Article

Single-nucleus sequencing reveals enriched expression of genetic risk factors in extratelencephalic neurons sensitive to degeneration in ALS

Francesco Limone [1,2,3,11] ✉, Daniel A. Mordes [1,2,4,11], Alexander Couto[1], Brian J. Joseph [1], Jana M. Mitchell[1,2], Martine Therrien [2,5], Sulagna Dia Ghosh [1,2,6], Daniel Meyer [2,6], Yingying Zhang[1,2], Melissa Goldman[2,6], Laura Bortolin[2,6], Inma Cobos [4], Beth Stevens[2,5,7], Steven A. McCarroll [2,6], Irena Kadiu[8], Aaron Burberry[1,2,9], Olli Pietiläinen[1,2,10] & Kevin Eggan[1,2] ✉

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder characterized by a progressive loss of motor function linked to degenerating extratelencephalic neurons/Betz cells (ETNs). The reasons why these neurons are selectively affected remain unclear. Here, to understand the unique molecular properties that may sensitize ETNs to ALS, we performed RNA sequencing of 79,169 single nuclei from cortices of patients and controls. In both patients and unaffected individuals, we found significantly higher expression of ALS risk genes in *THY1*+ ETNs, regardless of diagnosis. In patients, this was accompanied by the induction of genes involved in protein homeostasis and stress responses that were significantly induced in a wide collection of ETNs. Examination of oligodendroglial and microglial nuclei revealed patient-specific downregulation of myelinating genes in oligodendrocytes and upregulation of an endolysosomal reactive state in microglia. Our findings suggest that selective vulnerability of extratelencephalic neurons is partly connected to their intrinsic molecular properties sensitizing them to genetics and mechanisms of degeneration.

Amyotrophic lateral sclerosis (ALS) is a neuromuscular disease with survival limited to 2–5 years from onset, the most common motor neuron disease in aging and the neurodegenerative disease with one of the earliest onsets, in the mid-to-late 50s[1]. Although specific genetic causes have been identified, most cases are sporadic (~90%), have no family history and unknown etiology[2], thus rendering modeling of the disease difficult[3]. Variants in genes associated with ALS can contribute to a related disorder, frontotemporal dementia (FTD), leading to the view of ALS and FTD as clinical manifestations of shared molecular causes. Bulk RNA sequencing of ALS postmortem brains have identified differences[4] and similarities between sporadic and familial[5] cases and highlighted shared profiles[6–8]. While they provided valuable insights, these studies had limited resolution on the cell types altered by ALS.

The most striking feature in ALS–FTD are protein aggregates of TAR DNA/RNA-binding protein 43 (TDP-43) in over 95% of ALS cases and ~50% of FTD cases, mostly in neurons[9], providing one shared mechanism. It is still unknown how familial mutations and sporadic onset might converge on the formation of these aggregates and how it specifically affects classes of extratelencephalic corticospinal motor neurons (CSMNs), that is, Betz[10] and von Economo cells[11]. Moreover,

strong evidence demonstrated that cells other than neurons are key mediators of disease progression and it remains unclear how these might contribute to the disease[7,12,13].

Methods to study heterogeneity at a single-cell level have rapidly advanced and their application to human postmortem brain tissue is beginning to emerge. In this Article, we applied single-nucleus RNA sequencing (snRNAseq) and in vitro human induced pluripotent stem cell modeling to investigate changes in cortical cell types in sporadic ALS (sALS). Our profiling identified the intrinsically higher expression of ALS–FTD risk factors in a specific class of extratelencephalic excitatory neurons. In patients with ALS, these neurons and other subclasses of ETNs selectively express higher levels of genes connected to unfolded protein responses and RNA metabolism. We found that excitatory neuronal vulnerability is accompanied by a decrease in myelination-related transcripts in oligodendroglial cells and an upregulation of a reactive, proinflammatory state in microglia. We provide a preliminary, insightful view of disruptions triggered in human motor cortices in ALS and implicate aging-associated mechanisms that link altered proteostasis and inflammation to specific cell types in ALS.

Results

snRNAseq profiling of ALS cortex

We used snRNAseq to profile motor/premotor cortical gray matter from patients with sALS and age-matched controls with no neurological disease using Drop-seq technology[14] (Fig. 1a, Source Data Table 1 and Extended Data Fig. 1a–c). After screening for RNA quality, 79,169 barcoded droplets from eight individuals were analyzed ($n = 5$ sALS and $n = 3$ control), with a mean of 1,269 genes and 2,026 unique molecular identifiers (UMIs) (Extended Data Fig. 1d). We used Seurat[15], the single-cell analysis package, to cluster and annotate groups according to canonical markers of brain cell types[16]: excitatory and inhibitory neurons, oligodendrocytes, oligodendrocyte progenitor cells (OPCs), microglia, astrocytes and endothelial cells (Extended Data Fig. 1e,f). The observed cell type distribution corresponded to previous studies[16] and enabled robust categorization for downstream analysis. Cellular distribution was homogeneous between sexes and individuals, except for a modestly lower number of astrocytes in ALS samples (Extended Data Fig. 1g–i).

Elevated expression of ALS–FTD genes in a specific class of neurons

To potentially identify cell types underlying ALS pathophysiology, we examined the expression of known familial genes for ALS–FTD and variants identified as risk factors from genome-wide association studies (GWAS). These genes were expressed to a variable degree between cell types and many of them were ubiquitously expressed as already known[2,7] (Extended Data Fig. 2a). We then computed a 'module score' for this set of genes[17]; this metric generates a standardized z-score for the expression of each gene and sums it up as a total score for the gene set. A positive score suggests higher expression of the gene set compared with the average expression of the module across the dataset. We computed parallel module scores for lists from the latest GWAS for disorders that affect the cortex but not specifically Betz cells: Alzheimer's disease (AD)[18,19] and multiple sclerosis (MS)[20] (Fig. 1a and Source Data Table 2). No clear enrichment for the ALS–FTD gene list was identified (Fig. 1b), as anticipated by the scattered, ubiquitous expression pattern. On the other hand, AD and MS modules showed respective enrichment in microglia, as expected on the basis of the strong immune signature in these diseases[18–20] (Fig. 1c,d). These results corroborate knowledge in the field, underlying the strength of this analysis, and confirm our results in an unbiased, single-cell resolution.

Considering the selective loss of neurons in ALS[2], we further analyzed these cells. We found 32,810 nuclei from excitatory neurons with unbiased clustering identifying seven subgroups (Exc0–6) expressing known markers of different cortical layers[16] (Extended Data Fig. 2b–f).

Analysis of the ALS–FTD module in these cells showed a positive score in *THY1*-expressing subgroup Exc1 (normalized enrichment score of 1.834) (Fig. 1e and Extended Data Fig. 2g,h) and no significant enrichment for AD and MS modules (Fig. 1f,g). We decided to investigate the identity of these cells and the possibility of them being ETNs.

We identified three subgroups expressing markers of subcerebral projection neurons: Exc1, Exc5 and Exc6 (Fig. 2a). Exc5 and Exc6 expressed canonical genes *FEZF2*, *BCL11B* and *CRYM*[21]; Exc1 expressed *THY1*, enriched in human layer V[22] and used as a reporter for CSMNs[23], and high levels of neurofilament chains, markers of ETNs[24] (Fig. 2b). Recent reports dissected the transcriptomic identity of layer V extratelencephalic neurons in the human motor cortex[24]. We detected ETNs markers in these groups, with Exc1 expressing *SERPINE2* and *POU3F1*, specific of ETNs[24], and *NEFH* and *STMN2*, broad markers of motor neurons (MNs)[25] (Fig. 2c). Owing to the anatomical location of our samples and the presence of ETNs across motor-related areas, we plotted genes specific to layer V ETNs von Economo cells[26,27], affected in FTD and other long-range subcerebral projecting neurons (LR-SCPNs)[28] and confirmed that all three groups expressed these markers (Fig. 2d,e). To further characterize these patterns, we leveraged a publicly available single-cell, spatial dataset of control human dorsal cortex[29]. We confirmed that markers expressed in Exc1 *THY1*, *STMN2* and *SNCG* (Fig. 2f) are specifically expressed in layer V (L5) (Fig. 2g,h and Extended Data Fig. 3a,b). This evidence suggests that Exc1, Exc5 and Exc6 express markers of layer V extratelencephalic neurons of cortical areas affected by ALS–FTD.

To further confirm that *THY1*^high neurons expressed higher levels of ALS–FTD genes, we ran a module score analysis in two datasets that identified *THY1*^high excitatory neurons[22,30]. In these studies, *THY1* neurons expressed ETN markers, layer V, von Economo and LR-SCPN markers (Extended Data Fig. 3c–i) and expressed higher levels of the ALS–FTD module (Extended Data Fig. 3l–m). Analysis of the spatial transcriptomic dataset[29] confirmed that the top ten ALS–FTD-associated genes most highly expressed in Exc1 (Extended Data Fig. 2g) are expressed in deeper layers of the cortex and specifically in layer V (Fig. 2i,j and Extended Data Fig. 3n). Larger cohorts of patients and validations at the protein level are needed to confirm the degree of dependency of ETNs on this gene set but studies in human[31] and mouse[32] showed that deep-layer neurons have a higher propensity to form TDP-43 aggregates, the hallmark of ALS–FTD. Here, we provide a possible link to their specific vulnerability.

Higher cellular burden in deeper-layer excitatory neurons

We next examined whether the enriched expression of ALS–FTD genes might relate to changes that occur in excitatory neurons in response to ALS. We conducted differential gene expression (DGE) analysis between neurons from patients and controls, across all excitatory cells and within each subgroup (Fig. 3a). We selected genes significantly upregulated in patients globally (DGEall) and within each subgroup (DGE0–6), calculated module scores for each set and investigated whether certain neuronal subtypes might have similar responses to ALS (Source Data Table 3). We found a correlation between scores in groups expressing deep-layer markers and the global changes identified in patients (Fig. 3b), suggesting that pathology in lower cortical layers might be driving observed alterations. For instance, groups expressing ETN markers (Exc1, Exc5 and Exc6) shared many upregulated genes with each other and with the global signature (Fig. 3c). Intriguingly, this class of genes is, like genetic risk factors, constitutively expressed at higher levels in Exc1 ETNs (Fig. 3b), advocating for a proposed interplay between genetics and molecular pathways that sensitizes ETNs to ALS[33].

Owing to the small cohort size and because of the diverse etiology of ALS–FTD[6], we decided to test how shared these changes were and test whether they were driven by a single individual. We started by averaging gene expression by sample within groups of ETNs and ran principal component analysis (PCA) on the individual-aggregated matrices that
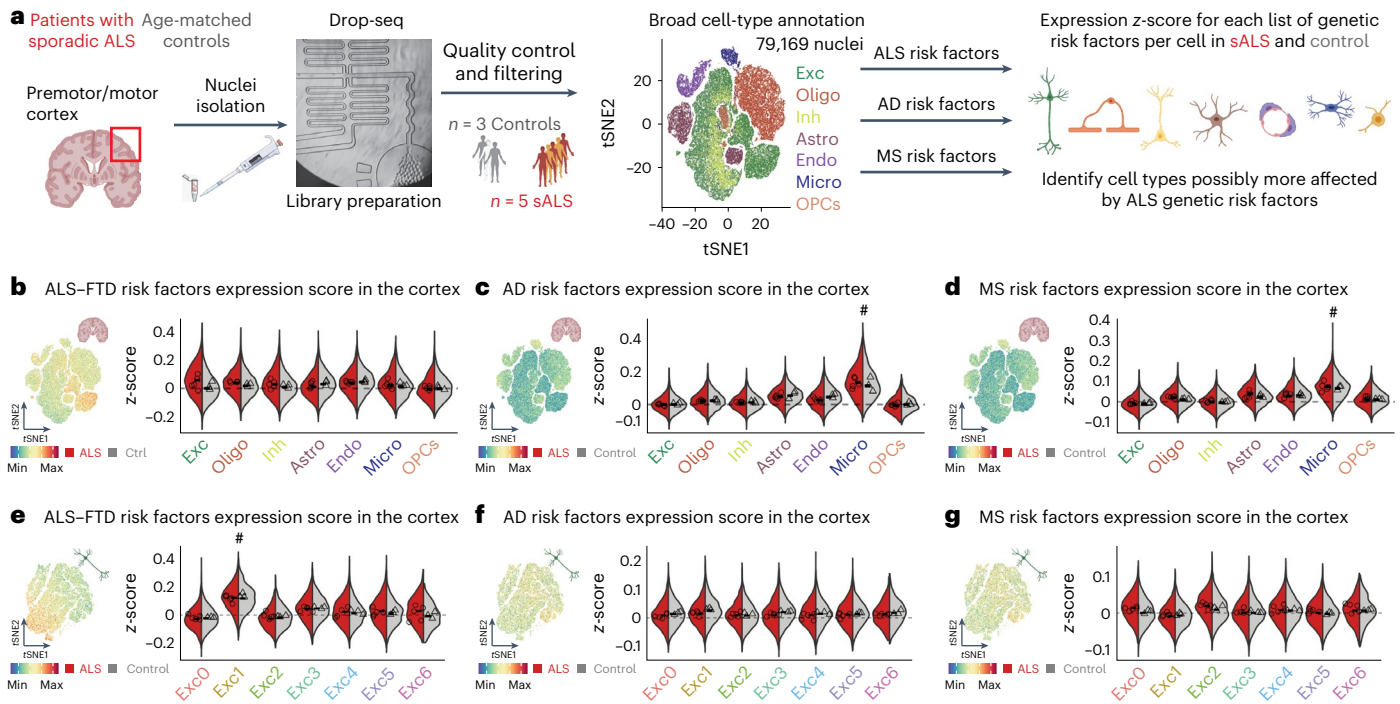
Fig. 1 | Cellular susceptibility to ALS–FTD in the human postmortem cortex. **a**, Diagram of the workflow for isolation of nuclei from cortices of patients with ALS and age-matched controls followed by snRNAseq and assessment of expression of gene modules associated with neurodegenerative diseases. Exc, excitatory neurons; Inh, inhibitory neurons; Oligo, oligodendrocytes; OPCs, oligodendrocyte progenitor cells; Micro, microglia; Astro, astrocytes; Endo, endothelial cells. **b**–**d**, tSNE projections and violin plots of z-scores for expression of genes associated with ALS–FTD (**b**), AD (**c**) and MS (**d**) in the different cell types identified (the bars denote median for each side of the violin and the symbols indicate the average score per individual). **e**–**g**, tSNE projections and violin plots of z-scores for expression of genes associated with the ALS–FTD (**e**), AD (**f**) and MS (**g**) in the different subtypes of excitatory neurons (the bars denote median for each side of the violin and the symbols indicate the average score per individual) (in **a**–**g**, n = 3 control individuals and n = 5 patients with sALS).

showed widespread diversity between both patients and controls with partial segregation by diagnosis (Extended Data Fig. 4a). This could be explained by a compound effect of diversity of human population, sampling of different cells per individual or different disease etiology. To avoid differential gene expression analyses confounded by this effect, we carried out several preanalyses to corroborate the strength of our study and come to more confident conclusions.

