# Hexanucleotide repeat expansions in C9orf72 alter microglial responses and prevent a coordinated glial reaction in ALS

Pegah Masrori[1,2,3#], Baukje Bijnens[4,5#], Kristofer Davie[6], Suresh Kumar Poovathingal[7], Annet Storm[2], Nicole Hersmus[2], Laura Fumagalli[4,5], Ludo Van Den Bosch[1,2], Mark Fiers[8], Dietmar Rudolf Thal[9,10], Renzo Mancuso[4,5*], Philip Van Damme[1,2,3*]

1. KU Leuven - University of Leuven, Department of Neurosciences, Leuven Brain institute (LBI), Leuven, Belgium.
2. Laboratory of Neurobiology, VIB Center for Brain & Disease Research, Leuven, Belgium
3. University Hospitals Leuven, Department of Neurology, Leuven, Belgium.
4. Microglia and Inflammation in Neurological Disorders (MIND) Lab, VIB Center for Molecular Neurology, VIB, Antwerp, Belgium.
5. Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium.
6. Single Cell Bioinformatics Expertise Unit (CBD), VIB Center for Brain & Disease Research, Leuven, Belgium.
7. Single Cell Analytics & Microfluidics Core, VIB Center for Brain & Disease Research, Leuven, Belgium.
8. VIB Center for Brain & Disease Research, Leuven, Belgium.
9. Laboratory for Neuropathology, Department of Imaging and Pathology, Leuven Brain Institute (LBI) KU Leuven, Leuven, Belgium.
10. Department of Pathology, University Hospitals Leuven, Leuven, Belgium.

Co-corresponding authors: philip.vandamme@kuleuven.be; renzo.mancuso@uantwerpen.vib.be
# These authors contributed equally
* These authors contributed equally

## Summary

Neuroinflammation is an important hallmark in amyotrophic lateral sclerosis (ALS). Experimental evidence has highlighted a role of microglia in the modulation of motor neuron degeneration. However, the exact contribution of microglia to both sporadic and genetic forms of ALS is still unclear. We generated single nuclei profiles of spinal cord and motor cortex from sporadic and *C9orf72* ALS patients, as well as controls. We particularly focused on the transcriptomic responses of both microglia and astrocytes. We confirmed that *C9orf72* is highly expressed in microglia and shows a diminished expression in carriers of the hexanucleotide repeat expansion (HRE). This resulted in an impaired response to disease, with specific deficits in phagocytic and lysosomal transcriptional pathways. Astrocytes also displayed a dysregulated response in *C9orf72* ALS patients, remaining in a homeostatic state. This suggests that C9orf72 HRE alters a coordinated glial response, which ultimately would increase the risk for developing ALS. Our results indicate that *C9orf72* HRE results in a selective microglial loss-of-function, likely impairing microglial-astrocyte communication and preventing a global glial response. This is relevant as it indicates that sporadic and familial forms of ALS may present a different cellular substrate, which is of great importance for patient stratification and treatment.

## Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease characterized by progressive motor neuron loss (MNs) in the motor cortex, brainstem, and spinal cord (1). Motor neuron loss is accompanied by neuroinflammation, and changes in both microglia and astrocytes (2). ALS typically presents with focal muscle weakness, which spreads to neighboring body regions. Extramotor manifestations, such as cognitive and behavioral impairment, occur in a subset of patients. Effective therapies are lacking, and the median survival after symptom onset is limited to 3 years (1). The widespread nature of this disease, with contribution from distant central nervous system (CNS) regions, represents a challenge when trying to elucidate the underlying cellular mechanisms. To achieve a deep understanding of ALS pathophysiology, it is essential to investigate cell-specific disease mechanisms acting both in the spinal cord and the motor cortex.

ALS has a strong genetic component, where the most common monogenic cause is the GGGGCC repeat expansions in the *chromosome 9 open reading frame 72* gene (*C9orf72*) (3), with a repeat length ranging from 38 to approximately 1600 repeats (4, 5). *C9orf72* presents with multiple transcript variants that are differentially enriched across cell types in the brain (6, 7). Both loss-of-
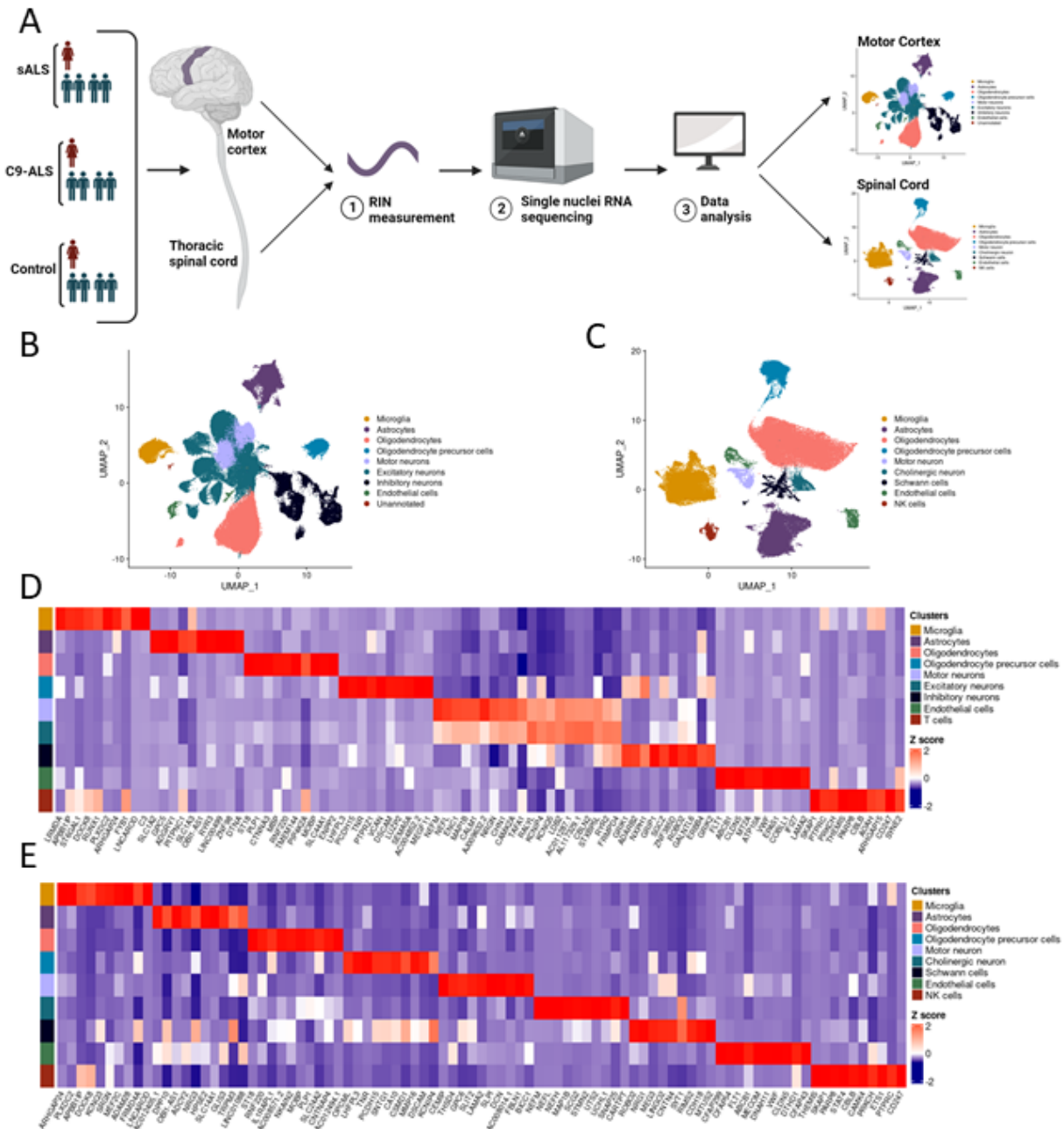
**Figure 1| sn-RNA sequencing of human post mortem motor cortex and spinal cord of ALS patients.**
(**a**) experimental outline of this study. (**b**) UMAP plot visualizing 93,176 nuclei isolated from post-mortem cortical samples of ALS patients (C9-ALS and sALS) and controls. Cells are colored according to their assigned cluster identity: microglia, astrocytes, oligodendrocytes, oligodendrocyte precursor cells, motor neurons, excitatory neurons, inhibitory neurons, endothelial cells and T-cells. (**c**) UMAP plot visualizing 63,076 nuclei isolated from post-mortem spinal cord samples of ALS patients (C9-ALS and sALS) and controls. Cells are colored according to their assigned cluster identity: microglia, astrocytes, oligodendrocytes, oligodendrocyte precursor cells, motor neurons, cholinergic neurons, Schwann cells, endothelial cells and an unannotated cluster. Heatmap showing the top 10 most expressed genes per cluster of the (**d**) motor cortex and (**e**) the spinal cord.

function (LOF) and toxic gain-of-function (GOF) mechanisms in neuronal and non-neuronal cells have been implicated in the disease pathogenesis (8). *C9orf72* is predominantly expressed in myeloid cells, including microglia (6). Whereas the exact mechanism by which the *C9orf72* mutations alters microglia function is not fully understood (9),

experimental studies have shown that its deletion leads to altered immune responses, exacerbated myeloid proliferation and lysosomal dysfunctions (10, 11).

In this study, we provide an unparalleled single-nuclei RNA sequencing resource from paired motor cortex and spinal cord samples from

sporadic ALS without any known gene mutation (sALS), ALS patients with a *C9orf72* mutation (C9-ALS) and controls. We show that mutations in *C9orf72* prevent the transition of microglia towards reactive cell states, and block a coordinated multi-cellular response to disease, when compared to sALS. While sALS and C9-ALS are clinically indistinguishable, our findings suggest that cellular substrate underlying disease might be fundamentally different. Also, our work allows interrogating the contribution of cellular subtypes involved in disease pathogenesis and cell-specific dysregulated pathways. We believe this unique resource and approach will lead to the identification of new neuroinflammatory biomarkers, better patient stratification and novel, more specific, disease-modifying strategies to tackle ALS.

## Results

We performed single nuclei RNA (snRNA)-sequencing on motor cortex and spinal cord samples from 5 sALS, 5 C9-ALS and 5 controls. The experimental setup is illustrated in Figure 1a. The demographic data of participants are shown in Supplementary Table 1. To minimize potential transcriptomic artefacts (12), we only used samples with a mean post-mortem interval of up to 9 hours.

### Optimization of single nuclei isolation and sequencing from post-mortem human samples

Recent advances in the isolation and sequencing of snRNA have demonstrated that, with optimal buffer/detergent compositions, the assay performance and gene detection sensitivity can be significantly enhanced. We first evaluated three different strategies for single nuclei isolation: the protocol by Habib et al. (13), Tween with salts and Tris (TST) (14), and Soma-seq (15) (Methods Online). After nuclei isolation, we performed snRNA sequencing of spinal cord and motor cortex from sporadic and C9orf72 ALS patients, as well as controls and compared the three methods for basic parameters including number of reads, genes and UMIs. Based on this analysis we concluded that the Soma seq method shows the best performance (Supplementary Figure 1c), and therefore we used this method for all the following experiments.

### Cellular diversity in the spinal cord and motor cortex

We investigated the cellular diversity across the motor cortex spinal cord. We analyzed 156,252

nuclei, after removal of low-quality nuclei and doublets. The full dataset from both pre- and post-batch correction is shown in Supplementary Figure 1.

In the spinal cord, we identified 9 major transcriptionally distinct nuclei clusters accounting for all the cell types present in the spinal cord (Figures 1c, e). We annotated these clusters into: microglia (*ARHGAP24*, *PLXDC2*, *KCNQ3*, *SRGN*)*,* oligodendrocytes (*ST18*, *IL1RAPL1*, *MOBP*, *SLC24A2*), astrocytes (*NRG3*, *DPP10*, *HPSE2*, *SCL14A1*), oligodendrocytes precursor cells (OPCs, *OPCML*, *TNR*, , *MMP16*), Schwann cells (*ROBO2*, *NRG1*, *MEG3*, *CNTN4*,), motor neurons (*CEMIP*, *SLT2*, *LAMA2*, *DCN*), cholinergic neurons (*NEFM*, *MAP1B*, *STMN2*, *SNAP25*), endothelial cells (*FLT1*, *ABCB1*, *CLDN5*, *CFAP43*), and Natural Killer cells (*THEMIS, SKAP1, PARP8, STAT4*).

