Acta Neuropathologica (2025) 150:10
https://doi.org/10.1007/s00401-025-02913-3

ORIGINAL PAPER

Single-cell transcriptomic landscape of the neuroimmune compartment in amyotrophic lateral sclerosis brain and spinal cord

John F. Tuddenham[1,2,3] · Masashi Fujita[1,2,3] · Emily Lee[1,3] · Nivedita Nimmagadda[1,3] · Anthony Khairallah[1,2] · Claire Harbison[1,2] · Xena E. Flowers[3,10] · Guillermo Coronas-Samano[3,10] · Silas Maniatis[9] · Aidan Daly[9] · Julie A. Schneider[4] · Andrew F. Teich[1,3,10] · Jean Paul G. Vonsattel[1,3] · Peter A. Sims[5,6] · Wassim Elyaman[1,3] · Elizabeth M. Bradshaw[1,3,12] · Lyle W. Ostrow[13] · Hemali Phatnani[1,7,8,9] · Neil A. Shneider[1,7] · David A. Bennett[4] · Philip L. De Jager[1,2,3] · Serge Przedborski[1,8,10,11] · Vilas Menon[1,2,3] · Marta Olah[1,3]

Received: 16 November 2024 / Revised: 13 July 2025 / Accepted: 14 July 2025

Abstract

Development of therapeutic approaches that target specific microglia responses in amyotrophic lateral sclerosis (ALS) is crucial due to the involvement of microglia in ALS progression. Our study identifies the predominant microglia subset in human ALS primary motor cortex and spinal cord as an undifferentiated phenotype with dysregulated respiratory electron transport. Moreover, we find that the interferon response microglia subset is enriched in donors with aggressive disease progression, while a previously described potentially protective microglia phenotype is depleted in ALS. Additionally, we observe an enrichment of non-microglial immune cell, mainly NK/T cells, in the ALS central nervous system, primarily in the spinal cord. These findings pave the way for the development of microglia subset-specific therapeutic interventions to slow or even stop ALS progression.

Keywords Microglia · Phenotypic heterogeneity · Amyotrophic lateral sclerosis · Single-cell RNA-sequencing

✉ Marta Olah
mo2738@cumc.columbia.edu

1 Department of Neurology, Divisions of Neuroimmunology/Translational Neurobiology/None, Columbia University Medical Center, New York, NY, USA

2 Center for Translational & Computational Neuroimmunology, Department of Neurology, Columbia University Irving Medical Center, New York, NY, USA

3 Taub Institute for Research On Alzheimer's Disease and the Aging Brain, Columbia University Medical Center, New York, NY, USA

4 Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA

5 Department of Systems Biology, Columbia University Medical Center, New York, NY, USA

6 Department of Biochemistry and Molecular Biophysics, Columbia University Medical Center, New York, NY, USA

7 Eleanor and Lou Gehrig ALS Center, Columbia University Medical Center, New York, USA

8 Center for Motor Neuron Biology and Disease, Columbia University Medical Center, New York, NY, USA

9 Center for Genomics of Neurodegenerative Disease, New York Genome Center, New York, NY, USA

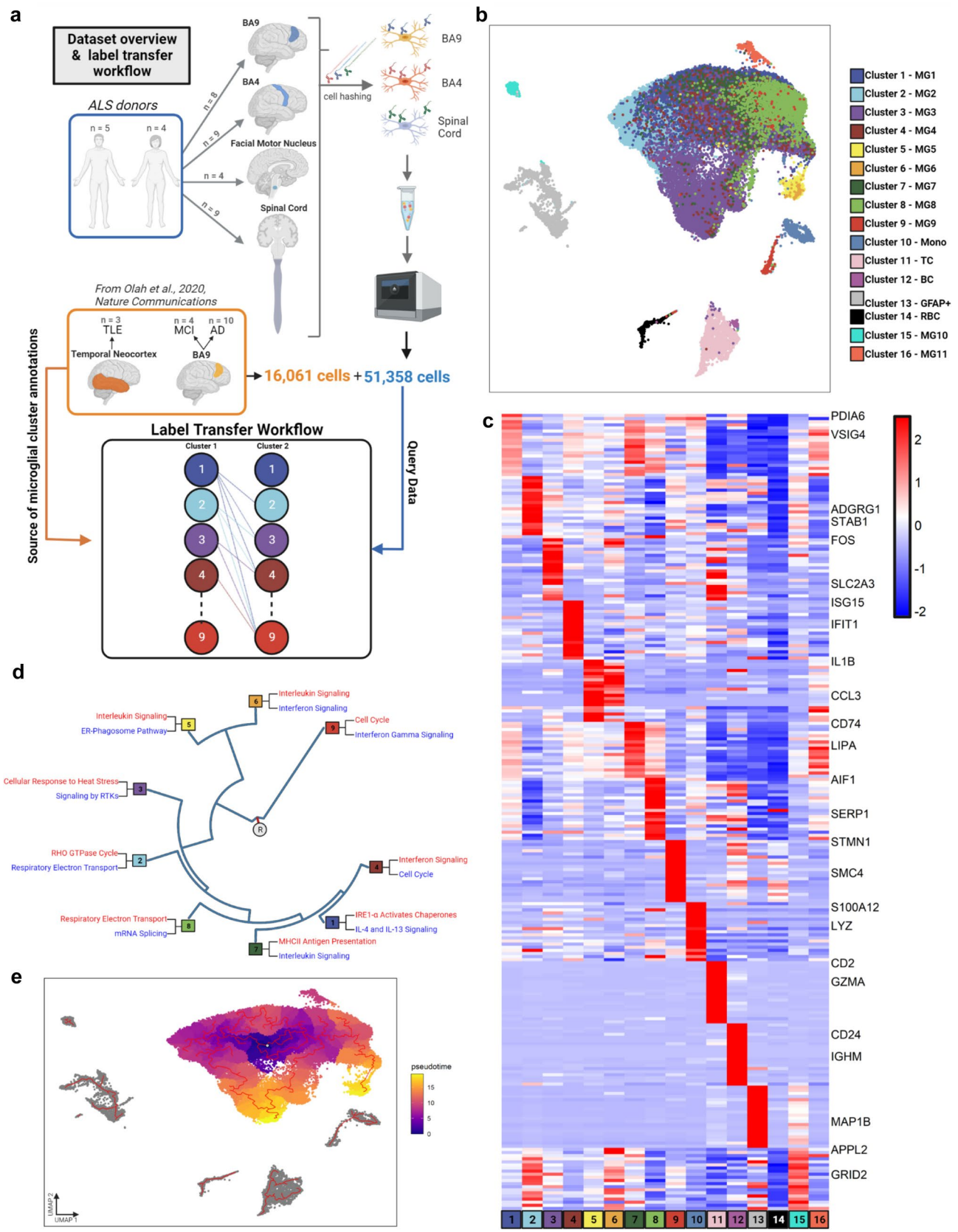
10 Department of Pathology and Cell Biology, Columbia University Medical Center, New York, NY, USA

11 Department of Neuroscience, Columbia University Medical Center, New York, NY, USA

12 The Carol and Gene Ludwig Center for Research On Neurodegeneration, Columbia University, New York, NY, USA

13 Department of Neurology, Lewis Katz School of Medicine at Temple University Hospital, Philadelphia, PA, USA

Published online: 29 July 2025


a
Dataset overview & label transfer workflow
ALS donors
n = 5
n = 4
n = 8
n = 9
n = 4
n = 9
BA9
BA4
Facial Motor Nucleus
Spinal Cord
cell hashing
BA9
BA4
Spinal Cord
From Olah et al., 2020, Nature Communications
n = 3
TLE
Temporal Neocortex
n = 4
MCI
n = 10
AD
BA9
16,061 cells + 51,358 cells
Source of microglial cluster annotations
Label Transfer Workflow
Cluster 1
Cluster 2
Query Data
b
Cluster 1 - MG1
Cluster 2 - MG2
Cluster 3 - MG3
Cluster 4 - MG4
Cluster 5 - MG5
Cluster 6 - MG6
Cluster 7 - MG7
Cluster 8 - MG8
Cluster 9 - MG9
Cluster 10 - Mono
Cluster 11 - TC
Cluster 12 - BC
Cluster 13 - GFAP+
Cluster 14 - RBC
Cluster 15 - MG10
Cluster 16 - MG11
c
PDIA6
VSIG4
ADGRG1
STAB1
FOS
SLC2A3
ISG15
IFIT1
IL1B
CCL3
CD74
LIPA
AIF1
SERP1
STMN1
SMC4
S100A12
LYZ
CD2
GZMA
CD24
IGHM
MAP1B
APPL2
GRID2
2
1
0
-1
-2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
d
Interleukin Signaling
Interferon Signaling
Interleukin Signaling
ER-Phagosome Pathway
Cell Cycle
Interferon Gamma Signaling
Cellular Response to Heat Stress
Signaling by RTKs
RHO GTPase Cycle
Respiratory Electron Transport
Interferon Signaling
Cell Cycle
Respiratory Electron Transport
mRNA Splicing
IRE1-α Activates Chaperones
IL-4 and IL-13 Signaling
MHCII Antigen Presentation
Interleukin Signaling
e
pseudotime
15
10
5
0
UMAP 2
UMAP 1

◂**Fig. 1** The population structure of microglia in human ALS brain and spinal cord. We used label transfer to explore the microglia (and other immune cell) phenotypes present in the central nervous system (CNS) of ALS donors. **a** Sample collection and label transfer workflow. Multiple CNS regions were sampled from 9 ALS donors with similar representation of both sexes. Single cell RNA sequencing data of live CD45 + immune cells was generated using the 10 × Chromium platform. The ALS data was mapped onto our previously published microglia population structure by utilizing a pairwise machine learning approach with random forest classifiers and consensus voting to identify final labels. **b** UMAP projection of the merged dataset. The merged dataset is plotted on the first two UMAP components following Harmony batch correction. Each dot is a single cell. Microglia cluster 15 (MG10) had no unique gene set, while cluster 16 (MG11) was present only in one donor, accordingly they were not included in downstream analysis. **c** ALS microglia subsets have unique marker gene sets. Heatmap representing Z-scored expression data. Each column is a single cluster and each row is a single gene. **d** Microglial subsets present in ALS have distinctive functional annotations. Hierarchical dendogram demonstrating the functional landscape of microglia subsets. Similarity of subsets was calculated using Euclidean distance across the average expression profiles of each cell subtype. R denotes the root node. Top up- and down-regulated terms were selected from Reactome pathway annotation to highlight unique aspects of each microglial subsets. Terms in red are upregulated in a given cluster while terms in blue are downregulated. **e** Microglia subsets in ALS reside along divergent state transition trajectories. A pseudotime trajectory was built with monocle3, setting the root point in the middle of cluster 1. The trajectory is highlighted in red, showing the shifts seen across different aspects of the microglial cloud, including different pseudotime endpoints. *BA* Brodmann area, *ALS* amyotrophic lateral sclerosis, *UMAP* uniform manifold approximation and projection, *R* root node

Introduction

Amyotrophic lateral sclerosis (ALS) is a rapidly progressing neurodegenerative disease characterized by the gradual loss of upper and lower motor neurons, leading to a decline in motor function [32]. The median survival time from onset of symptoms is typically only 3–5 years, depending on specific subtype of the disease [1]. Currently, there are no disease-modifying treatments available for ALS, making it uniformly fatal.