We proceeded to compute module scores for the global sALS signature (DGEall) and sALS signatures in ETNs (DGE1,5,6) and saw that, even though with intragroup variability, these signatures showed higher scores in ALS at single-cell and aggregated expression level with similarities between groups of ETNs (Fig. 3d and Extended Data Fig. 4b). To ensure that these signatures were driven by disease changes, we recalculated differentially expressed genes (DEGs) by down-sampling each group of ETNs to the smallest number of cells by diagnosis (that is, redoDGE1,5,6). This analysis confirmed the great overlap of DGE signatures and redoDGE signatures and that these signatures were higher in sALS at single-cell resolution and average level (Extended Data Fig. 4c,d).

Finally, to further test whether the results were driven by the transcriptome of a single individual, we repeated the DGE analysis by excluding cells from one patient at a time. This analysis shows that, even though different subsets of genes are detected in each analysis, the shared ones show a similar direction and amplitude in changes (Extended Data Fig. 4e). This comparison shows that singular gene changes might be sporadically driven by one individual and because of the small cohort size, it would be hard to discern between biological or technical outliers. However, the shared DEGs (>65% in this case; Extended Data Fig. 4f) and the general direction of the ALS-driven signature are maintained, with genes commonly unregulated in at least four out of five patients (ALLminus1 list) having >85% overlap with

DGEall (Extended Data Fig. 4g). We proceeded to use only genes shared by at least four individuals for subsequent analyses for all cell types.

We ran Gene Ontology (GO) analysis and showed that DEGs identified in classes of ETNs are connected to cellular stresses previously associated with ALS even from studies with hundreds of patients[7,34] (Fig. 3f). Interactome analysis revealed coordinated alterations in genes that function in translational machinery, mitochondria, protein folding, proteostasis and degradation pathways connected to the proteasome and shared many transcriptional changes with patients' excitatory cells as a whole (Extended Data Figs. 4k–m and 5). Interestingly, genes upregulated in upper layers of the cortex, a region relatively spared of pathology, shared less similarities with DGEall and were associated with synaptic biology (Extended Data Fig. 4h–j). Comparison of ALS-driven transcriptomic changes with other studies underlined similarities with genes upregulated in excitatory neurons from MS[22] but not from AD patients[35] (Extended Data Fig. 4n,o), suggesting that there are similar processes at the origin of neurodegeneration but that these changes are not universal. The analyses so far have highlighted the interindividual variability intrinsic of this kind of datasets and the particular attention studies need to divert into assuring reproducibility of the results. Nonetheless, we provide a snapshot of disruption in neuronal health in patients with ALS, in which lower layers of cortical excitatory neurons share higher levels of cellular stresses.

Next, we sought to determine what proportion of this transcriptomic signature may be associated with proteostatic stress specifically in neuronal cells. Presently, in vitro modeling of sporadic ALS requires high numbers of lines, high-throughput methods and needs further standardization[3,36]. To overcome these limitations, we implemented transient proteasome inhibition as a highly reproducible, dose-responsive, temporally controlled model to induce TDP-43 nuclear loss as seen in patients' Betz cells and other ALS-related
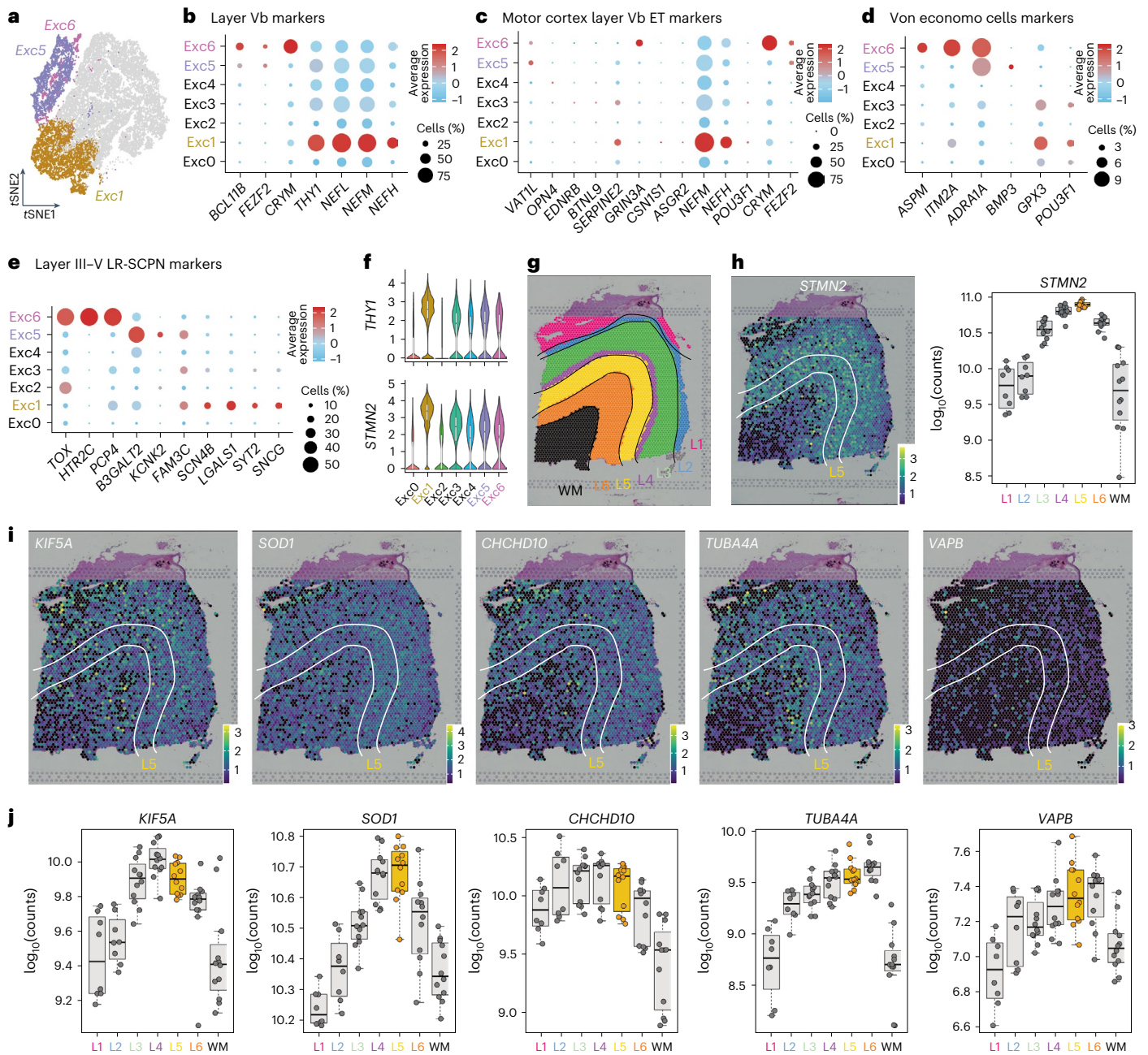
Fig. 2 | ALS–FTD susceptible neurons are layer V ETNs. a, *t*SNE projection of presumptive layer V neurons. **b**, A dot plot representing expression of layer V markers **c**, A dot plot for markers of LVb ETNs of human motor cortex. **d**, A dot plot representing expression of von Economo markers. **e**, A dot plot representing expression LR-SCPN markers. **f**, Representative violin plot for markers of layer V ETNs of human motor cortex (geometric box plots for median and interquantile ranges). **g**, Visual depiction of layers identified by Maynard et al. 2021 (*n* = 7, **g**–**j**, publicly available) (L, layer; WM, white matter). **h**, A spot plot depicting expression of layer Vb motor cortex marker, *STMN2*, identified as enriched in THY1-Exc1, with corresponding box plot quantification. **i**,**j**, Box plots (**j**) and corresponding spot plots (**i**) for the expression of the top five ALS–FTD-associated genes expressed in Exc1 (box plots for mean and interquartile ranges).

dysfunction in human neurons[4,37,38] (Fig. 4a and Extended Data Fig. 6a,b). Application of a proteasome inhibitor to human pluripotent stem (hPS) cell-derived neurons[37] induced nuclear loss of TDP-43 (Fig. 4b). Bulk RNA sequencing analysis showed widespread changes after treatment, with a significant overlap of upregulated genes between stressed hPS cell-derived neurons and sALS neurons (Fig. 4c), specifically proteasome subunits and heat-shock response-associated chaperonins, confirmed by GO analysis of shared genes (Fig. 4d and Extended Data Fig. 6d). Moreover, genes upregulated after inhibition show a significant overlap with transcripts misregulated after downregulation of TDP-43 in neurons[37] (Extended Data Fig. 6e). This

confirms that some changes identified in sALS neurons are connected to neuronally intrinsic proteostasis and are at least in part connected to alterations in TDP-43.

To confirm hindered proteostasis in ALS cortex, we selected a second cohort of patients with sALS and controls for biochemical evaluation. We extracted protein, confirmed increased insoluble TDP-43 in patients (Extended Data Fig. 6f,g) and showed that, despite the presence of core proteosomal subunits, pathology is accompanied by the accumulation of highly ubiquitinated proteins (Fig. 3g,h), the hallmark of impaired proteostasis. These findings suggest that the neuronal stress observed in ETNs from patients with ALS represents
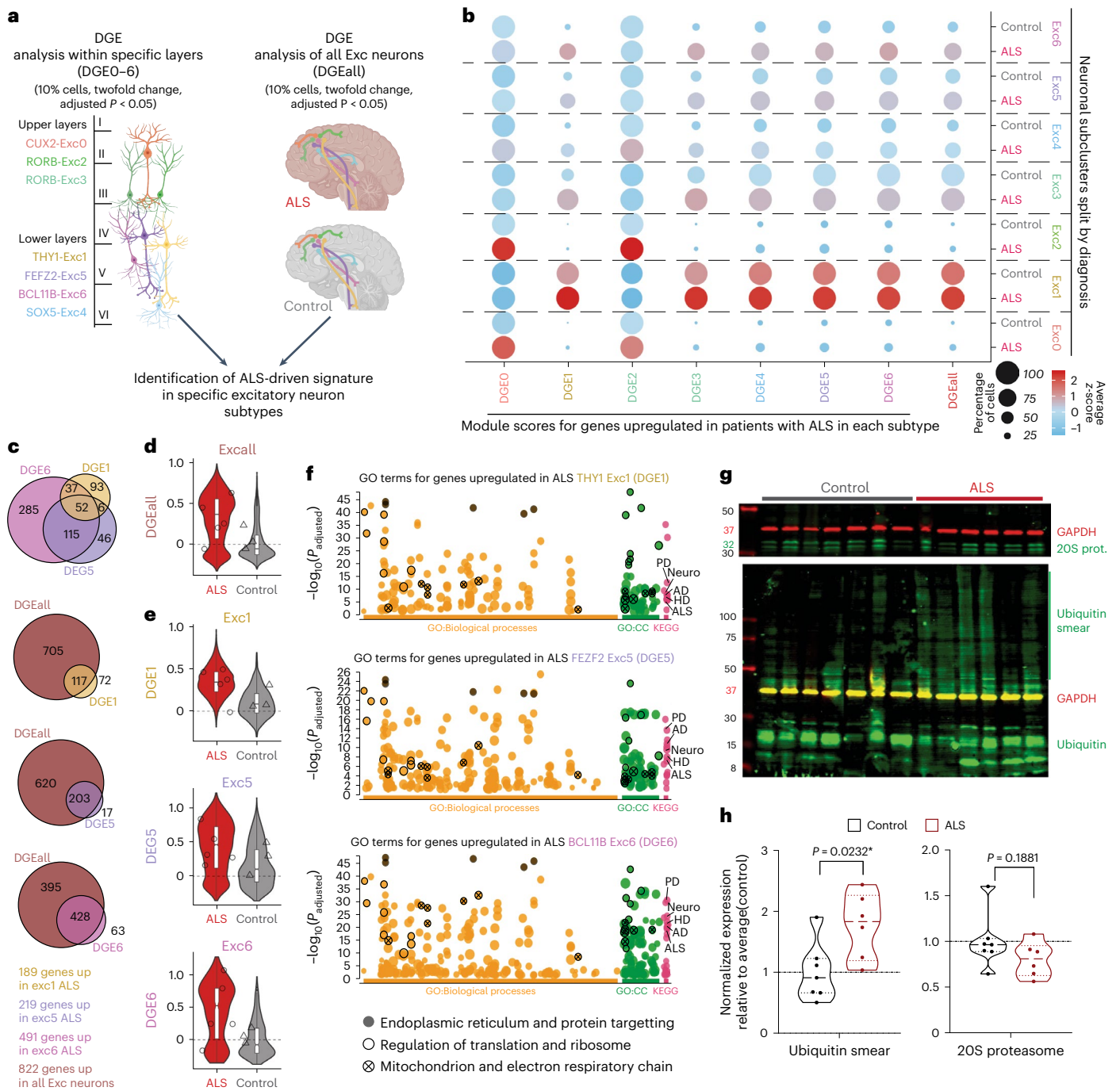
Fig. 3 | ALS excitatory neurons present increased expression of stress-related pathways. a, A schematic of DGE analysis. **b**, A dot plot representing scores for genes upregulated in each subgroup of Exc neurons (DGE0–6) and globally upregulated in all Exc (DGEall). **c**, Comparison of genes globally upregulated in ALS (DGEall) with genes upregulated in classes of L5-ETNs (genes expressed by >10% of cells, >2 FC and adjusted P value <0.05). **d**, Violin plots of z-scores for genes globally upregulated in all excitatory cells (DGEall) in all excitatory neurons (n = 3 controls, n = 5 patients with sALS; geometric box plots represent median and interquantile ranges and symbols indicate average score per individual). **e**, Violin plots of z-scores for genes upregulated in classes of L5 ETNs (DGE1, DGE5 and DGE6) in the three groups (geometric box plots represent median and interquantile ranges, symbols indicate average score per individual). **f**, GO analysis for genes upregulated in L5 ETNs classes (DGE1, DGE5 and DGE6); the highlighted terms are shared between the three. PD, Parkinson's Disease; AD, Alzheimer's Disease; HD, Huntington's disease; CC, cellular components; KEGG, Kyoto Encyclopedia of Genes and Genomes. **g**–**h**, Western blot quantification of ubiquitin accumulation and 20S proteasome (prot.) subunit from motor cortices of separate cohort of patients with ALS (n = 6) and age-matched controls (n = 7) (t-test).

an intrinsic susceptibility to proteostatic stress orchestrated, in part, by abnormal homeostasis of TDP-43.