On the other hand, we identified 9 major transcriptionally-distinct nuclei clusters in the motor cortex (Figures 1b, d), including oligodendrocytes (*ST18*, *PLP1*, *MBP*, *MOBP* )*,* astrocytes (*SLC1A2*, *SLC1A3*, *RYR3*, *DTNA*), excitatory neurons (*RALYL, CBLN2, KCNQ5, LDB2*), microglia (*LRMDA*, *ST6GAL1*, *RUNX1*, *C3*), motor neurons (*NEFM*, *NEFL*, *ENC1*, *MAP1B*), OPCs (*LHFPL3, TNR, VCAN, DSCAM*), inhibitory neurons (*MEGF11*, *GRIP1*, *ROBO2*, *GLANTL6*), endothelial cells (*ABCB1*, *CLDN5*, *MT2A*, *IFI27*) and T cell (*SKAP1*, *PTPRC*, *CD247*, *SYNE2*).

We did not observe any significant differences in the distribution of patients and experimental conditions across clusters (Supplementary Figures. 1d, e).

### The C9orf72 hexanucleotide expansion diminishes microglial activation

We subset, integrated and re-clustered the microglial nuclei from the spinal cord (n=13,980) and motor cortex (n=4,845 cells) (Supplementary Table 2). After removing peripheral and other CNS macrophages, as well as proliferating cells, (Figures 2a, b, and Supplementary Figure 2). The microglia subclusters were all characterized by a distinct transcriptomic signature (Figures 2c). Cluster MG1 closely resembled homeostatic microglia (*CX3CR1*, *P2RY12*) (16), and presented lower levels of known activation markers (*C1*, *IL1B*, *APOE*, *CTSB*) (17, 18). Cluster MG2 appeared to be a transitioning state that retained expression of homeostatic markers a
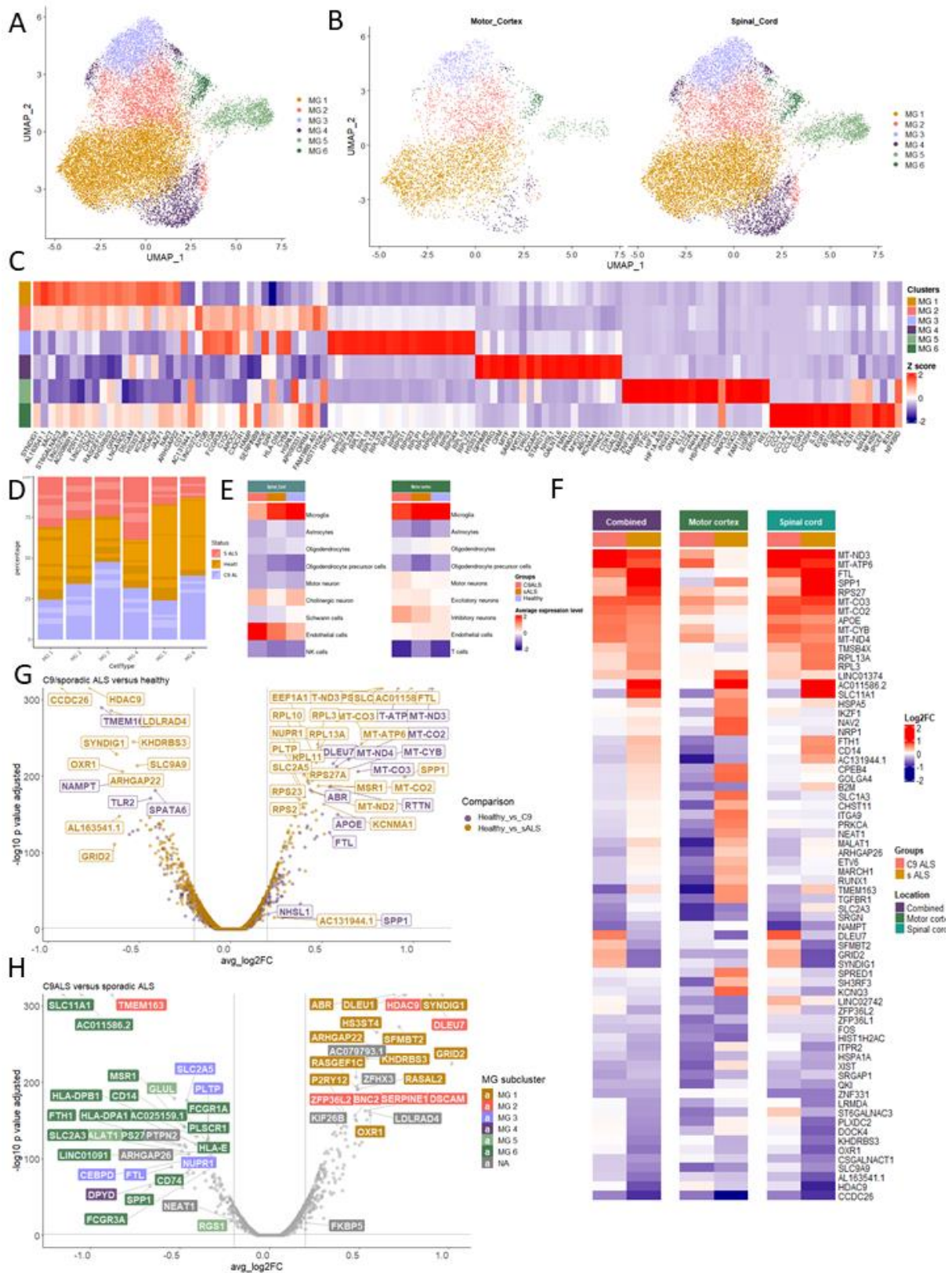
Figure legend on the next page

more intermediate state, with expression of homeostatic markers (*CX3CR1*, *P2RY12*) and intermediate levels of state transition markers (*SPP1*, *CD14*, *IL1B*, *C3*). We captured a microglia subset, MG3, very closely related to disease/damage associated microglia (DAM) described in multiple disease scenarios (17, 19). This cluster was characterized by higher levels of *CD74*, *APOE*, *TREM2*, *C1QA*, *CTSD.* Cluster MG4 was consistent with previously described axon tract

**Figure 2| C9orf72 HRE induces changes in the microglial transcriptomic profile compared to sALS.** (**a**) UMAP plot visualizing 16,763 microglia nuclei from post-mortem cortical samples of ALS patients (C9 and sALS) and controls. Microglia subsets are colored based on their assigned cluster into: MG1, MG2, MG3, MG4, MG5, MG6. (**b**) UMAP plot displaying the specific distribution of 4,216 nuclei from the motor cortex and 12,547 nuclei of the spinal cord. (c) heatmap showing the top 20 most expressed genes per cluster. (**d**) proportion of cells (y-axis) from each genotype in the different microglial subclusters, alternating shades indicate different individuals (x-axis). (**e**) heatmap depicting the Z-scored average expression of *C9orf72* in both motor cortex and spinal cord. (**f**) heatmaps showing the log2FC of differentially expressed genes (p-value adjusted <0.05) of both C9-ALS or sALS patients compared to healthy controls when comparing pooling the motor cortex and the spinal cord. (**g**) Volcano plot depicting 2 pairwise comparisons: between C9-ALS and healthy individuals, and sALS and healthy individuals for the combined microglial dataset (DESeq2 negative binomial distribution, p-values adjusted with Bonferroni correction based on the total number of genes in the dataset). Vertical lines depict Log2FC threshold of 0.2 whereas the horizontal sets the adjusted p-value threshold at < 0.05). (**h**) Volcano plot depicting the pairwise comparisons of C9-ALS vs. sALS for the combined microglial dataset (DESeq2 negative binomial distribution, p-values adjusted with Bonferroni correction based on the total number of genes in the dataset). Vertical lines depict Log2FC threshold of 0.2 and the horizontal line, the adjusted p-value threshold at < 0.05). Colors indicate marker genes of microglia clusters.

associated microglia (*LGAL3*, *GPNMB*, *MITF*, *CTSD*) (20). Cluster MG 5 displayed an expression profile related to hypoxia (*HIF1A*, *VEGFA*, *NAMPT*, *ELL2*, *CD83*), which has been previously described in brains with glioblastoma multiforme (21). Lastly, cluster MG 6 closely resembled cytokine response microglia (CRM) with key markers being *CCL2*, *CCL3* and *CCL4* (22). Distributions of these clusters over all three conditions are shown in Figure 2d.

Haploinsufficiency of C9orf72 protein, as a result of the hexanucleotide repeat expansion (HRE), is a key proposed disease mechanism - along with aggregating expanded RNAs and dipeptide repeat proteins (DPRs) - in ALS (23). We wondered whether HRE in the *C9orf72* had an impact on human microglia. We assessed the levels of *C9orf72* and confirmed the HRE leads to reduced *C9orf72* expression levels (Figure 2e). Next, we performed a separate differential expression analysis between C9-ALS vs. controls, and sALS vs. controls (Figure 2f and Supplementary Table 3). We observed strong differences between experimental groups, with C9-ALS microglia displaying a much lower extent of response to ALS pathology in the motor cortex and especially the spinal cord. Comparison between C9-ALS and sALS microglia showed that while sALS microglia display transcriptomics changes consistent with a reactive state (*SPP1, CD14, HLA, FCGR1A. FCGR3A*), C9-ALS microglia remained vastly homeostatic (*P2RY12*, *INPP5D*, *BLNK*, *SERPINE1*, *SORL1* (Figure 2g, h). This suggests that the HRE in *C9orf72* impairs microglial cell state transition, consistent with a partial loss of function phenotype.

We have recently mapped the transcriptomic changes occurring in human microglia during amyloid-b pathology using a xenotransplantation mouse model (24). To shed light on whether *C9orf72* HRE modifies the microglial phenotype, we performed a gene enrichment analysis and mapped those transcriptomic signatures in our dataset). We observed the C9-ALS microglia were impaired in the transition from homeostatic to DAM and antigen-presenting HLA phenotypes (Figure 3a). This is consistent with previous observations in mouse models of C9orf72 deficiency, supporting the idea of ablation of these specific phenotypes as a result of the C9orf72 haploinsufficiency (Supplementary Figure 3) (7, 25-27). These data also indicate that there might be strong similarities between the contribution of microglia across different neurodegenerative disorders (21, 24, 28-30). To confirm our findings, we repeated our gene enrichment analysis using several specific transcriptional pathways that account for different aspects of microglial function. Whereas sALS microglia showed increased transcription of genes involved in immune, lysosomal and phagocytic pathways, C9-ALS microglia remained similar to the controls (Figures 3b, c). On the contrary, C9-ALS microglia displayed an enrichment in genes related to ribosomal function (Figures 3b, c), which is consistent with previous data reporting similar findings in *C9orf72* derived ALS astrocytes (31). Interestingly, we observed that most of the changes occurred in the spinal cord, with only mild alterations in the motor cortex (Figures 3a-c). Whereas these alterations may reflect a different extent of pathology in both regions, it could also be
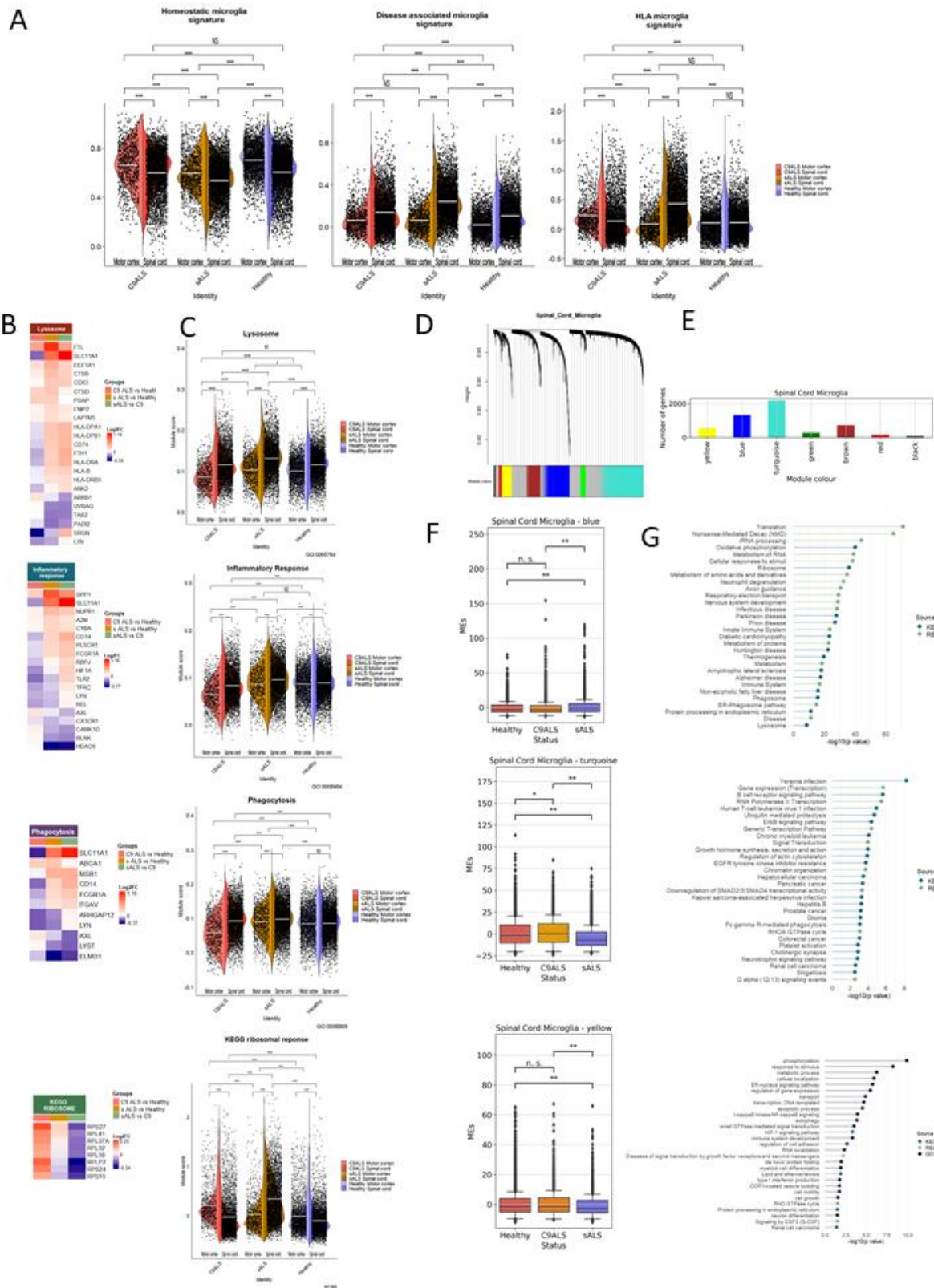
**Figure 3| Microglia from C9orf72 HRE patients show altered functional pathways. (a)** Seurat module scores of disease-associated (DAM), antigen-presenting HLA, and homeostatic microglia phenotypes as described by Mancuso et al., 2022. **(b)** heatmaps depicting Z-scored average expression of differentially expressed genes (log2FC < +/-0,20, p-value adjusted <0,05) of gene sets of lysosomes, inflammatory response, phagocytosis and KEGG ribosomes. **(c)**, Seurat module scores of GO terms from the lysosomes, inflammatory response, phagocytosis and KEGG ribosome. **(d)** dendrogram showing the result of unsupervised hierarchical clustering of modules identified by WGCNA. **(e)** barplots depicting the number of genes per module **(f)** boxplots visualizing the module eigengenes (y-axis) per genotype (x-axis) for the blue, turquoise and yellow cluster of the spinal cord. **(g)** pathway enrichment analysis of the WGCNA clusters blue, turquoise and yellow, with the associated GO term (y-axis) and –log10(p-value) (x-axis).