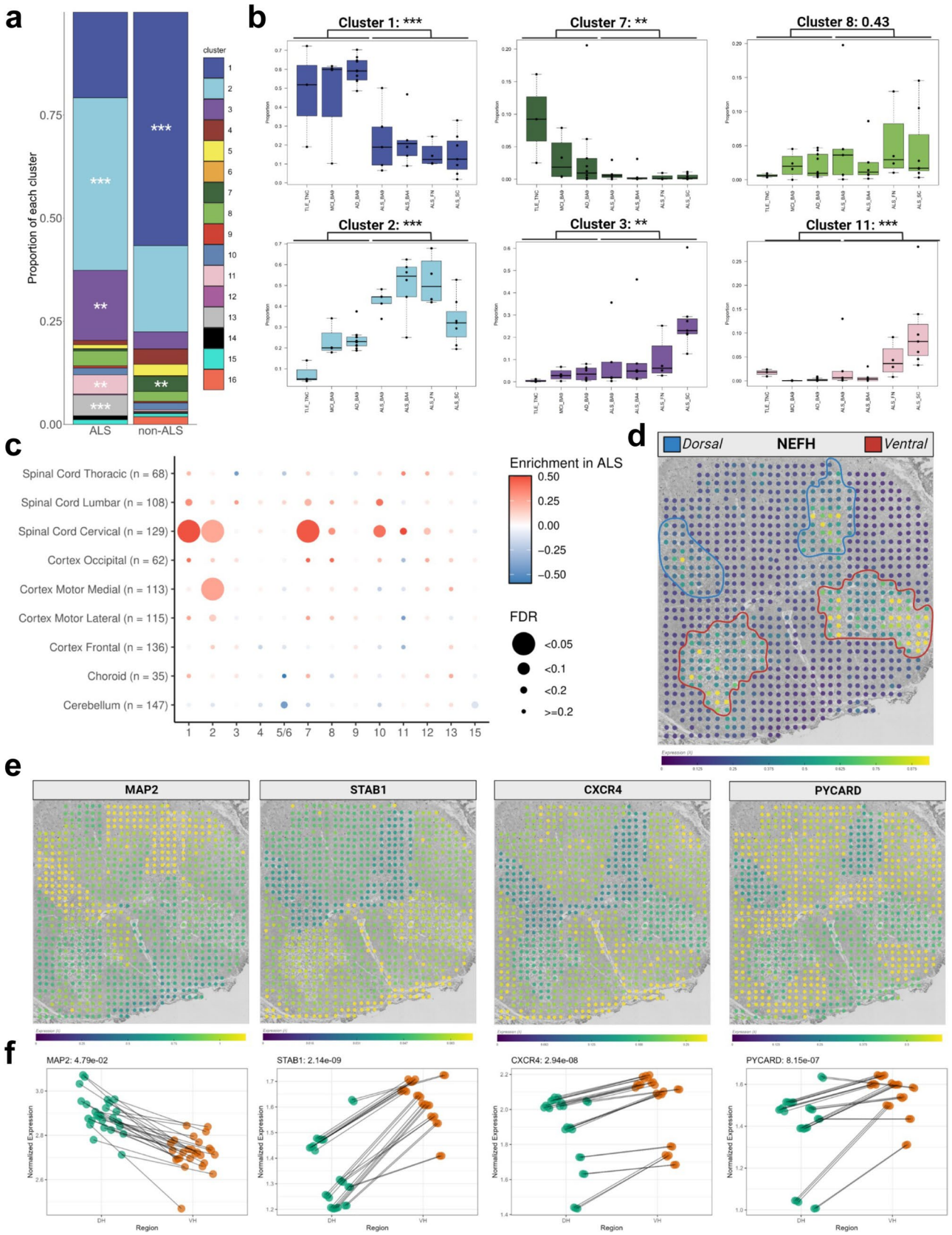
One of the most prominent neuropathological features of ALS, in addition to motor neuron degeneration, is neuroinflammation. Human positron emission tomography (PET) imaging studies using TSPO radioligands have revealed that neuroinflammation can be detected early in the course of ALS and persists up until the later stage of the disease [8]. Neuroinflammation is associated with the activation of microglia, which are the resident innate immune cells of the central nervous system (CNS). Several histopathological studies using post-mortem brain and spinal cord specimens have demonstrated microglial proliferation and activation in ALS, with the severity of microglia-associated pathology closely correlating with disease progression [5, 6]. While these studies have provided valuable insights, they mainly rely on morphological analysis or general markers for assessing microglial activation. As a result, our understanding of the complex molecular mechanisms underlying microglial involvement in ALS pathobiology remains limited. This knowledge gap is further emphasized by the failure of clinical trials that attempted to block general neuroinflammation in altering disease progression [41]. Surprisingly, therapies targeting neuroinflammation that demonstrated significant benefits in pre-clinical rodent models of ALS have not translated into successful outcomes in human clinical trials [41]. This disparity highlights the importance of studying the identity and role of microglia specifically in humans to gain a deeper understanding of ALS pathogenesis.

In our recent work, we made significant strides in this area by identifying nine human microglial subsets across different donor ages using single-cell RNA-sequencing [36]. Building upon this knowledge, we expanded our investigations to ALS donor samples using the same experimental and analytical pipeline. By comparing the results to our previous study, we successfully identified microglial and non-microglial immune cell subsets that are enriched in ALS. Furthermore, we explored the functional annotation and transcriptional regulatory networks of these ALS-associated microglia phenotypes, validating their relevance for ALS disease pathogenesis across independent datasets. Thus, our study provides a comprehensive account of the human microglial subsets associated with ALS, identified through unbiased transcriptomic approaches. Our findings will pave the way for developing therapeutic interventions that target specific microglia subsets and their unique pathways in the context of ALS.

Materials and methods

Experimental set up

For scRNA-seq analysis, "ALS" samples in Figs. 1, 2a, 3, 4 and Supplementary figures S5, S6, S8, S9, S11-S13 represent microglia and other immune cells isolated from donors and CNS regions shown in Table 1 and in the first tab of Supplementary data table S1. "Non-ALS" samples in in Figs. 1, 2a, 3, 4 and Supplementary figures S5, S6, S8, S9, S11-S13 represent microglia and other immune cells isolated from the temporal neocortex of donors undergoing temporal lobectomy for temporal lobe epilepsy and the prefrontal cortex of donors with mild cognitive impairment and Alzheimer's disease, as described in or previous publication [36]. For the meta-analysis of published bulk RNA-seq data in Fig. 2c and Supplementary figure S8a–i "control" refers to non-neurological controls described in detail in the original publication describing the dataset [51]. For the immunohistochemistry analysis shown in Supplementary figure S10


a
Proportion of each cluster
0.00
0.25
0.50
0.75
ALS
non-ALS
cluster
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
b
Cluster 1: ***
Cluster 7: **
Cluster 8: 0.43
Cluster 2: ***
Cluster 3: **
Cluster 11: ***
c
Spinal Cord Thoracic (n = 68)
Spinal Cord Lumbar (n = 108)
Spinal Cord Cervical (n = 129)
Cortex Occipital (n = 62)
Cortex Motor Medial (n = 113)
Cortex Motor Lateral (n = 115)
Cortex Frontal (n = 136)
Choroid (n = 35)
Cerebellum (n = 147)
1 2 3 4 5/6 7 8 9 10 11 12 13 15
Enrichment in ALS
0.50
0.25
0.00
−0.25
−0.50
FDR
<0.05
<0.1
<0.2
>=0.2
d
Dorsal
NEFH
Ventral
e
MAP2
STAB1
CXCR4
PYCARD
f
MAP2: 4.79e-02
STAB1: 2.14e-09
CXCR4: 2.94e-08
PYCARD: 8.15e-07
Normalized Expression
Region
DH
VH

◂**Fig. 2** ALS induces robust, region specific shifts in microglial subtype prevalence. **a** Global changes in microglia subset relative abundance in ALS. Stacked bar chart showing the overall changes in immune cell relative abundance. Each bar shows the proportion of the different microglial and non-microglia immune cell subsets in each condition: non-ALS and ALS. Asterisks denote significant difference between conditions, using the Wilcoxon rank-sum test and BH correction to determine significance. *** < 0.005, ** < 0.01, * < 0.05. **b** Region-specific changes in microglia cluster relative abundance in ALS. Boxplots showing the distribution of individual cluster relative abundances across disease-region pairings. Representative clusters with significant, region-specific changes are shown. Boxplots denote the 25th percentile, median, and 75th percentile, with whiskers capturing 1.5 IQR in both directions. **c** Orthogonal validation in an independent bulk RNA-seq dataset confirms consistent association of microglia Cluster 2 signature with ALS. Using a separate bulk RNA-seq cohort of 170 samples from ALS patients and non-ALS neurological disease patients or control patients, signatures of the top 20 genes per cluster were used to delineate the enrichment of different cluster signatures in ALS versus non-ALS samples. Notably, cluster 2 is the only cluster that shows significant enrichment in ALS in multiple regions, while cluster 1 and 7 enrichment likely capture the increase in overall microglia numbers in ALS. **d** Orthogonal validation in an independent spatial transcriptomic dataset confirms the upregulation of ALS associated microglial subset marker genes at the anatomical sites of motor neuron death in ALS. Spatial transcriptomic data was repurposed from Maniatas et al. A representative image from the slide viewer at https://als-st.nygenome.org/ is shown, displaying the neuronal gene NEFH, which is primarily found in the dorsal and ventral horns. Color bar is expression lambda calculated by the Splotch model. Dorsal and Ventral horns are demarcated in blue and red respectively. Representative images in the following panels are from the same section. **e**–**f** Marker genes of clusters 2 and 8 follow inverse patterns of upregulation in comparison to MAP2 in the dorsal and ventral horns of ALS patients. Representative images for each gene as in (**d**) are shown in (**e**). In **f**, dotplots compare the summed score of predicted counts for a given gene in all spots in the dorsal horn (green) to an identical score for that gene summed from all spots in the ventral horn (green) across a subset of donors with strong MAP2 detection. Testing for the significance of differences between regions for each gene was conducted with Welch's t-test and the Holm-Bonferroni correction. STAB1 is a defining gene for cluster 2, CXCR4 is a marker for cluster 3, and PYCARD is a defining gene for cluster 8. Abbreviations: BA Brodmann area, ALS amyotrophic lateral sclerosis, AD Alzheimer's disease, MCI mild cognitive impairment, TLE temporal lobe epilepsy, ALS amyotrophic lateral sclerosis, BA Brodmann area, TNC temporal neocortex, SC spinal cord, SN substantia nigra, FN facial nucleus, IQR interquartile range, FDR false discover rate, DH dorsal horn, VH ventral horn

"controls" are non-neurological controls described in detail in the second tab of Supplementary data table S1.

Source of fresh central nervous system autopsy specimens

All fresh ALS central nervous system specimens were obtained through the Eleanor and Lou Gehrig ALS Center's autopsy program (director: Neil A. Shneider) administered at Columbia University Medical Center (CUMC)/New York Brain Bank (NYBB). All procedures and research protocols were approved by the Institutional Review Board (IRB) of Columbia University Medical Center (protocol AAAR4962). For a detailed description of the brain regions sampled, clinical diagnosis, age and sex of the donors see Table S1. After weighing, the resected brain and spinal cord tissue was placed in ice-cold transportation medium (Hibernate-A medium (Gibco, A1247501) containing 1% B27 serum-free supplement (Gibco, 17504044) and 1% GlutaMax (Gibco, 35050061)) and transported from the autopsy suite at CUMC/NYBB to the wet lab at CUMC on ice (4 °C) for immediate processing and live microglia isolation.

