

Impact of Early-Life Inflammation and APOE Genotype on the Microglial Omic Profile During Alzheimer's Disease Pathology

Logan Larlham

Supervised by:
Rosalía Fernández Calle & Yiyi Yang
Tomas Deierborg

Master's in Bioinformatics, Course BINP52, 60cr

Email: LoganHLarlham@gmail.com

Supervisor email: rosalia.fernandez_calle@med.lu.se, yiyi.yang@med.lu.se, tomas.deierborg@med.lu.se
Experimental Neuroinflammation Laboratory, Department of Experimental Medical Sciences, Faculty of Medicine, Lund University, Lund, Sweden

1 Abstract

2 Alzheimer's disease (AD) is characterized by amyloid- β (A β) plaque deposition, tau pathology, and neuroinflammation, with genetic variation in the apolipoprotein E (APOE) gene representing the strongest risk factor for late-onset AD. Microglia, the brain's resident immune cells, are increasingly recognized as central to disease progression through their roles in neuroinflammation and A β clearance. In this study, we investigated how APOE genotype (E3 vs E4) and early-life inflammation interact to shape microglial states in a mouse model of A β pathology. Using single-cell RNA sequencing of microglia isolated from EFAD mice treated with lipopolysaccharide (LPS) or vehicle at postnatal day 9, we identified distinct microglial transcriptional states and assessed their proportions and gene expression profiles. E4FAD mice exhibited a higher proportion of disease-associated microglia (DAM) compared to E3FAD mice. Moreover, early-life inflammation induced long-lasting, genotype-specific alterations: LPS reduced the abundance of an NF- κ B-enriched Cytokine Response microglial population and downregulated pro-inflammatory genes in E4FAD, but not E3FAD, mice. These findings suggest that early-life immune events interact with APOE genotype to shape microglial responses to amyloid pathology and may offer targets for modulating disease-associated inflammation.

15 Introduction

16 *Alzheimer's Disease and Neuroinflammation*

17 Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the leading cause of dementia worldwide, affecting more than 55 million people as of 2019 (1). It is defined by the accumulation of extracellular amyloid- β (A β) plaques and intracellular tau neurofibrillary tangles, which are now considered central pathological hallmarks of the disease (2).

21 In addition to these proteinopathies, AD is characterized by widespread synaptic loss, neuronal degeneration, and a robust neuroinflammatory response. Microglia, the resident immune cells of the central nervous system (CNS), are key contributors to this inflammatory milieu. They participate in A β clearance, phagocytosis, synaptic pruning, and cytokine production, and have been implicated in both protective and pathogenic roles.

25 The increasing recognition of neuroinflammation as a core feature of AD has led to renewed interest in
26 microglial biology, particularly in the context of genetic and environmental risk factors that may shape microglial
27 responses during disease progression.

28 ***Microglia in Alzheimer's Disease***

29 Microglia are highly dynamic cells that continually survey the brain parenchyma. Their functions are diverse and
30 include responding to injury, clearing debris, and remodeling synapses. Early frameworks categorized microglia
31 as either "resting" or "activated," or applied the simplified 'M1/M2' designation. These binary models are now
32 considered oversimplified and inadequate for describing the true range of microglial phenotypes observed in
33 health and disease.

34 Advances in single-cell RNA sequencing (scRNA-seq) have transformed our understanding of microglial
35 heterogeneity. In particular, studies in mouse models of AD have identified a transcriptional program termed Dis-
36 ease-Associated Microglia (DAM), characterized by downregulation of homeostatic markers such as Tmem119
37 and P2ry12, and upregulation of genes involved in lipid metabolism, phagocytosis, and inflammation, including
38 Trem2, Apoe, Cst7, and Tyrobp (3), (4).

39 Subsequent work has revealed additional microglial subsets, including interferon-responsive, antigen-pre-
40 senting, and proliferative populations. However, there is limited consensus regarding nomenclature or functional
41 distinctions between these clusters. Moreover, it remains unclear how these states arise, how stable they are
42 across contexts, and how they relate to disease mechanisms *in vivo*.

43 ***Sex Differences in Alzheimer's Disease***

44 One important but often underappreciated aspect of AD is its differential impact on biological sex. Women
45 represent nearly two-thirds of AD patients, and several studies have reported that female individuals exhibit
46 more severe A β plaque burden, increased tau pathology, and faster cognitive decline compared to males (5). These
47 differences may be influenced by sex hormones, immune system regulation, and differential gene expression,
48 although the underlying mechanisms remain incompletely understood.

49 Sex also interacts with genetic risk factors. The APOE4 allele has been associated with a greater increase in
50 AD risk in women compared to men (6). In light of these findings, and because immune responses are known to
51 differ by sex, this study focuses exclusively on female mice. This choice enhances the interpretability of microglial
52 responses in a biologically relevant context and avoids the confounding influence of sex-dependent variability.

53 *APOE Genotype Modulates Microglial Function*

54 The apolipoprotein E (APOE) gene is the strongest known genetic risk factor for late-onset Alzheimer's disease
55 (AD). The three most common alleles, APOE2, APOE3, and APOE4, differ by two amino acid substitutions. These
56 differences give rise to distinct protein isoforms that vary in their effects on lipid metabolism and amyloid- β
57 ($\text{A}\beta$) clearance (7). Among them, APOE4 is linked to increased $\text{A}\beta$ aggregation, reduced clearance, disrupted
58 microglial lipid processing, and altered inflammatory signaling, while APOE2 is considered protective (8), (9).

59 Under normal physiological conditions, APOE is primarily produced by astrocytes. However, in the presence
60 of $\text{A}\beta$ pathology, APOE expression in microglia increases substantially. Experimental models using APOE
61 knockouts or humanized APOE alleles have shown that APOE regulates key microglial functions, including
62 transcriptional responses to $\text{A}\beta$, phagocytic activity, and the production of inflammatory mediators (4), (10).

63 These observations support the hypothesis that APOE isoforms play a critical role in directing microglial
64 behavior and may influence their progression into DAM or other disease-associated phenotypes.

65 *5xFAD and EFAD*

66 Studying the interplay between APOE genotype and microglial activation in humans is difficult due to limited
67 access to living brain tissue and the confounding effects of postmortem delay. As a result, transgenic mouse
68 models have become essential tools for investigating AD mechanisms.

69 The 5xFAD mouse model expresses five familial AD mutations in APP and PSEN1, leading to accelerated
70 $\text{A}\beta$ deposition, gliosis, and synaptic degeneration by six months of age (11). However, this model expresses the
71 murine Apoe gene, which differs from the human gene in structure and function.

72 To address this limitation, the EFAD model was created by crossing 5xFAD mice with human APOE knock-
73 in lines. EFAD mice express human APOE2, APOE3, or APOE4 alleles and recapitulate many of the genotype-

74 specific effects observed in humans. E4FAD mice, in particular, exhibit higher plaque load and more pronounced
75 neuroinflammatory signatures compared to E3FAD or E2FAD mice (12).

76 This model provides a powerful system for dissecting how human APOE isoforms influence microglial
77 responses to A β in vivo.

78 *Microglial Priming*

79 In addition to genetic risk, microglial states may be shaped by early-life environmental exposures. One mecha-
80 nism through which this occurs is innate immune memory, a process by which transient inflammatory stimuli
81 produce long-lasting changes in microglial function (13).

82 Previous studies have shown that systemic administration of lipopolysaccharide (LPS) during early devel-
83 opment can lead to persistent alterations in microglial gene expression, inflammatory responses, and behavior
84 later in life (14). However, these effects are context-dependent and may differ by genetic background, sex, and
85 timing of exposure.

86 Work from our laboratory has demonstrated that LPS injection at postnatal day 9 (P9) improves long-term
87 memory performance and enhances microglial internalization of A β in 5xFAD mice (15). These findings suggest
88 that early immune activation may modulate the trajectory of microglial aging and response to pathology, possibly
89 in a beneficial manner.

90 *Study Rationale and Objectives*

91 The current study was designed to test the hypothesis that early-life inflammation interacts with APOE genotype
92 to shape microglial transcriptional states in the context of A β pathology. To investigate this, we used EFAD mice
93 expressing human APOE3 or APOE4 alleles. At postnatal day 9, animals were treated with LPS or vehicle, and
94 at six months of age, microglia were isolated for single-cell RNA sequencing.

95 This approach allowed us to examine both cell-type proportions and transcriptional profiles across multiple
96 experimental conditions. Our aim was to determine whether APOE isoform modifies the microglial response to
97 early-life priming, with a particular focus on disease-associated and pro-inflammatory transcriptional states.

98 This study is part of a broader investigation integrating transcriptomics,  genomics, histology, and
99 behavioral analysis to understand how genetic and environmental risk factors converge to influence neuroin-
100flammation in AD.

101 Materials and methods

102 Animals

103 Humanized APOE3 and APOE4 knock-in mice were crossed with the 5xFAD mouse model to generate E3FAD and
104 E4FAD mice, as previously described (Youmans et al., 2012). Wild-type (WT) littermates were used as controls.
105 Mice were housed in a controlled environment (12-hour light/dark cycle, constant temperature and humidity)
106 with free access to food and water. All procedures were conducted in compliance with the European Union
107 laboratory regulations for animal experiments and were approved by the Animal Research Committee of Lund
108 University.

Table 1: Summary of Number of Animals by Genotype and Treatment

Genotype	Treatment	# of Animals
E3WT	Vehicle	4
	LPS	3
E3FAD	Vehicle	4
	LPS	3
E4WT	Vehicle	4
	LPS	3
E4FAD	Vehicle	4
	LPS	5
Total Animals		30

109 Experimental procedure and sample processing

110 Only female mice were used in this study, as described in the Introduction. At postnatal day 9 (P9), E3FAD, E4FAD,
111 and their wild-type (WT) littermates received a single intraperitoneal (i.p.) injection of either lipopolysaccharide
112 (LPS, 1 mg/kg) or vehicle (0.9% saline) (see Table 1 for group assignments). At six months of age, mice were
113 euthanized by transcardial perfusion. Brains were rapidly dissected, and the right hemisphere was used for
114 microglial cell isolation

115 Microglial enrichment was performed using magnetic-activated cell sorting (MACS) with CD11b Mi-
116 croBeads and LS columns. Enriched microglia were further purified by fluorescence-activated cell sorting (FACS)
117 using an ARIA III flow cytometer. Cells positive for CD45, CX3CR1, and CD11b were collected for downstream
118 analyses.