consequence of intrinsic regional differences in the microglia (32).

Next, we performed weighted gene co-expression network analysis (WGCNA) to confirm the alterations induced by *C9orf72* HRE in microglia. We focused on the spinal cord, where we found the most pronounced changes in our differential expression analysis. We identified 7 distinct co-expression clusters, many of them with differential significant enrichment across sALS, C9-ALS and controls (Figures 3d-f and Supplementary Figure 4). The blue cluster appeared particularly relevant, as it contained a high number of genes (Figure 3e) and was significantly enriched in sALS vs. both C9-ALS and controls (Figure 3f). Genes in this cluster were associated with functional pathways related to the immune system, phagocytosis, autophagy-lysosomal function, ubiquitin-proteasome system, oxidative stress and apoptosis confirming a strong immune component in the pathogenesis of sALS (Figure 3g). C9-ALS microglia showed no enrichment in these pathways, consistent with our findings on *C9orf72* HRE leading to an impaired microglial response (Figure 3g; Supplementary Figure 4). On the other hand, we observed an enrichment of clusters turquoise and yellow in the C9-ALS microglia compared to both sALS and controls (Figure 3f). These clusters encompass general biological pathways such as cell proliferation, migration, general metabolic processes, protein biogenesis and degradations are particularly involved, and as opposed to the blue cluster, without a predominant presence of the immune response (Figure 3g). Altogether, our data shows that the *C9orf72* HRE prevents the transition of microglia towards reactive cell states and induces perturbations in transcriptional pathways different than in sALS, therefore suggesting a different cellular mechanism underlying sporadic and inherited forms of ALS.

## Reduced microglial activation correlates with an impaired astrocytic reaction

Microglia-astrocyte interactions have been implicated in multiple neurodegenerative disorders and under inflammatory conditions (33). It is well accepted that microglia can interact and modulate the response of astrocytes (34-37), although the underlying molecular mechanism of this interplay remain unclear. As *C9orf72* is primarily expressed in the microglia (38), we hypothesize that in ALS,

*C9orf72* HRE induces direct alterations in the microglia, affecting their ability to properly communicate and modulate the activation of astrocyte. To address this, we first subset, integrated and re-clustered both spinal and cortical astrocytes to define their transcriptomic landscape (Supplementary Figure 5). We observed similar heterogeneity compared to the microglia, with 8 distinct transcriptional states across the spinal cord (n=12,978) and motor cortex (n= 9,657cells) (Figures 4a-c, and Supplementary Table 2). We found a number of clusters consistent to homeostatic astrocytes, with expression of genes related to synaptic support and glutamate metabolism (AS1: *SLC1A2*, *NRXN1*, *CABLES1*) (39), interaction with neurons and axonal guidance (AS3: *TNN*, *NRXN3*, *LAMA2*, *PLEKHA5*) (39), or myelin formation, synaptic formation and short term plasticity (AS6: *GRID2*, *CNTNAP5*, *ERFI1*, *SHISA9*, *KCND2*) (31). Cluster AS7 was linked to axon guidance, pruning of inappropriate synapses and development of neural circuits (*GNA14*, *NRP1*, *GFRA1*, *CHI3L1*, *CPNE8*) (40). We also found several transcriptomic phenotypes consistent with astrocyte activation. Cluster AS4 showed expression of stress markers (*HES1*, *FOS*, *JUN*, *ID3*) (41). Cluster AS5 displayed reactive astrocytes markers (*NAMPT*, *CADPS*, *LPAR1*, *SERPINA3*, *EMP1*, *CHI3L2*, *RASGEF1B*) (39, 42). Cluster AS8 was characterized by high expression levels of reactive astrocyte markers (*SERPINA3*, *EMP1*, *CHIL3L2*) as well as multiple immune-related genes (*CLU*, *VIM*, *B2M*, *APOE*, *AQPI*, *C3*) (43-45). AS2 appears to be a transitioning state that retains expression of AS1 markers but also displays an enrichment of mitochondrial and ribosomal transcripts, along with intermediate levels of AS3 (*TTN*, *NRXN3*, *LAMA2*, *PLEKHA5*) and AS8 markers (A*POE*, *B2M*, *AQP4*, *SERPINA3*) (Figures 4a-c). Distributions of these clusters over all three conditions are shown in Figure 4d.

We assessed for potential alterations in the response of astrocytes by performing a differential expression analysis between C9-ALS vs. controls, and sALS vs. controls (Supplementary Table 4). Similarly to what we observed in microglia, we found an overall lower extent of transcriptional changes in C9-ALS vs. sALS astrocytes (Figures 4f-e) sALS astrocytes exhibited an upregulation of genes associated to a reactive state (*SERPINA3, RASGEF, LPAR1, TPST1, CHI3L, EMP1*), as

Figure legend on the next page
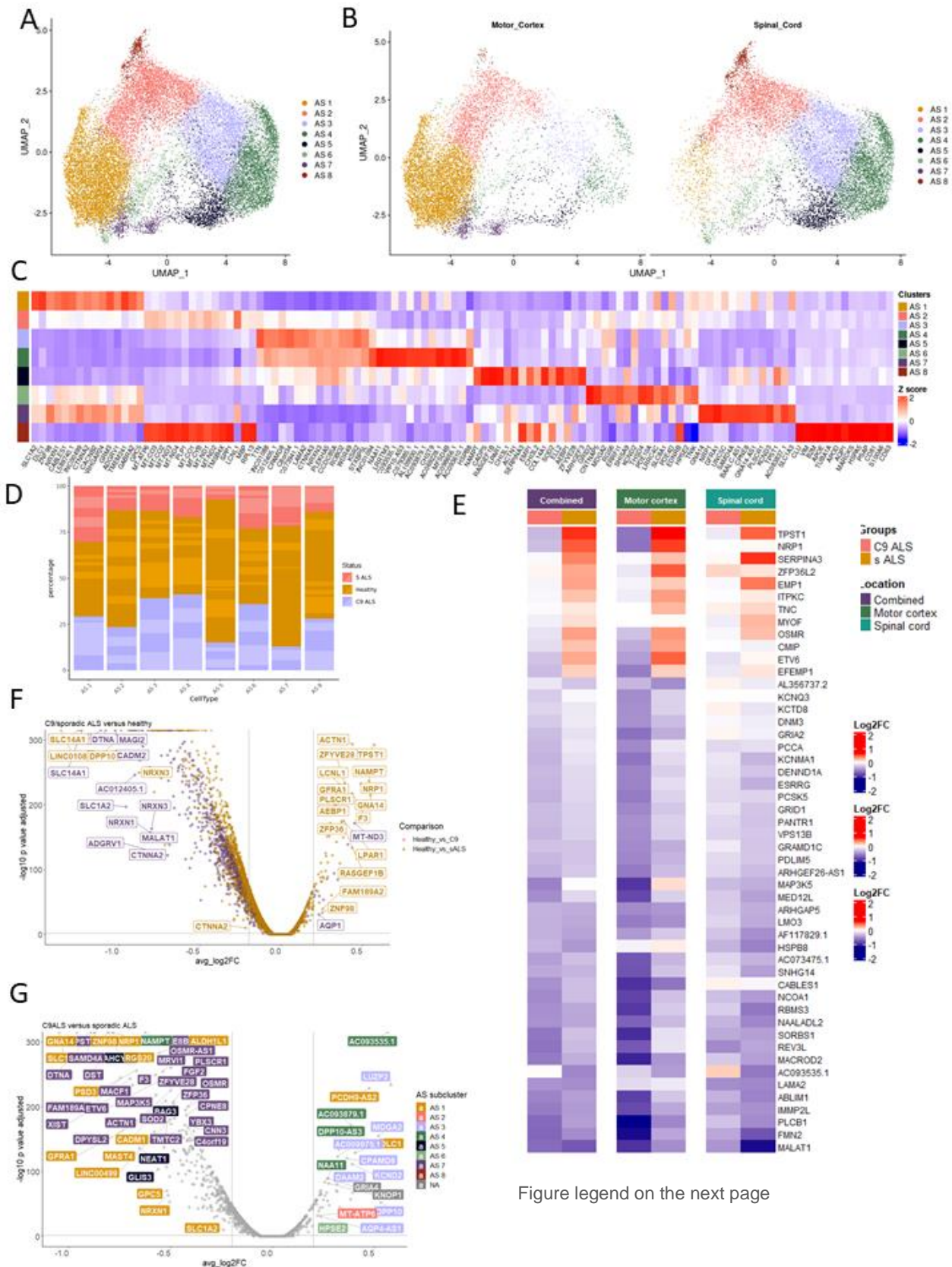
opposed to C9-ALS astrocytes which more often retained a homeostatic state (*AQP1, DLC1, SLC35F1*) (Figures 4f-g). These data indicate a *C9orf72* HRE-mediated defective engagement in a coordinated response in ALS pathology indicative of an aberrant microglia-astrocyte interaction.

**Figure 4| Astrocytes are differentially affected in C9-ALS compared to sALS**. **(a)** UMAP plot visualizing 22,635 astrocytes from post-mortem cortical samples of ALS patients (C9-ALS and sALS) and controls. Astrocyte subsets are colored based on their assigned cluster: AS1, AS2, AS3, AS4, AS5, AS6, AS7, AS8. **(b)** UMAP plot displaying the specific distribution of the 9,657 cortical and the 12,978 spinal nuclei. **(c)** Heatmap showing the top 15 most expressed genes per cluster. **(d)** proportion of cells (y-axis) from each genotype in the different astrocyte subclusters, alternating shades indicate different individuals (x-axis). **(e)**, heatmaps showing the log2FC of differentially expressed genes (p-value adjusted <0.05) of both C9-ALS or sALS compared to healthy controls. **(f)** Volcano plot depicting 2 pairwise comparisons between C9-ALS and controls, and sALS and controls for the combined astrocytes dataset (DESeq2 negative binomial distribution, p-values adjusted with Bonferroni correction based on the total number of genes in the dataset). Vertical lines depict Log2FC threshold of 0.2 and the horizontal line, the adjusted p-value threshold at <0.05. **(g)** Volcano plot showing a pairwise direct comparisons of C9-ALS vs. sALS for the combined astrocyte dataset (DESeq2 Wilcoxon Rank sumtest, p-values adjusted with Bonferroni correction based on the total number of genes in the dataset). Vertical lines depict Log2FC threshold of 0.2 and the horizontal line, the adjusted p-value threshold at < 0.05. Colors indicate marker genes of astrocyte clusters.

