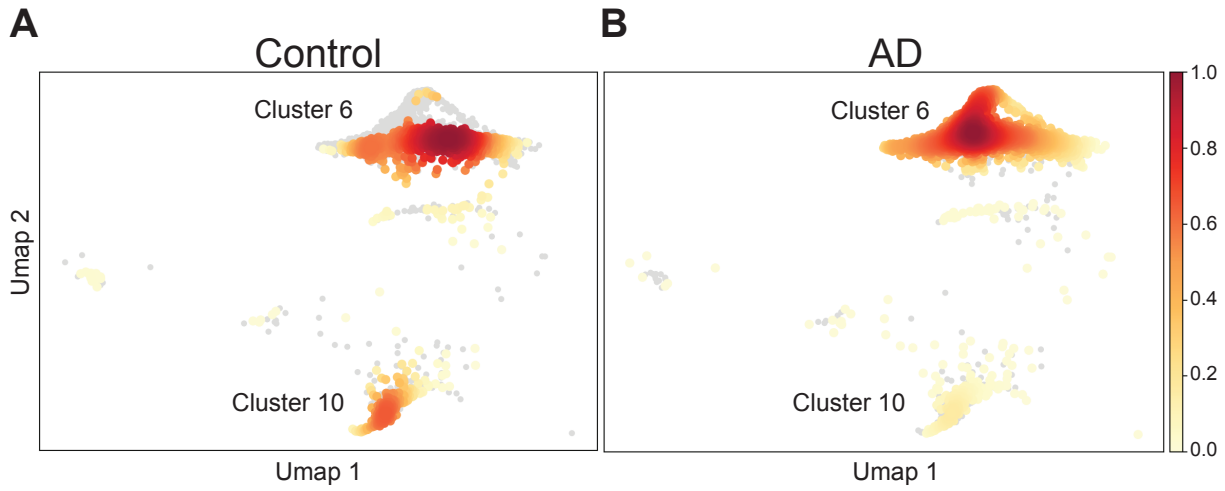


Figure 6. Cluster 6 is larger and Cluster 10 is depleted in AD brain. Normalized density of nuclei is displayed by color; dark red represents the highest density of nuclei on the UMAP, yellow the least dense collection of nuclei. Grey dots indicate nuclei from the opposite group (i.e. grey dots in controls indicate nuclei that are from AD individuals). **A)** Density of nuclei in Cluster 6 and Cluster 10 in control individuals compared to **B)** density of nuclei in Cluster 6 and Cluster 10 in AD individuals. Cluster 6 is more highly represented by AD nuclei, while Cluster 10 shows higher representation of controls.

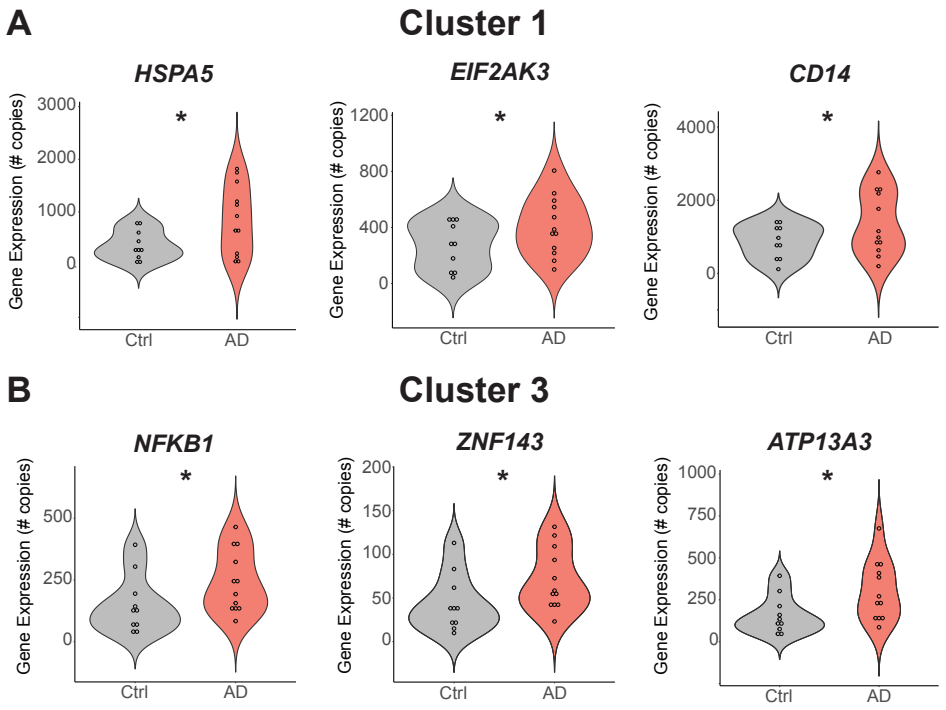


Figure 7. Unactivated microglia subpopulations show upregulation of inflammatory and stress-related gene expression in AD brain. A) Cluster 1 genes that were significantly upregulated in AD brain compared to Controls (Ctrl) included *HSPA5* (FDR corrected $q = 0.056$), *EIF2AK3* ($q = 0.083$), and *CD14* ($q = 0.055$) genes involved in inflammation and endoplasmic reticulum stress processes. **B)** Cluster 3 genes that were significantly upregulated in AD brain compared to Controls include *NFKB1* ($q = 0.052$), *ZNF143* ($q = 0.084$), and *ATP13A3* ($q = 0.078$) involved in NFκB signaling, response to DNA damage and recycling endosomes. Pro-inflammatory changes were not identified in more activated clusters, Cluster 6 and Cluster 8, suggesting that these are subpopulation-specific effects.