119 **Cell Sequencing**

120 From each experimental sample, approximately 2,500 sorted microglial cells were loaded for single-cell capture
121 and cDNA library preparation using the 10x Genomics Chromium Single Cell 3' v3 reagent kit and workflow.
122 Individual samples were barcoded using cell-hashing antibodies and pooled into groups of four samples per
123 sequencing run. Libraries were sequenced on an Illumina platform.

124 Raw sequencing data were processed with CellRanger (10x Genomics, v6.1.2) using the mouse reference
125 genome GRCm39 (mm39). This pipeline included alignment, filtering, barcode counting, and UMI (unique
126 molecular identifier) counting to generate feature-barcode matrices for each library.

127 **Quality Control, Normalization, and Data Correction**

128 Individual libraries were first filtered to remove cells identified as empty droplets or doublets using the HashSolo
129 algorithm (16). Additional quality control (QC) filters were applied to each library: cells with fewer than 2,000
130 total counts, fewer than 300 expressed genes, greater than 5% mitochondrial gene expression, or fewer than 5%
131 ribosomal gene expression were removed. The 5% mitochondrial threshold was selected based on recommenda-
132 tions by Osorio and Cai, who showed that this cutoff effectively excludes apoptotic cells in mouse scRNA-seq data
133 and improves downstream interpretability (17). Genes expressed in fewer than three cells were also excluded. QC
134 metrics and visualizations for each sample, both pre- and post-filtering, are included in Supplementary Figure 1
135 and Supplementary Table 1.

136 After filtering, individual libraries were merged into a single dataset using Anndata's concat function with
137 an outer join along the cell axis and a unique merge for the gene axis. Raw counts were saved as a separate layer
138 before normalization. Normalization was performed using Scanpy's normalize_total function followed by log-
139 transformation with log1p.

140 Following initial normalization, the merged dataset was examined for the presence of contaminating
141 non-microglial cells. Based on clustering, expression of canonical marker genes, and manual annotation, cells
142 identified as non-microglia were removed. The dataset was re-normalized after removal of non-microglial cells
143 using the same normalization strategy.

144 Highly variable genes (HVGs) were then identified using Scanpy's highly_variable_genes function with
145 sample ID as the batch key. These HVGs were used for subsequent dimensionality reduction steps, including
146 principal component analysis (PCA).

147 The merged dataset was assessed for batch effects using PCA. Several batch correction methods were tested,
148 including Harmony (integration via latent space correction), BBKNN (graph-based nearest neighbor correction),
149 and ComBat (empirical Bayes adjustment). Ultimately, batch effects were not observed to be significant, and no
150 batch correction was applied.

151 A high contribution of mitochondrial genes to variance across cells was observed during HVG and PCA
152 analysis. To reduce this effect, mitochondrial gene expression was regressed out using Scanpy's regress_out
153 function prior to downstream analyses.

Table 2: Summary of Datasets Before and After Quality Control

Library	Before		After	
	# of Cells	# of Genes	# of Cells	# of Genes
D1	3403	33989	2903	14987
D2	5041	33989	4698	15514
D3	5029	33989	4081	15463
D4	3892	33989	3558	15052
D5	3825	33989	3606	15260
D6	4408	33989	3603	15788
D7	3643	33989	3103	14935
D8	3903	33989	3601	15926
Merged	-	-	29153	17341

154 Clustering and Annotation

155 Principal component analysis (PCA) was computed on normalized, log-transformed data. Neighborhood graphs
156 were constructed using the top 25 principal components as input after analyzing the scree plot, and uniform

157 manifold approximation and projection (UMAP) embeddings were calculated with a min_dist parameter of 0.5
158 to preserve both local and global structures.

159 Leiden clustering (18) was performed iteratively across a range of resolutions (0.1–1.9 in steps of 0.1). UMAPs
160 colored by cluster identities were visually inspected at each resolution to identify regions of high heterogeneity
161 and monitor for under- and overclustering. Annotation started with the highest resolution clustering to capture
162 fine-grained heterogeneity and then merged clusters that showed high similarity in their gene expression
163 profiles. Overclustering was identified when newly created clusters did not significantly differ from neighboring
164 clusters in their top differentially expressed genes (DEGs).

165 DEGs were identified using the Wilcoxon rank-sum test (rank_genes_groups function). DEGs were calcu-
166 lated for each cluster against the total dataset and neighbouring clusters to merge clusters with insufficiently
167 distinct expression profiles. Expression profile annotations were assigned based on the combination of differen-
168 tial expression profiles, cluster relationships, and comparison to canonical gene signatures described in previous
169 studies, including Keren-Shaul et al. (3), Olah et al. (19), Prater et al. (20), Mancuso et al. (21), Sala-Frigerio et
170 al. (22), Mathys et al. (23), Millet et al. (24), (25), (26).

171 Clusters were annotated into 9 expression profiles; Homeostatic, DAM, MHC-II/Antigen Presentation DAM,
172 Cytokine Response, Cycling(G2M/Cdk1+), BAM-like, Neuronal surveillance/Ribosomal biogenesis, DAM with
173 Ribosomal upregulation, and Immediate Early Gene (IEG) Enriched Homeostatic. Naming of profiles was based
174 on similarity to previously described profiles in the literature, though should not be considered a formal defin-
175 ition.

176 Cluster Proportions

177 The proportion of each cluster was calculated as the number of cells in the cluster divided by the total number
178 of cells in the sample. A three-way ANOVA was performed to assess the effect of genotype (E3 vs E4), treatment
179 (LPS vs Vehicle), and disease status (FAD vs WT) on the proportion of each cluster. An ordinary least squares
180 (OLS) linear regression model was fit to the data with the proportion of each cluster as the dependent variable
181 and the variable main effects and interactions as predictors. Tukey's Honest Significant Difference (HSD) test

182 was used for post-hoc pairwise comparisons between conditions. All statistical analyses were performed using
183 the statsmodels package in Python.

184 **Pseudobulk Analysis**

185 Pseudobulk analysis was performed to more robustly identify DEGs between condition groups as it has been
186 shown to significantly reduce the number of false positives and increase the accuracy of subsequent enrichment
187 analyses as compared to the Wilcoxon rank-sum test (27). The Python package ‘Decoupler’ was used to generate
188 pseudobulk profiles for each individual mouse using the ‘sum’ method. The pseudobulk profiles were then used to
189 identify DEGs between groups using ‘PyDeseq2’, a python implementation of the DESeq2 algorithm (28,29). The
190 Wald test was used to generate p-values with cook’s filtering and correction for multiple testing using Benjamini-
191 Hochberg method. The model was fit with disease status (FAD vs WT), APOE genotype (E3 vs E4), and treatment
192 (LPS vs Vehicle) as factors with their interactions, and specific comparisons were extracted using the ‘results’
193 function. Genes were considered differentially expressed if they had an adjusted p-value < 0.05 and a log2 fold
194 change > |0.25|. The pseudobulk DEGs were then used for downstream analysis including Gene Ontology (GO)
195 enrichment analysis.

196 **Gene Set Enrichment Analysis (GSEA)**

197 Gene Set Enrichment Analysis (GSEA) is a statistical method used to determine whether selected sets of genes
198 show statistically significant differences between two biological states (30). Rather than focusing only on statis-
199 tically significant DEGs, as other methods do, GSEA evaluates the entire set of genes ranked by a metric which
200 reflects their different expression between conditions, thereby better capturing the pathway-level enrichment
201 differences.

202 For each contrast of interest, genes were ranked by their Wald test statistic derived from the pseudobulk
203 differential expression analysis. GSEA was then performed using the implementation provided in the Decoupler
204 package (28), with the MSigDB Hallmark gene sets (31) as the reference collection. This approach enabled the
205 identification of pathways significantly enriched in up- or down-regulated genes across our conditional contrasts
206 of interest.

207 **Proteomic Analysis**

208 *Sample Preparation*

209 50 µg of protein was digested on S-Trap micro spin columns (ProtiFi) with sequencing-grade trypsin (1:25)
210 overnight at 37 °C following the manufacturer's protocol.

211 *LC–MS/MS Acquisition*

212 Peptides (500 ng) were analyzed on an Evosep One–timsTOF HT system in diaPASEF mode (16 windows, 400–
213 1 200 m/z, 0.60–1.60 1/K₀; 1.8 s cycle), using 0.1 % formic acid mobile phases.

214 *Data Quantification*

215 Raw files were processed in Spectronaut v19 (directDIA) using default MS2 quantification and MaxLFQ. The
216 UniProt mouse reference proteome (17 184 entries), supplemented with human APOE3, APOE4, and 5×FAD APP/
217 PSEN1 sequences, served as the search database. Carbamidomethylation of cysteine and oxidation of methionine
218 were set as fixed and variable modifications, respectively; up to two missed cleavages were allowed. False
219 discovery rate was controlled at 1 % for both peptide and protein levels.

220 (Full protocol provided in Supplementary Methods)

221 *Data Analysis*

222 Log₂-transformed normalized label-free quantification (LFQ) values from Spectronaut v19 were imported into
223 R (v4.x), reshaped and annotated with sample metadata using tidyverse and dplyr [(32)]. Metadata fields included
224 sample ID, genotype, treatment, age, sex and batch. A SummarizedExperiment object was generated with DEP
225 (v2.0.0) [(33)] for downstream analyses.

226 Comprehensive quality control was conducted within DEP. Density and box plots of log₂ intensities were
227 inspected to confirm consistent normalization across samples. Protein detection frequency was assessed by
228 counting proteins detected in n samples, and the distribution was visualized to evaluate assay coverage. Missing
229 data patterns were explored by generating heatmaps of missingness and plotting the fraction of missing values
230 against intensity. On the basis of these diagnostics, missing values were classified as Missing Not At Random
231 (MNAR) and found to be enriched at low intensities. Imputation was performed using the MinProb algorithm (q

232 = 0.5) as implemented in DEP. Additional exploratory analyses included Principal Component Analysis (PCA),
233 hierarchical clustering, heatmaps of Pearson correlation coefficients between samples.

234 Differential abundance analysis was conducted with limma [(34)]. Linear models were fitted to \log_2 inten-
235 sities with genotype, treatment and their interaction specified as factors. An empirical Bayes procedure with
236 mean-variance trend fitting and robust hyperparameter shrinkage was applied. Contrasts were defined to test
237 the effect of LPS within APOE3, the effect of LPS within APOE4, and the interaction between genotype and
238 treatment. Proteins with nominal $P < 0.05$ were selected for visualization. Volcano plots were generated using
239 **ggplot2**, displaying $-\log_{10}(P)$ versus \log_2 fold change with top hits annotated. Concordance with transcriptomic
240 data was assessed by plotting proteomic \log_2 fold changes against matching RNA-seq contrasts and calculating
241 Spearman correlation coefficients.