## Discussion

Neuroinflammation is a key pathological hallmark for ALS (9, 46). Experimental (47, 48) and genetic evidence (46, 49, 50) indicate that microglia and neuroinflammation are active players - rather than simple bystanders - in the pathogenic events that lead to motor neuron degeneration in ALS. Our study confirms that *C9orf72* is mainly expressed in microglia in the human brain (6), and that *C9orf72* HRE leads to reduced expression. We report that *C9orf72* HRE impairs microglial cell state transition towards DAM and HLA cell states and induces alterations in transcriptomic pathways related to phagocytosis and lysosomal function. We also show that astrocytes from *C9orf72* HRE carriers display a general lower extent of transcriptomic change compared to ALS, and hypothesized that this could be the result of a defective microglia-astrocyte communication that may hinder a coordinated neuroinflammatory response to the disease. These findings suggest that patients with a *C9orf72* HRE may require a personalized intervention to modulate the neuroinflammatory response.

There are several proposed mechanisms explaining the causative role of *C9orf72* in ALS (3, 5, 51), and LOF due to reduced *C9orf72* mRNA and protein levels is one of them (23, 52, 53). Several studies have reported widespread alterations in microglia (and other myeloid cells) after conditional (54-57) or complete *C9orf72* deletion (10, 26, 58, 59), including transcriptional alterations, an exacerbated production of inflammatory cytokines, lysosomal dysfunction (6, 52, 60-62). Similarly, we also find that diminished levels of C9orf72 result in an impaired transition into DAM and HLA responses, and reduced

engagement in phagocytic and lysosomal transcriptional programs. This is not a unique phenotype of microglia harboring *C9orf72* mutations, and have been observed before in pathogenic genetic variants in *TREM2* and *APOE* (24). These findings indicate that the inflammatory substrate and microglial contribution is different in *C9orf72* and sALS pathology, and additional insights will be crucial to define tailored therapeutic strategies for groups of patients with different pathogenic contributors.

Recent studies have suggested a synergistic effect of *C9orf72* LOF and GOF (63). In vivo studies showed that *C9orf72* deficiency leads to exacerbation of motor phenotypes and premature death in C9orf72 GOF mouse models (64, 65). It has been proposed that reduced *C9orf72* levels might further enhance the pathological accumulation and toxicity of DPRs (66, 67) by altering autophagy (68). However, this potential synergism between LOF and GOF has been only investigated in neurons (65). Based on our observations, it is plausible that these two mechanisms act via different cell types instead, with GOF altering neuronal function, and LOF leading to a microglial defect, exacerbation of neuronal alterations and a clinical phenotype. Whereas the mechanisms underlying the cell specific impact of C9 HRE are unclear, it is plausible that different cell types use different *C9or72* transcript variants, therefore manifesting different functional consequences of carrying the pathogenic expansion. Supporting this idea, it has been reported that while neurons, oligodendroglia, and astrocytes use transcript variants containing exons 1a and 1b, microglia are enriched in transcripts including exon 1b (7), which is not associated with classical GOF mechanisms (i.e.

9

DRPs and RNA foci). Additionally, two novel specific transcript variants (V4 and V5) have been reported in CD14+ myeloid cells, containing an could explain a preference of LOF over GOF alterations in microglia (69).

Astrocyte activation has been reported in *in vitro* models of C9orf72 HRE, with somewhat conflicting results. Our data show a general transcriptional silencing in astrocytes of both sALS and *C9orf72* HRE carriers, slightly more pronounced in the latter. This mostly involved transcriptional pathways related to synaptic and neuronal support, and is consistent with *in vitro* data from *C9orf72* iPSC derived astrocytes, showing a downregulation of gene networks involved in synaptic transmission and neuronal support, and increased expression of rRNA processing and ribosomal subunits (70). Contrary to our findings, in organoid systems, *C9orf72* astrocytes show a dysregulated transcriptional profile with increased levels of autophagy (71).

alternative cryptic initial exon. These splice variants do not contain the HRE sequence and therefore

These inconsistencies may reflect intrinsic differences between human postmortem tissue, 2D and 3D models; but also the fact that microglia are missing in these cell culture platforms which may relevant aspects related to microglia-astrocytes crosstalk. Additional research using higher order organoids containing microglia, or xenotransplantation models will be useful to tackle these outstanding questions.

In summary, this study indicates that *C9orf72* HRE results in reduced *C9orf72* expression levels in the microglia and a subsequent reduction in their response to disease, consistent with deficits in phagocytic and lysosomal pathways. We hypothesize that this leads to a defective general glial response that extends to astrocytes, hindering a coordinated response that increases the risk for developing ALS.

## Acknowledgements

## Authors contribution

P.M. designed all experiments and led this project; P.M., D.R. T and P.V.D collected the post-mortem human samples; D.R.T performed neuropathological analysis and dissected brain/spinal cord samples

for analysis; P.M. performed RNA extraction with assistance of A.S.; S.K.P. performed the 10x genomics with the assistance of P.M., A.S. and N.H. B.B. and K.D performed the snRNAseq analysis; P.M., B.B., K.D, L.F., M.F., R.M. and P.V.D interpreted the snRNAseq data; P.M., B.B., R.M. and P.V.D. wrote the manuscript; L.F., L.V.D., M.F. and D.R.T critically revised the manuscript; all authors read and approved the final version of the paper.

## Competing interest statement

## References (for main text only)

1. Masrori P, Van Damme P. Amyotrophic lateral sclerosis: a clinical review. Eur J Neurol. 2020.

2. Ilieva H, Polymenidou M, Cleveland DW. Non-cell autonomous toxicity in neurodegenerative disorders: ALS and beyond. The Journal of cell biology. 2009;187(6):761-72.

3. Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron. 2011;72(2):257-68.

4. Dedeene L, Van Schoor E, Ospitalieri S, Ronisz A, Weishaupt JH, Otto M, et al. Dipeptide repeat protein and TDP-43 pathology along the hypothalamic-pituitary axis in C9orf72 and non-C9orf72 ALS and FTLD-TDP cases. Acta neuropathologica. 2020;140(5):777-81.

5. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron. 2011;72(2):245-56.

6. Rizzu P, Blauwendraat C, Heetveld S, Lynes EM, Castillo-Lizardo M, Dhingra A, et al. C9orf72 is differentially expressed in the central nervous system and myeloid cells and consistently reduced in C9orf72, MAPT and GRN mutation carriers. Acta Neuropathol Commun. 2016;4(1):37.

7. Lall D, Lorenzini I, Mota TA, Bell S, Mahan TE, Ulrich JD, et al. C9orf72 deficiency promotes microglial-mediated synaptic loss in aging and amyloid accumulation. Neuron. 2021;109(14):2275-91.e8.

8. Gendron TF, Petrucelli L. Disease Mechanisms of C9ORF72 Repeat Expansions. Cold Spring Harbor perspectives in medicine. 2018;8(4).

9. Masrori P, Beckers J, Gossye H, Van Damme P. The role of inflammation in neurodegeneration: novel insights into the role of the immune system in C9orf72 HRE-mediated ALS/FTD. Molecular neurodegeneration. 2022;17(1):22.

10. Atanasio A, Decman V, White D, Ramos M, Ikiz B, Lee H-C, et al. C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. Scientific reports. 2016;6:23204.

11. Amick J, Ferguson SM. C9orf72: At the intersection of lysosome cell biology and neurodegenerative disease. Traffic (Copenhagen, Denmark). 2017;18(5):267-76.

12. Marsh SE, Walker AJ, Kamath T, Dissing-Olesen L, Hammond TR, de Soysa TY, et al. Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. Nat Neurosci. 2022;25(3):306-16.

13. Habib N, Avraham-Davidi I, Basu A, Burks T, Shekhar K, Hofree M, et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. Nat Methods. 2017;14(10):955-8.

14. Slyper M, Porter CBM, Ashenberg O, Waldman J, Drokhlyansky E, Wakiro I, et al. A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. Nat Med. 2020;26(5):792-802.

15. Otero-Garcia M, Xue Y-Q, Shakouri T, Deng Y, Morabito S, Allison T, et al. Single-soma transcriptomics of tangle-bearing neurons in Alzheimer's disease reveals the signatures of tau-associated synaptic dysfunction. bioRxiv. 2020:2020.05.11.088591.

16. Masuda T, Sankowski R, Staszewski O, Prinz M. Microglia Heterogeneity in the Single-Cell Era. Cell Rep. 2020;30(5):1271-81.

17. Keren-Shaul H, Spinrad A, Weiner A, Matcovitch-Natan O, Dvir-Szternfeld R, Ulland TK, et al. A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. Cell. 2017;169(7):1276-90.e17.

18. Thrupp N, Sala Frigerio C, Wolfs L, Skene NG, Fattorelli N, Poovathingal S, et al. Single-Nucleus RNA-Seq Is Not Suitable for Detection of Microglial Activation Genes in Humans. Cell Rep. 2020;32(13):108189.

19. Srinivasan K, Friedman BA, Etxeberria A, Huntley MA, van der Brug MP, Foreman O, et al. Alzheimer's Patient Microglia Exhibit Enhanced Aging and Unique Transcriptional Activation. Cell Rep. 2020;31(13):107843.

20. Hammond TR, Dufort C, Dissing-Olesen L, Giera S, Young A, Wysoker A, et al. Single-Cell RNA Sequencing of Microglia throughout the Mouse Lifespan and in the Injured Brain Reveals Complex Cell-State Changes. Immunity. 2019;50(1):253-71.e6.

21. Sankowski R, Böttcher C, Masuda T, Geirsdottir L, Sagar, Sindram E, et al. Mapping microglia states in the human brain through the integration of high-dimensional techniques. Nat Neurosci. 2019;22(12):2098-110.

22. Mancuso R, Van Den Daele J, Fattorelli N, Wolfs L, Balusu S, Burton O, et al. Stem-cell-derived human microglia transplanted in mouse brain to study human disease. Nat Neurosci. 2019;22(12):2111-6.

23. Shi Y, Lin S, Staats KA, Li Y, Chang W-H, Hung S-T, et al. Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. Nat Med. 2018;24(3):313-25.

24. Mancuso R, Fattorelli N, Martinez-Muriana A, Davis E, Wolfs L, Van Den Daele J, et al. A multi-pronged human microglia response to Alzheimer's disease Aβ pathology. bioRxiv. 2022:2022.07.07.499139.

25. Shao Q, Liang C, Chang Q, Zhang W, Yang M, Chen JF. C9orf72 deficiency promotes motor deficits of a C9ALS/FTD mouse model in a dose-dependent manner. Acta Neuropathol Commun. 2019;7(1):32.

26. Burberry A, Suzuki N, Wang J-Y, Moccia R, Mordes DA, Stewart MH, et al. Loss-of-function mutations in the C9ORF72 mouse ortholog cause fatal autoimmune disease. Science translational medicine. 2016;8(347):347ra93.

27. Liu Y, Pattamatta A, Zu T, Reid T, Bardhi O, Borchelt DR, et al. C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD. Neuron. 2016;90(3):521-34.

28. McQuade A, Kang YJ, Hasselmann J, Jairaman A, Sotelo A, Coburn M, et al. Gene expression and functional deficits underlie TREM2-knockout microglia responses in human models of Alzheimer's disease. Nature communications. 2020;11(1):5370.

29. Olah M, Menon V, Habib N, Taga MF, Ma Y, Yung CJ, et al. Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. Nature communications. 2020;11(1):6129.

30. Masuda T, Sankowski R, Staszewski O, Böttcher C, Amann L, Sagar, et al. Spatial and temporal heterogeneity of mouse and human microglia at single-cell resolution. Nature. 2019;566(7744):388-92.

31. Ziff OJ, Clarke BE, Taha DM, Crerar H, Luscombe NM, Patani R. Meta-analysis of human and mouse ALS astrocytes reveals multi-omic signatures of inflammatory reactive states. Genome research. 2022;32(1):71-84.

32. Neurons shape the identity of microglia in the brain. Nature. 2022.

33.      Jha MK, Jo M, Kim J-H, Suk K. Microglia-Astrocyte Crosstalk: An Intimate Molecular Conversation. Neuroscientist. 2019;25(3):227-40.

34.      Pascual O, Ben Achour S, Rostaing P, Triller A, Bessis A. Microglia activation triggers astrocyte-mediated modulation of excitatory neurotransmission. Proc Natl Acad Sci U S A. 2012;109(4):E197-205.

35.      Rothhammer V, Borucki DM, Tjon EC, Takenaka MC, Chao CC, Ardura-Fabregat A, et al. Microglial control of astrocytes in response to microbial metabolites. Nature. 2018;557(7707):724-8.

36.      Liddelow SA, Guttenplan KA, Clarke LE, Bennett FC, Bohlen CJ, Schirmer L, et al. Neurotoxic reactive astrocytes are induced by activated microglia. Nature. 2017;541(7638):481-7.