Microglia isolation, cell hashing and sorting

The isolation of microglia was performed according to our published protocol [37], with minor modifications. In case of the cortical autopsy samples (BA9/46, BA4), the cortex (grey matter and the underlying white matter (subcortical white matter)) was dissected under a stereomicroscope. The subcortical white matter samples were not used in this study. The spinal cord sample (SC) was sampled at the level of the lumbar section and included both white and grey matter. All steps of the protocol were performed on ice. The dissected tissue was placed in HBSS (Lonza, 10-508F) and weighed. Subsequently, the tissue was homogenized in a 15 ml glass tissue grinder, 0.5 g at a time. The resulting homogenate was filtered through a 70 µm filter and spun down at 300 rcf for 10 min. The pellet was resuspended in 2 ml staining buffer (RPMI (Fisher, 72400120) containing 1% B27) per 0.5 g of initial tissue and incubated with anti-myelin magnetic beads (Miltenyi, 130–096-733) for 15 min according to the manufacturer's specification. The homogenate was then washed once with staining buffer and the myelin was depleted using Miltenyi large separation columns (Miltenyi, 130–042-202). The cell suspension was spun down and then incubated with anti-CD11b AlexaFluor488 (BioLegend, 301318) and anti-CD45 AlexaFluor647 (BioLegend, 304018) antibodies as well as 7AAD (BD Pharmingen, 559925) for 20 min on ice. Samples ALS1-ALS5 (see Table S1) were received and processed before the commercial availability of cell hashing antibodies. For samples ALS6-ALS9, the individual central nervous system regions were hashed using cell hashing antibodies along with anti-CD11b, anti-CD45 and 7AAD for 20 min on ice (for catalogue numbers of cell hashing antibodies see Table S1). Subsequently, the cell suspension was washed twice with staining buffer, filtered through a 70 µm filter and the CD45 + /7AAD- cells were sorted on a BD Influx cell sorter. Cells from each brain region were sorted in a separate A1 well of a 96-well PCR plate (Eppendorf, 951020401) containing 100 µl of PBS buffer with 0.3% BSA. For samples ALS1-ALS5 each sample/region was loaded independently on a 10 × Chromium channel. For samples ALS6-ALS9 following sorting cells

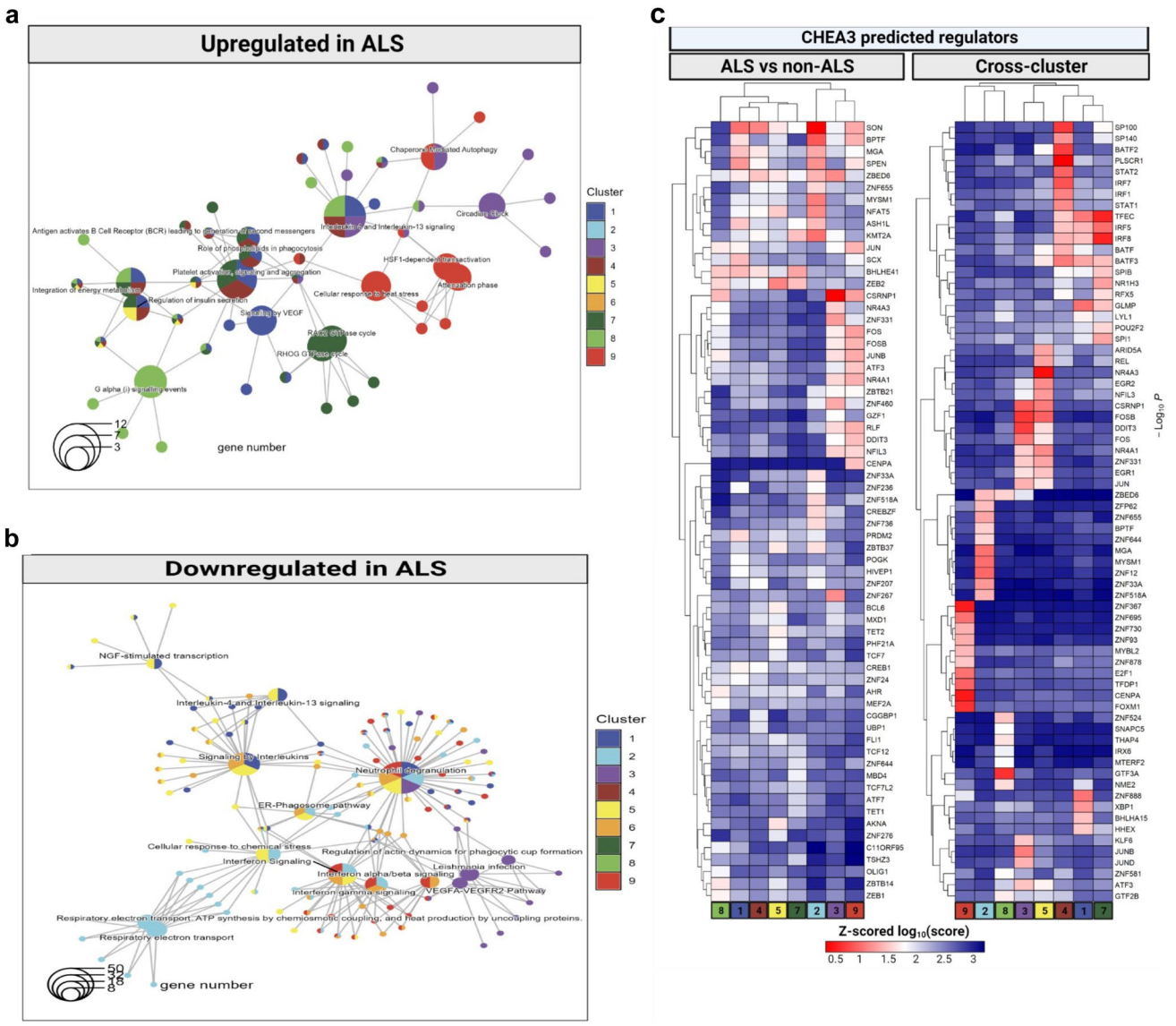


Fig. 3 ALS results in functional changes in microglia subsets. Annotating between subset and within-subset shifts in ALS microglia at the level of transcriptome, proteome, and epigenome. **a**–**b** Annotation of ALS associated transcriptional changes within each microglia cluster reveals functional shifts in ALS microglia. REACTOME pathway analysis of differentially expressed genes within each microglia cluster between ALS and non-ALS. Results are displayed as a connectivity plot, where central nodes represent REACTOME pathways, while terminal nodes represent genes associated with those REACTOME terms. Central nodes are colored based on the enrichment of the term in a given cluster, while terminal nodes are colored based on the presence of the gene in the differentially expressed gene list for the ALS vs. non-ALS comparison for a cluster. Genes/terms upregulated in ALS are shown in (**a**) and genes/terms downregulated in ALS are shown in (**b**). Please note that many of the pathways are differentially expressed in more than one microglia cluster. **c** Identification of differential and shared transcriptional regulators in cross-cluster and within-cluster cross-disease comparisons. Both heatmaps show Z-scored log-normalized scores from CHEA3. On the left panel, genes used for regulator calculation were the top 50 genes derived from within-cluster across-disease differential expression for each cluster. On the right panel, genes used as input were selected top 20 marker genes per cluster. Rows and columns are clustered hierarchically by absolute linkage. Each row shows the scores for a single regulator across all microglia clusters (columns)

from different brain regions were combined and immediately submitted to single cell capture)10 × Chromium channel), reverse transcription and library construction on the 10 × Chromium platform. All sorting was performed using a 100 µm nozzle. The sorting times varied according to the quality of the sample, but was on average between 10 and 20 min per sample. The sorting speed was kept between 8000 and 10000 events per second.

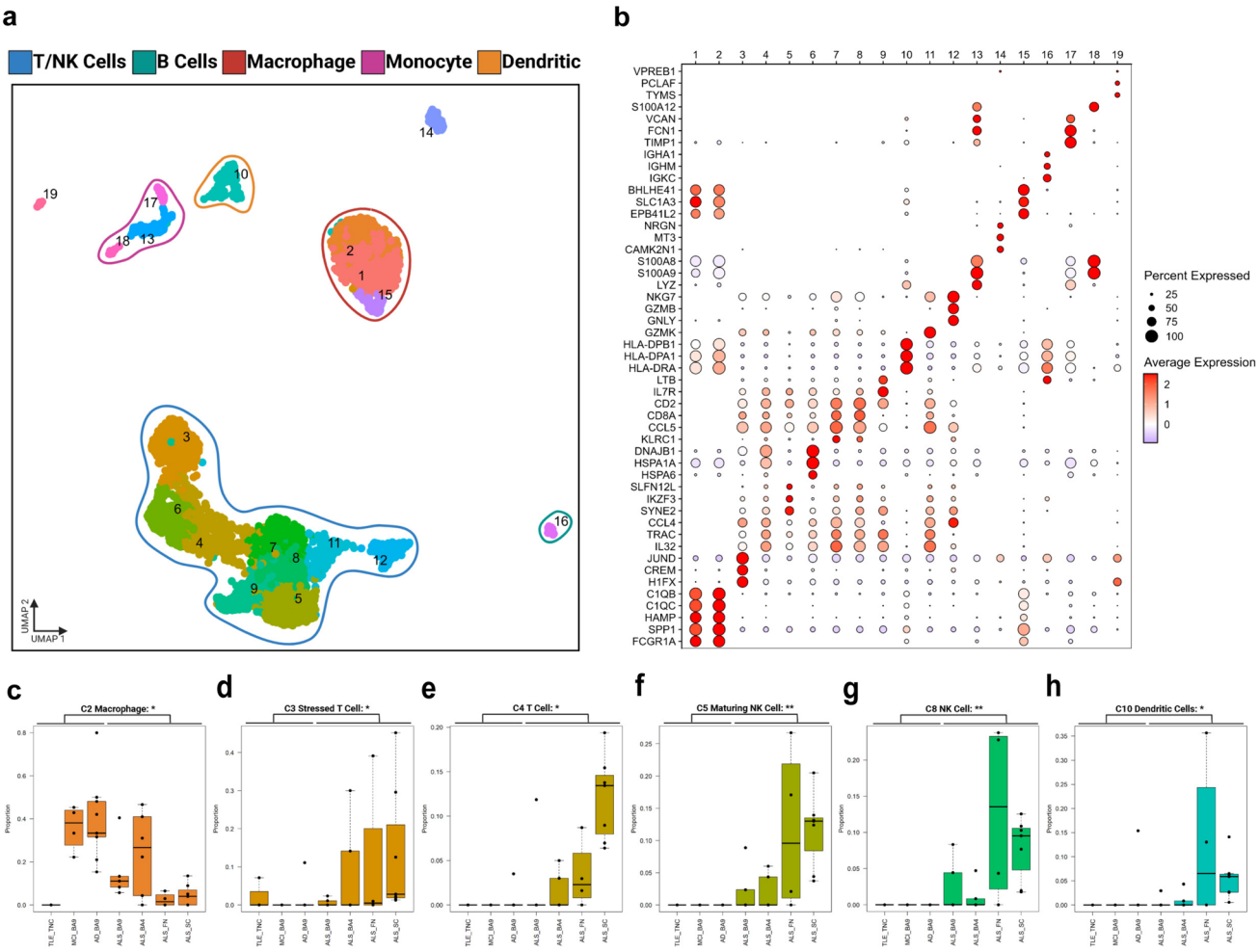


Fig. 4 Independent clustering of non-microglial immune cells identifies shifts in myeloid and adaptive immune cell populations in ALS. **a** Independent clustering of non-microglial immune cells. Non-microglial immune cells were isolated in silico and separately processed. Optimized clustering resolution was chosen with ChooseR. Clusters with greater than 10 cells are plotted on the UMAP. **b** Selected marker genes identify diverse immune populations. Each column represents one of the clusters, and each row represents the z-scored expression of a given gene. The plot is colored according to the level of expression and the size of each circle represents the percentage of cells in each cluster that express the gene. **c–h** Annotation identifies enrichment of T and NK cells in ALS, as well as enrichment of dendritic cells and depletion of macrophages. Boxplots denote the 25th percentile, median, and 75th percentile, with whiskers capturing 1.5 IQR in both directions. Asterisks denote the significance of the difference between conditions, using the Wilcoxon rank-sum test and BH correction to determine significance. * < 0.05. *ALS* amyotrophic lateral sclerosis, *NK* natural killer

10× Genomics chromium single-cell 3′ library construction

Cell capture, amplification and library construction on the 10× Genomics Chromium platform were performed according to the manufacturer's publicly available protocol. Briefly, viability was assessed by trypan blue exclusion assay, and cell density was adjusted to 175 cells per µl. 7000 cells were then loaded onto a single channel of a 10× Chromium chip for each sample. The 10× Genomics Chromium technology enables 3' digital gene expression profiling of thousands of cells from a single sample by separately indexing each cell's transcriptome. First, thousands of cells are partitioned into nanoliter-scale Gel Bead-In-EMulsions (GEMs). Within one GEM all generated cDNA share a common 10× barcode. Libraries were generated and sequenced from the cDNA, and the 10× barcodes were used to associate individual reads back to the individual partitions. To achieve single-cell resolution, the cells were delivered at a limiting dilution. Upon dissolution of the Single Cell 3' Gel Bead in a GEM, primers containing (i) an Illumina R1 sequence (read 1 sequencing primer), (ii) a 16 nucleotide 10× Barcode, (iii) a 10 nucleotide Unique Molecular Identifier (UMI), and (iv) a poly-dT primer sequence were released and mixed with cell lysate and Master Mix. Incubation of the GEMs then produced barcoded, full-length cDNA from polyadenylated mRNA. After

Table 1 Basic ALS donor demographics of the single-cell RNA-sequencing study

Donor ID	Age	Sex	Regions	SOD1mut	Race	Intake	Neuropath	BA4 TDP-43
ALS1	67	M	BA9, BA4, FN, SC	0	W	ALS	ALS	1
ALS2	89	F	BA9, BA4, FN, SC	1	W	ALS	PSP	0
ALS3	43	M	BA9, BA4, FN, SC	0	W	ALS	ALS	1
ALS4	33	F	BA9, BA4, FN, SC	1	W	ALS	ALS	0
ALS5	61	F	BA4, SC	0	W	ALS	ALS	1
ALS6	82	F	BA9, BA4, SC	1	W	ALS	ALS	0
ALS7	61	M	BA9, BA4, SC	0	W	ALS	ALS	1
ALS8	70	M	BA9, BA4, SC	0	W	ALS	ALS	1
ALS9	59	M	BA9, BA4, SC	0	W	ALS	ALS	1