242 Pathway enrichment analysis was performed using Camera [(35)] in limma against KEGG pathways from
243 the Molecular Signatures Database (MSigDB) C2:CP (curated canonical pathways) [(31)]. Camera's competitive
244 gene set test, which estimates inter-gene correlation from the dataset to control type I error, was executed with
245 default settings using the full vector of \log_2 fold changes as input. Pathways were ranked by false discovery rate
246 for interpretation.

247 **Results**

248 *Cluster Expression Profiles*

249 Following dimensionality reduction and Leiden clustering, we identified nine distinct microglial expression
250 profiles based on differentially expressed genes (DEGs) between clusters and reference to canonical expression
251 signatures described in previous studies. Cluster annotation was guided by comparisons to known microglial
252 states in both mouse and human brain studies, including those by Keren-Shaul et al. (3), Sala-Frigerio et al.
253 (22), Olah et al. (19), Mathys et al. (23), and others. The final annotated expression profiles were: Homeostatic,
254 DAM, MHC-II/Antigen Presentation DAM, Cytokine Response, Cycling (G2M/Cdk1+), BAM-like, Neuronal
255 Surveillance/Ribosomal Biogenesis, Ribosomal DAM, and Immediate Early Gene (IEG) Enriched Homeostatic.

These expression profiles are visualized in Figure 1. The UMAP plots (Figure 1a) show cluster organization across the entire dataset and by condition. Figure 1b displays the top 10 DEGs for each cluster when compared to the rest, and Figure 1c shows density plots of cluster distributions per condition.

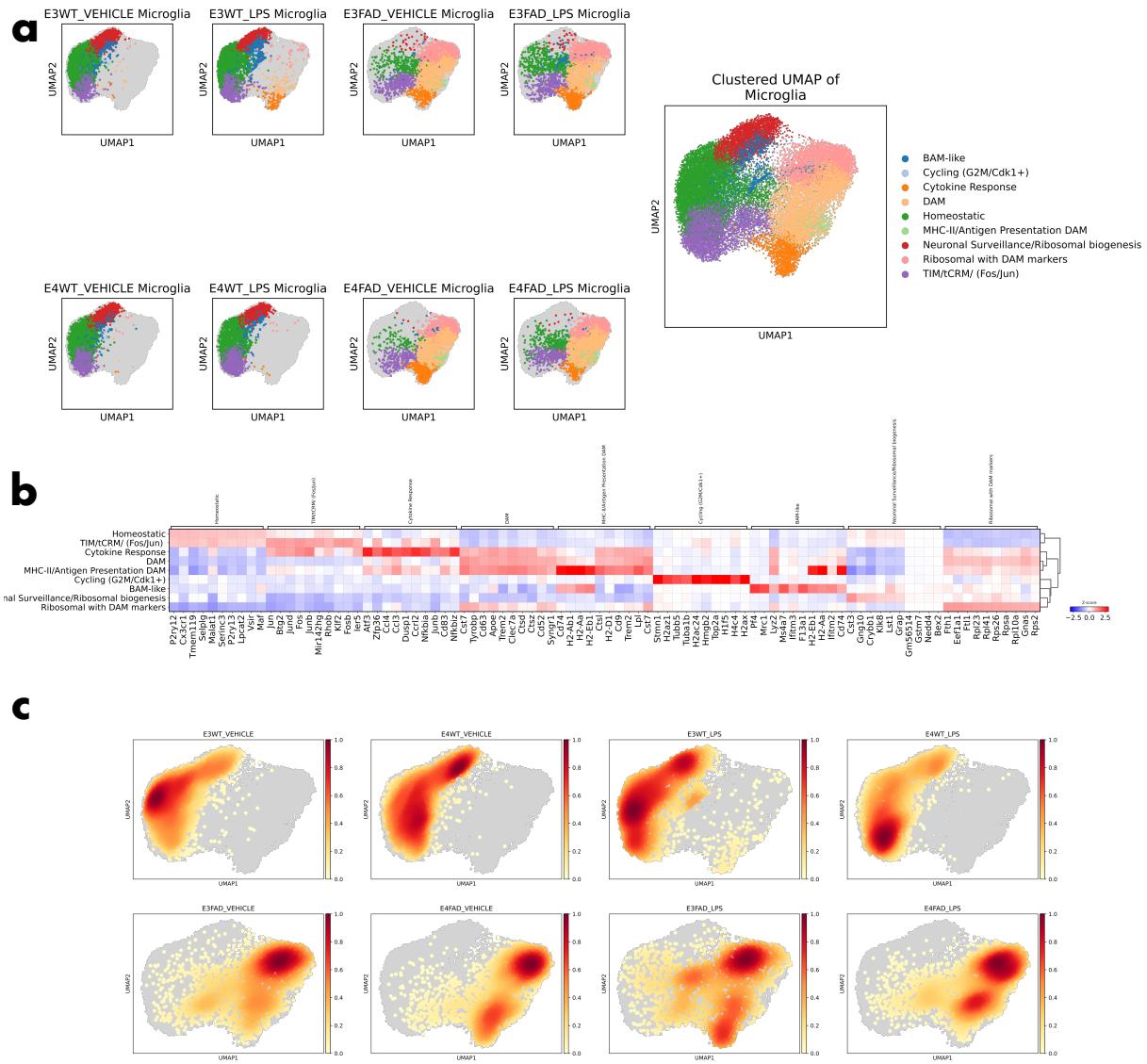


Figure 1: Distinct Expression Clusters and Their Proportions by Condition

a. UMAP visualizations of the dataset colored by manually annotated expression profiles, entire dataset, and by condition. **b.** Heatmap of the top 10 differentially expressed genes for each cluster expression profile showing scaled log-normalized expression values (z-scores) **c.** UMAP density plots for each condition.

259 The Homeostatic cluster was characterized by high expression of canonical microglial maintenance genes,
260 including P2ry12, Tmem119, and Cx3cr1, and served as a reference for identifying more activated or altered
261 states. The IEG cluster shared elevated expression of homeostatic markers but also showed upregulation of

262 immediate early genes such as Fos, Jun, and Egr1, a pattern that has been interpreted as both a stimulus-
263 responsive microglial state and, alternatively, as a dissociation-induced artifact (24,36).

264 Two small clusters were also identified. The Cycling cluster was defined by elevated expression of cell
265 cycle and proliferation-associated genes such as Stmn1, Top2a, and H2az1, consistent with previously reported
266 populations of proliferative microglia in AD models and aged brain tissue (22,26,37). The BAM-like (Border-
267 Associated Macrophage-like) cluster was characterized by high expression of perivascular macrophage markers
268 including Pf4, Mrc1, and Ms4a7. Although microglial enrichment was performed during FACS, this population
269 likely reflects a small number of CNS-associated macrophages retained during sorting.

270 The DAM (Disease-Associated Microglia) cluster showed a transcriptional profile closely matching that
271 originally described by Keren-Shaul et al. (3), including high expression of Trem2, Apoe, Cst7, Tyrobp, and Clec7a,
272 accompanied by suppression of homeostatic genes. Three additional clusters were identified with overlapping
273 DAM features but distinct gene expression signatures. The Cytokine Response cluster exhibited elevated expres-
274 sion of inflammatory chemokines and NF- κ B target genes such as Ccl3, Ccl4, and Il1b. The MHC-II/Antigen
275 Presentation DAM cluster was enriched for antigen presentation genes including Cd74, H2-Ab1, H2-Aa, and H2-
276 Eb1. The Ribosomal DAM cluster retained DAM-like features with additional upregulation of ribosomal genes.

277 ***Expression Profile Proportions Across Conditions***

278 After identifying different expression profiles in the dataset, the relative abundance of each cell type was
279 assessed across experimental conditions. The proportions observed in the eight experimental groups are shown
280 in Figure 2a, with absolute cell counts presented in Supplementary Figure 2 and corresponding values listed in
281 Supplementary Table 2. Disease versus wild-type is the main cause of variation in cell type composition, with
282 the most pronounced differences observed in the homeostatic and DAM clusters. Homeostatic cells accounted
283 for 39–56% of the population in WT mice, compared to only 3–17% in FAD mice, representing a statistically
284 significant reduction (ANOVA p < 1e-8; linear model coefficient = +46.2%, p < 0.001). The APOE genotype was
285 not associated with a statistically significant change in homeostatic cell proportions, nor was there a significant
286 difference with LPS treatment.

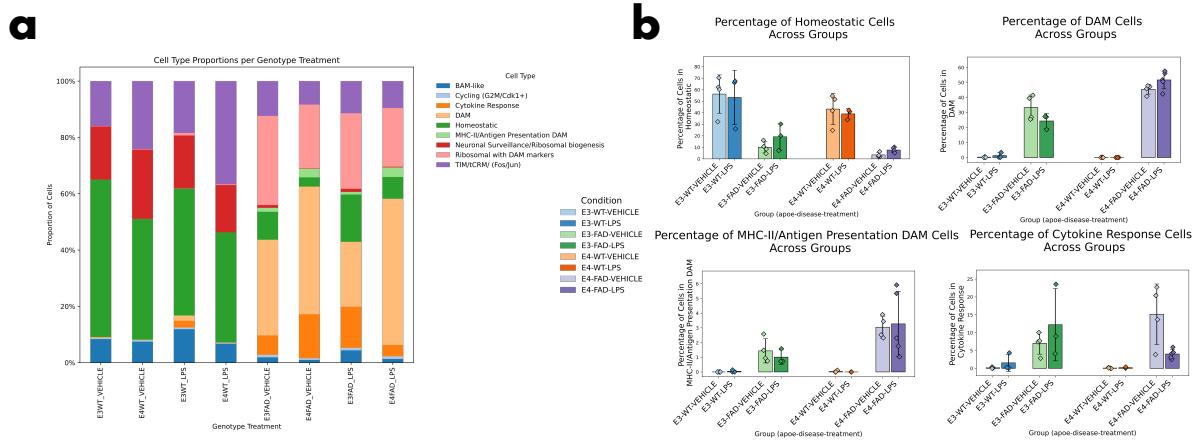


Figure 2: Changes in Cluster Proportions by Condition

a. Stacked Bar plot of the proportion of each cluster in each condition. b. Bar plots of the proportion of specific clusters in each condition, significant differences determine by three way anova and Tukey's post-hoc test and indicated by asterisks.