37.      Bezzi P, Domercq M, Brambilla L, Galli R, Schols D, De Clercq E, et al. CXCR4-activated astrocyte glutamate release via TNFalpha: amplification by microglia triggers neurotoxicity. Nat Neurosci. 2001;4(7):702-10.

38.      O'Rourke JG, Bogdanik L, Yáñez A, Lall D, Wolf AJ, Muhammad AKMG, et al. C9orf72 is required for proper macrophage and microglial function in mice. Science (New York, NY). 2016;351(6279):1324-9.

39.      Escartin C, Galea E, Lakatos A, O'Callaghan JP, Petzold GC, Serrano-Pozo A, et al. Reactive astrocyte nomenclature, definitions, and future directions. Nat Neurosci. 2021;24(3):312-25.

40.      Körner S, Thau-Habermann N, Kefalakes E, Bursch F, Petri S. Expression of the axon-guidance protein receptor Neuropilin 1 is increased in the spinal cord and decreased in muscle of a mouse model of amyotrophic lateral sclerosis. The European journal of neuroscience. 2019;49(11):1529-43.

41.      Um JW, Choi TY, Kang H, Cho YS, Choii G, Uvarov P, et al. LRRTM3 Regulates Excitatory Synapse Development through Alternative Splicing and Neurexin Binding. Cell Rep. 2016;14(4):808-22.

42.      Habib N, McCabe C, Medina S, Varshavsky M, Kitsberg D, Dvir-Szternfeld R, et al. Disease-associated astrocytes in Alzheimer's disease and aging. Nat Neurosci. 2020;23(6):701-6.

43.      MacLean M, López-Díez R, Vasquez C, Gugger PF, Schmidt AM. Neuronal-glial communication perturbations in murine SOD1(G93A) spinal cord. Communications biology. 2022;5(1):177.

44.      O'Leary LA, Davoli MA, Belliveau C, Tanti A, Ma JC, Farmer WT, et al. Characterization of Vimentin-Immunoreactive Astrocytes in the Human Brain. Front Neuroanat. 2020;14:31.

45.      Guttenplan KA, Weigel MK, Adler DI, Couthouis J, Liddelow SA, Gitler AD, et al. Knockout of reactive astrocyte activating factors slows disease progression in an ALS mouse model. Nature communications. 2020;11(1):3753.

46.      Beers DR, Appel SH. Immune dysregulation in amyotrophic lateral sclerosis: mechanisms and emerging therapies. The Lancet Neurology. 2019;18(2):211-20.

47.      Zhao W, Beers DR, Thonhoff JR, Thome AD, Faridar A, Wang J, et al. Immunosuppressive Functions of M2 Macrophages Derived from iPSCs of Patients with ALS and Healthy Controls. iScience. 2020;23(6):101192.

48.      Lorenzini I, Alsop E, Levy J, Gittings LM, Rabichow BE, Lall D, et al. Activated iPSC-microglia from C9orf72 ALS/FTD patients exhibit endosomal-lysosomal dysfunction. bioRxiv. 2020:2020.09.03.277459.

49.      Henkel JS, Beers DR, Zhao W, Appel SH. Microglia in ALS: the good, the bad, and the resting. Journal of neuroimmune pharmacology : the official journal of the Society on NeuroImmune Pharmacology. 2009;4(4):389-98.

50.      Kim G, Gautier O, Tassoni-Tsuchida E, Ma XR, Gitler AD. ALS Genetics: Gains, Losses, and Implications for Future Therapies. Neuron. 2020;108(5):822-42.

51.     Balendra R, Isaacs AM. C9orf72-mediated ALS and FTD: multiple pathways to disease. Nat Rev Neurol. 2018;14(9):544-58.

52.     Farg MA, Sundaramoorthy V, Sultana JM, Yang S, Atkinson RAK, Levina V, et al. C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. Human molecular genetics. 2014;23(13):3579-95.

53.     Gendron TF, Petrucelli L. Disease Mechanisms of C9ORF72 Repeat Expansions. Cold Spring Harbor perspectives in medicine. 2018;8(4).

54.     Zhou Q, Mareljic N, Michaelsen M, Parhizkar S, Heindl S, Nuscher B, et al. Active poly-GA vaccination prevents microglia activation and motor deficits in a C9orf72 mouse model. EMBO Mol Med. 2020;12(2):e10919.

55.     Hao Z, Liu L, Tao Z, Wang R, Ren H, Sun H, et al. Motor dysfunction and neurodegeneration in a C9orf72 mouse line expressing poly-PR. Nature communications. 2019;10(1):2906.

56.     Schludi MH, Becker L, Garrett L, Gendron TF, Zhou Q, Schreiber F, et al. Spinal poly-GA inclusions in a C9orf72 mouse model trigger motor deficits and inflammation without neuron loss. Acta neuropathologica. 2017;134(2):241-54.

57.     Chew J, Cook C, Gendron TF, Jansen-West K, Del Rosso G, Daughrity LM, et al. Aberrant deposition of stress granule-resident proteins linked to C9orf72-associated TDP-43 proteinopathy. Molecular neurodegeneration. 2019;14(1):9.

58.     Sudria-Lopez E, Koppers M, de Wit M, van der Meer C, Westeneng H-J, Zundel CAC, et al. Full ablation of C9orf72 in mice causes immune system-related pathology and neoplastic events but no motor neuron defects. Acta neuropathologica. 2016;132(1):145-7.

59.     Burberry A, Suzuki N, Wang JY, Moccia R, Mordes DA, Stewart MH, et al. Loss-of-function mutations in the C9ORF72 mouse ortholog cause fatal autoimmune disease. Sci Transl Med. 2016;8(347):347ra93.

60.     O'Rourke JG, Bogdanik L, Yanez A, Lall D, Wolf AJ, Muhammad AKMG, et al. C9orf72 is required for proper macrophage and microglial function in mice. Science. 2016;351(6279):1324-9.

61.     Nataf S, Pays L. Gene co-expression analysis unravels a link between C9orf72 and RNA metabolism in myeloid cells. Acta Neuropathol Commun. 2015;3:64.

62.     Prudencio M, Belzil VV, Batra R, Ross CA, Gendron TF, Pregent LJ, et al. Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. Nat Neurosci. 2015;18(8):1175-82.

63.     Pal A, Kretner B, Abo-Rady M, Glaβ H, Dash BP, Naumann M, et al. Concomitant gain and loss of function pathomechanisms in C9ORF72 amyotrophic lateral sclerosis. Life science alliance. 2021;4(4).

64.     Shao Q, Liang C, Chang Q, Zhang W, Yang M, Chen J-F. C9orf72 deficiency promotes motor deficits of a C9ALS/FTD mouse model in a dose-dependent manner. Acta Neuropathologica Communications. 2019;7(1):32.

65.     Zhu Q, Jiang J, Gendron TF, McAlonis-Downes M, Jiang L, Taylor A, et al. Reduced C9ORF72 function exacerbates gain of toxicity from ALS/FTD-causing repeat expansion in C9orf72. Nature neuroscience. 2020;23(5):615-24.

66.     Cykowski MD, Dickson DW, Powell SZ, Arumanayagam AS, Rivera AL, Appel SH. Dipeptide repeat (DPR) pathology in the skeletal muscle of ALS patients with C9ORF72 repeat expansion. Acta neuropathologica. 2019;138(4):667-70.

67.     Jiang J, Zhu Q, Gendron TF, Saberi S, McAlonis-Downes M, Seelman A, et al. Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. Neuron. 2016;90(3):535-50.

68.     Beckers J, Tharkeshwar AK, Van Damme P. C9orf72 ALS-FTD: recent evidence for dysregulation of the autophagy-lysosome pathway at multiple levels. Autophagy. 2021:1-17.

69.     Rizzu P, Blauwendraat C, Heetveld S, Lynes EM, Castillo-Lizardo M, Dhingra A, et al. C9orf72 is differentially expressed in the central nervous system and myeloid cells and consistently reduced in C9orf72, MAPT and GRN mutation carriers. Acta Neuropathol Commun. 2016;4(1):37.

70.     Zhao C, Devlin A-C, Chouhan AK, Selvaraj BT, Stavrou M, Burr K, et al. Mutant C9orf72 human iPSC-derived astrocytes cause non-cell autonomous motor neuron pathophysiology. Glia. 2020;68(5):1046-64.

71.     Szebényi K, Wenger LMD, Sun Y, Dunn AWE, Limegrover CA, Gibbons GM, et al. Human ALS/FTD brain organoid slice cultures display distinct early astrocyte and targetable neuronal pathology. Nat Neurosci. 2021;24(11):1542-54.

**Figure legends (for main text only)**

**Figure 1. sn-RNA sequencing of human post mortem motor cortex and spinal cord of ALS patients.** (**a**) experimental outline of this study. (**b**) UMAP plot visualizing 93,176 nuclei isolated from post-mortem cortical samples of ALS patients (C9-ALS and sALS) and controls. Cells are colored according to their assigned cluster identity: microglia, astrocytes, oligodendrocytes, oligodendrocyte precursor cells, motor neurons, excitatory neurons, inhibitory neurons, endothelial cells and T-cells. (**c**) UMAP plot visualizing 63,076 nuclei isolated from post-mortem spinal cord samples of ALS patients (C9-ALS and sALS) and controls. Cells are colored according to their assigned cluster identity: microglia, astrocytes, oligodendrocytes, oligodendrocyte precursor cells, motor neurons, cholinergic neurons, Schwann cells, endothelial cells and an unannotated cluster. Heatmap showing the top 10 most expressed genes per cluster of the (**d**) motor cortex and (**e**) the spinal cord.

**Figure 2: C9orf72 HRE induces changes in the microglial transcriptomic profile compared to sALS.** (**a**) UMAP plot visualizing 16,763 microglia nuclei from post-mortem cortical samples of ALS patients (C9 and sALS) and controls. Microglia subsets are colored based on their assigned cluster into: MG1, MG2, MG3, MG4, MG5, MG6. (**b**) UMAP plot displaying the specific distribution of 4,216 nuclei from the motor cortex and 12,547 nuclei of the spinal cord. (**c**) heatmap showing the top 20 most expressed genes per cluster. (**d**) proportion of cells (y-axis) from each genotype in the different microglial subclusters, alternating shades indicate different individuals (x-axis). (**e**) heatmap depicting the Z-scored average expression of C9orf72 in both motor cortex and spinal cord. (**f**) heatmaps showing the log2FC of differentially expressed genes (p-value adjusted <0.05) of both C9-ALS or sALS patients compared to healthy controls when comparing pooling the motor cortex and the spinal cord. (**g**) Volcano plot depicting 2 pairwise comparisons: between C9-ALS and healthy individuals, and sALS and healthy individuals for the combined microglial dataset (DESeq2 negative binomial distribution, p-values adjusted with Bonferroni correction based on the total number of genes in the dataset). Vertical lines depict Log2FC threshold of 0.2 whereas the horizontal sets the adjusted p-value threshold at < 0.05). (**h**) Volcano plot depicting the pairwise comparisons of C9-ALS vs. sALS for the combined microglial dataset (DESeq2 negative binomial distribution, p-values adjusted with Bonferroni correction based on the total number of genes in the dataset). Vertical lines depict Log2FC threshold of 0.2 and the horizontal line, the adjusted p-value threshold at < 0.05). Colors indicate marker genes of microglia clusters.

**Figure 3: Microglia from C9orf72 HRE patients show altered functional pathways**. (**a**) Seurat module scores of disease-associated (DAM), antigen-presenting HLA, and homeostatic microglia phenotypes as described by Mancuso et al., 2022. (**b**) heatmaps depicting Z-scored average expression of differentially expressed genes (log2FC < +/-0,20, p-value adjusted <0,05) of gene sets of lysosomes, inflammatory response, phagocytosis and KEGG ribosomes. (**c**), Seurat module scores of GO terms from the lysosomes, inflammatory response, phagocytosis and KEGG ribosome. (**d**) dendrogram showing the result of unsupervised hierarchical clustering of modules identified by WGCNA. (**e**) barplots depicting the number of genes per module (**f**) boxplots visualizing the module eigengenes (y-axis) per genotype (x-axis) for the blue, turquoise and yellow cluster of the spinal cord. (**g**) pathway enrichment analysis of the WGCNA clusters blue, turquoise and yellow, with the associated GO term (y-axis) and –log10(p-value) (x-axis).