Additional donor information can be found in Table S1. *ALS* amyotrophic lateral sclerosis; SOD1mut SOD1 mutation: carriers = 1, non-carriers = 0; *W* white, *PSP* progressive supranuclear palsy, Intake intake diagnosis; Neuropath final neuropathological diagnosis; BA4 TDP-43 TDP-43 pathological inclusions in neurons in Brodmann area 4 (primary motor cortex): 1 = present, 0 = absent

incubation, the GEMs were broken and the pooled fractions were recovered. Full-length, barcoded cDNA was then amplified by PCR to generate sufficient mass for library construction. Enzymatic fragmentation and size selection were used to optimize the cDNA amplicon size prior to library construction. R1 (read 1 primer sequence) were added to the molecules during GEM incubation. P5, P7, a sample index, and R2 (read 2 primer sequence) were added during library construction via end repair, A-tailing, adaptor ligation, and PCR. The final libraries contained the P5 and P7 primers used in Illumina bridge amplification. The described protocol produced Illumina-ready sequencing libraries. A Single Cell 3' Library comprises standard Illumina paired-end constructs which begin and end with P5 and P7. The Single Cell 3′ 16 bp 10 × Barcode and 10 bp UMI are encoded in Read 1, while Read 2 is used to sequence the cDNA fragment. Sample index sequences were incorporated as the i7 index read. Read 1 and Read 2 are standard Illumina sequencing primer sites used in paired-end sequencing. Sequencing the library produced a standard Illumina BCL data output folder. The BCL data includes the paired-end Read 1 (containing the 16 bp 10 × Barcode and 10 bp UMI) and Read 2 and the sample index in the i7 index read.

Batch structure and sequencing

Tissue specimens were processed upon receipt. The different central nervous system regions from the same donor were processed parallel. Samples from donors ALS1-ALS5 were loaded 1 region per 10 × Chromium channel. The different central nervous system region samples from donors ALS6-ALS9 were hashed, combined and loaded 1 donor per 10 × Chromium channel. Since each central nervous system region was processed from any given donor was processed parallel on the same day, each sample constitutes one batch for microglia isolation, cell capture and library construction. All sequencing was performed on an Illumina NovaSeq6000 machine. For specifics on sequencing and QC metrics regarding the generated reads see Table S1.

Single-cell RNA-seq data processing, alignment, and hashtag deconvolution

The majority of our downstream analysis was conducted using the R programming language (v4.0.5 for harmonization and clustering, v4.1.0 for annotation and downstream visualization) [42] and the RStudio [43] integrated development environment. CellRanger V3.1.0 with default parameters was used to demultiplex and align our barcoded reads with the Ensembl transcriptome annotation (downloaded March 2019, GRCh38.91). A recent report [38] suggested that filtering cells with greater than 10% mitochondrial reads is the preferred baseline for human tissue, and that for brain tissue a higher threshold may even be optimal. Thus, a mitochondrial percentage that was the higher of either 10% of reads or 2 absolute deviations above the median for mitochondrial reads within the sample was chosen as a threshold. Cells below this threshold with between 500 and 10000 UMIs were retained for downstream analysis. All ribosomal genes, mitochondrial genes, and pseudogenes were removed, as they interfered with the downstream differential gene expression. For samples where we used cell hashing to combine regions or subjects in a single sequencing run, droplets were demultiplexed using the following workflow. For each HTO, a mixture model with two components was fitted to the HTO counts using an EM algorithm. The component with the smaller mean (negative component) represents droplets that were not tagged with the HTO, whereas the component with the larger mean (positive components) represents droplets that were tagged. We then assign each droplet to either the negative or positive component based on its posterior probability. Droplets that were assigned to the negative component for all HTOs, as well as multiplets were discarded. Singlets with high uncertainty, i.e. without

confident assignment to either the negative or positive component, were discarded as well, leaving only high certainty singlets for downstream analysis. The method is implemented in the R package demuxmix, which is available on github: https://github.com/cu-ctcn/demuxmix. Some of our hashtag data had lower overall counts, and thus, the demuxmix model was unable to effectively segregate distributions for some hashtags in several samples. These samples were identified as having high percentages of negative/uncertain cells with demuxmix. In these cases, to try and recover cells for further analysis, the problematic hashtags were reclassified using one of two different algorithms, a demixing algorithm developed for MULTI-seq [34] or HTOdemux from Seurat v3.2.0 [50]. Hashtag classifications were merged, and doublet/negative/uncertain cell removal proceeded as before.

Label transfer to annotate ALS single-cell RNA-sequencing data

Cluster labels were assigned to the cells obtained from ALS donors, as follows:

Assigning non-microglial identities.

(1) Expression data for all cells from all donors were first normalized using the SCTransform algorithm in Seurat, then integrated using the Harmony R package with 30 PCs and theta = 5 and each donor as a batch, and subsequently clustered using 15 Harmony dimensions using the standard Louvain algorithm in the Seurat package (version 4) with resolution = 0.5.
(2) T-cells, B-cells, monocytes, red blood cells, and GFAP + clusters were identified using a combination of known gene markers, as well as the composition of clusters 10–14 in the cells from non-ALS donors, as described in Olah et al., 2020 [36].
(3) In this process, we also identified two additional clusters that were assigned cluster labels 15 and 16, since they did not correspond to any of the 14 clusters previously identified in Olah et al., 2020 [36].
(4) The remaining cells, which were putatively assigned microglial identity, were then assigned identities using a bootstrapped random forest approach, as described below.

Assigning microglial cluster identities.

(5) A training set of 200 cells randomly selected from each of the microglial clusters in Olah et al., 2020 [36] was constructed. For clusters with fewer than 200 cells, sampling was done with replacement.
(6) For each cluster (1–9), differential genes distinguishing that cluster from all other cells were identified using a Mann–Whitney test on the SCTransformed cells from the training set, with nominal p-value < 0.01
(7) All 9 gene sets from step 6 above were then combined (removing duplicates) to form the master gene set for random forest training.
(8) A random forest classifier was then constructed on the training set using the combined gene set in step 7, and subsequently run on all of the ALS donor cells, none of which were included in the training set.
(9) Steps 5–8 were run 20 times, with a different random seed each time.
(10) Cells from ALS cells were then assigned to the most commonly classified cluster identity (plurality voting) over the 20 runs.

After assignment of both the non-microglial and microglial cells, the entire data set was visualized using the same Harmony-based integration described in step 1 above, followed by standard UMAP implementation in Seurat (v4) with default parameters and 15 Harmony dimensions.

The final cluster labels thus included the 9 microglial and 5 non-microglial clusters from Olah et al., 2020 [36], which now included cells both from that original data set as well as the new cells from ALS donors, and 2 additional clusters that did not emerge in the clustering from our original data set.

10× chemistry correction

Striking differences were observed in the distributions of UMI counts between 10X v2 and v3 chemistry. As this was driving differential clustering, count matrices from v3 samples were down-sampled by 50% using the DropletUtils [29] package in R to achieve comparable UMI distributions across the two technologies. This was done prior to downstream visualization of gene expression and computation of differentially expressed genes between ALS and non-ALS microglia.

Identification of cluster-defining gene sets

To identify cluster-defining gene sets, the *FindMarkers* function in Seurat was used to implement a pairwise testing approach. We prioritized differentially expressed genes that could best delineate a given cluster from each other cluster in our dataset. To do so, MAST [11] was applied to normalized count data from the "RNA" assay of the Seurat object to find differentially expressed genes between every combination of pairs of clusters. Within each cluster, all the differentially expressed genes that were identified with this approach were filtered to only include those that were only found to be differentially expressed in one direction (either up or down). Any genes that were found to be upregulated

in comparison to some clusters but downregulated in comparison to other clusters or vice versa were removed from our downstream analysis. Furthermore, to ensure that the specific cluster-defining genes were prioritized, upregulated genes were ranked by the number of comparisons in which they were upregulated, and only those upregulated in 3 or more comparisons were used for downstream analyses. An identical process was applied for downregulated genes. Full marker gene lists are reported in Table S2.

Identification of differentially expressed genes between ALS and non-ALS microglia

To identify cluster-defining gene sets, each cluster was separated into ALS and non-ALS microglial cells, then the *FindMarkers* function in Seurat was used, leveraging MAST to identify differentially expressed genes between the ALS and non-ALS conditions. The down-sampled (see Sect. "10 × chemistry correction") normalized count data was used for this purpose to mitigate technical bias.

Constructing phylogenetic tree of microglial subtypes

To construct a phylogenetic tree evaluating the similarity of different microglial subtypes, the *BuildClusterTree* function in Seurat was used to build a phylogenetic tree on a distance matrix of "average" cells for each of the identity classes computed from the first 30 principal components. Results were visualized with the ggtree [62] package, as in Fig. 1D.

Functional annotation of microglial clusters

To perform functional annotation of microglial clusters, the top 50 up- or down-regulated genes in the differentially expressed gene lists for each cluster were taken for analysis. Annotation of these gene lists was performed with Reactome [20] pathway analysis using clusterProfiler [63]. For all functional analysis, the Benjamini–Hochberg correction [2] was used to correct *p*-values for multiple testing. Corrected *p*-values below a threshold of 0.05 were chosen as significant for Reactome results. This type of annotation was also performed on the ALS vs. non-ALS differentially expressed gene lists for each microglial cluster, as well as in the non-microglial immune cell analysis.

Monocle3 pseudotime analysis

As an orthogonal method of evaluating the continuity of different microglial states in our cluster structure, the monocle3 algorithm was used to build a pseudotime trajectory across our dataset as shown in Fig. 1E. Using the Seurat interface to monocle3 found in SeuratWrappers, the Seurat object was converted into a monocle data object, and a pseudotime trajectory was derived using the *learn_graph* function, retaining the final cell identity assignments from our original label transfer and clustering pipeline (see Sect. "Label transfer to annotate ALS single-cell RNA-sequencing data").To establish an originating point, the pseudotime root was placed on the border of clusters 1 and 2, as these subtypes were described in our prior paper [36] as being the two most prevalent microglial subtypes. Interestingly, this state was best captured by choosing cells with maximal *AVP* expression, a marker of hematopoietic stem cells [64] that is frequently used to mark the root cells in hematopoietic pseudotime tracing.

Identifying differentially represented subsets between ALS and non-ALS samples

To compare the relative abundance of different subsets between ALS and non-ALS samples, proportions of different cell types or subtypes were aggregated at the level of donor-region pairings. For example, ALS donor 1 would be represented by ALS1-Spinal Cord, ALS1-BA9, and ALS1-BA4. From here, each donor-region is treated as one sample for the purposes of comparing proportions of specific subtypes between ALS and non-ALS samples. A Wilcoxon rank-sum test was used to compute the significance of differences in the proportion of each subset between ALS and non-ALS regions, and BH correction was used to correct *p*-values for multiple testing.

Identifying differentially represented subsets between regions in ALS samples

Similar to above, samples were aggregated at the level of donor-regions and examined for differences between different groups. In this case, the generalization of the Wilcoxon, the Kruskal–Wallis test, was used to assess the significance of differences in proportion between different regions in ALS, and the BH correction was used to account for multiple testing.

Annotating single-cell microglial signatures in bulk RNA sequencing data from ALS donors

Bulk RNA-seq BAM files of the Target ALS Foundation were downloaded from the New York Genome Center (NYGC) in September 2020. These were 1063 BAM files of spinal cords and brains from 208 unique donors. BAM files were reverted to FASTQ files using the SamToFastq function of Picard tools (v2.17.4). RNA-seq reads were mapped onto the reference human genome GRCh38 using the STAR aligner (v2.5.3a) with 2-pass mapping mode. Gene expression levels were quantified using RSEM (v1.2.31) with

the Ensembl human gene model (release 91). Following samples were excluded at the quality control stage: one of samples sequenced twice, samples that have > 3% ribosomal RNA, > 40% duplication rate, < 4 RNA integrity number (RIN), and a sample whose library preparation method (manual vs. automated) was missing from metadata. Donors were retained if they were diagnosed as ALS or non-neurological control. Tissues that had less than 10 samples were also excluded. After these filtering steps, 913 samples from 170 unique donors were included for subsequent analysis. The voom function of the R package limma was used to compute $\log_2$ counts-per-million-mapped-reads (CPM) adjusted for RIN and library preparation methods. For each gene, $\log_2$CPM expression levels were further normalized into z-scores. Microglial signature scores were computed by averaging z-scores of $\log_2$CPM over member genes.