287 The DAM cluster was absent in all WT groups except for a small proportion (1.9%) in the E3WT LPS group
 288 and ranged from 23-52% in FAD mice. Additionally, E4FAD mice have a larger proportion of DAM microglia
 289 than E3FAD mice (interaction coefficient = +12.3%, p = 0.012), which is consistent with previous single cell
 290 studies. The three clusters representing more “activated” states similar to DAM (Cytokine Response, MHC-II/
 291 Antigen Presentation DAM, and Ribosomal DAM), also show zero or very low proportions in WT mice, and are
 292 significantly increased in FAD mice.

293 The Cytokine Response cluster shows an interesting pattern in which the E4FAD LPS group has a significantly
 294 lower proportion of cells compared to the E4FAD vehicle group (Tukey p-adj = 0.028), while the E3FAD
 295 LPS group is not significantly different from the E3FAD vehicle group (Tukey p-adj = 0.807). This suggests an
 296 APOE isoform-specific effect of LPS priming on this microglial population. This interpretation is supported by
 297 a statistically significant three-way interaction term in the ANOVA (p = 0.038) and a notable effect size in the
 298 linear model (coefficient = -15.1%, p = 0.038), indicating that LPS treatment reduces the proportion of Cytokine
 299 Response cells specifically in E4FAD mice.

300 ***E4FAD but Not E3FAD or WT Mice Exhibit Reduced NfkB-related Expression after LPS Priming***

302 To more robustly identify differentially expressed genes (DEGs) between conditions, pseudobulk analysis was
303 performed on the dataset. This involves summing the expression each gene in all cells from an individual mouse
304 to create a single expression profile, aiding in overcoming the sampling variability in single cell RNA sequencing.
305 The pseudobulk profiles can then be used to identify DEGs with methods typically used for bulk RNA sequencing
306 data, such as DESeq2 which produces fewer false positives than the more common Wilcoxon rank-sum test
307 used in single cell analysis. DEGs were identified between the LPS and vehicle treatment groups in the four
308 genetic backgrounds, the results of which are visible in the volcano plots in Figure 3a. There are no more than 1
309 significant DEG which is shared between any of the comparisons. However, as a general pattern it can be seen
310 that the FAD mice have a much larger number of downregulated genes in the LPS treatment group compared to
311 the vehicle group, while the WT mice show a much smaller number of downregulated genes.

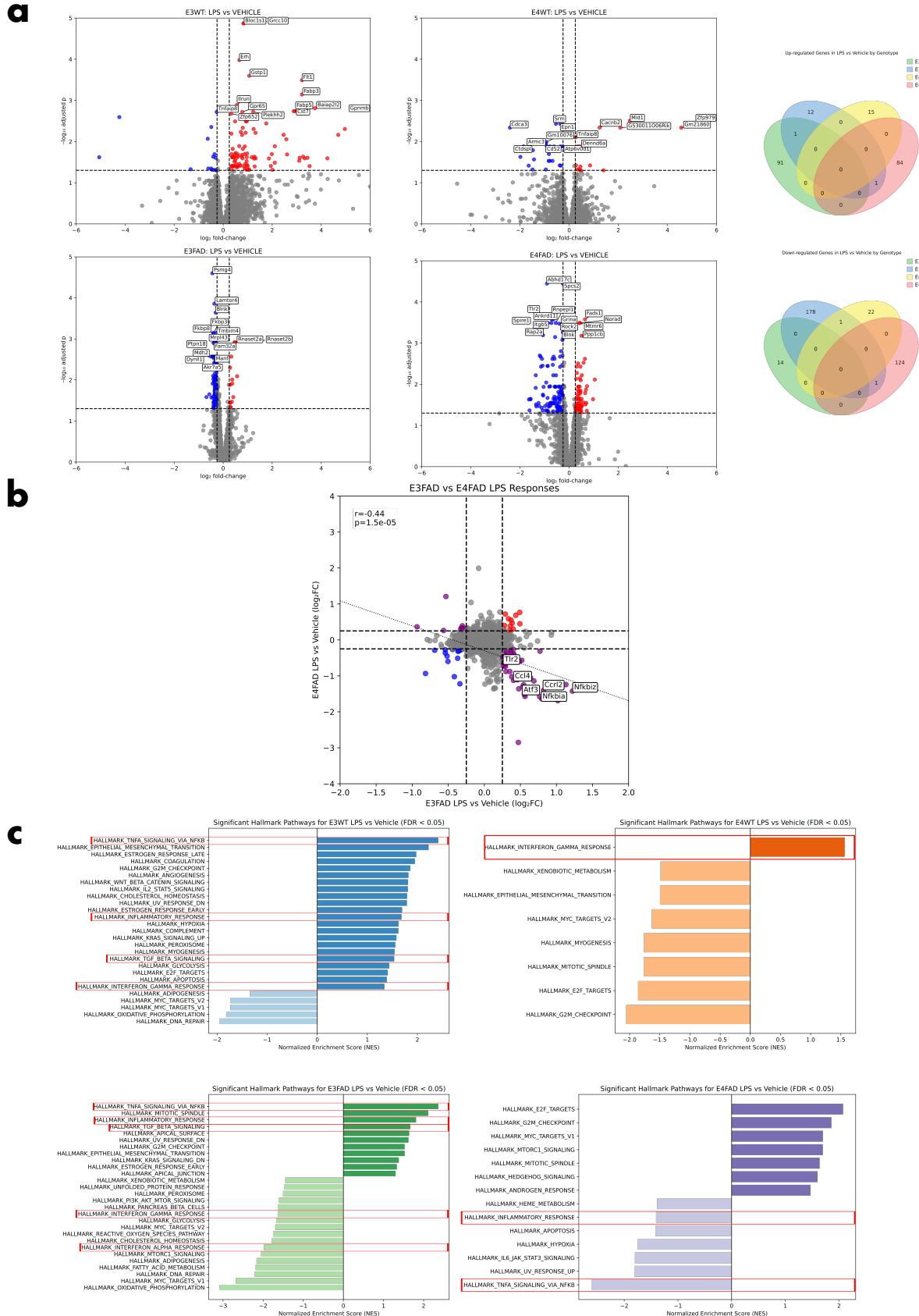


Figure 3: Differentially expressed Genes and Pathways in LPS versus Vehicle Treated Mice

a. Volcano plots of pseudobulk DEGs in LPS versus vehicle comparisons in the four genotypes with top genes labelled. b. Correlation plot of LogFC values of genes in LPS versus vehicle treatment in E3FAD compared to E4FAD. Pearson's correlation computed on genes above the logFC threshold of 0.25 c. GSEA of significant pathways in the LPS versus vehicle comparisons computed on Wald test statistics.

312 When the E3FAD and E4FAD LPS treatment group DEGs are compared, a statistically significant correlation is
313 observed (Pearson's $r = -0.44$, $p = 1.5\text{e-}5$), indicating that genes that are upregulated in E3FAD mice are likely
314 to be downregulated in E4FAD mice. A number of these genes are those associated with the NfkB pathway,
315 including Ccl4, Tlr2, and Nfkbia. Based on the pseudobulk DEGs, Gene Set Enrichment Analysis (GSEA) was
316 performed to identify pathways that were significantly enriched in the LPS treatment groups. The results (Fig 3c)
317 show that the top upregulated pathways in both E3WT and E3FAD groups are TNF-a signaling via NfkB, while
318 it is the top downregulated pathway in the E4FAD group. It does not appear in the significant pathways in the
319 E4WT group, however the top upregulated pathway in this group is Interferon gamma response. Additionally,
320 the E3WT and E3FAD groups show significant upregulation of the Inflammatory response pathway and TGF
321 beta signaling, while the E4FAD group shows downregulation of the Inflammatory response pathway and no
322 upregulation of inflammatory pathways.

323 *Cytokine Response Microglia*

324 Given that the dampening of NF- κ B-related pathways in the E4FAD LPS group coincided with a reduced
325 proportion of Cytokine Response microglia, the transcriptional profile of this cluster was examined to assess its
326 potential contribution to the observed pseudobulk differences.

327 The Cytokine Response cluster was marked by elevated expression of genes associated with pro-inflam-
328 matory signaling and canonical NF- κ B pathway activation. These included the chemokines Ccl3 and Ccl4, the
329 interleukins Il1a and Il1b, and the TNF family member Tnf. Several key regulators and targets of NF- κ B signaling,
330 including Nfkbia, Nfkbiz, and Tnfaip3, were also upregulated. Additional genes enriched in this cluster were
331 involved in stress response (Atf3, Gadd45b), immune regulation (Cd83, Zfp36, Bcl2a1b), pattern recognition
332 (Tlr2), cytokine signaling (Csf1), and chemokine receptor activity (Ccrl2).

333 These expression patterns are shown in Figure 4. UMAP projections reveal that these transcripts are highly
334 expressed within the Cytokine Response cluster, with limited expression in other microglial states. Violin plots
335 further illustrate the cluster-specific enrichment of these genes.