Oligodendroglia respond with a neuronally engaged state

To extend deep into the cord, ETNs are dependent on robust axonal integrity[39]. As others detected changes in myelination in ALS[12] and in

FTD[40], we analyzed 19,151 nuclei from myelinating cells that clustered into five groups: one of OPCs—oliglia3, and four of oligodendrocytes—oliglia0, 1, 2, 4 (Fig. 5a–c and Extended Data Fig. 7a,b). We noted a significant shift of ALS nuclei representation in oliglia0 versus oliglia1 and oliglia4 (Fig. 5d). Control-enriched oliglia0 were characterized by GO terms connected to oligodendrocyte development and myelination
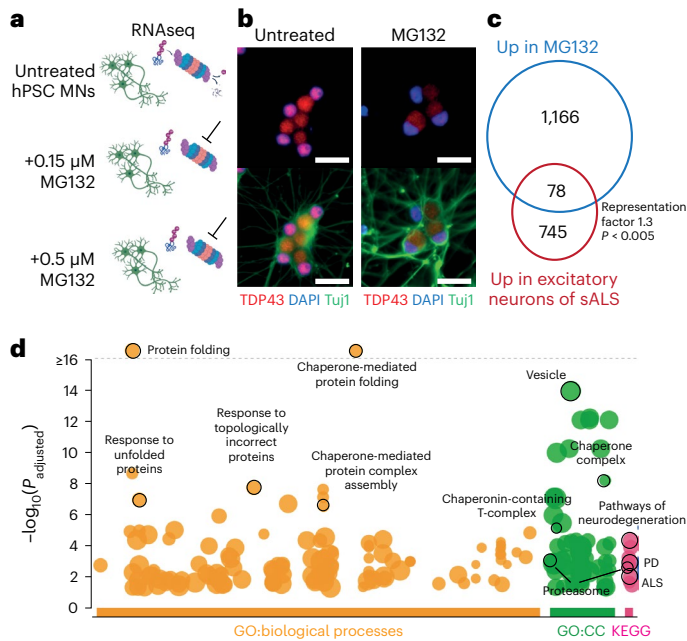
Fig. 4 | Proteostatic stress in hPS cell-derived neurons resembles changes in excitatory neurons from brain of patients with ALS. a, Diagram of neuronal differentiation from PSCs and treatment with proteasome inhibitors for bulk RNA sequencing. **b**, Immunofluorescence of TDP-43 localization after treatment. **c**, Venn diagram depicting shared upregulated genes between treated hPS cell-derived neurons and excitatory neurons from patients with ALS. **d**. GO analysis for shared genes in **c**, highlighted terms involved in protein folding and neurodegenerative diseases. CC, cellular components.

and expressed higher levels of myelinating genes, for example, *CNP*, *OPALIN* and *MAG* (Fig. 5e and Extended Data Fig. 7c,d). ALS-enriched oligilia1 showed terms for neurite morphogenesis, synaptic organization and higher expression of synaptic-related genes *DLG1*, *DLG2* and *GRID2* (Fig. 5f and Extended Data Fig. 7e,f). Intriguingly, expression of neuronal enriched transcripts has been found in oligodendrocytes in primate motor cortex[24].

While intragroup variability was apparent between aggregated oligodendroglial transcriptomic signatures, there was a demarcated separation of individuals by diagnosis with global gene expression conserved at single-cell and aggregated levels (Extended Data Fig. 7g,h). We proceeded to calculate DEGs by excluding one individual at a time (ALLminus1 approach described above), selected genes commonly unregulated in at least four out of five patients and compiled gene lists for subsequent analyses (Extended Data Fig. 7i–k). Altogether, DGE and GO analyses support a shift of sALS oligodendrocytes from myelinating to neuronally engaged states with upregulation of genes involved in synapse modulation and decrease of regulators of myelination (Fig. 5g–i and Extended Data Fig. 7l–o). Loss of myelination is exemplified by the changes in G-protein coupled receptors (GPRCs) marking developmental milestones: *GPR56*, expressed in OPCs[41], and *GPR37*, expressed in myelinating cells[42], were lowly expressed in ALS-enriched groups and globally downregulated (Extended Data Fig. 7p).

To further explore these changes, we compared them to published reports that identified shifts in oligodendrocytes in MS (Source Data Table 4)[43]. Comparison of Jäkel et al.[43] with our study revealed that control-related oligilia0 most closely resembled highly myelinating *OPALIN*⁺ cells from Jäkel6 (Extended Data Fig. 8a–c), while ALS-associated oligilia1 and oligilia4 aligned to not-actively myelinating Jäkel1 (Extended Data Fig. 8d–h). To confirm this shift, we ran validations on protein extracts from motor cortices and showed that oligodendrocyte-specific, myelin-associated proteins CNP and MBP are downregulated in patients (Fig. 5j–k), consistent with studies

identifying demyelination in patients with sALS[12] and with bulk RNA sequencing studies that identified a decrease in myelinating markers[7]. The data so far show how activation of stress pathways in deep-layer neurons is juxtaposed to a shift in oligodendrocytes from active myelination to oligo-to-neuron contact (Fig. 5l).

Microglia activate an endolysosomal response

Mouse models[44], patient samples[5] and function of ALS-related genes in myeloid cells[45] have demonstrated the importance of microglia as modifiers of disease. In the 1,452 nuclei from microglia (Fig. 6a), comparative DGE analyses showed intragroup variability in magnitude but conserved directionality of changes and robustness of the core common biology (Extended Data Fig. 9a–e). We identified 159 genes upregulated in patients and, remarkably, many were associated with endocytosis and exocytosis, previously implicated in ALS[45] (Fig. 6b,c). Several of these genes were also associated with microglial activation and neurodegenerative disorders (*CTSD, SPP1, CPM* and *APOE*) (Fig. 6c,d) and interestingly with ALS–FTD (*TREM2, OPTN, SQSTM1/p62* and *GRN*) (Fig. 5e,f). GO analysis for upregulated genes confirmed a proinflammatory state highlighting activation of endolysosomal pathways, secretion and immune cell degranulation previously associated with myeloid cells in ALS[45] (Fig. 6g,h). Further subclustering identified three groups: homeostatic Micro0, 'disease-associated microglia (DAM)'-like Micro1 and cycling Micro2 (Extended Data Fig. 9f–h). Notably, genes that characterized Micro1 were also upregulated in sALS (Extended Data Fig. 9i,j), in conjunction with a downregulation of homeostatic genes and upregulation of reactive pathways (Extended Data Fig. 9k–m).

To identify modulators of this signature, we used the Connectivity Map pipeline[46], which contains gene expression data of nine human cell lines after thousands of perturbations, allowing association between a given transcriptomic signature and a specific alteration. This analysis revealed that genes dysregulated in microglia positively correlated with regulators of cell cycle and senescence, suggesting an exhaustion of microglial proliferation. We also found a negative correlation with a type I interferon-associated responses (*IFNB1*), often targeted in treatments for neurological diseases to reduce inflammation (Extended Data Fig. 10a).

By comparing our results with published snRNAseq studies[35,47], we identified dysregulation of lipid metabolism (*APOE, APOC1* and *SPP1*) as a common feature in microglia, genes associated with DAMs shared between ALS and MS (*CTSD, GPNMB, CPM* and *LPL*) and ALS and AD (for example, *TREM2*), as inferred by bulk RNA sequencing studies[7] (Fig. 5i). Genes specifically upregulated in ALS were related to vesicle trafficking, myeloid cell degranulation and the lysosome (for example, *SQSTM1/p62, LGALS3, GRN, ASAH1* and *LRRK2*). This evidence suggests the induction of a shared microglial reactive state, yet in ALS these changes are connected to endolysosomal pathways.

Given the stress signature identified in neurons, we wondered whether these transcriptomic changes were driven by neuronal distress. We differentiated induced microglia-like cells (iMGLs)[48] and neurons[38,49] from hPS cells, triggered neuronal apoptosis and then introduced apoptotic neurons to iMGLs in vitro[48] (Fig. 7a,b and Extended Data Fig. 10b). Quantitative assessment by quantitative reverse trascription polymerase chain reaction (RT–qPCR) confirmed the treatment lead to significant downregulation of homeostatic genes, upregulation of genes involved in endolysosomal trafficking (*CTSD, ITGAX, LGALS3* and *SQSTM1/p62*) and downregulation of actively cycling cells markers (Fig. 7c,d and Extended Data Fig. 10c), suggesting that changes identified in microglia from patients are, at least in part, a response to neuronal apoptosis.

Discussion

A key question in the study of neurodegeneration is why certain cell types are more susceptible to different diseases. In this study, we identified the enrichment for ALS–FTD-associated genes in a class of ETNs,
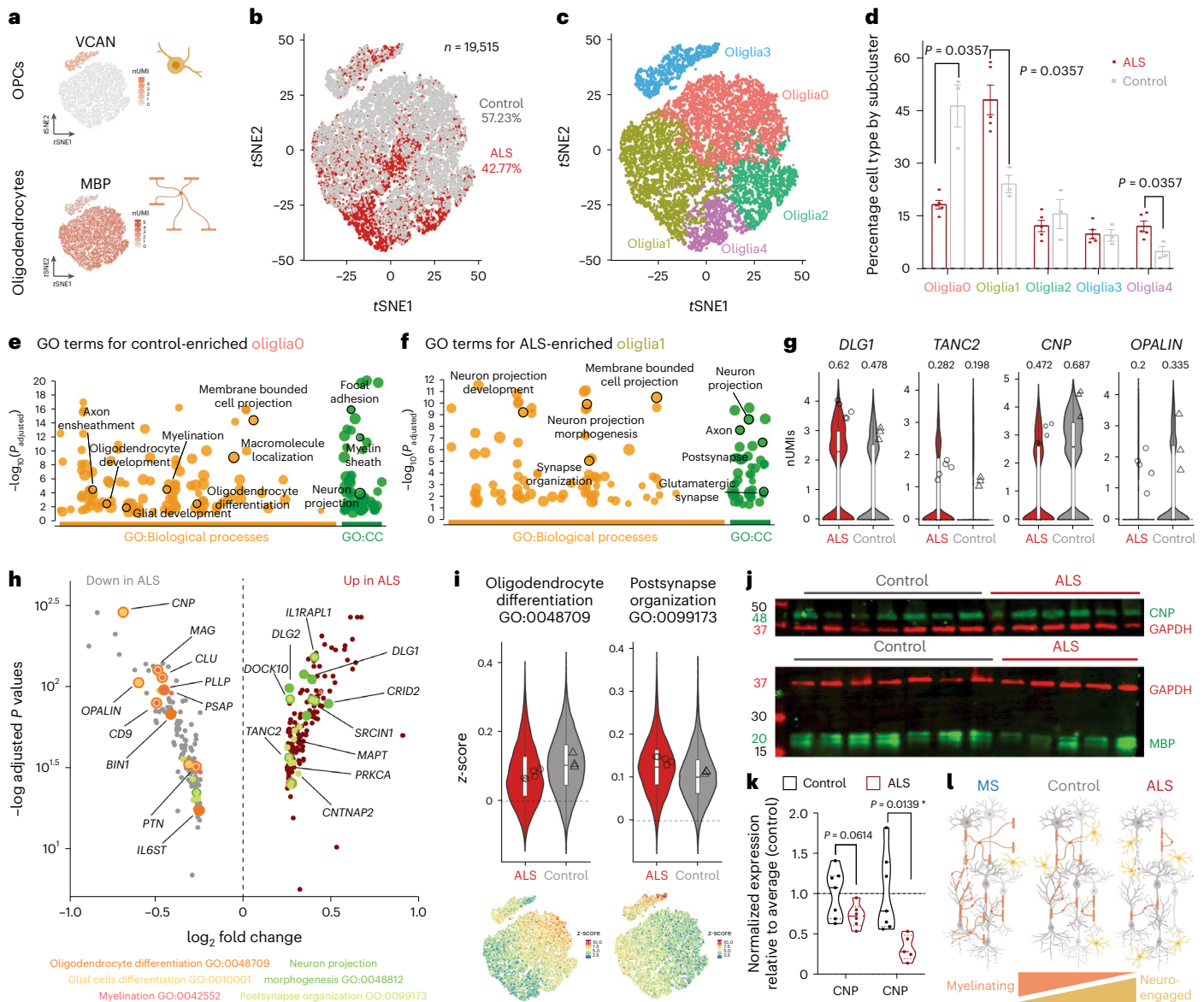
Fig. 5 | In ALS, oligodendroglial cells decrease their myelinating machinery in favor of a neuro-engaged state. a, *t*SNE projection of OPCs and oligodendrocytes markers. **b**, *t*SNE projection of oligodendroglia (ALS *n* = 8,372 nuclei and control *n* = 11,168 nuclei). **c**, *t*SNE projection of subclusters within oligodendroglia (Wilcoxon–Mann–Whitney). **d**, Distribution of subclusters by diagnosis) (mean ± s.e.m. **e**, GO analysis for genes characteristic of control-enriched oliglia0 highlighted terms involved in myelination. CC, cellular components. **f**, GO analysis for genes characteristic of ALS-enriched oliglia1 highlighted terms involved in neuro-engaged functions. **g**, Violin plots of representative genes for neuro-supportive functions (left) and myelination (right) (geometric box plots for media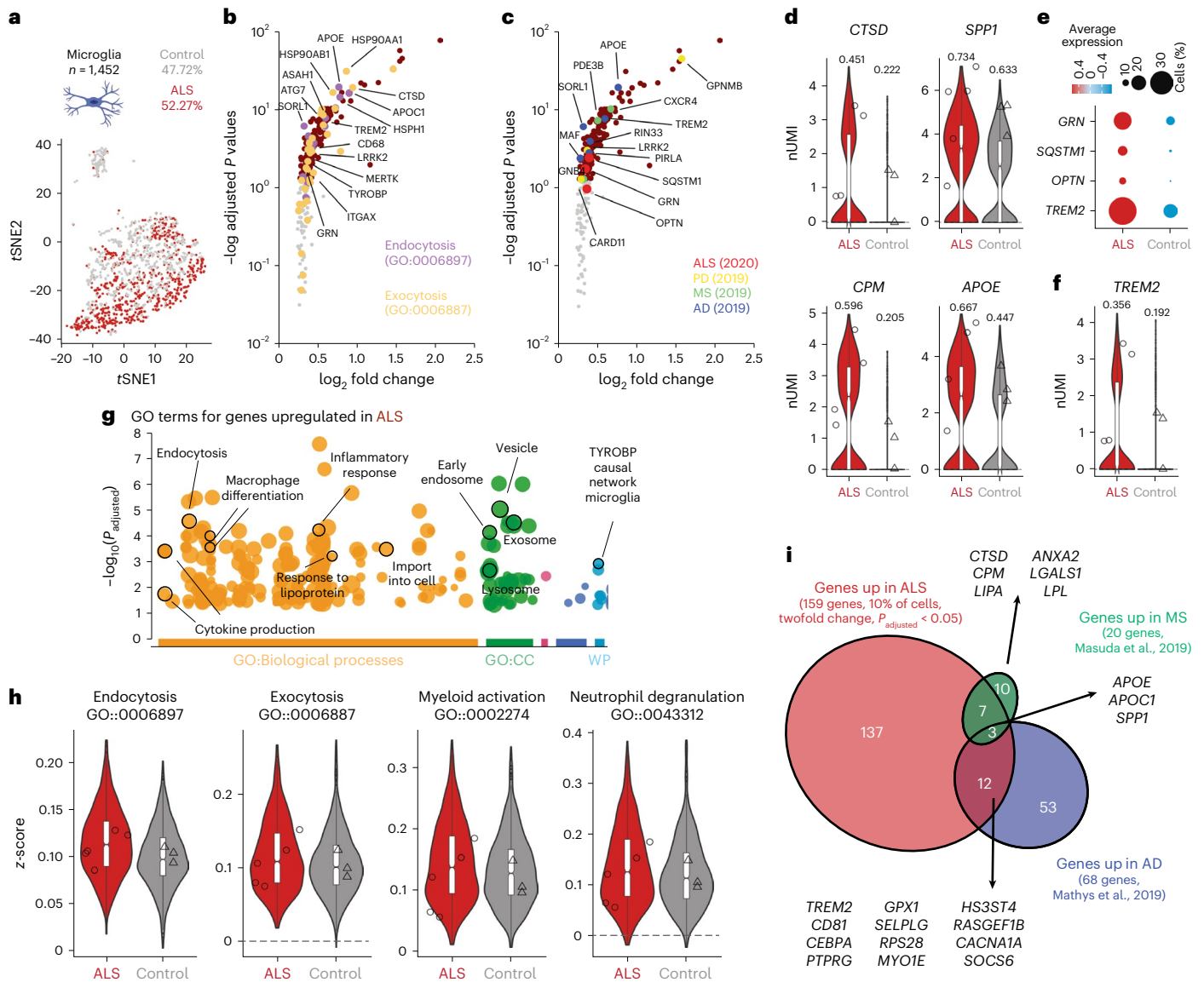n and interquartile ranges; symbols indicate log₂(average expression) per individual (fraction of cell expressing). **h**, Volcano plot of DEGs in oligodendroglia. Highlighted genes identified in GO terms related to myelination (orange) and neuro-engaged functions (green). **i**, Violin plots representing *z*-score for selected GO terms and related *t*SNE projection (boxplot representing median and interquantile ranges; symbols indicate average score per individual). **j,k**, Western blot (**j**) and quantification (**k**) of CNPase and MBP from motor cortices of patients with ALS and age-matched controls (*t*-test). **l**, Diagram illustrates shift of oligodendrocytes states (*t*-test) (for **a–l**, *n* = 3 control and *n* = 5 patients with sALS). nUMI - normalized Unique Molecular Identifier.