Computing pseudobulk PCA on single-cell microglial data

After down-sampling to mitigate technical bias, gene expression was aggregated at the donor-region level (i.e. summing all counts for cells derived from a specific donor-region combination). Principal components analysis was computed on this pseudo-bulk expression data, and the first two components were plotted.

Analyzing ALS spinal cord spatial transcriptomic data

Initial analysis of the spatial transcriptomic data from Maniatis et al. 2019 [30] showed infrequent expression of microglial genes and cluster markers due to the sparsity of the data. As such, the posterior lambda values that result from the use of the Splotch model were obtained from the authors of the original paper and used for further analysis. These posterior values were converted to predicted counts with the formula:

$$counts = lambda^{*}\ spot\_depth\ /\ median\_spot\_depth$$

where *median_spot_depth* was set to 1203. Samples with good quality *MAP2* signal were manually selected from the slide viewer at https://als-st.nygenome.org/. 26 sections were used in downstream analysis. Individual ST spots were assigned to specific spinal cord regions based on expert annotation from the original publication. Analysis of microglial signature genes was restricted to the ventral horn *versus* dorsal horn of the spinal cord, as the ventral horn (anatomical site of motor neuron cell bodies) is known to undergo higher levels of neuronal cell death in ALS compared to the dorsal horn (site of sensory neuron cell bodies) [59]. To compare levels of expression of genes between the two regions, predicted counts for all genes were normalized on a per-section basis. Next, for each gene, the summed score for all spots in the dorsal horn was compared to the summed score of all spots in the ventral horn for the same gene in the same section. Testing was conducted with Welch's t-test with the Holm-Bonferroni correction [17, 56], setting a threshold of 0.05 for significance.

Identifying regulators of different microglial subtypes with CHEA3

To identify regulators, we leveraged the recently published CHEA3 [22] tool using the web interface. We performed two primary analyses: comparison of different clusters, and comparison of ALS and non-ALS microglia within individual clusters. Gene sets for input were identified as above. For the ALS vs. non-ALS analysis, we used the top 50 genes per cluster that were upregulated in ALS. Conversely, for the cross-cluster analysis, we used curated gene lists that composed of signature genes for individual clusters. All analyses were conducted using only microglial clusters. Results of the analysis per cluster were sorted by "Score". For visualization, the top 15 predicted regulators per cluster were chosen. After removing duplicates, the Z-scored regulator scores per cluster were described in heatmaps. Clustering of regulators and clusters was done with hierarchical clustering using absolute linkage.

Annotation of non-microglial immune subtypes in ALS single-cell RNA-sequencing data.

Non-microglial immune subtypes (clusters 10, 11, and 12) were separated from other data and re-run through a clustering/dimensionality reduction pipeline. In brief, we used SCTransform with default parameter and harmony to integrate across batches using a thetaval of 2, and used 20 principal components for downstream analysis. To identify the optimal number of clusters for downstream analysis, we used the recently described ChooseR tool to identify an optimal clustering resolution for our dataset. Next, enrichment of different immune subtypes in ALS vs. non-ALS samples and in different regions was computed as described in sect. "Identifying differentially represented subsets between ALS and non-ALS samples" and "Identifying differentially represented subsets between regions in ALS samples". Next, cluster-defining gene subsets were identified using the approach described in Sect. "Identification of cluster-defining gene sets". Cluster identities were assigned with the use of canonical markers.

Statistical analysis and data visualization

Statistical analysis was conducted as described in the associated methods sections above. Specific *p*-values (both significant and not), if not found in the figures, may be found in the Supplementary Information tables before and after testing for multiple correction. *T*-values and degrees of freedom are also provided where relevant. Unless otherwise noted, all measurements are taken from distinct samples. In general, statistical methods were not used to re-calculate or predetermine sample sizes. All plots were created in R v4.1.0 using either base R visualization packages, ggplot2 [57] with ggrepel [47], ggfortify [52], patchwork [40], cowplot [58], and ggsci [60], or packages mentioned in the methods text. Heatmaps were made with the pheatmap [24] package. Volcano plots were made with the EnhancedVolcano [4] package. Figures were created using BioRender.com. All boxplots denote the 25th percentile, median, and 75th percentile, with whiskers representing 1.5 times the IQR in both directions. Outliers, if any, are represented as circles beyond the whiskers.

Immunohistochemistry and image analysis

Tissue acquisition, immunohistochemistry, imaging and image analysis were performed as follows. Formalin fixed paraffin embedded (FFPE) tissue sections of prefrontal cortex (Brodmann area 9, BA9), primary motor cortex (Brodmann area 4, BA4), motor nucleus of the facial nerve, and spinal cord were obtained from The Johns Hopkins ALS Postmortem Tissue Core and the Target ALS Postmortem Core (through Dr. Ostrow) and through the brain donation program of Dr. Shneider. Tissue was processed according to standard protocol. Briefly, glass slide mounted FFPE sections were de-paraffinized in CitrisSolv (Cat. No.: 1601H) for 20 min, subjected to decreasing ethanol series (ethanol 100%, ethanol 100%, ethanol 70%) for 1 min each, rehydrated in distilled water for 1 min, which was followed by 1 × phosphate buffered saline wash (PBS) (Cat. No.: 21–040-CM) for 5 min. Slides were then submerged in antigen retrieval low pH (10X) (Cat. No.: 00-4955-58) solution and heated using a water bath for 25 min. Slides were cooled in distilled water for 1 min and 1 × PBS for 5 min, each at room temperature, then dried using Kimwipes. A liquid-repellent slide marker pen (A-PAP Pen) was used to draw boundaries around each tissue section. The tissue was rehydrated using PBS and was washed 3 times at 5-min intervals in a humidified chamber. Unspecific antibody blocker solution consisting of 3% bovine serum albumin (BSA)(Cat. No.: B4287-25G) and 0.1% Triton X (Cat. No.: T8787-50ML) in PBS was added to each slide and incubated in a humidified chamber for 1 h. Primary antibody solution was prepared in a 1.7 mL Eppendorf tube by adding the primary antibodies to the BSA-containing buffer. Added 200 µL of primary antibody solution and incubated slides in the humidified chamber in a 4 °C cold room overnight. The next morning, the tissue sections were washed 3 times at 5-min intervals with 1 × PBS. Secondary antibody solutions consisted of Alexa Fluor Plus fluorochrome-conjugated secondary antibodies in their respective dilutions in the BSA Buffer. 200 µL of secondary antibody solution was added to each slide, which were then incubated in the humidified chamber for 1 h. Then the tissue sections were washed 3 times at 5-min intervals with 1 × PBS. Next, we added 200 µL of 70% ethanol and Sudan Black (Cat. No.: 23007) in a 1:20 dilution to tissue sections, incubating in the humidified chamber for 10 min to quench the endogenous autofluorescence. The tissue sections were washed 4 times at 5-min intervals with 1 × PBS, the slides were dipped in distilled water and mounted using 50 µL of ProLong Gold antifade reagent containing DAPI (Cat. No.: P36931). The slides were covered with microscope cover glass (Cat. No.: 12541042) and sealed with clear nail polish. The tissue sections were imaged on a Leica epifluorescence microscope with a 20 × objective. For each donor and brain region, 5 randomly selected, independent regions of interest (ROIs) were imaged. Images were analyzed using CellProfiler. The primary antibodies used were: goat anti human Iba1 (Wako, Cat.No. 011–27991), mouse anti human NeuN (Cell Signaling, Cat.No. 94403), rabbit anti human activated Caspase 3 (Cell Signaling, Cat.No. 9579S), rabbit anti human CD68 (Cell Signaling, Cat.No. 76437), rat anti human pTDP43 (Ser409/410) (Biolegend, Cat.No. 829901), anti human SON (Sigma-Aldrich, Cat.No. HPA023535). The secondary antibodies used were donkey anti-mouse IgG (H + L) Alexa Fluor Plus 555 (Invitrogen, Cat.No. A32773), donkey anti-goat IgG (H + L) Alexa Fluor Plus 488 (Invitrogen, Cat.No. A32814), donkey anti-rabbit IgG (H + L) Alexa Fluor Plus 647 (Invitrogen, Cat.No. A32795), Alexa Fluor 488 donkey anti-rat IgG (H + L) (Invitrogen, Cat.No. A21208).

Results

Establishing a microglia population structure in human ALS brain and spinal cord

To profile the microglial phenotypes present in the CNS of sporadic ALS in an unbiased fashion, we utilized our optimized pipeline for the cold, non-enzymatic isolation and single-cell RNA-sequencing (scRNA-seq) of live microglia and non-microglial immune cells (all $CD45^+$) from autopsy brain and spinal cord tissue[37]. This approach has been shown to be superior to other single-cell and single-nucleus RNA-sequencing approaches in terms of capturing the full extent of microglia phenotypes without inducing a stress

response signature [31, 53]. To have a good coverage of anatomical sites along the neuraxis, we sampled multiple CNS regions from sporadic ALS donors of both sexes: prefrontal cortex (BA9), primary motor cortex (BA4), motor nucleus of the facial nerve (FN), and lumbar spinal cord (SC) (Fig. 1A, Table 1 and Table S1). In situ each donor showed classical signs of microglia activation in the affected CNS regions (BA4, FN, SC), such as increased CD68 expression in microglia and reduced ramified microglia morphology, as compared to brain region that was not affected by ALS pathology (BA9) (Figure S1a, b and Figure S2). Pathological phosphorylation of TDP43 at Ser409/410 was confirmed in all donors in BA4, FN and SC regions using immunohistochemistry and fluorescence microscopy (Figure S3). While we found evidence of caspase 3 activation in neurons is some of the donors, neuronal apoptosis was not a general and widespread feature of these end-stage ALS tissue samples (Figure S4). Using our fresh autopsy tissue-derived live cell-based single-cell RNA-sequencing and analytical pipeline, which includes stringent quality control (see Table S1), our dataset contained > 51000 single-cell transcriptomes of microglia and non-microglial immune cells. To facilitate comparison to the microglia phenotypes that we recently identified in aged and young donors [36], we applied a label transfer to assign the ALS-associated immune cells to the clusters that we had defined previously (microglial clusters 1–9 and non-microglial clusters 10–14) (Fig. 1b and Figure S5a and b). Each microglial cluster is marked by distinct genes that robustly distinguish them from one another (Fig. 1c, Table S2). The identified microglial subsets are also characterized by non-overlapping functional annotations (Fig. 1d) and reside along divergent state transition trajectories (Fig. 1d and 1e), highlighting their unique, non-overlapping identities.

ALS is characterized by robust, region-specific shifts in microglial subtype prevalence

We first aimed to identify the global changes in microglia subset relative abundance in ALS, using our published dataset of young and old non-ALS microglia [36] as a reference. This reference dataset contains scRNA-seq data of microglia isolated from the temporal neocortex and prefrontal cortex of young ($N = 3$, mean age = 37.67 years; SD = 19.86 years) and aged ($N = 14$, mean age = 92.07 years; SD = 5.31 years) donors, respectively (Fig. 1a). All "non-ALS" designation in the figures and the Sect. "Results" and Sect. "Discussion" sections below refer to this reference dataset published previously by our group [36]. We observed robust, statistically significant shifts in the population structure of microglia in ALS (Fig. 2a, Fig. 2b, Figure S5c, Table S3). These changes are dominated by an inverse relationship between the two homeostatic microglia subsets MG1 and MG2. While MG1 is the most abundant microglial cluster in non-ALS samples, MG2 is the predominant microglia cluster in ALS. Furthermore, ALS samples also had higher representation of stress-related microglia cluster MG3, and lower proportions of the MHC-pathway enriched microglial cluster MG7. The higher prevalence of MG2 in ALS was prominent in all CNS regions examined but was less pronounced in the spinal cord (Fig. 2b, Figure S6b, c—per donor brakedown can be seen in Figure S1c). In contrast, the enrichment of MG3 was more dramatic in ALS spinal cord samples and less so in ALS brain samples (Fig. 2b, Figure S6c), while the microglia subset associated with respiratory electron transport (MG8) showed a trend to be enriched in ALS, especially in the subcortical samples. The non-microglial clusters 11 (Fig. 2b), identified as T cells, were also found to be significantly enriched in ALS.