18

**Figure 4: Astrocytes are differentially affected in C9-ALS compared to sALS**. (**a**) UMAP plot visualizing 22,635 astrocytes from post-mortem cortical samples of ALS patients (C9-ALS and sALS) and controls. Astrocyte subsets are colored based on their assigned cluster: AS1, AS2, AS3, AS4, AS5, AS6, AS7, AS8. (**b**) UMAP plot displaying the specific distribution of the 9,657 cortical and the 12,978 spinal nuclei. (**c**) Heatmap showing the top 15 most expressed genes per cluster. (**d**) proportion of cells (y-axis) from each genotype in the different astrocyte subclusters, alternating shades indicate different individuals (x-axis). (**e**), heatmaps showing the log2FC of differentially expressed genes (p-value adjusted <0.05) of both C9-ALS or sALS compared to healthy controls. (**f**) Volcano plot depicting 2 pairwise comparisons between C9-ALS and controls, and sALS and controls for the combined astrocytes dataset (DESeq2 negative binomial distribution, p-values adjusted with Bonferroni correction based on the total number of genes in the dataset). Vertical lines depict Log2FC threshold of 0.2 and the horizontal line, the adjusted p-value threshold at <0.05). (**g**) Volcano plot showing a pairwise direct comparisons of C9-ALS vs. sALS for the combined astrocyte dataset (DESeq2 Wilcoxon Rank sumtest, p-values adjusted with Bonferroni correction based on the total number of genes in the dataset). Vertical lines depict Log2FC threshold of 0.2 and the horizontal line, the adjusted p-value threshold at < 0.05). Colors indicate marker genes of astrocyte clusters.

19

**Tables**

Supplementary Table 1. The demographic data of participants.

Supplementary Table 2. Overview of Microglia and Astrocytes gene markers.

Supplementary Table 3. Overview of DEG's microglia in three conditions

Supplementary Table 4. Overview of DEG's astrocytes in three conditions

## Methods (online)

### Human donor tissue

The brain and spinal cord samples from were obtained from the Department of Pathology at University Hospitals Leuven (UZ Leuven), the Netherlands Brain Bank (NBB), Netherlands Institute for Neuroscience (NIN), Amsterdam and Hospital Clinic de Barcelona (HCB) and the d'Investgiacions Biomédiques August Pi I Sunyer (IDIBAPS). All material has been collected from donors for or from whom a written informed consent for a brain autopsy and the use of the material and clinical information for research purposes had been obtained by the UZ Leuven (S65097, S59292, S60803) NBB and HCB-IDIBAPS.

### Experiments

*RNA isolation and assessing of RNA integrity of post-mortem human brain and spinal cord tissues*

1ml TriPure per 50-100 mg post-mortem tissue was added into a vial containing beads and post-mortem tissue. The sample was homogenized in tissue homogenizer (Magnalyser) 3 x 30 sec. at 6500 rpm. The homogenate was transferred to a new RNase free tube. 200 µl chloroform (= $1/5^{th}$ volume of TriPure) was added. The sample was mixed by inversion during 15 sec. and incubated at room temperature between 2 and 10 min. Next, the samples were centrifuged for 15 min. at 12000 rpm, 4°C. 3 phases became visible: aqueous phase (colorless) containing RNA, interphase (white color) containing DNA and organic phase (red color) containing protein. The aqueous phase was transferred to a new RNase free tube. Only in case of low expected yields of RNA, 1 µl glycogen was added and precipitated with 500 µl isopropanol (= 1/2th volume of TriPure). The sample was mixed by inversion (15 sec.), incubated at room temperature between 5 and 10 min., and centrifuged for 10 min. at 12000 rpm, 4°C. The supernatant was discarded. The pellet was washed with 1ml 75% EtOH (same volume as TriPure) and centrifuged for 5 min. at 12000 rpm, 4°C. Again, the supernatant was discarded. After drying of the pellet, RNase free water was added. Per sample, an amount of 1 µg of total RNA was used as input. RNA concentration and purity were determined spectrophotometrically using the NanoDrop ND-1000 (NanoDrop Technologies) and RNA integrity was assessed using a Bioanalyzer 2100 (Agilent).

*Nuclei isolation of post-mortem human brain and spinal cord tissues*

Recent advances in the isolation and sequencing of snRNA have demonstrated that, with optimal buffer/detergent compositions, the assay performance and gene detection sensitivity can be significantly enhanced. We used three different strategies including the protocol by standard nuclei isolation protocol Habib et al. (13), Tween with salts and Tris (TST) (14), and Soma-seq (15).

*Standard nuclei isolation protocol*

We used tissue pieces smaller than 50 mg for nuclei. They were transferred immediately in homogenizer containing 1 mL of ice-cold homogenization buffer (HB) (320 mM Sucrose, 5 mM $CaCl_2$, 3 mM $Mg(OAc)_2$, 10 mM Tris 7.8, 0.1 mM EDTA, 0.1% IGEPAL CA-360, 0.1 mM Phenylmethylsulfonyl fluoride, 1 mM β-mercaptoethanol with 5µl Rnasin Plus). The tissues were homogenized with 10 manual gentle strokes (pestle A) + 10 manual gentle strokes (pestle B). The homogenate was filtered through 70 µm cell mesh and was washed strainer with 1.65 ml HB and was added to a final volume of 2.65 ml. The nuclei was homogenated in the HB with 2.65 ml of Gradient Medium (5 mM $CaCl_2$, 50% Optiprep, 3 mM $Mg(OAc)_2$, 10 mM Tris 7.8, 0.1 mM Phenylmethylsulfonyl fluoride, 1 mM β-mercaptoethanol). 29% cushion was prepared by dilution of Optiprep with Optiprep Diluent Medium (150 mM KCl, 30 mM $MgCl_2$, 60 mM Tris pH 8.0, 250 mM sucrose). The nuclei suspension in the HB + GM mix was

21

layered over the 29 % cushion and centrifuged in an SW 41 Ti Swinging-Bucket Rotor (Beckman Coulter), at 7.700 rpm and 4°C for 30 min. The supernatant was removed with a Pasteur pipette andlower supernatant was removedwith P200. Nuclei was resuspended in 50 uL of Resuspension Buffer (PBS, 1% BSA) and transferred to a new tube. Resuspended nuclei was counted using LUNA dual florescence cell counter (Logos Biosystems).

*TST nuclei isolation protocol*

Similar to standard protocol above, tissue chunks less than 50 mg were used for nuclei isolation and transferred immediately in homogenizer containing 1 mL of ice-cold homogenization buffer (146 mM NaCl, 10 mM Tris 7.5, 1 mM $CaCl_2$, 21 mM $MgCl_2$, 0.03% Tween-20, 1% BSA, 25 mM KCl, 250 mM sucrose, 1 mM β-mercaptoethanol, 0.5X cOmplete™ Protease Inhibitor (Roche), 0.2U/ul RNasin® Plus Ribonuclease Inhibitor (Promega)). The tissues were homogenized with (n=15) manual gentle strokes with pestle A only. Nuclei suspension was filtered using 40 µm cell strainer. The filtrate was transferred to a low bind eppendorf and centrifuged at 500xg for 5 min. The supernatant was discarded. The pellet was resuspended  thoroughly in 1 ml of HB w/o Tween buffer (146 mM NaCl, 10mM Tris 7.5, 1 mM $CaCl_2$, 21 mM $MgCl_2$, 1% BSA, 25 mM KCl, 250 mM sucrose, 1 mM β-mercaptoethanol, 0.5X cOmplete™ Protease Inhibitor (Roche), 0.2U/ul RNasin® Plus Ribonuclease Inhibitor (Promega), and transferred to 15 ml falcon. Additional 1.65 mL HB w/o tween buffer was added to a final volume of 2.65 ml. Rest of the protocol for density centrifugation and nuclei enrichment was identical to the standard protocol.

*Soma-seq nuclei isolation protocol*

The frozen tissue chunks were immediately transfered to homogenizer containing 1 mL of ice-cold homogenization buffer (10 mM Tris 8.0, 5 mM $MgCl_2$, 25 mM KCl, 250 mM sucrose, 1 mM β-mercaptoethanol, 0.5X cOmplete™ Protease Inhibitor (Roche), 0.2U/ul RNasin® Plus Ribonuclease Inhibitor (Promega). The tissues were homogenized with (n=15) manual gentle strokes with pestle A only. The nuclei suspension was filtered using 70µm cell strainer. The the filtrate was transfered to a low bind eppendorf and centrifuged at 400xg for 5 min. The supernatant was discarded. RThe pellet was resuspended thoroughly in 1 ml of HB buffer, and transfered to 15 ml falcon. Additional 1.65mL HB buffer was added to a final volume of 2.65 ml. For this protocol as well the density centrifugation and nuclei enrichment is identical to what was done with standard nuclei isolation protocol.

*Library preparation*

Library preparations for the scRNA-seq was performed using 10X Genomics Chromium Single Cell 3' Kit, v3.1 NextGEM chemistry (10X Genomics, Pleasanton, CA, USA). The cell count and the viability of the samples were assessed using LUNA dual fluorescence cell counter (Logos Biosystems) and a targeted cell recovery of 6000 cells were aimed for each of the samples. After cell count and quality control (QC), the samples were immediately loaded on to the Chromium Controller. Multiple samples (up to three patients) were pooled at equal counts and were loaded to a single lane of 10X. Single cell RNAseq libraries were prepared using manufacturers recommendations (Single cell 3' reagent kits v3.1 user guide; CG000204 Rev D), and at the different check points the library quality was accessed using Qubit (ThermoFisher) and Bioanalyzer (Agilent). cDNA libraries were sequenced with an expected coverage of 50,000 reads per cell, either on Illumina's NovaSeq 6000 platform using paired-end sequencing workflow and with recommended 10X; v3.1 read parameters (28-8-0-91 cycles).

## Data analysis

### Data pre-processing

Raw sequencing data was first demultiplexed using the cellranger (5.0.1) mkfastq command, the resulting fastq files were processed using the cellranger count command with the --include-introns option enabled and using the GRCh38-2020-A index provided on the 10X Genomics website. Cells passing the default cellranger filter were kept for subsequent analyses.

Samples that were multiplexed for droplet generation were demultiplexed into patient specific matrices using the following procedure. First, a library of SNPs present in the general population identified by the 1000 Genomes Project was obtained from http://ftp.1000genomes.ebi.ac.uk specifically, files ALL.chr(1-22,X,Y)_GRCh38_sites.20170504.vcf.gz were obtained and merged into a single vcf. Chromosomes were renamed to match the scheme used in the index provided by 10X genomics using bcftools annotate. These SNPs were then filtered to remove rare variants as well as to only include SNPs present in the regions of the genome included in the cellranger index using bcftools filter with the parameters --regions-file (with a bed file containing the above mentioned regions) and -i "INFO/AF[0] < 0.90 && INFO/AF[0] > 0.10". This final vcf file was then used to run freemuxlet from the popscle package [https://github.com/statgen/popscle] with default parameters other than nsample = 2 and using the tools from https://github.com/aertslab/popscle_helper_tools to speed up computation. As freemuxlet only provides clusters of cells with no labels, we used verifyBamID [http://griffan.github.io/VerifyBamID/] to compare each genotype provided by freemuxlet to the bulk RNA-seq data unique to each patient, the highest scoring match was used to assign patient labels to freemuxlet clusters.

### Data processing and cell type annotation

Libraries were individually processed using Scanpy 1.7.2 (72). First, patient labels, and other associated metadata were added to the cells, with any doublets identified by freemuxlet being removed. Next, scrublet was used to remove any remaining identifiable doublets (threshold = 0.25) and all samples were combined into two main datasets based on their tissue of origin; motor cortex and spinal cord. Nuclei with less than 250 genes or more than 5% of total counts associated to mitochondrial genes were removed, the resulting matrices were normalised (to 10000 counts), log transformed, and highly variable genes were identified using default parameters. Both number of UMIs and percentage of mitochondrial reads were regressed out and the data was scaled to a maximum of 10. A principal component analysis (PCA) was used to reduce the dimensions of the data, Harmony (73) was used to integrate the PCA components based on the "Patient" batch key and the pcacv workflow from vsn-pipelines [https://doi.org/10.5281/zenodo.3703108] was then used to determine the number of principal components (PCs) to continue with (76 for Motor Cortex and 82 for Spinal Cord). Neighborhood embedding, UMAP and tSNE were calculated using the number of PCs listed above and clusters were identified at various resolutions (0.4, 0.8, 1.0, 1.2, 1.6, 2.0, 4.0, 8.0) using the leiden algorithm. The resulting data was exported to a loom file and uploaded to SCope (74), cell type annotations were provided via SCope by experts based on marker gene expression. All plots were generated with ggplot2 (v 3.3.5) in R v4.1.2.

### Processing of the microglia and astrocytes

To investigate microglia subsets, we isolated microglial nuclei from the single-nucleus data set (Masrori_Postmortem) for both the motor cortex and spinal cortex and merged them with Seurat (v4.1.0) under R version 4.1.2. The data was log normalized, highest variable genes were identified followed by scaling of the data and PCA. The data was batch corrected with Harmony (v0.1.0) based on patient

identifiers and location. Cells were clustered using the first 20 principal components and UMAP plots were generated with Seurat standard methods. Non- microglial cells were excluded to perform final clustering of pure microglia.