which provides a connection between this neuronal subtype and its propensity to accumulate TDP-43 aggregates[31] leading to their gradual loss in ALS–FTD[10]. This enrichment is not recapitulated for risk factors connected to AD and MS, related to immune processes and more enriched in microglia[18–20]. One study suggested that ALS-associated variants connected to autophagy and protein clearing are most highly expressed in glutamatergic neurons[33], and these findings add to the importance of axonal dynamics and ribonucleotide metabolism[34]; here, we provide a more detailed dissection of which subtype that might be.

Additionally, we identified a broadly shared transcriptomic signature of cellular stress pathways in classes of deep-layer excitatory

neurons. These alterations in RNA translation and proteostasis have been implicated in models of ALS[1,2]; our study highlights their cell type specificity and links them to rare mutations in regulators of these pathways in familial forms of ALS[4]. These molecular mechanisms are confirmed to be connected to proteasomal function and proteostasis by human neuronal in vitro models, a system that is already being used to identify therapeutic candidates funneled into clinical trial pipelines[50]. The nuclear nature and the low coverage of this kind of sequencing methodology, but also the small sample size in our study, obviates further, confident dissection of the neuronal-specific changes in RNA biology identified in in vitro models and patient samples[37]. Moreover,

Fig. 6 | DAM signature in ALS. a, *t*SNE projection of microglia (ALS *n* = 759 nuclei and control *n* = 693 nuclei). **b,c**, Volcano plot of genes upregulated in microglia from ALS. Genes identified in GO terms for endocytosis and exocytosis (**b**) and genes associated with neurodegenerative diseases (**c**). **d**, Violin plots of representative genes upregulated in patients with ALS associated with reactive microglia (geometric box plots represent median and interquantile ranges; symbols indicate log₂(average expression) per individual) (fraction of cell expressing). **e**, A dot plot representing expression of genes associated with ALS–FTD pathogenesis. **f**, Violin plots of representative ALS–FTD genes upregulated in ALS (geometric box plots represent median and interquantile ranges; symbols

indicate log₂(average expression) per individual) (fraction of cell expressing). **g**, GO analysis for genes upregulated in ALS microglia, highlighted terms involved in myeloid cells biology and/or pathogenesis of ALS (WP, WikiPathways). **h**, Violin plots representing *z*-scores for selected, statistically significant GO terms from **f** (geometric box plots represent median and interquantile ranges; symbols indicate average score per individual). **i**, Comparison of genes upregulated in microglia from ALS with genes upregulated in microglia in other neurodegenerative diseases (for **a**–**i**, *n* = 3 control individuals and *n* = 5 patients with sALS). nUMI, normalized Unique Molecular Identifier.

the phenotypic and potentially genetic heterogeneity of sporadic ALS is probably reflected in our dataset and should be considered. This study is not powered to distinguish the relative contribution of technical effects versus intrinsic heterogeneity across individuals with sALS, but it opens the field to important considerations for future studies that should include greater numbers of patients. Nonetheless, our report highlights the importance of RNA metabolism and proteostasis and their specific misregulation in ETNs. Mouse models where these pathways are specifically altered in CSMNs might shed a light on their interplay in this specific neuronal type.

We suggest two mechanisms by which ETNs are rendered more susceptible to sALS: (1) the intrinsically higher expression of risk factors and (2) processes of degeneration in classes of ETNs that might

exacerbate/contribute to their vulnerability in a combinatorial effect. Recent snRNAseq studies unraveled susceptibility of specific neuronal types in other diseases: mid-layer *RORB* neurons in AD[51,52], upper-layer *CUX2* neurons in MS[22], dopaminergic neurons in Parkinson's disease[53] and ETNs affected in ALS–FTD as described by our study and spinal cord motor neurons as recently suggested in ALS[54]. Impairment of proteostatic mechanisms seems to be a common theme in degenerating neurons regardless of the disease; however, only in ALS are these changes specifically connected to upregulation of transcripts connected to RNA metabolism, a trend that appears to go in the opposite direction in AD[52]. Integrative analyses of these studies might mark the beginning of a new understanding of the mechanisms behind selective neuronal vulnerability to different diseases.
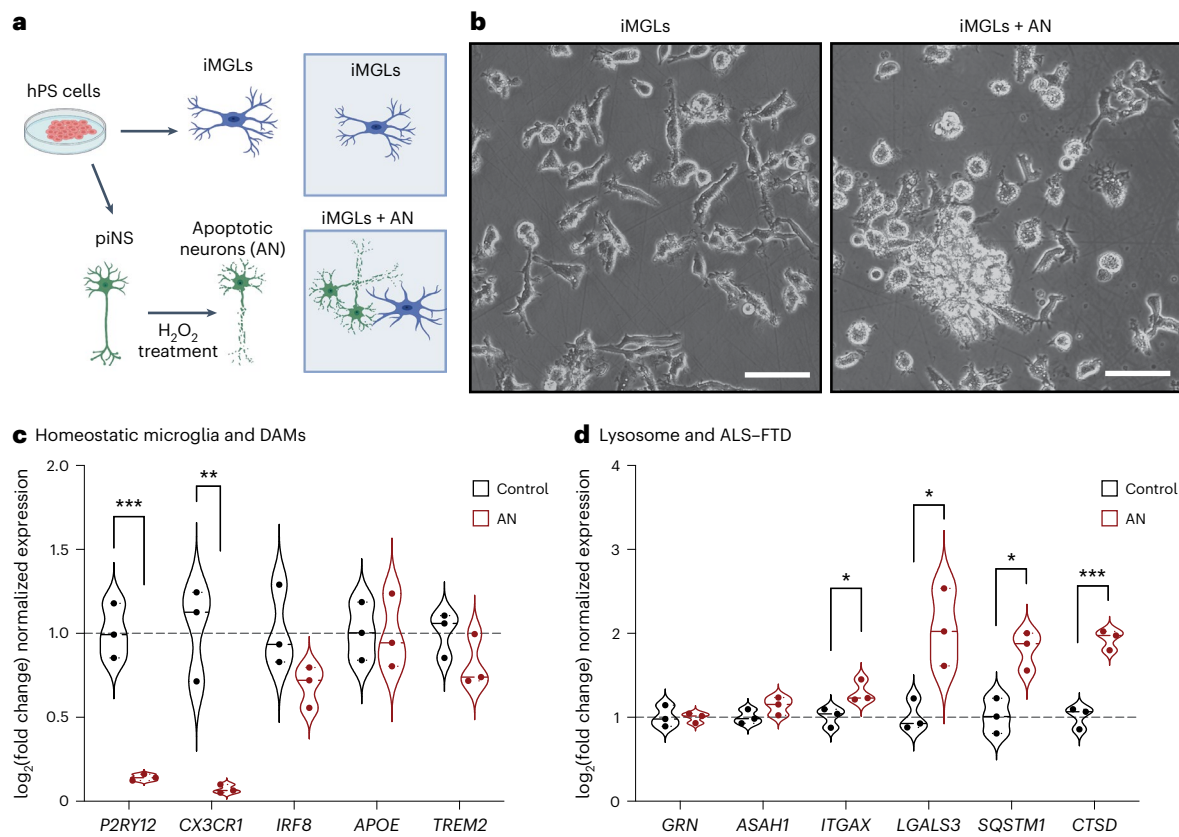
Fig. 7 | Apoptotic neurons upregulate lysosomal genes in microglia.
a, Diagram of microglia and neuronal differentiation from PSCs and induction of apoptosis neurons and feeding to iMGLs (piNs, patterned induced neurons). **b**, Brightfield images of untreated day 40 iMGLs and day 40 iMGLs that were fed apoptotic neurons for 24 h. **c**, RT–qPCR quantification of homeostatic and DAM genes after feeding (AN, apoptotic neurons). **d**, RT–qPCR quantification of selected ALS–FTD-associated and lysosomal genes 24 h after feeding iMGLs with apoptotic neurons (AN, apoptotic neurons). (t-test, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, $n = 3$ biological replicates). piNS, induced neurons from hPS cells.

Emerging studies have shown that glial cells are important modifiers in ALS–FTD[44]. We show that changes in processes involved in oligodendrocyte differentiation and myelination may contribute to degeneration and/or be a coordinated response to ALS and appear to contrast with those described in MS[43]. We revealed perturbations in key myelin regulators, such as *OPALIN*, *CNP* and *MAG*, across oligodendrocyte clusters but in these cells only, as opposed to AD where myelination-related changes were present across multiple cell types[35,55]. Given the similarities in the stress signature identified in neurons with changes in MS lesions but not in AD, it is puzzling how changes in myelination might be a consequence or cause of neuronal degeneration.

Intriguingly, recent work showed expression of neuronal RNA in oligodendrocytes in human motor cortex[24]. Upregulation of synaptic transcripts in this cell type in patients with sALS might represent phagocytic activity[56] or the need for synaptic proteins during deposition of myelin sheath[57]. These speculations are interesting if coupled with the upregulation of synaptic machinery in upper layer neurons and the documented loss of postsynaptic molecules in ETNs in ALS[58]. Recent snRNAseq studies of FTD cortices identified changes in myelinating cells in response to neuronal loss and underlined the importance of cell-to-cell communication[40], also GWAS studies have pointed at excitatory neurons, myelinating cells and inhibitory neurons' sensitivity to genetic risks for ALS[34]. These observations suggest a coordinated response of the motor circuit in an attempt to compensate for loss of inputs to the cord. Further investigations could focus on shifting oligodendroglial states in mouse models and determine changes during disease progression to complement efforts aimed at controlling neuronal activity[59].

Finally, we found distinct perturbations in ALS-associated microglia, particularly in endolysosomal pathways. We and others have

implicated ALS–FTD-associated gene *C9orf72* in endosomal trafficking and secretion in myeloid cells[44,45] and the upregulation of lysosomal constituents, for example, *CTSD*, was identified in this study and by others in patients[60]. Coupled with the upregulation of ALS–FTD-associated genes *SQSTM1/p62*, *OPTN*, *TREM2* and *GRN*, this suggests a mechanistic convergence on vesicle trafficking and inflammatory pathways that may initiate/exacerbate the homeostatic-to-DAM transition in ALS[7]. The interferon response-related changes we delineate, as identified by others in *C9orf72* ALS[61], provide a parallel between sporadic and familial ALS. Overall, these changes had partial overlap with microglia in AD[35] and in MS[47], suggesting that drugs modulating myeloid cells in neurodegenerative diseases may provide a basis for new therapeutic approaches. Recent reports showed how DAM might be beneficial in disease contexts[62], but others have inferred that microglial activation might result in poor disease outcomes. Studies manipulating microglial states might elucidate their 'friend or foe' role in sALS[48].