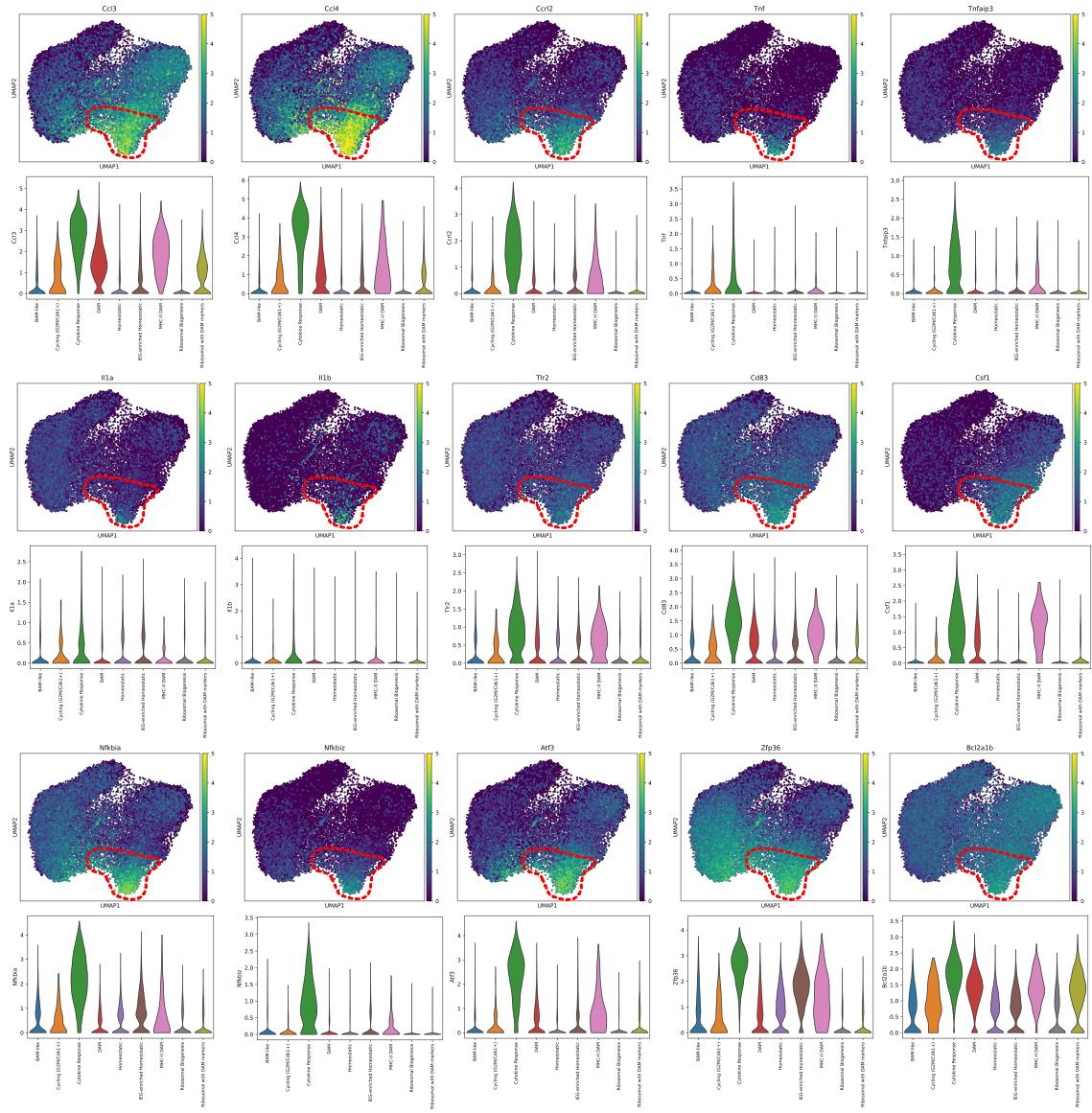


Figure 4: Genes in the Cytokine Response Cluster

UMAP plots of the dataset colored by the expression of specific genes which are differentially upregulated in the Cytokine Response cluster and which are part of the NF- κ B pathway paired with the violin plots in all clusters.

336 The selective expression of inflammatory genes within this cluster, combined with its reduced abundance in
 337 E4FAD mice following LPS treatment, suggests that changes in this microglial state contribute to the largest
 338 observed pathway-level differences in the more conservative pseudobulk analysis.

339 In addition to the upregulation of pro-inflammatory signaling genes, the Cytokine Response cluster was
 340 found to show elevated expression of several immediate early genes (IEGs), including Fos, Jun, Egr1, and
 341 Dusp1. This overlap prompted further comparison with the IEG-enriched cluster, which also showed increased

342 expression of these same IEGs relative to the rest of the dataset. Despite this similarity, clear differences were
343 observed between the two clusters. The IEG-enriched cluster retained high expression of canonical homeostatic
344 markers such as Tmem119, Cx3cr1, and P2ry12, which were largely absent in the Cytokine Response cluster. In
345 contrast, the Cytokine Response cluster was enriched for DAM-associated genes including Apoe, Clec7a, Trem2,
346 and Ctsd, both in comparison to the total population and when directly compared to the IEG-enriched cluster.

347 When compared to the DAM cluster, the Cytokine Response cluster remained distinct. The DAM cluster did
348 not show enrichment for IEGs or NF- κ B pathway-associated transcripts, suggesting that the Cytokine Response
349 cluster represents a distinct activation state with overlapping features of both DAM and IEG-enriched microglia.

350 These relationships are visualized in the dot plot in figure 5.

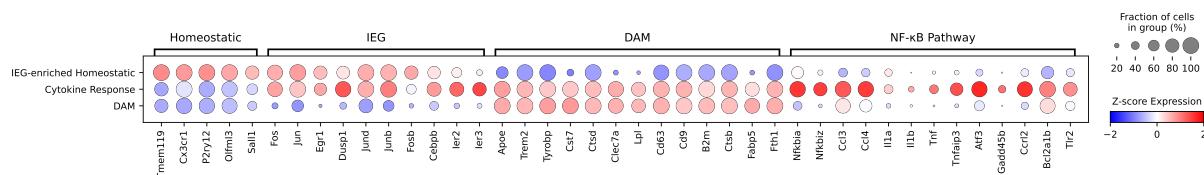


Figure 5: Comparison of Overlapping and Distinct Features Between Microglial States

Dot plot showing the expression of immediate early genes (IEGs), homeostatic markers, DAM-associated genes, and NF- κ B pathway-related transcripts across the Cytokine Response, IEG-enriched, and DAM clusters. Dot size represents the fraction of cells expressing each gene, calculated using a threshold of log-normalized expression > 0 (equivalent to raw counts ≥ 1). Color indicates average expression level scaled across clusters (Z-score normalized).

351 *Proteomic Analysis*

352 The relationship between mRNA abundance and protein levels in vivo is complex and influenced by translational
353 regulation, post-translational modifications, and protein degradation. Therefore, transcript levels often do not
354 directly predict protein abundance, with reported correlations varying widely across tissues and cell types and
355 frequently being low (Liu et al. 2016). To corroborate and extend our transcriptomic findings, mass spectrometry-
356 based proteomic analysis was performed on the same individuals used for single-cell sequencing.

357 A total of 8 065 proteins were quantified across all samples. UniProt mouse identifiers were mapped to
358 gene symbols (GRCm39) using the R package org.Mm.eg.db, allowing direct comparison with the transcriptomic
359 dataset. The proteomic cohort comprised the female whole-cortex samples from LPS and vehicle groups of E3FAD
360 and E4FAD mice ($n = 5$ per condition), plus additional samples not included in the transcriptomics (yielding

361 up to five replicates per group). Unlike the microglia-enriched preparations used for transcriptomics, proteomic
362 profiling was performed on whole cortex. Although wild-type and male cohorts were also analyzed, only E3FAD
363 and E4FAD mice are included here to align with the transcriptomic analysis.

364 Principal component analysis (PCA) of the log₂-transformed LFQ intensities revealed clear separation by
365 APOE genotype but no discernible clustering by treatment (Fig 6a), indicating that LPS priming did not induce
366 a major shift in the cortical proteome. Consistent with this, differential abundance analysis using limma on
367 imputed LFQ values identified no proteins with significant changes (FDR < 0.05) between LPS and vehicle in
368 either genotype.

369 To assess concordance between transcriptomic and proteomic responses to LPS, we compared the log₂ fold
370 changes of differentially expressed genes (DEGs) from the RNA-seq analysis to the corresponding values for their
371 protein products (n = 7 051 matched identifiers). Spearman correlation coefficients were $\rho = 0.032$ ($p = 0.007$) for
372 E3FAD and $\rho = 0.013$ ($p = 0.028$) for E4FAD (Fig 6c), indicating extremely weak correlation between mRNA and
373 protein fold-change directions.

374 Despite the lack of individually significant proteins and the low global correlation, differentially abundant
375 proteins can be visualized in a clustered heatmap (Fig 6b). To determine whether pathway-level signals might
376 emerge, we applied Camera to test KEGG Medicus Reference pathways for enrichment against the full set of log₂
377 fold changes. This approach can detect changes in pathway enrichment even when individual proteins fail to
378 reach statistical significance by leveraging coordinated shifts across gene sets.

379 After multiple-testing correction, the “TNF NFKB signaling pathway” was the only KEGG pathway signif-
380 icantly enriched (FDR < 0.05) in the upregulated direction for E3FAD mice treated with LPS (Fig 6d). This
381 finding aligns with our transcriptomic GSEA results, which also highlighted TNF-NF-κB activation in E3FAD LPS
382 samples. In the E4FAD LPS group, the pathway showed a trend toward downregulation but did not reach FDR
383 < 0.05. The interaction contrast, however, met FDR < 0.25 for downregulation, pointing to genotype-dependent
384 changes in NF-κB signaling after LPS treatment.