We next performed orthogonal validation of our findings in an independent bulk RNA-seq dataset consisting of 170 samples from ALS patients, non-ALS neurological disease patients, and control donors [51]. Signatures of the top 20 genes per microglial cluster (Table S4) were used to delineate the enrichment of different cluster signatures in ALS versus non-ALS samples. This investigation confirmed that MG2-enriched genes had significantly higher expression in ALS motor cortex and cervical spinal cord when compared to controls (Fig. 2c). Interestingly, MG1- and MG7-enriched genes also showed higher expression in the bulk RNA-sequencing dataset, but not in our scRNA-seq dataset. This may be because signature gene sets of MG1 and MG7 are particularly enriched in pan-microglial marker genes expressed by all microglia (Figure S7). Consequently, the enrichment of the MG1 and MG7 signatures in the bulk RNA-seq data likely reflects the expansion of the microglial population in the cervical spinal cord of ALS patients; a phenomenon that has been documented previously [18]. Interestingly, relative abundance of microglia (detected in the bulk RNA-seq dataset through the signatures of MG1 and MG7) positively correlated with family history of ALS as well as disease duration (Figure S8b and d), but not with sex, C9orf72 repeat expansion status, age at death, or age at onset (Figure S8a, c, e and f, respectively). MG2 was significantly positively associated with a positive family history of ALS, as were MG8 and MG9 and the monocytic cluster 10 (Figure S8b). Intriguingly, MG8 was negatively associated with disease duration along with MG4 (Figure S8d) and was significantly enriched in donors who had rapidly progressing disease (less then 2 years survival following diagnosis).

To further validate our scRNA-seq findings, we examined the spatial expression pattern of microglia subset-specific genes in an independent spatial transcriptomic dataset of ALS spinal cord[30]. Specifically, we investigated the differences in the expression levels of microglia cluster markers between the dorsal horn, which harbors intact neurons, and

the ventral horn of the spinal cord, the site of motor neuron cell death. This dataset reliably showed reduced gene expression of MAP2, a neuronal marker, in the ventral horn of ALS spinal cord when compared to the dorsal horn (Fig. 2d, e, f, Table S5), consistent with extensive motor neuron demise in this area. Importantly, marker genes of MG2 (*STAB1*), MG3 (*CXCR4*) and MG8 (*PYCARD*) were inversely correlated with MAP2 expression in the dorsal and ventral horns of ALS spinal cord (Fig. 2e and 2f), suggesting their association with motor neuron death in the ventral horn of ALS spinal cord. Patterns of expression for other microglial subset-specific genes in the spatial transcriptomic dataset are shown in Figure S8**j**.

ALS results in functional changes in microglia subsets

A central question in the analysis of microglial single-cell transcriptomics signatures is the degree to which the identified subsets are functionally distinct. We first established the non-overlapping functional identity of our microglial subsets using their unique signature gene sets and the REACTOME pathway analysis tool (Fig. 1d, and Figures S9a and b, Table S6). We have previously reported that MG1 and MG2 are both homeostatic/undifferentiated microglia subsets [36]. In the combined dataset, they maintained this annotation, with additional enrichment of 'IRE1-a activated chaperons' in MG1 and 'RHO GTPase cycle' in MG2. Similarly, the other microglia subsets also maintained their top functional annotation in the combined dataset: 'Cellular response to stress' for MG3, 'Interferon Signaling' for MG4, 'Interleukin Signaling' for MG5 & MG6, 'MHCII Antigen Presentation' for MG7, 'Respiratory Electron Transport' for MG8 and 'Cell Cycle' for MG9 (Fig. 1d and Figure S9a and b).

To better understand microglial phenotypic changes in ALS, we used REACTOME for functional annotation [20]. First, we investigated the within-cluster transcriptional changes between ALS and non-ALS samples (Fig. 3a**,** Table S7). The REACTOME terms that were upregulated within each cluster included pathways such as: 'Chaperone mediated autophagy' (upregulated in MG3 and MG9), 'Role of phospholipids in phagocytosis' (upregulated in MG1, MG4 and MG7), 'Integration of energy metabolism' (MG1, MG4, MG7 and MG8) and 'G alpha (i) signaling events' (upregulated in MG8) (Fig. 3a). Most of these terms were upregulated in more than one microglia subset (e.g. 'Interleukin-4 and Interleukin-13 signaling' in MG1, MG3, MG4 and MG8), suggesting more global phenotypic changes in response to the ALS microenvironment. However, no responses were common to every cluster, suggesting that even in the face of a relatively similar microenvironment, the different microglial clusters may have different adaptive responses. Similarly, among the downregulated pathways, 'Neutrophil degranulation' was shared between 6 independent microglia clusters (MG1, MG2, MG3, MG5, MG6 and MG9), while others were shared between three microglia subsets (e.g. 'Interferon gamma signaling' in MG2, MG6 and MG9) or two subsets (e.g. 'NGF-stimulated transcription' in MG1 and MG5). On the other hand, downregulation of the 'Respiratory electron transport chain' pathway was only observed in MG2 while 'Regulation of actin dynamics for phagocytic cup formation' was specific to MG3 (Fig. 3b). These results suggest microglia subset-specific metabolic and functional shifts that occur in the ALS microenvironment.

Next, we aimed to identify the regulators of gene expression governing microglial subset-specific and ALS-induced transcriptional programs. Using ChEA3 [22], we first predicted transcriptional regulators for the individual microglia subsets (right panel of Fig. 3c, Table S8). As proof-of-concept, MG4, which showed strong enrichment in the 'Interferon Signaling' pathway (and is accordingly also known as interferon response microglia or IRM phenotype), was predicted to be regulated by the transcription factors STAT1, STAT2, IRF1 and IRF7, while the stress-associated MG3 showed regulation by JUNB and JUND. ZNF888 seems to be a specific regulator of the MG1 transcriptional identity, BPTF, MYSM1 and MGA were specific to MG2, NR4A3 to MG5, GTF3A appeared to be specific to MG8, while E2F1 and CENPA target genes were enriched in MG9. Many of the transcriptional regulators of MG7 are shared with other clusters, especially with clusters MG1 and MG4, suggesting shared ontology for these microglia subsets (Fig. 1d). In contrast, most of the ALS-associated gene expression changes within clusters were regulated by SON, a member of the nuclear speckle pre-mRNA processing machinery [28] (left panel of Fig. 3c, Table S9). One exception was MG8, which had very few transcriptional regulators that were shared with other clusters in this comparison (Fig. 3c, left panel).

Given that one of the enriched genes in microglia Cluster 2, that was significantly more abundant in the ALS samples compared to our previous study, was SON (Figure S10a), we found it intriguing that SON also came up as a prominent regulator of transcriptional profiles of ALS microglia subsets in the CHEA analysis (Fig. 3c). Accordingly, we have decided to utilize SON as a marker gene for Cluster 2 microglia to confirm in situ the changes in Cluster 2 distribution in ALS samples. For this end we utilized tissue sections from prefrontal cortex (FCtx, BA9), primary motor cortex (MCtx, BA4) and spinal cord (SC) of 3 ALS and 3 non-neurological control donors from the Target ALS collection (see Table S1for donor demographics), and performed immunohistochemistry for Iba1 and SON on them, followed by fluorescence microscopy and automated image analysis using CellProfiler. In all investigated brain regions and donors, protein expression was nuclear and punctate

in nature (Figure S10b). Most cells expressed SON, but there was a clear distribution of SON expression ranging from no expression to high expression in all cells, including microglia (Figure S10b arrows). To investigate the distribution of SON expression in microglia, we measured Son expression in each Iba1 + cell using CellProfiler. We found a robust change in the distribution of SON expression in microglia in the ALS cortical samples (Figure S10c), compared to controls. This bimodal distribution of SON protein expression in the nucleus of ALS cortical microglia is in line with our findings in single-cell RNA-sequencing dataset, where Cluster 2 microglia (most abundant in ALS) had more pronounced SOD mRNA expression, when compared to all other clusters (Figure S10a). These changes were not so pronounced in the spinal cord (Figure S10c), in line with Cluster 2 not being the dominant cluster in the spinal cord samples of our ALS single-cell RNA-sequencing dataset.

Independent clustering of non-microglial immune cells identifies shifts in myeloid and adaptive immune cell populations in ALS

Since our dataset included other non-microglial immune cells (all $CD45^+$ cells) isolated from CNS tissue, we performed additional de novo clustering exclusively on these cells to better understand their identity and relevance in ALS. We identified 19 non-microglial immune cell clusters that could be broadly classified as T cells, NK cells, macrophages, monocytes, B cells, and dendritic cells (Fig. 4a, Table S10). These peripheral immune cells profiled from the brain parenchyma had non-overlapping signature gene sets (Fig. 4b, Table S11) and showed divergent enrichment in ALS (Fig. 4c-h, Table S12). Interestingly, the macrophage cluster C2 was depleted from most ALS CNS regions, except for BA4, while the stressed T cells (C3), T cells (C4), maturing NK cells (C5), NK cells (C8) and dendritic cells clusters (C10) were robustly enriched in ALS samples (Fig. 4c-h). CD8 + (C7) and CD4 + T cells (C9) were also enriched in ALS when compared to Alzheimer's' disease (AD) and mild cognitive impairment (MCI) samples, but not when compared to samples from young donors (Figure S11). These findings were further corroborated by mapping the non-microglia immune cells to the Azimuth reference atlas[15] and examining the relative abundance of the different non-microglial immune cell subsets in our dataset (Figure S12a–h and Figure S13, Table S13). Similar to the de novo clustering, CD4 positive central memory T cells (T_{CM}s) and CD8 T_{CM}s identified through this reference mapping approach were also relatively enriched in ALS samples, although a substantial number of CD4 T_{CM}s were observed in TLE samples as well (Figure S12c and e). The reference mapping approach also revealed the enrichment of γδT cells in ALS samples as well as the depletion of plasmablasts in ALS (Figure S12d and g, respectively).

Discussion

In this study, we present a comprehensive assessment of human microglia phenotypes in ALS using scRNA-seq of live microglia. Our study by design incorporates a variety of ALS donors in terms of age, sex, TDP43 and genetic status. We designed our experimental set up as such with the aim to capture a common theme in terms of microglia population structure in ALS—one that transcends other factors that can affect the relative abundance of microglia subsets in this disease. This approach carried the risk of not finding such an overarching pattern of ALS-specific microglia population structure in the set of donors included in this study. Importantly, this was not the case, as all donors showed very similar patterns of microglia subset relative abundances (Figure S1c), that was significantly different from the microglia population structure we observed in young brains and in aged donors with MCI and AD [36]. Future, population-level studies will be able to address in a systematic way the modifying effect of covariates on microglia subset relative abundances in ALS. Nonetheless, in our experience, these effects will likely be smaller in size, compared to the ALS-specific changes reported here.

One characteristic of studies requiring fresh human CNS specimens is the lack of control samples as they are, understandably, very difficult to procure. To circumvent this drawback, we utilized a variety of approaches to best describe the microglia phenotypes in ALS. First, to maximize comparability with our previous study [36], we mapped the transcriptomes from the ALS samples onto previously established microglia subset identities. This allowed us to directly compare shifts in microglia subset relative abundances between young, aged MCI and AD and ALS brains. Second, we explored the disease associations of our microglia subset signatures in a large bulk transcriptomic dataset (from the NYGC), that contained samples from both ALS donors as well as non-neurological controls, and finally, our in situ confirmation pipeline also followed the conventional case–control setup. This multi-pronged approach allowed us to establish the microglia population structure in ALS and confirm the disease relevance of the observed shifts in the relative abundance of ALS-associated microglia subsets.