In summary, we show that classes of ETNs require the expression of a collection of genetic risk factors for ALS–FTD with pivotal roles in proteostasis. This intrinsically higher expression of disease-associated genes might be at the bottom of a 'first over the line' mechanism leading to disruptions in groups of deep layer excitatory neurons. These alterations trigger a cascade of responses in glia: oligodendroglia shift from a myelinating to a neuronally engaged state and microglia activate a proinflammatory signature. Our study offers a view in which neurocentric disease vulnerability might spark responses in other cell types, but it also shows that enrichment of ALS–FTD-related genes in ETNs is coupled with processes engaging these genes in other cells too, that is, microglia. This view is a first insight into the disruptions of cortical biology in ALS and provides a connection between changes

in cellular components and mechanisms associated with ALS. Future investigations should consider multicellular disruptions in ALS–FTD, where the survival of the neuron is unmistakably pivotal, but targeting other cells to reduce inflammation, promote myelination and bolster neuronal circuitry may re-establish a neuroprotective environment.

Limitations of this study

One limitation is the small size of the cohort. ALS is a heterogeneous disease[7] and smaller cohort sizes might not fully recapitulate its etiological diversity. Only recently have biobanks started to collect enough samples[6–8], and we hope that increased sample availability and affordability of single-nucleus technology will allow a more comprehensive view of alterations in ALS. A larger cohort would also enable a more stringent analysis of differentially expressed transcripts that incorporates more sophisticated analytical tool.

We recognize that consensus on best practices in snRNAseq is still being reached, including for re-analysis of published studies[63,64]. We acknowledge that using single cells as a variable and not pseudobulked individuals might yield larger gene sets with possible confounding factors. We also recognize that our additional analyses revealed that some specific genes might be derived from outlier-driven effects and because of that we limited follow-up analyses only to genes shared by at least four out of five individuals. This is why we highlighted common features in wider biological pathways disrupted in cell types in patients with ALS followed up by validations at the protein level in a separate cohort of patients' samples or in our in vitro studies. We hope that new reports will take into account the need for a more stringent investigation of reproducibility and we advocate for a more transparent conveying of results.

We also recognize that our study would benefit from additional validation at the RNA and/or protein level. This would elucidate some of the intriguing questions we raised. For example, are oligodendrocytes really expressing higher levels of neuronal genes or is this an artifact[65]? Nonetheless, we believe that this study provides novel insights in the involvement of different cell types in ALS and a different view in the motor cortex of patients with ALS. The increase in cohort sizes, more sophisticated analyses and new technologies, such as spatial transcriptomics, might further enrich the understanding of neurodegeneration in ALS.

Methods

Human donor tissue sources and ethics

Frozen postmortem human cortical samples from cases of patients with sporadic ALS and age-matched controls were obtained from the Target ALS Neuropathology Core that drew upon the repositories of five institutions. Specimens from the medial, lateral or unspecified motor cortex were grouped together. Deidentified postmortem brain tissue samples were obtained from the Massachusetts Alzheimer's Disease Research Center at Massachusetts General Hospital (MGH) and the Target ALS Multicenter Human Postmortem Tissue Core, which integrates five academic tissue repositories for ALS research. The protocols of the Massachusetts Alzheimer's Disease Research Center for brain donations and the collection of postmortem tissue and clinical information for research purposes were approved by the Institutional Review Board of Partners Healthcare (currently Mass General Brigham) at MGH. Informed consent for brain autopsy and the use of postmortem tissue for research was provided by the legal next-of-kin in compliance with local and institutional guidelines at all brain tissue repositories involved. Use of postmortem deidentified tissue samples followed 'Not Human Subjects Research' determinations by the Harvard Faculty of Arts and Sciences and Partners (MGH) and considered exempt from the instituational research board given lack of interaction with living individuals/participants. The study protocol was further approved by Harvard Stem Cell and Regenerative Biology Department, Harvard University. Informed consent and study protocol for human stem cell work were provided by Stanley Center for Psychiatric Research at Broad Institute of MIT and Harvard and the Harvard Stem Cell and Regenerative Biology Department at Harvard University.

Isolation of nuclei

RNA quality of brain samples was assessed by running bulk nuclear RNA on an Agilent TapeStation for RNA integrity number scores. Extraction of nuclei from frozen samples was performed as previously described[66]. Briefly, tissue was dissected and minced with a razor blade on ice and then placed in 4 ml ice-cold extraction buffer (wash buffer (82 mM Na2SO4, 30 mM K2SO4, 5 mM MgCl2, 10 mM glucose and 10 mM HEPES, pH adjusted to 7.4 with NaOH) containing 1% Triton X-100 and 5% Kollidon VA64). Tissue was homogenized with repeated pipetting, followed by passing the homogenized suspension twice through a 26.5 gauge needle on a 3 ml syringe (prechilled), once through a 20 mm mesh filter and once through a 5 mm filter using vacuum. The nuclei were then diluted in 50 ml ice-cold wash buffer, split across four 50 ml tubes and centrifuged at 500g for 10 min at 4 °C. The supernatant was discarded, the nuclei pellet was resuspended in 1 ml cold wash buffer.

10× loading and library preparation

Nuclei were 4,6-diamidino-2-phenylindole (DAPI) stained with Hoechst, loaded onto a hemocytometer and counted using brightfield and fluorescence microscopy. The solution was diluted to -176 nuclei μl⁻¹ before proceeding with Drop-seq, as described[14]. Complementary DNA amplification was performed using around 6,000 beads per reaction with 16 PCR cycles. The integrity of both the complementary DNA and tagmented libraries were assessed for quality control on the Agilent Bioanalyzer. Libraries were sequenced on a Nova-seq S2, with a 60 bp genomic read. Reads were aligned to the human genome assembly (hg19). Digital gene expression files were generated with the Zamboni drop-seq analysis pipeline, designed by the McCarroll group[66,67].

Filtering of expression matrices and clustering of single nuclei

A single matrix for all samples was built by filtering any barcode with less than 400 genes and resulting in a matrix of 27,600 genes across 119,510 barcodes. This combined UMI matrix was used for downstream analysis using Seurat (v3.0.2)[15]. A Seurat object was created from this matrix by setting up a first filter of min.cells=20 per genes. After that, barcodes were further filtered by number of genes detected nFeature_RNA > 600 and nFeature_RNA < 6,000. Distribution of genes and UMIs were used as parameters for filtering barcodes. The matrix was then processed via the Seurat pipeline: log normalized by a factor of 10,000, followed by regressing UMI counts (nCount_RNA) and scaled for gene expression.

After quality filtering, 79,830 barcodes and 27,600 genes were used to compute shared nearest-neighbor graphs and t-distributed stochastic neighbor embedding (tSNE) projections using the first ten statistically significant principal components. As previously described[44,68], tSNE projection was used to determine minimum number of clusters at Resolution=0.2 (FindClusters). Broad cellular identities were assigned to groups on the basis of DEGs as calculated by Wilcoxon rank sum test in FindAllMarkers(min.pct=0.25, logfc.threshold=0.25). One subcluster with a specifically high ratio of UMIs/genes was filtered out resulting in 79,169 barcodes grouped in seven major cell types. Markers for specific cell types were identified in previously published small conditional RNA sequencing studies[16].

Analysis of cellular subtypes was conducted by subsetting each group. Isolated barcodes were renormalized and scaled and relevant principal components were used for clustering as a separate analysis. Newly scaled matrices were used for DGE analysis with MAST algorithm in Seurat, as previously reported[38,43,44,53,68], with parameters FindAllMarkers(min.pct=0.10, logfc.threshold=0.25) and subclustering for identification of subgroups. DGEs for downsampled excitatory neurons groups were computed by adding parameters to the functions

described above, FindAllMarkers(min.pct=0.10, logfc.threshold=0.25, min.cells = 'minimum number in the group', random.seed=TRUE), and re-iterative lists were generated with >95% of overlap (data not shown). PCA by individual was performed using matrices generated by AverageExpression(celltype, group.by = "ID", return.seurat=TRUE); these were normalized and scaled and PCA was run using the DGEs with RunPCA(celltype.averages, features=c(DGEs), npcs=7). To further confirm that changes in DGEs were not driven by one individual only, we ran DGE analysis by excluding cells from one ALS individual at a time (control versus ALS-1) in a re-iterative manner and showed that magnitude of changes in DGE expression were still similar and that identified DGEs were shared by all five or at least four out of five individuals; these gene lists were used for GOy and protein–protein interaction analyses. Gene scores for different cellular subclusters were computed in each renormalized, rescaled submatrix using the AddModuleScore function in Seurat v3.0.2.

Re-analysis of publicly available datasets was performed using matrices and metadata available. Only barcodes with available metadata concerning their cellular identity were selected to use identities assigned by peer review publication[22,30]. The available barcodes were then loaded into Seurat v4.0.1 (ref. 69). Gene scores for different cellular subclusters were computed in each renormalized, rescaled submatrix using the AddModuleScore function, as previously described. Re-analysis of spatial transcriptomic from Maynard et al. was performed using publicly available data and codes from publication itself[29].

GO, interactome and GSEAs

For GO terms analysis, we selected statistically significant upregulated or downregulated genes identified in each subcluster as described before (adjusted P values <0.05, log fold change (FC) of 2). These lists were fed in the gProfiler pipeline[70] with the following settings: use only annotated genes, g:SCS threshold of 0.05, GO cellular components and GO biological processes (26 May 2020 to 9 December 2021), only statistically significant pathways are highlighted. Only statistically significant upregulated genes identified in each subcluster as described before (adjusted P values < 0.05, log FC of 2) were used for GO analysis. The interactome map was built using STRING[71] protein–protein interaction networks, all statistically significant upregulated genes were used, 810 were identified as interacting partners using 'experiments' as interaction sources and a medium confidence threshold (0.400), only interacting partners are shown in Extended Data Fig. 6. Gene set enrichment analysis (GSEA) was performed using GSEA software designed by UC San Diego and the Broad Institute (v4.0.3)[72]. Briefly, gene expression matrices were generated in which for each subcluster each individual was a metacell, lists for disease-associated risk genes were compiled using available datasets (PubMed, ALS–FTD; Source Data Table 2) or recently published GWAS for AD[18,19] and MS[20].

Generation of microglia-like cells

Microglial-like cells were differentiated as described[48]. Briefly, WA01-H1 hPS cells were cultured in E8 medium (Stemcell Technologies) on Matrigel (Corning), dissociated with accutase (Stemcell Technologies), centrifuged at 300g for 5 min and resuspended in E8 medium with 10 μM Y-27632 ROCK inhibitor, with 2 M cells transferred to a low-attachment T25 flask in 4 ml of medium and left in the suspension for 24 h. The first 10 days of differentiation are carried out in iHPC medium: Iscove's modified Dulbecco's medium (50%, Stemcell Technologies), F12 (50%, Stemcell Technologies), ITS-G-X 2% vol/vol (Thermo Fisher), L-ascorbic acid 2-phosphate (64 μg ml⁻¹, Sigma), monothioglycerol (400 mM, Sigma), PVA (10 mg ml⁻¹; Sigma), glutamax (1×, Stemcell Technologies), chemically defined lipid concentrate (1×, Stemcell Technologies) and nonessential amino acids (Stemcell Technologies). After 24 h (day 0), cells were collected, and differentiation is started in iHPC medium supplemented with fibroblast growth factor 2 (FGF2) (PeproTech, 50 ng ml⁻¹), bone morphogenetic protein (PeproTech,

50 ng ml⁻¹), activin A (PeproTech, 12.5 ng ml⁻¹), Y-27632 ROCK inhibitor (1 μM) and LiCl (2 mM) and transferred in a hypoxic incubator (20% O_2, 5% CO_2 at 37 °C). On day 2, the medium was changed to iHPC medium plus FGF2 (PeproTech, 50 ng ml⁻¹) and vascular enothelial growth factor (PeproTech, 50 ng ml⁻¹) and returned to hypoxic conditions. On day 4, cells were resuspended in iHPC medium supplemented with FGF2 (PeproTech, 50 ng ml⁻¹), vascular endothelial growth factor (PeproTech, 50 ng ml⁻¹), TPO (PeproTech, 50 ng ml⁻¹), SCF (PeproTech, 10 ng ml⁻¹), interleukin (IL)-6 (PeproTech, 50 ng ml⁻¹) and IL-3 (PeproTech, 10 ng ml⁻¹), and placed into a normoxic incubator (20% O_2, 5% CO_2 at 37 °C). Expansion of hematopoietic progenitors was continued by supplementing the flasks with 1 ml of iHPC medium with small molecules every 2 days. On day 10, cells were collected and filtered through a 40 mm filter. The single-cell suspension was counted and plated at 500,000 cells per well in a six-well plate coated with Matrigel (Corning) in microglia differentiation medium: Dulbecco's modified Eagle medium/F12 (Stemcell Technologies), ITS-G 2% vol/vol (Thermo Fisher Scientific), B27 (2% vol/vol, Stemcell Technologies), N_2 (0.5% vol/vol, Stemcell Technologies), monothioglycerol (200 mM, Sigma), glutamax (1×, Stemcell Technologies), nonessential amino acids (1×, Stemcell Technologies), supplemented with macrophage colony-stimulating factor (25 ng ml⁻¹, PeproTech), IL-34 (100 ng ml, PeproTech) and transforming growth factor β-1 (50 ng ml⁻¹, PeproTech). iMGLs were kept in this medium for 20 days with change three times a week. On day 30, cells were collected and plated on poly-D-lysine/laminin-coated dishes in microglia differentiation medium supplemented with CD200 (100 ng ml⁻¹, Novoprotein) and CX3CL1 (100 ng ml⁻¹, PeproTech), macrophage colony-stimulating factor (25 ng ml, PeproTech), IL-34 (100 ng/ml, PeproTech) and transforming growth factor β-1 (50 ng ml⁻¹, PeproTech) until day 40.

Feeding of apoptotic neurons to microglia-like cells

For feeding assays, neurons were generated from human iPSCs using an NGN2 overexpression system, as described previously[38,49,73]. Day-30 WA01-H1 hiPSC-derived neurons were treated with 2 μM H_2O_2 for 24 h to induce apoptosis. Apoptotic neurons were gently collected from the plate and the medium containing the apoptotic bodies was transferred into wells containing day 40 iMGLs. After 24 h, iMGLs subjected to apoptotic neurons and controls were collected for RNA extraction.