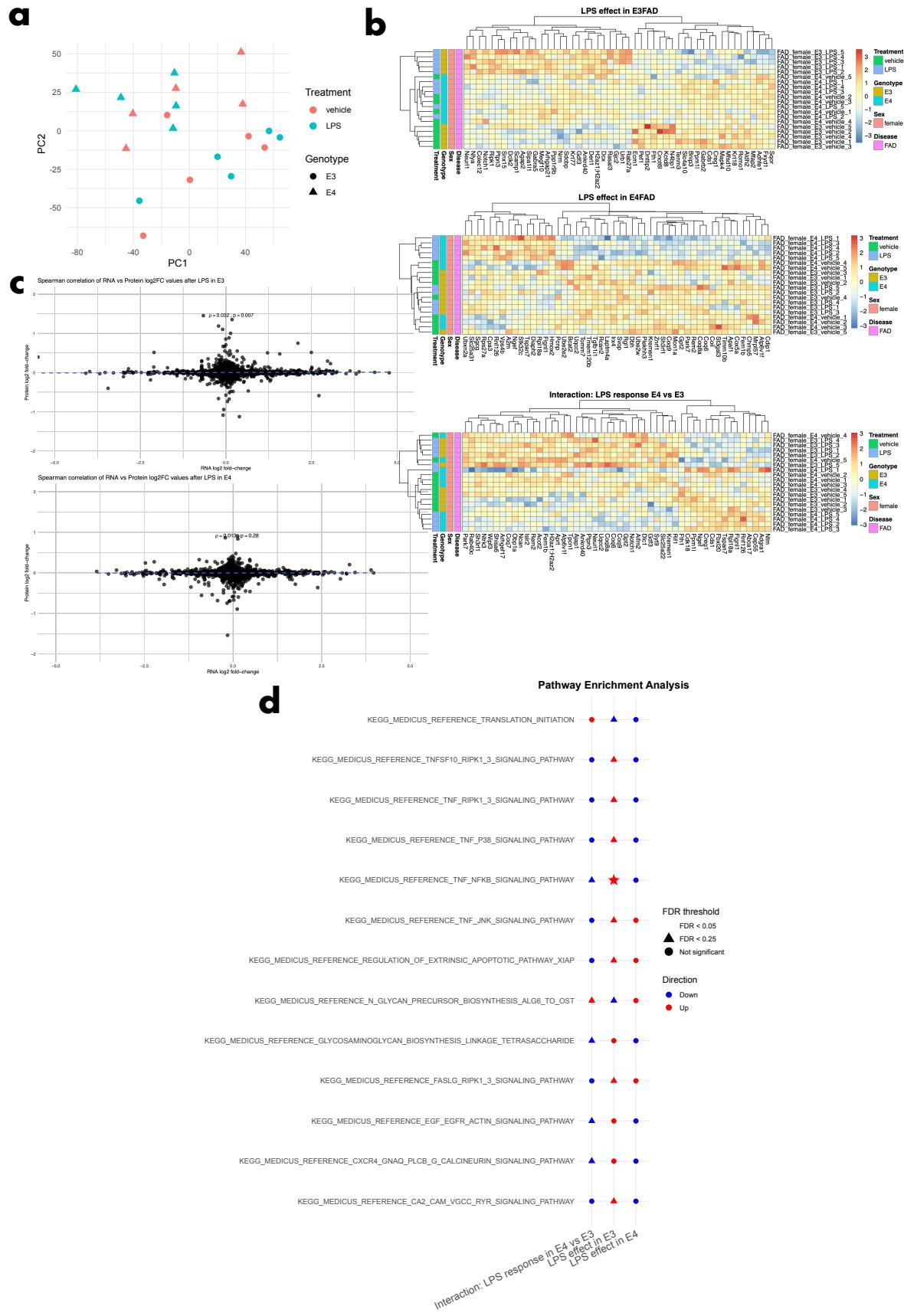


Figure 6: Proteomic Analysis of LPS effect in E3/E4FAD

a. PCA plot of proteomic data for E3FAD and E4FAD female mice in LPS and Vehicle treatment groups. **b.** Heatmap of Z-score scaled logFC of proteins with smallest nominal P values for the LPS treatment effect in both genotypes, and the interaction term which indicates those which

have the biggest different treatment effect between the two genotypes. Rows correspond to individual samples. both Rows and Columns are ordered by heirarchical clustering. **c.** Scatter plots comparing Logfc values after LPS treatment in both genotypes in the proteomic data against the transcriptomic data. Only proteins/genes which could be matched between both were kept (n = 7051). Spearman correlation calculated (E3FAD: $\rho=0.032$, p = 0.007; E4FAD: $\rho=0.013$, p=0.028) **d.** Dot plot of all pathways which have a FDR < 0.25 in atleast one contrast. Dot coloured by direction of enrichment and dot shape corresponds to FDR values.

385 Discussion

386 *Microglial Heterogeneity in Alzheimer's Disease*

387 Neuroinflammation has become recognized as a central phenomenon in the pathogenesis of Alzheimer's disease
388 (AD), both regulating and responding to amyloid- β production, plaque formation, tau hyperphosphorylation, and
389 synaptic and neuronal loss, within a complex and dynamic system that remains incompletely understood (38).

390 While all central nervous system (CNS) cell types have been implicated, accumulating evidence has positioned
391 microglia—the resident innate immune cells of the brain—as central players in these processes. As research into
392 microglial involvement in AD has expanded, so too has the understanding of their diverse functional states.
393 However, the molecular mechanisms governing transitions between these states remain poorly defined.

394 Advances in laboratory techniques and in particular single-cell transcriptomics have given us an entirely
395 new view of microglial heterogeneity. The outdated binary classification of microglia into “resting” or “activated”
396 states has been replaced by an understanding of multiple distinct transcriptional states influenced by disease
397 progression, brain region, age, microenvironment, and genetic background. This shift has been driven by
398 studies describing specific microglial activation profiles, among which the Disease-Associated Microglia (DAM)
399 program, described by Keren-Shaul et al. (3), remains the most influential. This transcriptional program was
400 proposed to occur in two stages, dependent on TREM2 signaling, and in response to amyloid deposition.

401 Subsequent studies have identified a broader array of microglial states, including cycling, antigen-presenting,
402 phagocytic, and various pro-inflammatory states. The current challenge in this area of research is integrating
403 a rapidly expanding body of evidence produced with varied model systems, methodologies, nomenclatures and
404 sometimes conflicting interpretations.

405 APOE genotype remains the strongest genetic risk factor for late-onset Alzheimer's disease, and in the
406 context of microglial heterogeneity, its importance has become increasingly apparent. APOE is highly expressed

407 by microglia in response to amyloid pathology and has been shown to modulate a number of microglial func-
408 tions including phagocytosis, lipid metabolism, and inflammatory signaling. The data presented here provide
409 additional evidence for APOE-dependent modulation of microglial state. As expected, homeostatic microglia
410 represented the majority population in wild-type mice and were nearly absent in the presence of amyloid
411 pathology. FAD mice showed a marked shift toward DAM and related activated states, with a significantly higher
412 proportion of DAM microglia observed in E4FAD compared to E3FAD mice. This finding is consistent with
413 previous studies using humanized APOE models.

414 *Early-Life Immune Priming Alters Microglia Trajectory*

415 LPS priming of microglia has been shown to influence their function both acutely and over extended periods
416 following the initial insult (39). Unpublished data from members of our lab have demonstrated that LPS adminis-
417 tration at postnatal day 9 alters the microglial response to amyloid pathology later in life. Changes were observed
418 in the transcriptional levels of specific cytokines, and a microglial population not present in vehicle-treated mice
419 was identified ((15), preprint). These findings, combined with the well-established role of APOE in modulating
420 microglial state and inflammatory signaling, motivated the present analysis to investigate how human APOE
421 isoforms may interact with early-life immune priming to shape microglial phenotypes in the context of AD-
422 related pathology.

423 The data presented here provide additional evidence for the impact of APOE genotype on microglial pheno-
424 types and the modulatory effect of LPS priming. As expected, a majority of microglia in non-disease (WT) mice
425 were assigned to the homeostatic cluster, with near-zero representation in the disease-associated transcriptomic
426 profiles. In contrast, FAD mice exhibited a substantial shift, with large proportions of their microglial populations
427 classified into disease-associated clusters and a corresponding reduction in homeostatic cells. An increased
428 proportion of microglia in the DAM cluster was observed in E4FAD mice compared to their E3FAD counterparts,
429 consistent with previous studies using mice expressing human APOE alleles.

430 *Genotype-Specific Modulation of Inflammatory Signal*

431 The analysis shows a genotype-specific priming effect on the expression of particular cytokines and inflamma-
432 tory genes, primarily those in the canonical NF-κB signaling pathway. In mice treated with LPS, a downregulation

433 of these genes was observed specifically in E4FAD animals, while an upregulation was seen in the LPS-treated
434 E3FAD group (Fig. 3). The genes associated with these pathways were not expressed uniformly across the
435 microglial population, indicating that the observed transcriptional changes were not global but instead reflect the
436 behavior of a specific microglial subset. This subset, identified as the Cytokine Response cluster, was not present
437 in non-disease mice, indicating that its emergence is dependent on amyloid pathology. Cluster-level analysis
438 suggests that the observed pathway-level effect in E4FAD mice treated with LPS is driven by a reduction in the
439 proportion of microglia occupying this state.

440 To determine whether these APOE- and treatment-specific transcriptional signatures translate to protein
441 abundance, we performed LC–MS/MS–based proteomics on the matching cortex samples. PCA of LFQ intensities
442 separated samples by APOE genotype but not by LPS treatment, and limma analysis revealed no individual
443 proteins with FDR < 0.05. Nonetheless, pathway enrichment on the fold-change values identified the TNF-NF-
444 kB signalling cascade as significantly upregulated in E3FAD + LPS (FDR < 0.05) and trending down in E4FAD +
445 LPS (interaction FDR < 0.25), mirroring the observed transcriptomic pattern between genotypes.

446 *Contextualizing the Cytokine Response Cluster*

447 A Cytokine Response cluster, defined by increased expression of cytokines, chemokines, and NF- κ B-related
448 genes, has been observed in other studies. Lee et al. (10) used E3FAD and E4FAD mice challenged with LPS and
449 performed single-cell sequencing 24 hours later; a cluster enriched for NF- κ B–related transcripts was identified,
450 with higher abundance in E3 compared to E4. Mancuso et al. (21) identified a transcriptionally similar population
451 in their xenograft model, significantly enriched in the presence of amyloid pathology, although no APOE
452 genotype–specific differences were reported. Hammond et al. (25) analyzed healthy mice across a range of ages
453 and identified microglial clusters with similar chemokine and cytokine expression profiles, which increased in
454 abundance with age. This microglial phenotype has now been identified across multiple models and conditions;
455 however, its functional role in Alzheimer’s disease remains unclear.

456 Beyond the identification of this cluster, cytokine signaling more broadly remains poorly characterized
457 in the context of AD. Cytokines have been proposed to contribute to increased amyloid- β production and
458 accumulation (40), and microglial phagocytosis of amyloid- β has been shown to induce cytokine release (41).

459 Other studies have reported that interleukin-1 β reduces microglial clearance of amyloid- β (42) and that cytokine
460 production is associated with neuronal loss. However, there is also evidence that LPS-induced inflammation can
461 promote amyloid- β clearance (43). These findings highlight the complexity of cytokine signaling in AD, which
462 is governed by context-specific feedback loops that depend critically on the specific cytokines, chemokines, and
463 signaling pathways involved, as well as their timing and intensity

464 Continued investigation of these microglial populations will be important for understanding their function,
465 how they develop and progress, and how they transition across brain regions and in response to disease states.
466 The analysis presented here suggests that this population can be modulated by pre-symptomatic intervention,
467 which may represent a viable therapeutic strategy. This transcriptomic analysis forms part of a larger study
468 examining the interaction between early-life LPS priming and APOE genotype. Unpublished data from this study
469 indicate improved short-term memory performance in LPS-treated E4FAD mice, an effect not observed in their E3
470 counterparts. Taken together, these findings suggest that modulating microglial phenotype at an early timepoint
471 may yield beneficial effects on Alzheimer's disease-related outcomes.

472 *Biological vs. Artifactual: Evaluating Immediate Early Gene Expression*

473 One of the subclusters identified in our dataset shows upregulation of immediate early genes (IEGs), most
474 prominently from the Fos and Jun families, while also retaining expression of canonical homeostatic microglial
475 markers. This pattern, referred to here as IEG-enriched, has been previously reported by Marsh et al. (36) who
476 showed that it can be induced by enzymatic dissociation during single-cell processing. They found that the use
477 of mechanical dissociation or the addition of transcriptional inhibitors, such as actinomycin D, during enzymatic
478 dissociation could prevent the emergence of this expression pattern. Based on these findings, they concluded that
479 this IEG-rich profile represents an ex vivo stress response, and advocated for the use of transcriptional inhibitors
480 as a standard practice for microglial single-cell experiments.