As it is common with studies utilizing postmortem CNS tissue samples of fast-progressing human diseases, the disease stage of all of the donors at death was terminal. Consequently, the cause of death for all (but one) of the donors was end-stage ALS. One donor (ALS2, Table 1 and Table S1) was diagnosed with ALS at the clinic, however, neuropathological work-up revealed a primary taupathy, progressive

supranuclear palsy (PSP), to be the dominant neuropathology. Despite the final PSP diagnosis, we have decided to keep this donor in the dataset, given that 1) it carried a SOD1 mutation, 2) while in this donor BA4 was devoid of pathological TDP43 inclusions in neurons, we found evidence that the motor nucleus of the facial nerve (FN) and the lumbar spinal cord (SC) did contain neurons with pTDP43 inclusions (Figure S3), and 3) microglia activation was present in FN as well as SC, as assessed by classical microglia activation markers in situ (CD68 expression and microglia morphology, Figure S1a and b, respectively)—suggesting complex, unconventional neuropathology in this donor, that is worth exploring further in a separate study.

Importantly, we document significant deviation from the observed changes of the microglia population structure in mouse models of this disease. To our surprise, MG7, the human microglia cluster that most closely resembled the murine disease-associated microglia (DAM) phenotype [36] that was found to be enriched at every symptomatic stage in the mouse SOD1 model [7, 23], was significantly depleted in human ALS, similarly to what we observed in AD [36]. Our findings suggest that the depletion of the putatively protective MG7 microglia phenotype might be a common feature of late-stage neurological diseases in humans and highlight the limitations of mouse models to study microglia phenotypes associated with age-related neurodegeneration due to species differences in microglia aging [13, 48, 49]. Additionally, here we document a robust, region-specific reorganization of microglia population structure in this disease. In contrast to AD, MCI, and temporal lobe epilepsy (TLE), the dominant microglia phenotype in ALS brain and spinal cord was MG2, a homeostatic/undifferentiated microglia subset characterized by a profound restructuring of the oxidative phosphorylation machinery, through the downregulation of several electron transport chain subunits. Using an independent bulk tissue transcriptomic dataset [18, 51] we confirmed the significant enrichment of the MG2 expression signature in ALS motor cortex and spinal cord, when compared to healthy controls. Importantly, dysregulated microglia energy metabolism has been shown previously to be a dominant feature of microglia in the motor cortex of ALS donors compared to healthy controls in a study that used pseudobulk assessment of microglia snRNA-seq data without identifying microglia subsets [45]. In our dataset, a prominent feature of the ALS spinal cord was the strong enrichment of the stress- and early-response associated microglia cluster, MG3, in this CNS region. In an independent spatial transcriptomic dataset [30] gene signatures of both MG2 and MG3 showed an inverse relationship with the expression of MAP2, a neuronal marker depleted from ALS spinal cord ventral horn, suggesting that these subsets may localize specifically to areas of motor neuron demise. While MG2 and MG3 signatures did not show an association with disease duration, two smaller microglia subsets, MG4 and MG8, were found to be significantly enriched in the spinal cord of donors with aggressive disease progression. Interferon response, the functional annotation of the MG4 microglia cluster, has been previously shown to be the dominant pathogenic feature of C9orf72-/- myeloid cells and microglia [25, 33]. We also found MG8 to be enriched in the ALS brain and spinal cord in our dataset. This small microglia cluster was a mirror image of MG2 in terms of functional annotation (upregulated genes in MG2 were downregulated in MG8, and vice versa). Intriguingly, signature markers of the ALS enriched protein co-expression module [55] were expressed at the highest level in MG8 in our dataset, as were many of the known interacting partners of TDP-43 [12], including SOD1 and CHCHD10; both implicated in familial forms of ALS [3, 21] (Figure S14). Furthermore, the novel ALS cerebrospinal fluid biomarker candidate CAPG [35] was most prominently expressed in MG8 among the different microglia clusters (Figure S14). Importantly, none of these transcriptomic microglia phenotypes could be deduced from bulk tissue ALS RNA sequencing data [9, 18, 51], nor were they identified in smaller scale ALS single-nucleus RNA sequencing datasets [26, 27, 45, 46], confirming the importance of ex-vivo (whole cell) microglia profiling for the study of transcriptional heterogeneity in this cell type [53]. With the exception of MG4 (also known as IRM), none of the ALS relevant microglia phenotypes (MG2, MG8) identified here through unsupervised clustering have been previously identified as in independent microglia subset in either mouse or human [39].

An examination of non-microglial immune cells in ALS revealed a preponderance of T cells, as well as relative enrichment of dendritic cells and NK cells in the ALS spinal cord. Interestingly, annotation of this subset of our dataset with the Azimuth PBMC reference [15] highlights enrichment of γδT cells in ALS, which have been reported to be increased in the peripheral blood of ALS patients, but have not been explored yet in the CNS [14]. Our findings open novel avenues for further investigations into the role of the different infiltrating peripheral immune cells in the ALS spinal cord to understand whether they are contributing to the pathobiology directly or through their interaction with local microglia.

Despite the challenges associated with ex-vivo studies of human microglia in ALS, we believe that the approach described here is crucial to understanding of their role in ALS disease pathobiology. Most of the ALS-enriched human microglia phenotypes presented here (e.g. MG2, MG8) have not been recapitulated in murine or in vitro model systems yet. Accordingly, mechanistic studies addressing their role in ALS models are challenging at this moment. Nonetheless, our data set will enable the establishment of translational studies that aim to generate

novel in vitro and preclinical model systems in which drug screens targeting specific microglia subsets and/or subset-specific gene/pathways can be performed to halt or slow down ALS disease progression.

Our dataset offers multiple such potential starting points to pursue for the development of therapeutic strategies targeting the neuroimmune component of ALS. To demonstrate the utility of our dataset and provide examples of potential genes/pathways for follow-up, we explored whether these microglia phenotype-specific marker gene sets (Table S4) contain any druggable targets using dgidb.org. For this investigation, we focused on two microglia subsets, that were either enriched in ALS (MG2) or had a significant association with disease duration (MG4 and MG8). This inquiry revealed that MG2, the metabolically dysregulated microglia subset that was the most enriched in ALS samples and potentially represents a "burned-out" phenotype, has several druggable targets (with approved inhibitors) among its marker genes—including IL17RA. Finding IL17RA as a marker gene of the most enriched microglia cluster in ALS was intriguing. The IL-17 axis is yet unexplored as a potential therapeutic target in the context of ALS, despite the documented elevated levels of IL-17A in the serum of sporadic ALS patients [10] and the pro-inflammatory effect of IL-17 on microglia [19]. The availability of an approved inhibitor for IL17RA makes this target specifically noteworthy for devising therapeutic strategies to modulate MG2 in ALS. Interferon response transcriptomic signatures have been identified to contribute to early neuroinflammatory changes in animal models of ALS [16]. Accordingly, we also found that the interferon response microglia subset MG4 showed strong enrichment in ALS donors with an aggressive disease course. One of the promising druggable targets for MG4 was BST2. BST2 has been suggested as a biomarker for microglia activation in ALS, as it was found to be upregulated in an ALS mouse model even in the presymptomatic stages of the disease [61]. Future studies will establish whether BST2 (and other druggable markers of MG4, such as ISG15) show similar disease associations in the CSF and serum of human ALS patients. In our study, the microglia subset MG8, which was enriched in translation and inflammasome-related genes, was specifically abundant in donors with an aggressive disease course. Intriguingly, one of the druggable targets (with available inhibitor) for this subset was GPNMB. GPNMB has been recently identified as a biomarker for ALS, and was found to be most elevated in the CSF of short-lived ALS patients [65]—in line with our own findings presented here. While the exact biological role of GPNMB in microglia is not yet known, accumulating evidence suggests that it is an important potential therapeutic target, not only in ALS but also in other neurodegenerative diseases and CNS malignancies [44]. Given their disease association and therapeutic potential, further investigations into the cell biology of MG2, MG4 and MG8 microglia phenotypes and their contribution to disease pathogenesis in ALS are warranted.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00401-025-02913-3.

Acknowledgements We are grateful to the donors and their families who made this study possible with the donation of brain and spinal cord tissue at autopsy. Some of the experiments (cell sorting) reported in this publication were performed in the CCTI Flow Cytometry Core, supported in part by the Office of the Director, National Institutes of Health under awards S10OD020056. Single-cell RNA sequencing was performed at the Columbia University Medical Center Genome Center Single Cell Analysis Core, which is funded in part through the NIH/NCI Cancer Center Support Grant P30CA013696 as well as by the National Center for Advancing Translational Sciences/NIH through grant number UL1TR001873. SP is supported by NS117583 and NS107442. Data acquisition and analysis were also supported by a Chan-Zuckerberg Neurodegeneration Challenge Network grant (CS-02018-191971) and NIH RF1 AG057473, P30 AG066462, and U01 AG061356.

Author contribution Conceptualization: J.F.T., M.O.; data curation: J.F.T., M.F., S.M., A.D., V.M., M.O.; formal analysis: J.F.T., M.F., V.M., M.O.; funding acquisition: N.S., H.P., P.L.D., S.P., M.O.; investigation: J.F.T., A.K., C.H., M.O.; methodology: J.F.T., A.K., C.H., M.O.; project administration: M.O.; resources: N.S., A.F.T., H.P., P.S., J.A.S., D.A.B., P.L.D.; software: J.F.T., V.M.; supervision: V.M., O.M.; validation: J.F.T., V.M., M.O.; visualization: J.F.T., M.F.; writing – original draft: J.F.T. & M.O.; writing – review & editing: all authors.

Funding National Institute of Health,USA,NS107442,AG057473,NS117583,Chan-Zuckerberg Neurodegeneration Challenge Network,CS-02018-191971

Data availability Data Availability. Raw scRNA-seq data (fastq files) generated from CD45+ cells isolated from ALS autopsy samples were deposited to GEO (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE204704. It is also available at Synapse, with ID number syn61001870, as part of large series [54]. Table S1 contains the GUID (NeuroGUID, https://ncri1.partners.org/GUID/) information for all donors (including the ones used for scRNA-seq and immunohistochemistry), based on which further clinical and neuropathological data can be requested.