When starting the reanalysis of the microglia extracted from the entire dataset, 5 other cell subsets were identified (peripheral macrophages n=898 , synaptosome n= 821, oligodendrocytes n= 200 and proliferating cells n= 143). The clusters contained low expression of microglia markers CXCR3 and P2RY12 with subset specific markers for peripheral macrophages (*CD163*, *MRC1*, *F13A1*), oligodendrocytes (*MBP*, *SHROOM4*, *PLP1*, *PCDH9*, *EDIL3*) and synaptosomes/astrocytes (*SNAP25*, *RORA*, *DPP10*, *ROBO2*, *LSAMP*). The oligodendrocytes were most likely doublets of microglia and oligodendrocytes, which is why we still found overlap with the microglia markers. The same for the synaptosomes. Additionally, a cluster of proliferating cells (TOP2A, MIK67, MELK, KIF11) was excluded from the final analysis. While all the these subsets were initially defined as microglial cells in the large transcriptomic landscape of both the motor cortex and spinal cord, due to their transcriptomic similarity to microglia, the reanalysis of the microglial subset made these differences more apparent, driving us to exclude them.

Next, we extracted the astrocytes (n= 23483) from the spinal cord and motor cortex and re-identified the astrocyte population. Apart from astrocytes, two other clusters were identified: one of oligodendrocytes, n= 500 (*ELMO1*, *IL1RAPL1*, *SLC24A2*, *SHROOM4*) and one with doublets/low quality n =348 cells (*PCLO*, *WARS*, *RUNX1*), the latter being a cluster with markers of multiple cell types that could not be confidently annotated.

All plots were generated with ggplot2 (v 3.3.5) inR v4.1.2.

*Differential expression*

Differential expression testing was performed using *DESEQ2* in Seurat Findmarkers. Genes with a p-value adjusted of <0.05 and log2FC of > +/- 0.20 were considered as differentially expressed. Heatmaps were generated with ComplexHeatmap_2.10.0.

*Gene enrichment analysis*

Specific microglia functions were assessed by extracting human specific genes (taxon: 9606) 3 GO terms (GO:0005764, GO:0006954 and GO:006906) and 1 pathway in KEGG (M189). Gene signatures were calculated in Seurat with AddModuleScore.

*WGCNA*

Matrices for microglia and astrocytes were extracted from each tissue and were used to perform WGCNA (75) to identify modules of co-expressed genes using the hdWGCNA package (v0.1.1.9005) (76). Genes were filtered to include only those expressed within at least 5% of the cells within a dataset using the SetupForWGCNA function, and MetacellsByGroups was used to create metacells by the combination of ALS Status, Patient and Library of origin using values of 25 for the number of cells to be combined and a maximum number of shared cells of 10, finally, metacells were normalized using NormalizeMetacells. Next, the groups of interest (Healthy, sALS and C9-ALS) were selected to build the co-expression network using SetDatExpr and TestSoftPowers was used per dataset to select a soft threshold (Spinal Cord – Astrocytes: 5, Spinal Cord – Microglia: 7, Motor Cortex – Astrocytes: 5, Motor Cortex – Microglia: 4). ConstructNetwork was used with default parameters to generate the co-expression networks and modules. Finally, module eigengenes were corrected for between-patient batch effects using harmony.

24

*WGCNA Modules-Trait Relationships*

To identify WGCNA genes modules which may be correlated with ALS status, a one-way ANOVA (via the scipy python package v.1.7.3) was performed using the module eigengene values for each module across all datasets, p-values across all tests were corrected using the Benjamini/Hochberg correction (via the statsmodels python package v.0.12.2). A TukeyHSD test was then performed (also using statsmodels) on modules discovered in the ANOVA as having at least one significantly different mean with the goal of identifying the specific statuses which are different.

*Gene Ontology Enrichment Analysis*

Gene ontology enrichment was performed using the g:OSt functionality of gProfiler (77) via the gprofiler-official python package (v1.0.0).

**Data availability**

Processed data can be visualized via the interactive single cell data viewer SCope at https://scope.aertslab.org/#/Masrori_Postmortem.

Raw sequencing data have been deposited in the European Genome-phenome Archive (EGA) under and with data accession no. EGA… (to access the data itself under restricted access). Requests for accessing raw sequencing data will be reviewed by the VIB data access committee (dac@vib.be). Any data shared will be released via a Data Transfer Agreement that will include the necessary conditions to guarantee protection of personal data (according to European GDPR law).

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Philip Van Damme (philip.vandamme@kuleuven.be)

**Methods (only references)**

13.     Habib N, Avraham-Davidi I, Basu A, Burks T, Shekhar K, Hofree M, et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. Nat Methods. 2017;14(10):955-8.

14.     Slyper M, Porter CBM, Ashenberg O, Waldman J, Drokhlyansky E, Wakiro I, et al. A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. Nat Med. 2020;26(5):792-802.

15.     Otero-Garcia M, Xue Y-Q, Shakouri T, Deng Y, Morabito S, Allison T, et al. Single-soma transcriptomics of tangle-bearing neurons in Alzheimer's disease reveals the signatures of tau-associated synaptic dysfunction. bioRxiv. 2020:2020.05.11.088591.

72.     Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome biology. 2018;19(1):15.

73.     Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019;16(12):1289-96.

74.     Davie K, Janssens J, Koldere D, De Waegeneer M, Pech U, Kreft Ł, et al. A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. Cell. 2018;174(4):982-98.e20.

75.     Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics. 2008;9:559.

76.     Morabito S, Miyoshi E, Michael N, Shahin S, Martini AC, Head E, et al. Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. Nat Genet. 2021;53(8):1143-55.

77.     Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 2019;47(W1):W191-w8.
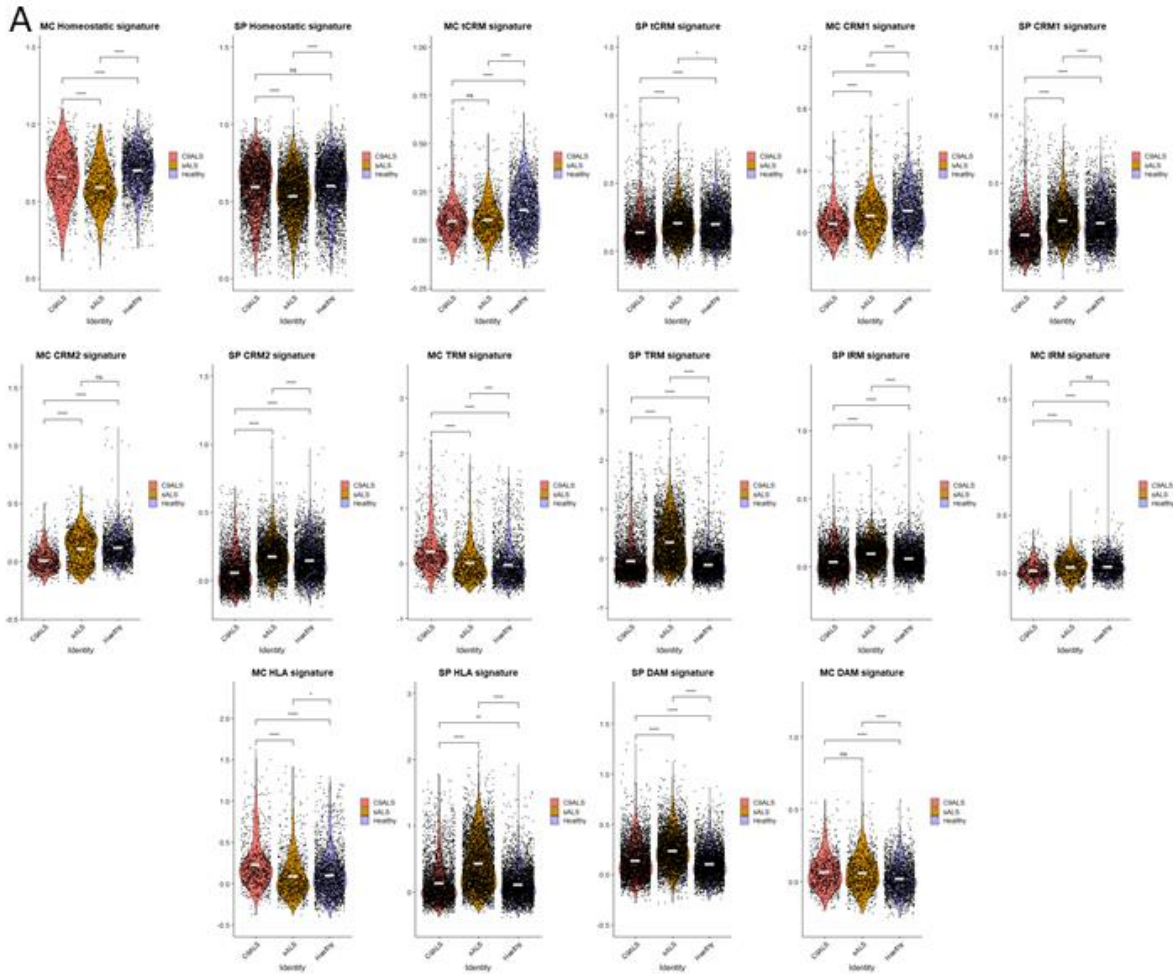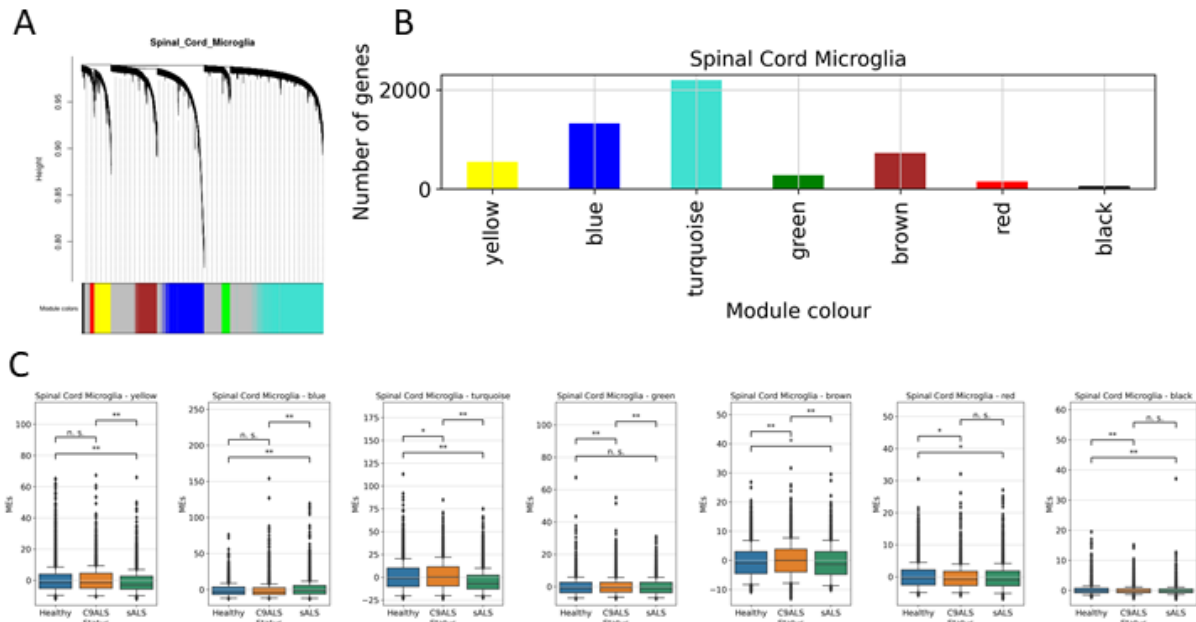
**Supplementary figures**
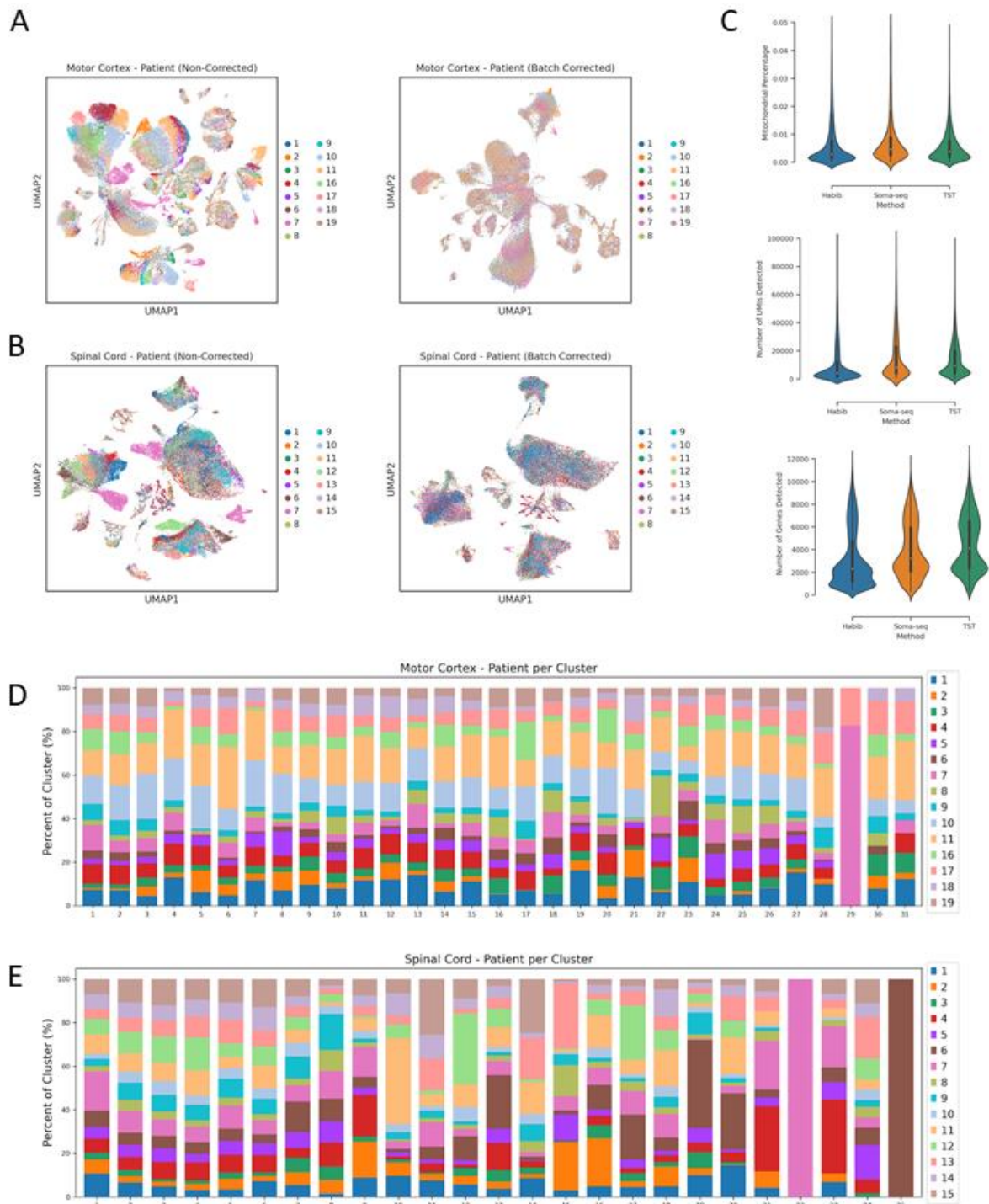