481 However, more recent studies suggest that this signature may not be purely artifactual. Mancuso et al.
482 (21) generated a single-cell microglial transcriptomic dataset using the App⁺NL-G-F amyloid mouse model with
483 transplanted human microglia expressing APOE2, APOE3, or APOE4, and included transcriptional inhibitors
484 specifically to address the ex vivo activation signature reported by Marsh. Despite this, they still identified a

485 cluster they termed “Transitioning Cytokine Response Microglia” (tCRM), which they describe as “...show high
486 levels of homeostatic genes but also express CRM markers.” The top 15 DEGs for this cluster (FOS, JUN, DUSP1,
487 KLF2, HSPA1A1A, IER2, IER3, RHOB, JUNB, Ch25H, HSPA1AB, JUND, FOSB, CEBPD, RGS1) are primarily
488 immediate early genes and closely resemble the ex vivo stress signature identified by Marsh. While Mancuso et
489 al. did not report any APOE isoform-specific differences in this population, they found that it was significantly
490 reduced in APOE knockout mice, suggesting that despite the use of inhibitors, this expression profile may reflect
491 a real APOE-dependent microglial state.

492 Further evidence for the biological relevance of this type of expression pattern comes from Millet et al.
493 (24), who performed scRNA-seq on microglia from E3 and E4 × 5xFAD mice aged to 96 weeks. They identified
494 a population they termed “Terminally Inflammatory Microglia” (TIM), which was enriched in aged animals and
495 which they describe as “marked by concomitant expression of inflammatory genes such as S100a8 and S100a9
496 and immediate early response genes such as Fos, Jun, and Egr1.” Although their protocol did not include
497 transcriptional inhibitors and they acknowledge resemblance to the ex vivo stress profile reported by Marsh, they
498 argue that TIMs represent a bona fide microglial state. They support this by showing that similar populations
499 are seen in human snRNAseq datasets prepared using both mechanical and enzymatic dissociation, and that the
500 TIM cluster is more prevalent in aged brains and modulated by APOE genotype.

501 A more recent study by Mulenge et al. (44) extended the work of Marsh by evaluating the effects of both
502 dissociation and sorting. They confirmed that transcriptional inhibitors can suppress IEG expression but also
503 showed that FACS sorting itself induces an ex vivo activation signature. Specifically, they found that Zfp36,
504 Dusp1, Jun, Ccl4, Ccl3, Tnf, Socs3, and Hspa1a were consistently upregulated across sorted datasets. However,
505 they noted that this expression was generally elevated across the entire dataset and not confined to specific
506 clusters.

507 In our data, IEG and cytokine-related transcripts were not uniformly elevated across all microglia but
508 confined to the Cytokine Response and IEG-enriched clusters. This spatial restriction suggests a bona fide
509 subpopulation rather than a global dissociation artifact. Importantly, bulk proteomic profiling of the same

510 cortex samples independently highlighted the TNF–NF-κB signalling cascade—central to cytokine responses—as
511 significantly enriched in E3FAD LPS–treated mice (Fig. 6d). The concordant pathway-level signal at both RNA
512 and protein layers, together with parallels in Mancuso and Millet’s studies, reinforces the conclusion that these
513 clusters reflect true biological microglial states.

514 *Conclusions and Future Directions*

515 Continued investigation of these microglial populations will be important for understanding their function,
516 how they develop and progress, and how they transition across brain regions and in response to disease states.
517 The analysis presented here suggests that this population can be modulated by pre-symptomatic intervention,
518 which may represent a viable therapeutic strategy. This transcriptomic analysis forms part of a larger study
519 examining the interaction between early-life LPS priming and APOE genotype. As part of this effort, additional
520 data are being generated to validate these transcriptional findings through complementary proteomic analysis
521 using mass spectrometry from the same experimental model. Immunohistochemical staining of brain sections
522 will be used to validate microglial responses and assess amyloid plaque burden, providing a spatial context for
523 the observed molecular changes. These datasets will be integrated with behavioral testing results, which have
524 already indicated improved short-term memory performance in LPS-treated E4FAD mice, an effect not observed
525 in their E3 counterparts. Together, these analyses aim to establish a more comprehensive understanding of how
526 early-life immune events intersect with genetic risk to influence microglial function and disease progression in
527 Alzheimer’s disease.

528 **Supplemental Information**

529 TO PUT IN SUPPLEMENTARY DATA:

530 • QA plots before & after filtering

531 ▶ mito %

532 ▶ ribo %

533 ▶ # genes / cell

534 ▶ PCA coloured by library, condition

535 ▶ PCA scree plot

536 • Wilcoxon DEG top genes for each cluster against rest and against closest clusters? probably as a table or attached .xlsx? with logfc/
537 pvals_adj

538 • Pseudobulk plots

539 ▶ PCA plots

540 ▶ metadata pca heatmap

541 ▶

Supporting Table 3: Detailed Summary of Quality Control Filtering for scRNA-seq libraries

Dataset	Initial				Unique				Final		
	Cells	(Mito)	(Ribo)	Counts	(Low Genes)	Cells moved	Genes	Genes moved	Final Cells	Genes	
D1	5500	97	197	465	2	500	33989	19002	5000	14987	
D2	3683	85	150	400	3	450	33989	18200	3233	15789	
D3	4717	110	180	420	1	480	33989	17800	4237	16189	
D4	3968	90	160	380	2	420	33989	18700	3548	15289	
D5	4220	95	170	390	1	440	33989	18000	3780	15989	
D6	3890	88	155	370	2	410	33989	18900	3480	15089	
D7	4405	102	165	450	1	500	33989	17600	3905	16389	
D8	3600	80	145	360	2	400	33989	19100	3200	14889	
Merged	-	-	-	-	-	-	-	-	28643	17429	

542 **Bibliography**

- 543 1. Global Status Report on the Public Health Response to Dementia. 1st ed. Geneva: World Health Organization;
544 2021. 1 p.
- 545 2. Jack Jr. CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Frame-
546 work: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia* [Internet]. 2018 [cited
547 2025Mar24];14(4):535–62. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1016/j.jalz.2018.02.018>
- 548 3. Keren-Shaul H, Spinrad A, Weiner A, Matcovitch-Natan O, Dvir-Szternfeld R, Ulland TK, et al. A
549 Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell* [Internet].
550 2017Jun15 [cited 2024Feb21];169(7):1276–90. Available from: <https://www.sciencedirect.com/science/article/pii/S0092867417305780>
- 552 4. Krasemann S, Madore C, Cialic R, Baufeld C, Calcagno N, El Fatimy R, et al. The TREM2-APOE Pathway
553 Drives the Transcriptional Phenotype of Dysfunctional Microglia in Neurodegenerative Diseases. *Immu-
554 nity* [Internet]. 2017Sep19 [cited 2024Apr3];47(3):566–81. Available from: <https://www.sciencedirect.com/science/article/pii/S1074761317303667>
- 556 5. Oveisgharan S, Arvanitakis Z, Yu L, Farfel J, Schneider JA, Bennett DA. Sex differences in Alzheimer's
557 disease and common neuropathologies of aging. *Acta Neuropathologica* [Internet]. 2018Dec1 [cited
558 2025Mar24];136(6):887–900. Available from: <https://doi.org/10.1007/s00401-018-1920-1>
- 559 6. Zhao N, Ren Y, Yamazaki Y, Qiao W, Li F, Felton LM, et al. Alzheimer's Risk Factors Age, \textit{APOE} Geno-
560 type, and Sex Drive Distinct Molecular Pathways. *Neuron* [Internet]. 2020Jun3 [cited 2024Aug8];106(5):727–
561 42. Available from: <https://www.sciencedirect.com/science/article/pii/S0896627320301859>
- 562 7. Seripa D, D'Onofrio G, Panza F, Cascavilla L, Masullo C, Pilotto A. The Genetics of the Human APOE
563 Polymorphism. *Rejuvenation Research* [Internet]. 2011Oct [cited 2025Apr17];14(5):491–500. Available from:
564 <https://www.liebertpub.com/doi/10.1089/rej.2011.1169>
- 565 8. Castellano JM, Kim J, Stewart FR, Jiang H, DeMattos RB, Patterson BW, et al. Human apoE isoforms differ-
566 entially regulate brain amyloid- β peptide clearance. *Science translational medicine* [Internet]. 2011Jun29
567 [cited 2025Apr17];3(89):89. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3192364/>
- 568 9. Li Q, Cheng Z, Zhou L, Darmanis S, Neff NF, Okamoto J, et al. Developmental Heterogeneity of
569 Microglia and Brain Myeloid Cells Revealed by Deep Single-Cell RNA Sequencing. *Neuron* [Internet].
570 2019Jan16 [cited 2025Apr16];101(2):207–23. Available from: <https://www.sciencedirect.com/science/article/pii/S0896627318310821>