Code availability Code used to perform preprocessing, clustering, cluster validation, and label transfer of scRNA-seq data in the current study are available publicly at https://github.com/jtuddenham/ALS_sc_microglia.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Ahmed RM, Devenney EM, Strikwerda-Brown C, Hodges JR, Piguet O, Kiernan MC (2020) Phenotypic variability in ALS-FTD and effect on survival. Neurology 94:e2005–e2013. https://doi.org/10.1212/WNL.0000000000009398
2. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Statistical Soc 57(1):289–300
3. Berdynski M, Miszta P, Safranow K, Andersen PM, Morita M, Filipek S et al (2022) SOD1 mutations associated with amyotrophic lateral sclerosis analysis of variant severity. Sci Rep 12:103. https://doi.org/10.1038/s41598-021-03891-8
4. Blighe K, Rana S, Lewis M (2022) EnhancedVolcano: publication-ready volcano plots with enhanced colouring and labeling
5. Brettschneider J, Libon DJ, Toledo JB, Xie SX, McCluskey L, Elman L et al (2012) Microglial activation and TDP-43 pathology correlate with executive dysfunction in amyotrophic lateral sclerosis. Acta Neuropathol 123:395–407. https://doi.org/10.1007/s00401-011-0932-x
6. Brettschneider J, Toledo JB, Van Deerlin VM, Elman L, McCluskey L, Lee VM et al (2012) Microglial activation correlates with disease progression and upper motor neuron clinical symptoms in amyotrophic lateral sclerosis. PLoS ONE 7:e39216. https://doi.org/10.1371/journal.pone.0039216
7. Chiu IM, Morimoto ET, Goodarzi H, Liao JT, O'Keeffe S, Phatnani HP et al (2013) A neurodegeneration-specific gene-expression signature of acutely isolated microglia from an amyotrophic lateral sclerosis mouse model. Cell Rep 4:385–401. https://doi.org/10.1016/j.celrep.2013.06.018
8. Corcia P, Tauber C, Vercoullie J, Arlicot N, Prunier C, Praline J et al (2012) Molecular imaging of microglial activation in amyotrophic lateral sclerosis. PLoS ONE 7:e52941. https://doi.org/10.1371/journal.pone.0052941
9. Eshima J, O'Connor SA, Marschall E, Consortium NA, Bowser R, Plaisier CL et al (2023) Molecular subtypes of ALS are associated with differences in patient prognosis. Nat Commun 14:95. https://doi.org/10.1038/s41467-022-35494-w
10. Fiala M, Chattopadhay M, La Cava A, Tse E, Liu G, Lourenco E et al (2010) IL-17A is increased in the serum and in spinal cord CD8 and mast cells of ALS patients. J Neuroinflammation 7:76. https://doi.org/10.1186/1742-2094-7-76
11. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK et al (2015) MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol 16:278. https://doi.org/10.1186/s13059-015-0844-5
12. Freibaum BD, Chitta RK, High AA, Taylor JP (2010) Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. J Proteome Res 9:1104–1120. https://doi.org/10.1021/pr901076y
13. Galatro TF, Holtman IR, Lerario AM, Vainchtein ID, Brouwer N, Sola PR et al (2017) Transcriptomic analysis of purified human cortical microglia reveals age-associated changes. Nat Neurosci 20:1162–1171. https://doi.org/10.1038/nn.4597
14. Gustafson MP, Staff NP, Bornschlegl S, Butler GW, Maas ML, Kazamel M et al (2017) Comprehensive immune profiling reveals substantial immune system alterations in a subset of patients with amyotrophic lateral sclerosis. PLoS ONE 12:e0182002. https://doi.org/10.1371/journal.pone.0182002
15. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A et al (2021) Integrated analysis of multimodal single-cell data. Cell 184(3573–3587):e3529. https://doi.org/10.1016/j.cell.2021.04.048
16. Hiew JY, Lim YS, Liu H, Ng CS (2025) Integrated transcriptomic profiling reveals a STING-mediated Type II interferon signature in SOD1-mutant amyotrophic lateral sclerosis models. Commun Biol 8:347. https://doi.org/10.1038/s42003-025-07790-w
17. Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6:65–70
18. Humphrey J, Venkatesh S, Hasan R, Herb JT, de Paiva LK, Kucukali F et al (2023) Integrative transcriptomic analysis of the amyotrophic lateral sclerosis spinal cord implicates glial activation and suggests new risk genes. Nat Neurosci 26:150–162. https://doi.org/10.1038/s41593-022-01205-3
19. Iitani Y, Miki R, Imai K, Fuma K, Ushida T, Tano S et al (2024) Interleukin-17A stimulation induces alterations in Microglial microRNA expression profiles. Pediatr Res 95:167–173. https://doi.org/10.1038/s41390-023-02825-6
20. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A et al (2020) The reactome pathway knowledgebase. Nucleic Acids Res 48:D498–D503. https://doi.org/10.1093/nar/gkz1031
21. Johnson JO, Glynn SM, Gibbs JR, Nalls MA, Sabatelli M, Restagno G et al (2014) Mutations in the CHCHD10 gene are a common cause of familial amyotrophic lateral sclerosis. Brain 137:e311. https://doi.org/10.1093/brain/awu265
22. Keenan AB, Torre D, Lachmann A, Leong AK, Wojciechowicz ML, Utti V et al (2019) ChEA3: transcription factor enrichment analysis by orthogonal omics integration. Nucleic Acids Res 47:W212–W224. https://doi.org/10.1093/nar/gkz446
23. Keren-Shaul H, Spinrad A, Weiner A, Matcovitch-Natan O, Dvir-Szternfeld R, Ulland TK et al (2017) A unique microglia type associated with restricting development of Alzheimer's disease. Cell 169(1276–1290):e1217. https://doi.org/10.1016/j.cell.2017.05.018
24. Kolde R (2019) pheatmap: pretty heatmaps. R package version 1.0.12. https://CRAN.R-project.org/package=pheatmap
25. Lall D, Lorenzini I, Mota TA, Bell S, Mahan TE, Ulrich JD et al (2021) C9orf72 deficiency promotes microglial-mediated synaptic loss in aging and amyloid accumulation. Neuron 109(2275–2291):e2278. https://doi.org/10.1016/j.neuron.2021.05.020
26. Li J, Jaiswal MK, Chien JF, Kozlenkov A, Jung J, Zhou P et al (2023) Divergent single cell transcriptome and epigenome alterations in ALS and FTD patients with C9orf72 mutation. Nat Commun 14:5714. https://doi.org/10.1038/s41467-023-41033-y
27. Limone F, Mordes DA, Couto A, Joseph BJ, Mitchell JM, Therrien M et al (2024) Single-nucleus sequencing reveals enriched expression of genetic risk factors in extratelencephalic neurons sensitive to degeneration in ALS. Nat Aging 4:984–997. https://doi.org/10.1038/s43587-024-00640-0
28. Lu X, Ng HH, Bubulya PA (2014) The role of SON in splicing, development, and disease. Wiley Interdiscip Rev RNA 5:637–646. https://doi.org/10.1002/wrna.1235
29. Lun ATL, Riesenfeld S, Andrews T, Dao TP, Gomes T, participants in the 1st Human Cell Atlas J et al (2019) EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. Genome Biol 20:63. https://doi.org/10.1186/s13059-019-1662-y
30. Maniatis S, Aijo T, Vickovic S, Braine C, Kang K, Mollbrink A et al (2019) Spatiotemporal dynamics of molecular pathology in amyotrophic lateral sclerosis. Science 364:89–93. https://doi.org/10.1126/science.aav9776
31. Marsh SE, Walker AJ, Kamath T, Dissing-Olesen L, Hammond TR, de Soysa TY et al (2022) Dissection of artifactual and

confounding glial signatures by single-cell sequencing of mouse and human brain. Nat Neurosci 25:306–316. https://doi.org/10.1038/s41593-022-01022-8
32. McCauley ME, Baloh RH (2019) Inflammation in ALS/FTD pathogenesis. Acta Neuropathol 137:715–730. https://doi.org/10.1007/s00401-018-1933-9
33. McCauley ME, O'Rourke JG, Yanez A, Markman JL, Ho R, Wang X et al (2020) C9orf72 in myeloid cells suppresses STING-induced inflammation. Nature 585:96–101. https://doi.org/10.1038/s41586-020-2625-x
34. McGinnis CS, Patterson DM, Winkler J, Conrad DN, Hein MY, Srivastava V et al (2019) MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. Nat Methods 16:619–626. https://doi.org/10.1038/s41592-019-0433-8
35. Oeckl P, Weydt P, Thal DR, Weishaupt JH, Ludolph AC, Otto M (2020) Proteomics in cerebrospinal fluid and spinal cord suggests UCHL1, MAP2 and GPNMB as biomarkers and underpins importance of transcriptional pathways in amyotrophic lateral sclerosis. Acta Neuropathol 139:119–134. https://doi.org/10.1007/s00401-019-02093-x
36. Olah M, Menon V, Habib N, Taga MF, Ma Y, Yung CJ et al (2020) Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. Nat Commun 11:6129. https://doi.org/10.1038/s41467-020-19737-2
37. Olah M, Patrick E, Villani AC, Xu J, White CC, Ryan KJ et al (2018) A transcriptomic atlas of aged human microglia. Nat Commun 9:539. https://doi.org/10.1038/s41467-018-02926-5
38. Osorio D, Cai JJ (2021) Systematic determination of the mitochondrial proportion in human and mice tissues for single-cell RNA-sequencing data quality control. Bioinformatics 37:963–967. https://doi.org/10.1093/bioinformatics/btaa751
39. Paolicelli RC, Sierra A, Stevens B, Tremblay ME, Aguzzi A, Ajami B et al (2022) Microglia states and nomenclature: A field at its crossroads. Neuron 110:3458–3483. https://doi.org/10.1016/j.neuron.2022.10.020
40. Pedersen TL (2020) patchwork: the composer of plots
41. Petrov D, Mansfield C, Moussy A, Hermine O (2017) ALS clinical trials review: 20 years of failure. Are we any closer to registering a new treatment? Front Aging Neurosci 9:68. https://doi.org/10.3389/fnagi.2017.00068
42. R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/
43. RStudio Team (2020) RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. http://www.rstudio.com/
44. Saade M, Araujo de Souza G, Scavone C, Kinoshita PF (2021) The role of GPNMB in inflammation. Front Immunol 12:674739. https://doi.org/10.3389/fimmu.2021.674739
45. Shen D, Ji Y, Qiu C, Wang K, Gao Z, Liu B et al (2024) Single-Cell RNA sequencing analysis of microglia dissected the energy metabolism and revealed potential biomarkers in amyotrophic lateral sclerosis. Mol Neurobiol 61:4473–4487. https://doi.org/10.1007/s12035-023-03806-w
46. Shi Y, Zhu R (2023) Analysis of damage-associated molecular patterns in amyotrophic lateral sclerosis based on ScRNA-seq and bulk RNA-seq data. Front Neurosci 17:1259742. https://doi.org/10.3389/fnins.2023.1259742
47. Slowikowski K(2021) ggrepel: automatically position non-overlapping text labels with 'ggplot2'. R package version 0.9.6.9999. https://github.com/slowkow/ggrepel
48. Smith AM, Dragunow M (2014) The human side of microglia. Trends Neurosci 37:125–135. https://doi.org/10.1016/j.tins.2013.12.001
49. Streit WJ, Xue QS (2010) The brain's aging immune system. Aging Dis 1:254–261
50. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd et al (2019) Comprehensive integration of single-cell data. Cell 177(1888–1902):e1821. https://doi.org/10.1016/j.cell.2019.05.031
51. Tam OH, Rozhkov NV, Shaw R, Kim D, Hubbard I, Consortium NA et al (2019) Postmortem cortex samples identify distinct molecular subtypes of ALS: retrotransposon activation, oxidative stress, and activated glia. Cell Rep 29:1164–1177. https://doi.org/10.1016/j.celrep.2019.09.066
52. Tang Y, Horikoshi M, Li W (2016) ggfortify: unified interface to visualize statistical results of popular R packages. R J 8(2):474-485. https://doi.org/10.32614/RJ-2016-060
53. Thrupp N, Sala Frigerio C, Wolfs L, Skene NG, Fattorelli N, Poovathingal S et al (2020) Single-nucleus RNA-seq is not suitable for detection of microglial activation genes in humans. Cell Rep 32:108189. https://doi.org/10.1016/j.celrep.2020.108189
54. Tuddenham JF, Taga M, Haage V, Marshe VS, Roostaei T, White C et al (2024) A cross-disease resource of living human microglia identifies disease-enriched subsets and tool compounds recapitulating microglial states. Nat Neurosci. https://doi.org/10.1038/s41593-024-01764-7
55. Umoh ME, Dammer EB, Dai J, Duong DM, Lah JJ, Levey AI et al (2018) A proteomic network approach across the ALS-FTD disease spectrum resolves clinical phenotypes and genetic vulnerability in human brain. EMBO Mol Med 10:48–62. https://doi.org/10.15252/emmm.201708202
56. Welch BL (1947) The generalisation of student's problems when several different population variances are involved. Biometrika 34:28–35. https://doi.org/10.1093/biomet/34.1-2.28
57. Wickham H (2016) ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York
58. Wilke CO (2025) cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. R package version 1.2. 0.9000. {Link: GitHub. https://github.com/wilkelab/cowplot}
59. Williams DB, Windebank AJ (1991) Motor neuron disease (amyotrophic lateral sclerosis). Mayo Clin Proc 66:54–82. https://doi.org/10.1016/s0025-6196(12)61175-6
60. Xiao N (2024). ggsci: Scientific Journal and Sci-Fi Themed Color Palettes for 'ggplot2'. R package version 3.2.0. https://github.com/nanxstats/ggsci
61. Xu X, Zhang J, Li S, Al-Nusaif M, Zhou Q, Chen S et al (2021) Bone marrow stromal cell antigen 2: Is a potential neuroinflammation biomarker of SOD1(G93A) mouse model of amyotrophic lateral sclerosis in pre-symptomatic stage. Front Neurosci 15:788730. https://doi.org/10.3389/fnins.2021.788730
62. Xu S, Li L, Luo X, Chen M, Tang W, Zhan L, Dai Z, Lam TT, Guan Y, Yu G (2022) Ggtree: a serialized data object for visualization of a phylogenetic tree and annotation data. iMeta 1(4):e56. https://doi.org/10.1002/imt2.56
63. Yu G, Wang LG, Han Y, He QY (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16:284–287. https://doi.org/10.1089/omi.2011.0118
64. Zheng S, Papalexi E, Butler A, Stephenson W, Satija R (2018) Molecular transitions in early progenitors during human cord blood hematopoiesis. Mol Syst Biol 14:e8041. https://doi.org/10.15252/msb.20178041
65. Zhu S, Wuolikainen A, Wu J, Ohman A, Wingsle G, Moritz T et al (2019) Targeted multiple reaction monitoring analysis of CSF identifies UCHL1 and GPNMB as candidate biomarkers for ALS. J Mol Neurosci 69:643–657. https://doi.org/10.1007/s12031-019-01411-y