RNA extraction and RT–qPCR analysis

RNA was extracted with the miRNeasy Mini kit (Qiagen, 217004). cDNA was produced with iScript kit (Bio-Rad) using 50 ng of RNA. RT–qPCR reactions were performed in triplicates using 20 ng of cDNA with SYBR Green (Bio-Rad) and were run on a CFX96 Touch PCR Machine for 39 cycles at 95 °C for 15 s, 60 °C for 30 s and 55 °C for 30 s.

Generation of hiPSC-derived neurons for bulk RNA sequencing

Human embryonic stem cells were cultured in mTeSR (Stemcell Technologies) on Matrigel (Corning). Motor neurons were generated from HuES-3-Hb9:GFP based on the differentiation protocol previously described[74]. On completion of differentiation, cells were sorted via flow cytometry based on green fluorescent protein (GFP) signal intensity to yield GFP-positive neurons that were plated on PDL/laminin-coated plates (Sigma, Life Technologies). Neurons were maintained in neurobasal medium (Life Technologies) and supplemented with N_2 (STEMCELL Technologies), B27 (Life Technologies), glutamax (Life Technologies), nonessential amino acids (Life Technologies) and neurotrophic factors (BDNF, GDNF and CNTF), and were grown for 28 days before the application of the proteasome inhibitors MG132 for 24 h.

RNA was extracted using RNeasy Plus kit (Qiagen), libraries were prepared using the Illumina TruSeq RNA kit v2 according to the manufacturer's directions and sequenced at the Broad Institute core with samples randomly assigned between two flow chambers. The total

population RNA sequencing FASTQ data was aligned against ENSEMBL human reference genome (build GRCh37/hg19) using STAR (v.2.4.0). Cufflinks (v.2.2.1) was used to derive normalized gene expression in fragments per kilo base per million. The read counts were obtained from the aligned BAM-files in R using Rsubread[73]. Differential gene expression was analyzed from the read counts in DESeq2 using a Wald's test for the treatment dosage and controlling for the sequencing flow cell[73].

Western blot analysis

As previously described, tissue was minced, lysed in RIPA buffer with protease inhibitors (Roche) and sonicated[75]. After centrifugation, the supernatant was collected as soluble fraction and the insoluble pellet was resuspended in 8 M urea buffer (Bio-Rad, 1632103). After protein quantification by the bicinchoninic acid assay (Thermo Fisher), 10 µg of protein was preheated in Laemmli's buffer (Bio-Rad) and loaded in 4–20% mini-PROTEAN TGX precast protein gels (Bio-Rad), and then gels were transferred to a polyvinylidene fluoride membrane. Membranes were blocked in Odyssey blocking buffer (Li-Cor) and incubated overnight at 4 °C with primary antibodies. After washing with TBS-T, membranes were incubated with IRDye secondary antibodies (1:10,000, Li-Cor) for 1 h and imaged with the Odyssey CLx imaging system (Li-Cor). Primary antibodies used 1:1,000 dilution: TDP-43 (PeproTech 10782-2-AP), GAPDH (Millipore cat. no. MAB374; CST 2118 (14C10)), MBP (Thermo Fisher PA-1-10008), CNP (Abcam ab6319(11-5b)), 20S (Enzo BML-PW8195-0025) and ubiquitin (CST 3936 T (P4D1)). IRDye provided by Licor was used at 1:10,000 dilution.

Immunofluorescence assays

Cells were washed once with phosphate-buffered saline (PBS), fixed with 4% paraformaldehyde for 20 min, washed again in PBS and blocked for 1 h in 0.1% Triton in PBS with 10% donkey serum. Fixed cells were then washed and incubated overnight with primary antibodies at 4 °C. Primary antibody solution was washed and cells were subsequently incubated with secondary antibodies (1:2,000, AlexaFluor, Life Technologies) at room temperature for 1 h, washed with PBS and stained with DAPI. Primary antibodies used were Tuj1 (1:250, R&D, MAB1195) and TDP-43 (1:200, PeproTech 10782-2-AP). Images were analyzed using FIJI.

Proteasome activity assay

Neurons were sorted in 96-well plates and, after 2 weeks of maturation, treated for 24 h. Cells were washed with 1× PBS, exposed to ProteasomeGlo (Promega, G8660) and incubated for 30 min at room temperature. Fluorescence was measured using a Cytation 3 reader (BioTek).

Statistics and reproducibility

No statistical method was used to predetermine sample size. No statistical methods were used to predetermine sample sizes but our sample sizes are similar to those reported in previous publications[22,24,26,29,30,40,43,44,47,48,51,52,54,68]. No data were excluded from the analyses. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. Data distribution was assumed to be normal but this was not formally tested. Data collection and analysis were not performed blind to the conditions of the experiments. Software used for analyses were Licor (Image Studio version 2.1), GraphPad Prism (version 7 and above), ImageJ (FIJI version 2.14) with Nikon NIS Elements version 4.0, Seurat version 3.0.2 and version 4.0.1, and GSEA version 4.0.3.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Transcriptomic raw data have been deposited in the Gene Expression Omnibus database under accession number GSE226753. All other data supporting the findings of this study are available as source data files or from the corresponding author upon reasonable request. All other data analyzed from previously published sources will be available at publication references in the manuscript (for Schirmer et al. Sequence Read Archive (SRA), under accession number PRJNA544731 and NCBI Bioproject ID: 544731; for Velmeshev et al. Sequence Read Archive, accession number PRJNA434002; for Maynard et al. available via GitHub at https://github.com/LieberInstitute/HumanPilot and https://github.com/LieberInstitute/spatialLIBD).

References

1. Taylor, J. P., Brown, R. H. Jr & Cleveland, D. W. Decoding ALS: from genes to mechanism. *Nature* **539**, 197–206 (2016).
2. Brown, R. H. & Al-Chalabi, A. Amyotrophic lateral sclerosis. *N. Engl. J. Med.* **377**, 162–172 (2017).
3. Giacomelli, E. et al. Human stem cell models of neurodegeneration: from basic science of amyotrophic lateral sclerosis to clinical translation. *Cell Stem Cell* **29**, 11–35 (2022).
4. Mordes, D. A. et al. Dipeptide repeat proteins activate a heat shock response found in C9ORF72-ALS/FTLD patients. *Acta Neuropathol. Commun.* **6**, 55 (2018).
5. D'Erchia, A. M. et al. Massive transcriptome sequencing of human spinal cord tissues provides new insights into motor neuron degeneration in ALS. *Sci. Rep.* **7**, 10046 (2017).
6. Tam, O. H. et al. Postmortem cortex samples identify distinct molecular subtypes of ALS: retrotransposon activation, oxidative stress, and activated glia. *Cell Rep.* **29**, 1164–1177 e1165 (2019).
7. Humphrey, J. et al. Integrative transcriptomic analysis of the amyotrophic lateral sclerosis spinal cord implicates glial activation and suggests new risk genes. *Nat. Neurosci.* **26**, 150–162 (2023).
8. Eshima, J. et al. Molecular subtypes of ALS are associated with differences in patient prognosis. *Nat. Commun.* **14**, 95 (2023).
9. Neumann, M. et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).
10. Hammer, R. P. Jr., Tomiyasu, U. & Scheibel, A. B. Degeneration of the human Betz cell due to amyotrophic lateral sclerosis. *Exp. Neurol.* **63**, 336–346 (1979).
11. Seeley, W. W. et al. Early frontotemporal dementia targets neurons unique to apes and humans. *Ann. Neurol.* **60**, 660–667 (2006).
12. Kang, S. H. et al. Degeneration and impaired regeneration of gray matter oligodendrocytes in amyotrophic lateral sclerosis. *Nat. Neurosci.* **16**, 571–579 (2013).
13. Boillee, S. et al. Onset and progression in inherited ALS determined by motor neurons and microglia. *Science* **312**, 1389–1392 (2006).
14. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
15. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 e1821 (2019).
16. Lake, B. B. et al. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat. Biotechnol.* **36**, 70–80 (2018).
17. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
18. Kunkle, B. W. et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019).
19. Jansen, I. E. et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019).

20. International Multiple Sclerosis Genetics Consortium Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science* **365**, eaav7188 (2019).

21. Arlotta, P. et al. Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. *Neuron* **45**, 207–221 (2005).

22. Schirmer, L. et al. Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature* **573**, 75–82 (2019).

23. Ozdinler, P. H. et al. Corticospinal motor neurons and related subcerebral projection neurons undergo early and specific neurodegeneration in hSOD1G(9)(3)A transgenic ALS mice. *J. Neurosci.* **31**, 4166–4177 (2011).

24. Bakken, T. E. et al. Comparative cellular analysis of motor cortex in human, marmoset and mouse. *Nature* **598**, 111–119 (2021).

25. Guerra San Juan, I. et al. Loss of mouse Stmn2 function causes motor neuropathy. *Neuron* https://doi.org/10.1016/j.neuron.2022.02.011 (2022).

26. Hodge, R. D. et al. Transcriptomic evidence that von Economo neurons are regionally specialized extratelencephalic-projecting excitatory neurons. *Nat. Commun.* **11**, 1172 (2020).

27. Cobos, I. & Seeley, W. W. Human von Economo neurons express transcription factors associated with layer V subcerebral projection neurons. *Cereb Cortex* **25**, 213–220 (2015).

28. Zeng, H. et al. Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. *Cell* **149**, 483–496 (2012).

29. Maynard, K. R. et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat. Neurosci.* **24**, 425–436 (2021).

30. Velmeshev, D. et al. Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* **364**, 685–689 (2019).

31. Nana, A. L. et al. Neurons selectively targeted in frontotemporal dementia reveal early stage TDP-43 pathobiology. *Acta Neuropathol.* **137**, 27–46 (2019).

32. Porta, S. et al. Patient-derived frontotemporal lobar degeneration brain extracts induce formation and spreading of TDP-43 pathology in vivo. *Nat. Commun.* **9**, 4220 (2018).

33. van Rheenen, W. et al. Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology. *Nat. Genet.* **53**, 1636–1648 (2021).

34. Saez-Atienzar, S. et al. Genetic analysis of amyotrophic lateral sclerosis identifies contributing pathways and cell types. *Sci. Adv.* **7**, eabd9036 (2021).

35. Mathys, H. et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019).

36. Limone, F., Klim, J. R. & Mordes, D. A. Pluripotent stem cell strategies for rebuilding the human brain. *Front. Aging Neurosci.* **14**, 1017299 (2022).

37. Klim, J. R. et al. ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat. Neurosci.* **22**, 167–179 (2019).

38. Limone, F. et al. Efficient generation of lower induced motor neurons by coupling Ngn2 expression with developmental cues. *Cell Rep.* **42**, 111896 (2023).

39. Tomassy, G. S. et al. Distinct profiles of myelin distribution along single axons of pyramidal neurons in the neocortex. *Science* **344**, 319–324 (2014).

40. Gerrits, E. et al. Neurovascular dysfunction in GRN-associated frontotemporal dementia identified by single-nucleus RNA sequencing of human cerebral cortex. *Nat. Neurosci.* **25**, 1034–1048 (2022).

41. Giera, S. et al. The adhesion G protein-coupled receptor GPR56 is a cell-autonomous regulator of oligodendrocyte development. *Nat. Commun.* **6**, 6121 (2015).

42. Yang, H. J., Vainshtein, A., Maik-Rachline, G. & Peles, E. G protein-coupled receptor 37 is a negative regulator of oligodendrocyte differentiation and myelination. *Nat. Commun.* **7**, 10884 (2016).

43. Jakel, S. et al. Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature* **566**, 543–547 (2019).

44. Limone, F. et al. Myeloid and lymphoid expression of C9orf72 regulates IL-17A signaling in mice. *Sci. Transl. Med.* **16**, eadg7895 (2024).

45. Burberry, A. et al. C9orf72 suppresses systemic and neural inflammation induced by gut bacteria. *Nature* **582**, 89–94 (2020).

46. Subramanian, A. et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* **171**, 1437–1452 e1417 (2017).

47. Masuda, T. et al. Spatial and temporal heterogeneity of mouse and human microglia at single-cell resolution. *Nature* **566**, 388–392 (2019).

48. Dolan, M. J. et al. Exposure of iPSC-derived human microglia to brain substrates enables the generation and manipulation of diverse transcriptional states in vitro. *Nat. Immunol.* **24**, 1382–1390 (2023).

49. Nehme, R. et al. Combining NGN2 programming with developmental patterning generates human excitatory neurons with NMDAR-mediated synaptic transmission. *Cell Rep.* **23**, 2509–2523 (2018).

50. Morimoto, S. et al. Phase 1/2a clinical trial in ALS with ropinirole, a drug candidate identified by iPSC drug discovery. *Cell Stem Cell* **30**, 766–780 e769 (2023).

51. Leng, K. et al. Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. *Nat. Neurosci.* https://doi.org/10.1038/s41593-020-00764-7 (2021).

52. Otero-Garcia, M. et al. Molecular signatures underlying neurofibrillary tangle susceptibility in Alzheimer's disease. *Neuron* **110**, 2929–2948 e2928 (2022).

53. Kamath, T. et al. Single-cell genomic profiling of human dopamine neurons identifies a population that selectively degenerates in Parkinson's disease. *Nat. Neurosci.* **25**, 588–595 (2022).

54. Yadav, A. et al. A cellular taxonomy of the adult human spinal cord. *Neuron* **111**, 328–344 e327 (2023).

55. Sadick, J. S. et al. Astrocytes and oligodendrocytes undergo subtype-specific transcriptional changes in Alzheimer's disease. *Neuron* **110**, 1788–1805 e1710 (2022).

56. Falcao, A. M. et al. Disease-specific oligodendrocyte lineage cells arise in multiple sclerosis. *Nat. Med.* **24**, 1837–1844 (2018).

57. Hughes, A. N. & Appel, B. Oligodendrocytes express synaptic proteins that modulate myelin sheath formation. *Nat. Commun.* **10**, 4125 (2019).

58. Genc, B. et al. Apical dendrite degeneration, a novel cellular pathology for Betz cells in ALS. *Sci. Rep.* **7**, 41765 (2017).

59. Wainger, B. J. et al. Effect of ezogabine on cortical and spinal motor neuron excitability in amyotrophic lateral sclerosis: a randomized clinical trial. *JAMA Neurol.* **78**, 186–196 (2021).

60. O'Rourke, J. G. et al. C9orf72 is required for proper macrophage and microglial function in mice. *Science* **351**, 1324–1329 (2016).