**Supplementary Figure 1. Overview of the full dataset and quality control.** UMAP plot visualization of the single nuclei from (**a**) motor cortex and (**b**) prior and after batch correction. (**c**) Quality control of the three isolation protocols (Habib, Soma-seq, TST): percentage of mitochondrial RNA, number of molecules detected per cell, number of genes detected per cell. Contribution of each patient (x-axis) to each cluster of the (**d**) motor cortex and (**e**) spinal cord.
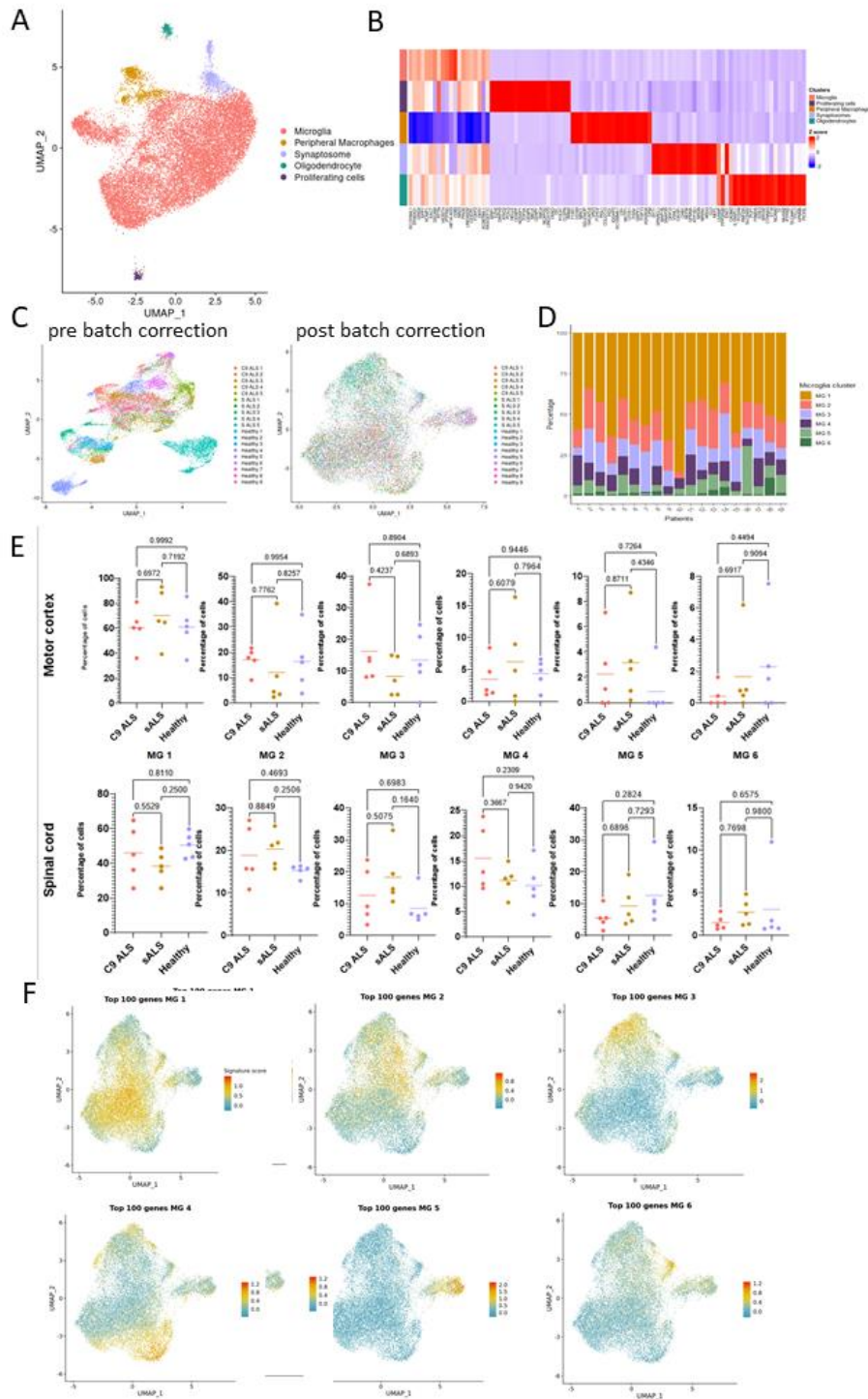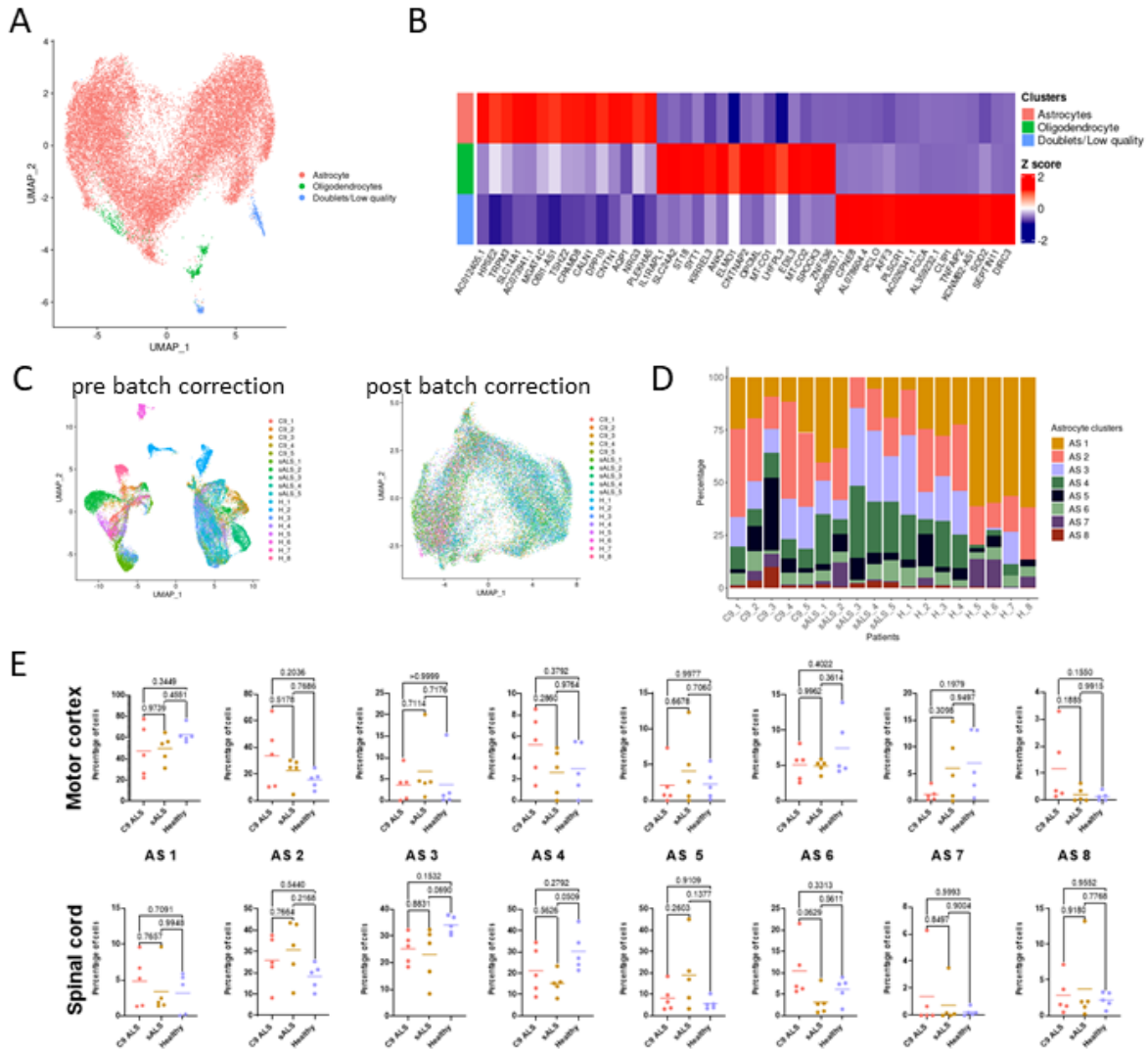
**Supplementary Figure 2: Extended clustering and quality control in the microglial subset.** (**a**) UMAP plot visualizing 4,845 nuclei from the motor cortex 13,980 nuclei from the spinal cord. Cell populations are colored based on their assigned into the following clusters: microglia, peripheral macrophages, synaptosomes, oligodendrocytes and proliferating cells. (**b**) Heatmap showing the top 20 most expressed genes per cluster. (**c**) UMAP plot visualizing the selected microglia before and after batch correction. (**d**) Proportion of cells (y-axis) from each individual (x-axis) in the different microglia subsets. (**e**) proportion of cells (y-axis) from each genotype (per individual) in the different microglia subsets across both tissue locations. (f) Seurat module scores of the top 100 marker genes from the different microglia subsets.

**Supplementary Figure 3: Extended analysis of the microglia transcriptional signatures.** (**a**) Seurat module scores of homeostatic, transitioning cytokine response (tCRM), cytokine response 1 (CRM1), cytokine response 2 (CRM2), translational response (TRM), interferon response (IRM), antigen-presenting (HLA), disease-associated (DAM), and homeostatic microglia phenotypes as described by Mancuso et al., 2022.

29

**Supplementary Figure 4: Extended WGCNA analysis.** (**a**) Dendrogram showing result of unsupervised hierarchical clustering of modules identified by WGCNA. (**b**) Barplots depicting the number of genes per module. (**c**) Boxplot visualizing the module eigengenes (y-axis) per genotype (x-axis).

**Supplementary Figure 5: Extended clustering and quality control in the astrocytic subset.** (**a**) UMAP plot visualizing 10,119 nuclei from the motor cortex 13,364 nuclei from the spinal cord. Cell populations are colored based on their assigned cluster into: astrocytes, oligodendrocytes, doublets/low quality. (**b**) heatmap showing the top 20 most expressed genes per cluster. (**c**) UMAP plot visualizing the selected astrocytes prior and after batch correction. (**d**) Proportion of cells (y-axis) from each individual (x-axis) contributing to the different astrocyte clusters. (**e**) Proportion of cells (y-axis) from each genotype (per individual) in the different astrocyte subsets across locations.