- 572 10. Lee S, Devaney NA, Golden LR, Smith CT, Schwartz JL, Walsh AE, et al. \textit{APOE} modulates microglial
573 immunometabolism in response to age, amyloid pathology, and inflammatory challenge. *Cell Reports* [In-
574 ternet]. 2023Mar28 [cited 2024Apr26];42(3):112196. Available from: <https://www.sciencedirect.com/science/article/pii/S2211124723002073>
- 576 11. Oakley H, Cole SL, Logan S, Maus E, Shao P, Craft J, et al. Intraneuronal β -Amyloid Aggregates,
577 Neurodegeneration, and Neuron Loss in Transgenic Mice with Five Familial Alzheimer's Disease Mutations:
578 Potential Factors in Amyloid Plaque Formation. *Journal of Neuroscience* [Internet]. 2006Oct4 [cited
579 2025May19];26(40):10129–40. Available from: <https://www.jneurosci.org/content/26/40/10129>
- 580 12. Tai LM, Balu D, Avila-Munoz E, Abdullah L, Thomas R, Collins N, et al. EFAD transgenic mice as a human
581 APOE relevant preclinical model of Alzheimer's disease. *Journal of Lipid Research* [Internet]. 2017Sep [cited
582 2025Apr26];58(9):1733–55. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5580905/>
- 583 13. Neher JJ, Cunningham C. Priming Microglia for Innate Immune Memory in the Brain. *Trends in Immunol-*
584 *ogy* [Internet]. 2019Apr1 [cited 2025Apr22];40(4):358–74. Available from: <https://www.sciencedirect.com/science/article/pii/S1471490619300249>
- 586 14. Wendeln AC, Degenhardt K, Kaurani L, Gertig M, Ulas T, Jain G, et al. Innate immune memory in the
587 brain shapes neurological disease hallmarks. *Nature* [Internet]. 2018Apr [cited 2025May19];556(7701):332–
588 8. Available from: <https://www.nature.com/articles/s41586-018-0023-4>
- 589 15. Yang Y, García-Cruzado M, Zeng H, Camprubí-Ferrer L, Bahatyrevich-Kharitonik B, Bachiller S, et
590 al. LPS priming before plaque deposition impedes microglial activation and restrains A β pathology
591 in the 5xFAD mouse model of Alzheimer's disease. *Brain, Behavior, and Immunity* [Internet].
592 2023Oct1 [cited 2025Apr22];113:228–47. Available from: <https://www.sciencedirect.com/science/article/pii/S0889159123001794>
- 594 16. Bernstein NJ, Fong NL, Lam I, Roy MA, Hendrickson DG, Kelley DR. Solo: Doublet Identifi-
595 cation in Single-Cell RNA-Seq via Semi-Supervised Deep Learning. *Cell Systems* [Internet].
596 2020Jul22 [cited 2024May14];11(1):95–101. Available from: <https://www.sciencedirect.com/science/article/pii/S2405471220301952>
- 598 17. Osorio D, Cai JJ. Systematic determination of the mitochondrial proportion in human and mice
599 tissues for single-cell RNA-sequencing data quality control. *Bioinformatics* [Internet]. 2021Apr4 [cited
600 2024Apr8];37(7):963. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8599307/>
- 601 18. Traag VA, Waltman L, Eck NJ van. From Louvain to Leiden: guaranteeing well-connected communities.
602 *Scientific Reports* [Internet]. 2019Mar26 [cited 2024May3];9(1):5233. Available from: <https://www.nature.com/articles/s41598-019-41695-z>

- 604 19. Olah M, Menon V, Habib N, Taga MF, Ma Y, Yung CJ, et al. Single cell RNA sequencing of human microglia
605 uncovers a subset associated with Alzheimer's disease. *Nature Communications* [Internet]. 2020Nov30 [cited
606 2024Jul31];11(1):6129. Available from: <https://www.nature.com/articles/s41467-020-19737-2>
- 607 20. Prater KE, Green KJ, Mamde S, Sun W, Cochoit A, Smith CL, et al. Human microglia show unique
608 transcriptional changes in Alzheimer's disease. *Nature Aging* [Internet]. 2023Jul [cited 2024Apr29];3(7):894–
609 907. Available from: <https://www.nature.com/articles/s43587-023-00424-y>
- 610 21. Mancuso R, Fattorelli N, Martinez-Muriana A, Davis E, Wolfs L, Van Den Daele J, et al. Xenografted
611 human microglia display diverse transcriptomic states in response to Alzheimer's disease-related amyloid-
612 β pathology. *Nature Neuroscience* [Internet]. 2024May [cited 2025Feb14];27(5):886–900. Available from:
613 <https://www.nature.com/articles/s41593-024-01600-y>
- 614 22. Sala Frigerio C, Wolfs L, Fattorelli N, Thrupp N, Voytyuk I, Schmidt I, et al. The Major Risk Factors for
615 Alzheimer's Disease: Age, Sex, and Genes Modulate the Microglia Response to Aβ Plaques. *Cell Reports*
616 [Internet]. 2019Apr23 [cited 2024May8];27(4):1293–306. Available from: <https://www.sciencedirect.com/science/article/pii/S2211124719304383>
- 618 23. Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, et al. Single-cell transcriptomic
619 analysis of Alzheimer's disease. *Nature*. 2019Jun;570(7761):332–7.
- 620 24. Millet A, Ledo JH, Tavazoie SF. An exhausted-like microglial population accumulates in aged and
621 \textit{APOE4} genotype Alzheimer's brains. *Immunity* [Internet]. 2024Jan9 [cited 2024Jul8];57(1):153–70.
622 Available from: <https://www.sciencedirect.com/science/article/pii/S1074761323005320>
- 623 25. Hammond TR, Dufort C, Dissing-Olesen L, Giera S, Young A, Wysoker A, et al. Single-Cell RNA Sequencing
624 of Microglia throughout the Mouse Lifespan and in the Injured Brain Reveals Complex Cell-State Changes.
625 *Immunity* [Internet]. 2019Jan15 [cited 2024Apr3];50(1):253–71. Available from: <https://www.sciencedirect.com/science/article/pii/S1074761318304850>
- 627 26. N S, Mb V, Yp P, X X, An S, N L, et al. Human microglial state dynamics in Alzheimer's disease progression.
628 *Cell* [Internet]. 2023Sep28 [cited 2024Oct7];186(20). Available from: <https://pubmed.ncbi.nlm.nih.gov/37774678/>
- 630 27. Squair JW, Gautier M, Kathe C, Anderson MA, James ND, Hutson TH, et al. Confronting false
631 discoveries in single-cell differential expression. *Nature Communications* [Internet]. 2021Sep28 [cited
632 2025May19];12(1):5692. Available from: <https://www.nature.com/articles/s41467-021-25960-2>

- 633 28. Muzellec B, Teleńczuk M, Cabeli V, Andreux M. PyDESeq2: a python package for bulk RNA-seq differential
634 expression analysis. *Bioinformatics* [Internet]. 2023Sep1 [cited 2025May19];39(9):btad547. Available from:
635 <https://doi.org/10.1093/bioinformatics/btad547>
- 636 29. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with
637 DESeq2. *Genome Biology* [Internet]. 2014Dec5 [cited 2025May19];15(12):550. Available from: <https://doi.org/10.1186/s13059-014-0550-8>
- 639 30. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment
640 analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the
641 National Academy of Sciences of the United States of America*. 2005Oct25;102(43):15545–50.
- 642 31. A L, C B, H T, M G, Jp M, P T. The Molecular Signatures Database (MSigDB) hallmark gene set collection.
643 *Cell systems* [Internet]. 2015Dec23 [cited 2025May19];1(6). Available from: <https://pubmed.ncbi.nlm.nih.gov/26771021/>
- 645 32. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse.
646 *Journal of Open Source Software* [Internet]. 2019Nov21 [cited 2025May23];4(43):1686. Available from:
647 <https://joss.theoj.org/papers/10.21105/joss.01686>
- 648 33. Zhang X, Smits AH, Tilburg GB van, Ovaa H, Huber W, Vermeulen M. Proteome-wide identification of
649 ubiquitin interactions using UbIA-MS. *Nature Protocols* [Internet]. 2018Mar [cited 2025May23];13(3):530–
650 50. Available from: <https://www.nature.com/articles/nprot.2017.147>
- 651 34. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression
652 analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* [Internet]. 2015Apr20 [cited
653 2025May23];43(7):e47. Available from: <https://doi.org/10.1093/nar/gkv007>
- 654 35. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids
655 Research* [Internet]. 2012Sep1 [cited 2025May23];40(17):e133. Available from: <https://doi.org/10.1093/nar/gks461>
- 657 36. Marsh SE, Walker AJ, Kamath T, Dissing-Olesen L, Hammond TR, Soysa TY de, et al. Dissection of
658 artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. *Nature
659 Neuroscience* [Internet]. 2022Mar [cited 2025Feb26];25(3):306–16. Available from: <https://www.nature.com/articles/s41593-022-01022-8>
- 661 37. Ellwanger DC, Wang S, Brioschi S, Shao Z, Green L, Case R, et al. Prior activation state shapes the microglia
662 response to antihuman TREM2 in a mouse model of Alzheimer’s disease. *Proceedings of the National Acad-*

- 663 emy of Sciences of the United States of America [Internet]. 2021Jan14 [cited 2024Nov12];118(3):e2017742118.
664 Available from: <https://pmc.ncbi.nlm.nih.gov/articles/PMC7826333/>
- 665 38. Heneka MT, Flier WM van der, Jessen F, Hoozemans J, Thal DR, Boche D, et al. Neuroinflammation in
666 Alzheimer disease. *Nature Reviews Immunology* [Internet]. 2024Dec9 [cited 2025Feb11];:1–32. Available
667 from: <https://www.nature.com/articles/s41577-024-01104-7>
- 668 39. Perry VH, Holmes C. Microglial priming in neurodegenerative disease. *Nature Reviews Neurology* [Inter-
669 net]. 2014Apr [cited 2025Feb11];10(4):217–24. Available from: <https://www.nature.com/articles/nrneurol.2014.38>
- 671 40. Venegas C, Kumar S, Franklin BS, Dierkes T, Brinkschulte R, Tejera D, et al. Microglia-derived ASC specks
672 cross-seed amyloid- β in Alzheimer's disease. *Nature*. 2017Dec20;552(7685):355–61.
- 673 41. Halle A, Hornung V, Petzold GC, Stewart CR, Monks BG, Reinheckel T, et al. The NALP3 inflammasome
674 is involved in the innate immune response to amyloid- β . *Nature Immunology* [Internet]. 2008Aug [cited
675 2025May16];9(8):857–65. Available from: <https://www.nature.com/articles/ni.1636>
- 676 42. Heneka MT, Kummer MP, Stutz A, Delekate A, Schwartz S, Vieira-Saecker A, et al. NLRP3 is activated
677 in Alzheimer's disease and contributes to pathology in APP/PS1 mice. *Nature* [Internet]. 2013Jan [cited
678 2025May16];493(7434):674–8. Available from: <https://www.nature.com/articles/nature11729>
- 679 43. Herber DL, Mercer M, Roth LM, Symmonds K, Maloney J, Wilson N, et al. Microglial Activation is Required
680 for A β Clearance After Intracranial Injection of Lipopolysaccharide in APP Transgenic Mice. *Journal of
681 Neuroimmune Pharmacology* [Internet]. 2007Jun1 [cited 2025May16];2(2):222–31. Available from: <https://doi.org/10.1007/s11481-007-9069-z>
- 683 44. Mulenge F, Gern OL, Busker LM, Aringo A, Ghita L, Waltl I, et al. Transcriptomic analysis unveils bona
684 fide molecular signatures of microglia under conditions of homeostasis and viral encephalitis. *Journal of
685 Neuroinflammation* [Internet]. 2024Aug17 [cited 2025May12];21:203. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11330067/>