61. McCauley, M. E. et al. C9orf72 in myeloid cells suppresses STING-induced inflammation. *Nature* **585**, 96–101 (2020).

62. Ennerfelt, H. et al. SYK coordinates neuroprotective microglial responses in neurodegenerative disease. *Cell* **185**, 4135–4152 e4122 (2022).

63. Murphy, A. E., Fancy, N. & Skene, N. Avoiding false discoveries in single-cell RNA-seq by revisiting the first Alzheimer's disease dataset. *eLife* https://doi.org/10.7554/eLife.90214 (2023).

64. Gautier, O. et al. Challenges of profiling motor neuron transcriptomes from human spinal cord. *Neuron* **111**, 3739–3741 (2023).

65. Marsh, S. E. et al. Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. *Nat. Neurosci.* **25**, 306–316 (2022).

66. Krienen, F. M. et al. Innovations present in the primate interneuron repertoire. *Nature* **586**, 262–269 (2020).

67. Mitchell, J. M. et al. Mapping genetic effects on cellular phenotypes with 'cell villages'. Preprint at *bioRxiv* https://doi.org/10.1101/2020.06.29.174383 (2020).

68. Rapino, F. et al. Small-molecule screen reveals pathways that regulate C4 secretion in stem cell-derived astrocytes. *Stem Cell Rep.* **18**, 237–253 (2023).

69. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 e3529 (2021).

70. Raudvere, U. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).

71. Szklarczyk, D. et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).

72. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

73. Pietilainen, O. et al. Astrocytic cell adhesion genes linked to schizophrenia correlate with synaptic programs in neurons. *Cell Rep.* **42**, 111988 (2023).

74. Fukuda, A. et al. De novo DNA methyltransferases DNMT3A and DNMT3B are essential for XIST silencing for erosion of dosage compensation in pluripotent stem cells. *Stem Cell Rep.* **16**, 2138–2148 (2021).

75. Hao, J. et al. Loss of TBK1 activity leads to TDP-43 proteinopathy through lysosomal dysfunction in human motor neurons. Preprint at *bioRxiv* https://doi.org/10.1101/2021.10.11.464011 (2021).

Author contributions

The study was designed by F.L. and D.A.M and directed and coordinated by K.E. and S.A.M. with input from B.S. and I.K. Manuscript writing was performed by F.L. and D.A.M., with support from O.P. and A.B. F.L. performed bioinformatics analysis with the help of S.D.G., D.M. and D.A.M. D.A.M. and I.C. supported obtaining postmortem samples and carried out nuclei isolation and RNA sequencing with M.G. and L.B.; Y.Z., F.L., M.T., O.P., A.B., A.C. and B.J.J. performed bioinformatics analyses of bulk RNA sequencing and helped with protein and RNA validation with cellular models; F.L. and J.M.M. performed analysis of published datasets; F.L., M.T. and B.S. contributed to microglial biology section; and K.E. acquired primary funding.

Competing interests

I.K. is an employee at UCB Pharma and holds stock options. K.E. is a cofounder of Q-State Biosciences, Quralis, Enclear Therapies and is group vice-president at BioMarin Pharmaceutical. The remaining authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s43587-024-00640-0.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43587-024-00640-0.

Correspondence and requests for materials should be addressed to Francesco Limone or Kevin Eggan.

Peer review information *Nature Aging* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

¹Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA. ²Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. ³Neuroscience Institute, NYU Grossman School of Medicine, New York, NY, USA. ⁴Department of Pathology, Massachusetts General Hospital, Boston, MA, USA. ⁵FM Kirby Neurobiology Center, Boston Children's Hospital, Boston, MA, USA. ⁶Department of Genetics, Harvard Medical School, Boston, MA, USA. ⁷Howard Hughes Medical Institute, Boston, MA, USA. ⁸Neuroinflammation Focus Area, UCB Pharma, Braine-l'Alleud, Belgium. ⁹Department of Pathology, School of Medicine, Case Western Reserve University, Cleveland, OH, USA. ¹⁰Neuroscience Center, Helsinki Institute of Life Science, University of Helsinki, Helsinki, Finland. ¹¹These authors contributed equally: Francesco Limone, Daniel A. Mordes. ✉e-mail: francesco.limone@nyulangone.org; kevin.eggan@bmrn.com
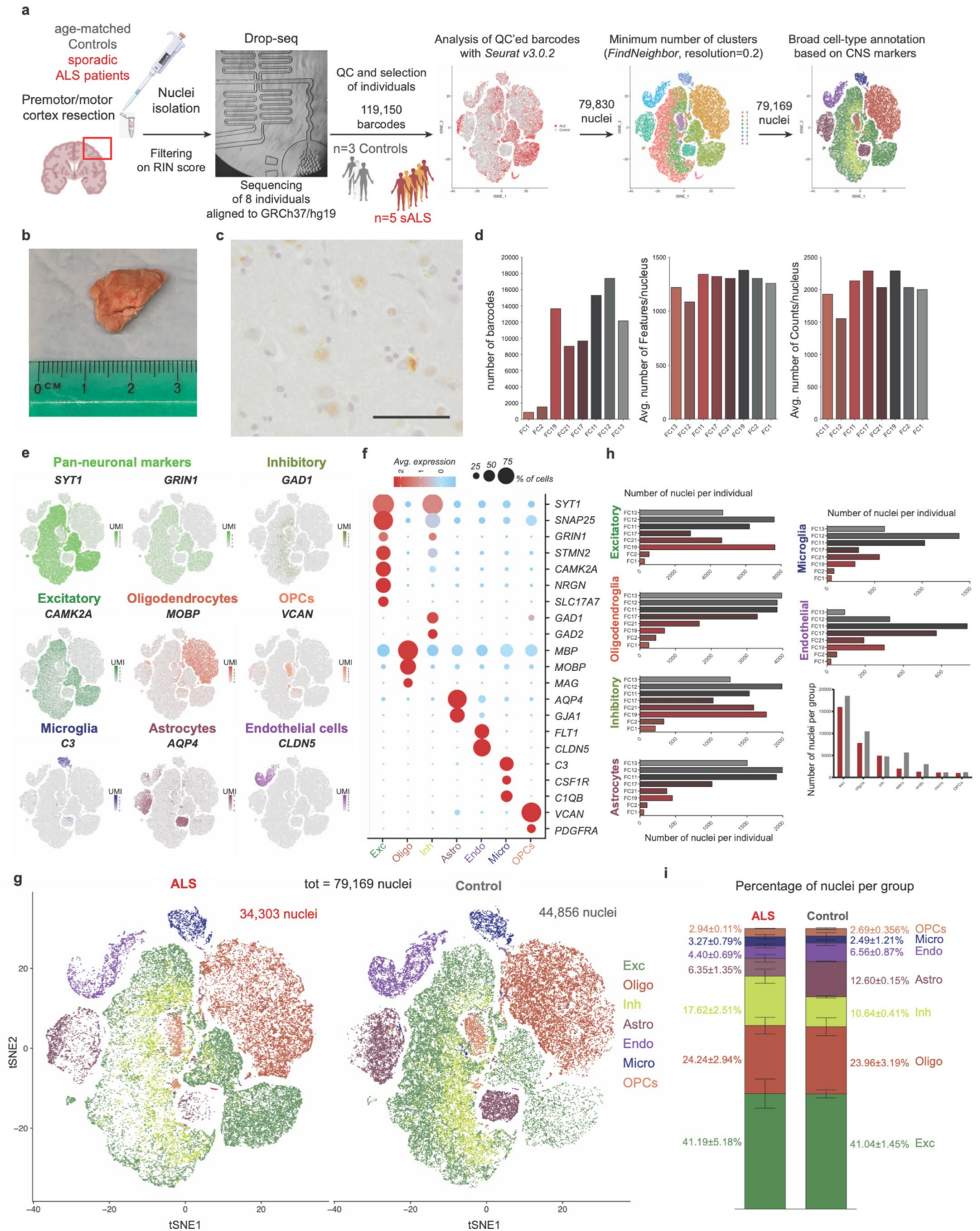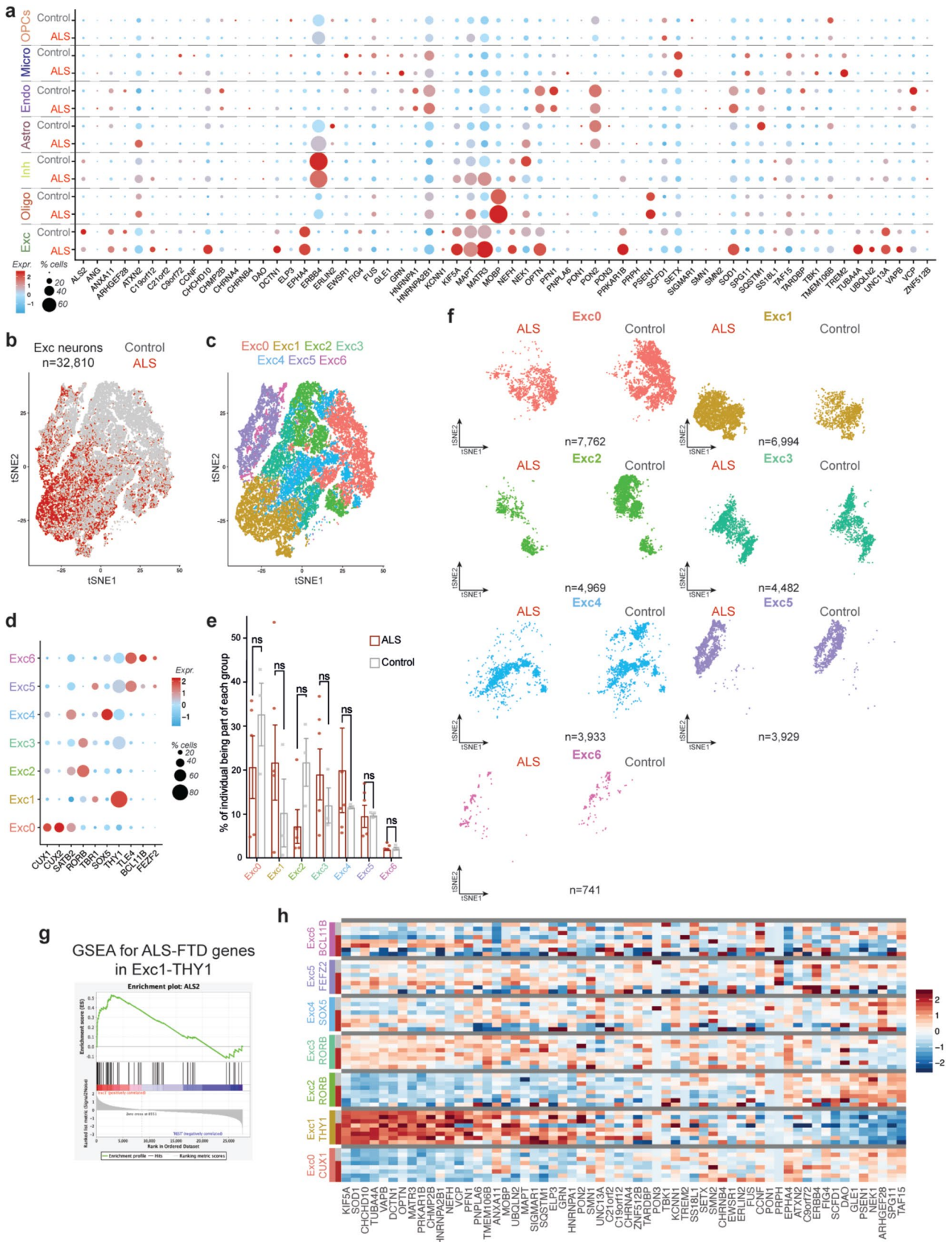
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Technical parameters of snRNAseq and cell-type distribution across individuals. a. Schematic diagram of workflow for isolation of nuclei from cortices of ALS patients and age-matched controls followed by single-cell RNA sequencing by DropSeq, library generation and Quality Controls for analysis with Seurat 3.0.2 **b**. Frozen tissue from one of the individuals. **c**. Staining for TDP-43 in one of the patient sample, note neuron with skein-like inclusions and faint nuclear staining (scale bar 25μm, n = 3 patients analysed). **d**. Quality controls post-filtering (FC – Frontal Cortex): number of total nuclei detected (barcodes), average number of genes per nucleus (nFeatures), and average number of UMIs (Unique Molecular Identifiers) per nucleus (nCounts).

e. t-SNE projections of the whole cohort with expression of broad cell type markers. **f**. Dotplot representing percentage of cells expressing additional cell type specific markers. **g**. t-SNE distribution of whole cohort with annotated cell types split by diagnosis (ALS patients n = 5, age-matched Controls n = 3, n = 79,169 total nuclei). **h**. Quality controls post-filtering (FC – Frontal Cortex): number of total nuclei detected (barcodes), average number of genes per nucleus (nFeatures), and average number of UMIs (Unique Molecular Identifiers) per nucleus (nCounts). **i**. Fraction of each cell types identified in whole cohort split by diagnosis (mean ± SEM, ALS patients n = 5, age-matched Controls n = 3).

Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Expression of ALS-FTD associated genes in different cellular subtypes and excitatory neurons subtypes. a. Dotplot representing expression of gene associated with the ALS-FTD spectrum in each cell type identified in the whole cortex split by diagnosis. **b**. *t*-SNE projection of excitatory neurons clusters (ALS *n* = 15,227 nuclei, Control *n* = 17,583 nuclei). **c**. *t*-SNE projection of subclusters identified in excitatory neurons represents different, biologically relevant neuronal layers (*FindNeighbor*(res=0.2)). **d**. Dotplot representing percentage of cells expressing broad markers for different cortical layers. **e**. Distribution of excitatory neurons in subclusters by diagnosis (mean ± SEM). **f**. *t*-SNE projection of excitatory neurons by clusters and by diagnosis. **g**. Gene Set Enrichment Analysis for the ALS-FTD associated genes in Exc1 excitatory neuron subtype. **h**. Heatmap representing expression of gene associated with the ALS-FTD spectrum in each excitatory neurons identified split by diagnosis.
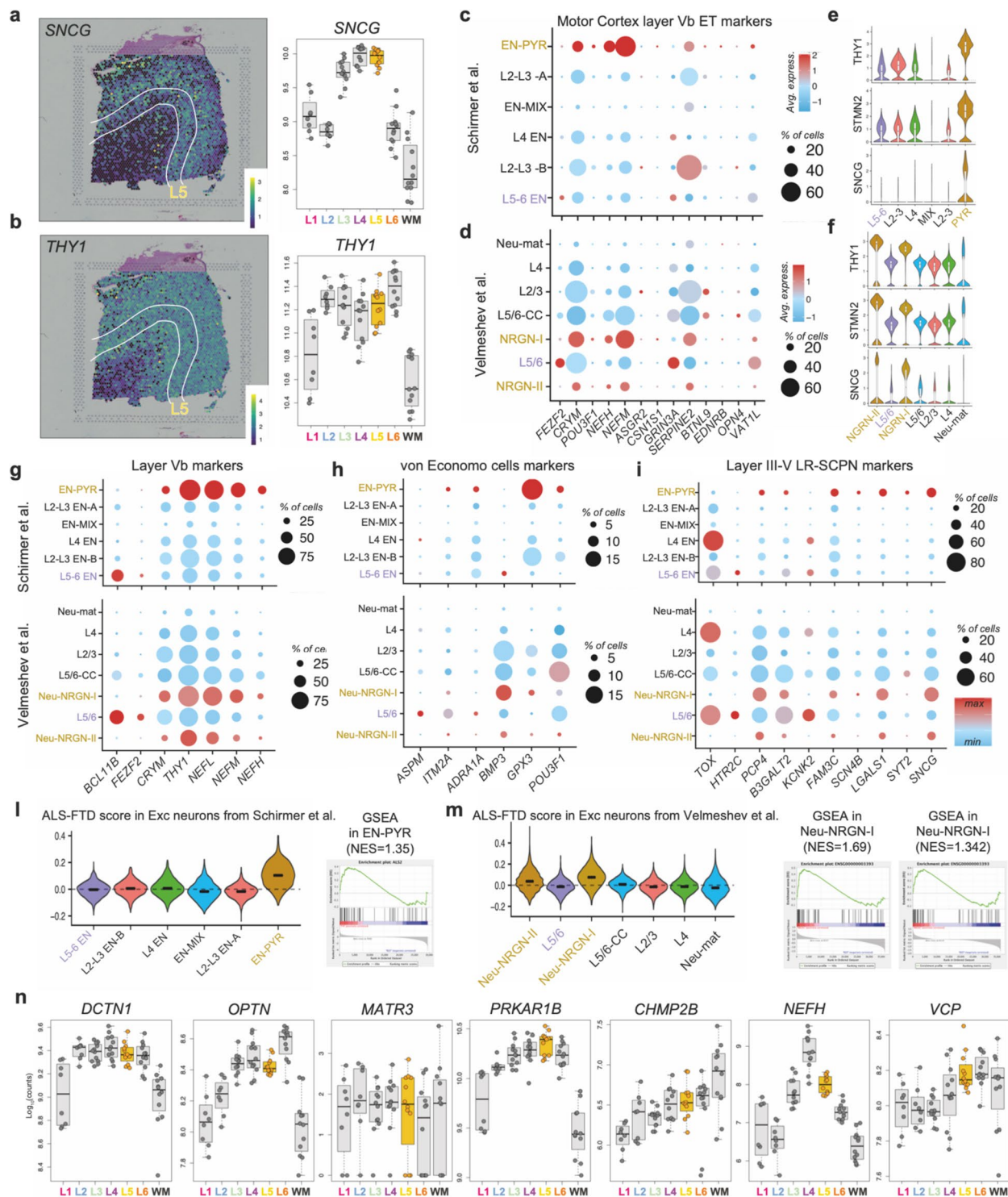
Extended Data Fig. 3 | L5-ETNs/CSMNs-like neurons express higher levels of ALS-FTD related genes. a,b, Spotplot and corresponding boxplot from Maynard et al. for the expression of layer Vb Motor Cortex marker, *SNCG* and *THY1*, identified as enriched in Exc1 (boxplots for mean and interquartile ranges + -SD). **c,d**. Dotplot and representative Violin plots for markers of L5 ExtraTelencephalic neurons of human Motor Cortex in Schirmer et al. **e,f**. Dotplot and representative Violin plots for markers of L5 ExtraTelencephalic neurons of human Motor Cortex in Velmeshev et al. (geometric boxplots for median and interquartile ranges) **g-i**. Dotplot representing expression of Layer V markers (d), von Economo markers

(e), LR-SCPN markers (f) in Schirmer et al. and Velmeshev et al. **l**. Violin plots and corresponding Gene Set Enrichment Analysis of z-scores for expression of ALS-FTD-associated genes in THY1-neurons identified by Schimer et al. (bars denote median). **m**. Violin plots and corresponding Gene Set Enrichment Analysis of z-scores for expression of ALS-FTD-associated genes in THY1-neurons identified by Velmeshev et al. (bars denote median). **n**. Boxplot from Maynard et al. for the expression of additional top ALS/FTD associated genes identified in Exc1 (geometric boxplots for mean and interquartile ranges + -SD) *(n=multiple individuals as reported in publicly available datasets)*
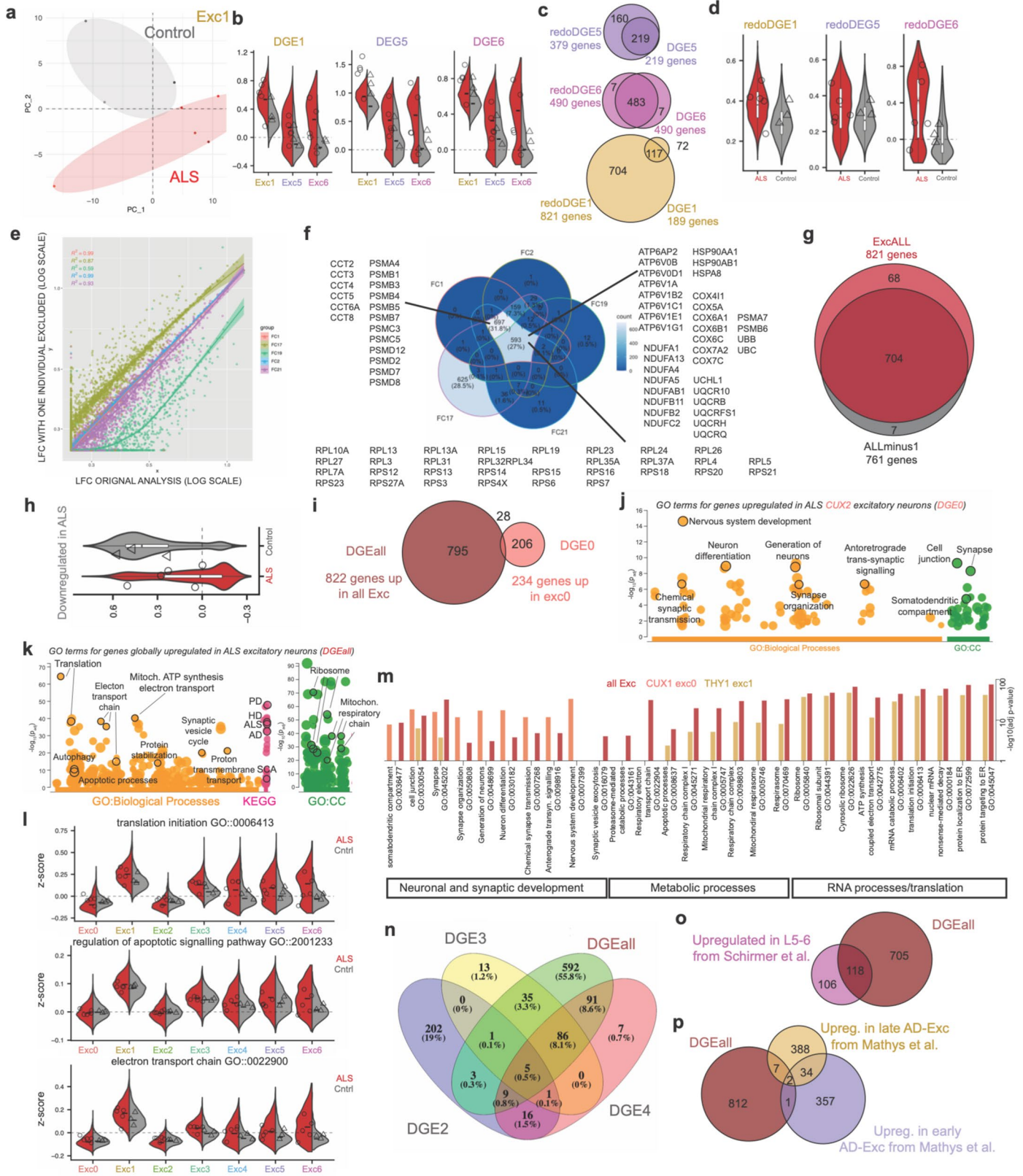
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Classes of L5-ETNs express higher levels of stress pathways. a. DGE-based PCA plots showing separation of individuals by diagnosis in Exc1. **b.** Violin plots of z-scores for genes upregulated in each class of L5-ETNs (DGE1, DGE5, DGE6) in the three groups split by diagnosis (bars denote median – symbols: average score per individual). **c**, Comparison of DGEs signature (>10% cells, >2-FC, adj.p < 0.05) and redoDGE signatures identified in each subgroup with random seeding of equal cell numbers per diagnosis (>10% cells, >2-FC, adj.p < 0.05, max.cells.per.ident=lowest.number.per.group). **d**. Violin plots of z-scores for redoDGE signatures (median and interquantile ranges – symbols: average score per individual). **e**. Correlation plot of Log Fold Changes in ALLminus1 comparisons. **f**. Venn diagram comparing genes shared between ALLminus1 analyses. **g**. Venn diagram comparing genes in common between ALLminus1 analysis and ExcAll. **h**. Violin plots of z-scores for genes globally downregulated in Excall (geometric boxplots represent median and interquantile ranges – symbols: average score per individual). **i**. Comparison of genes globally upregulated in ALS (DGEall) with genes upregulated in CUX2-exc0 (>10% of cells, >2-FC, adj.p < 0.05) **j**. Gene Ontology analysis of terms for DGE0 highlighted terms involved in synaptic biology (CC=Cellular Components). **k.** Gene Ontology analysis of terms for DGEall highlighted terms involved in synaptic biology (CC=Cellular Components). **l,m.** Selected GO terms for DGE0, DGE1 and DGEall. **n.** Venn Diagram for shared upregulated genes between other excitatory neurons and global signature. **o,p.** Venn Diagrams for shared upregulated genes between treated MS excitatory neurons from Schirmer et al. and AD excitatory neurons from Mathys et al. with changes found in this study.
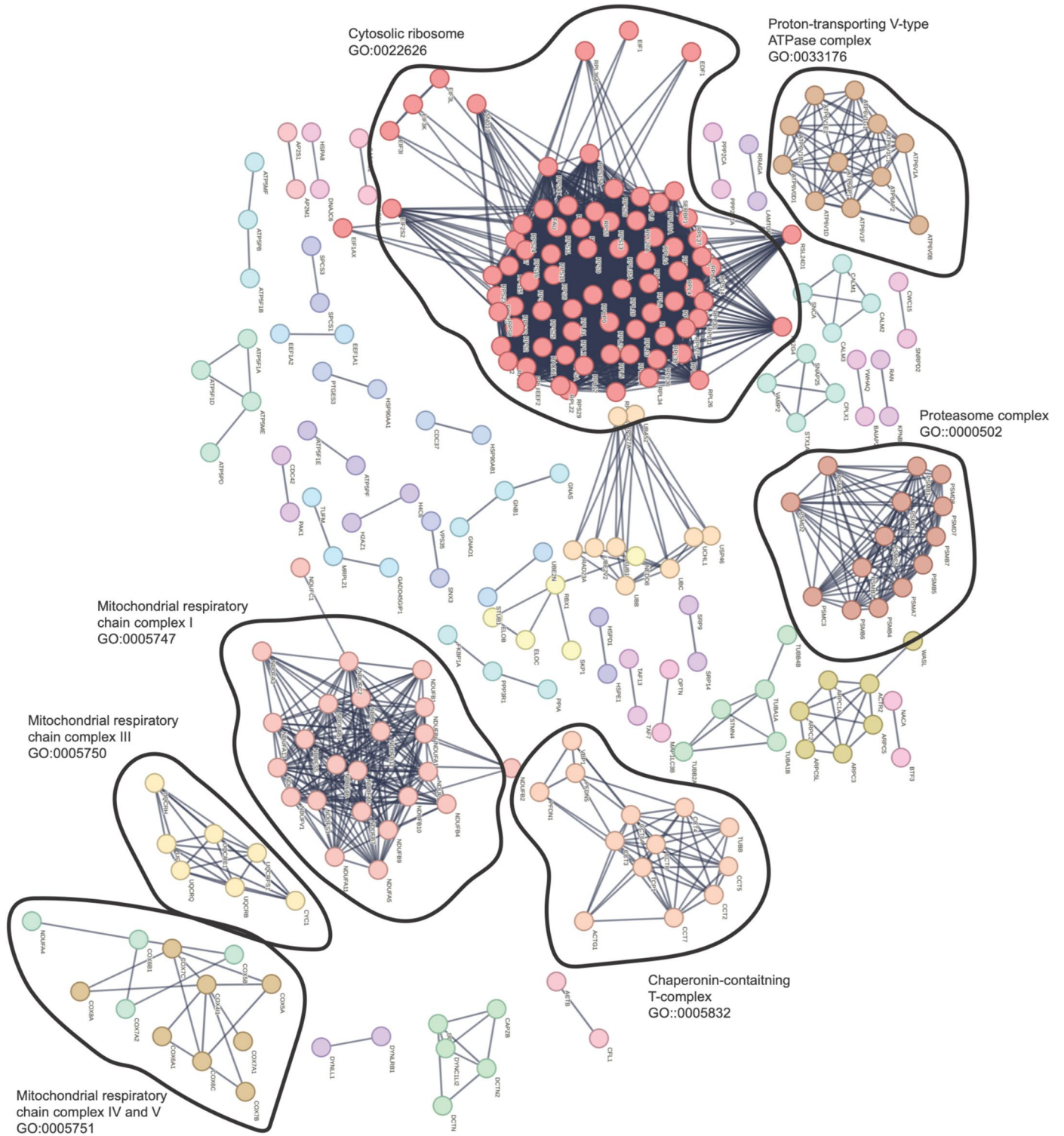
Extended Data Fig. 5 | Global protein-protein interaction network for genes upregulated in ALS excitatory neurons. Color-coding derived from MCL clustering to identified closely related groups of proteins.
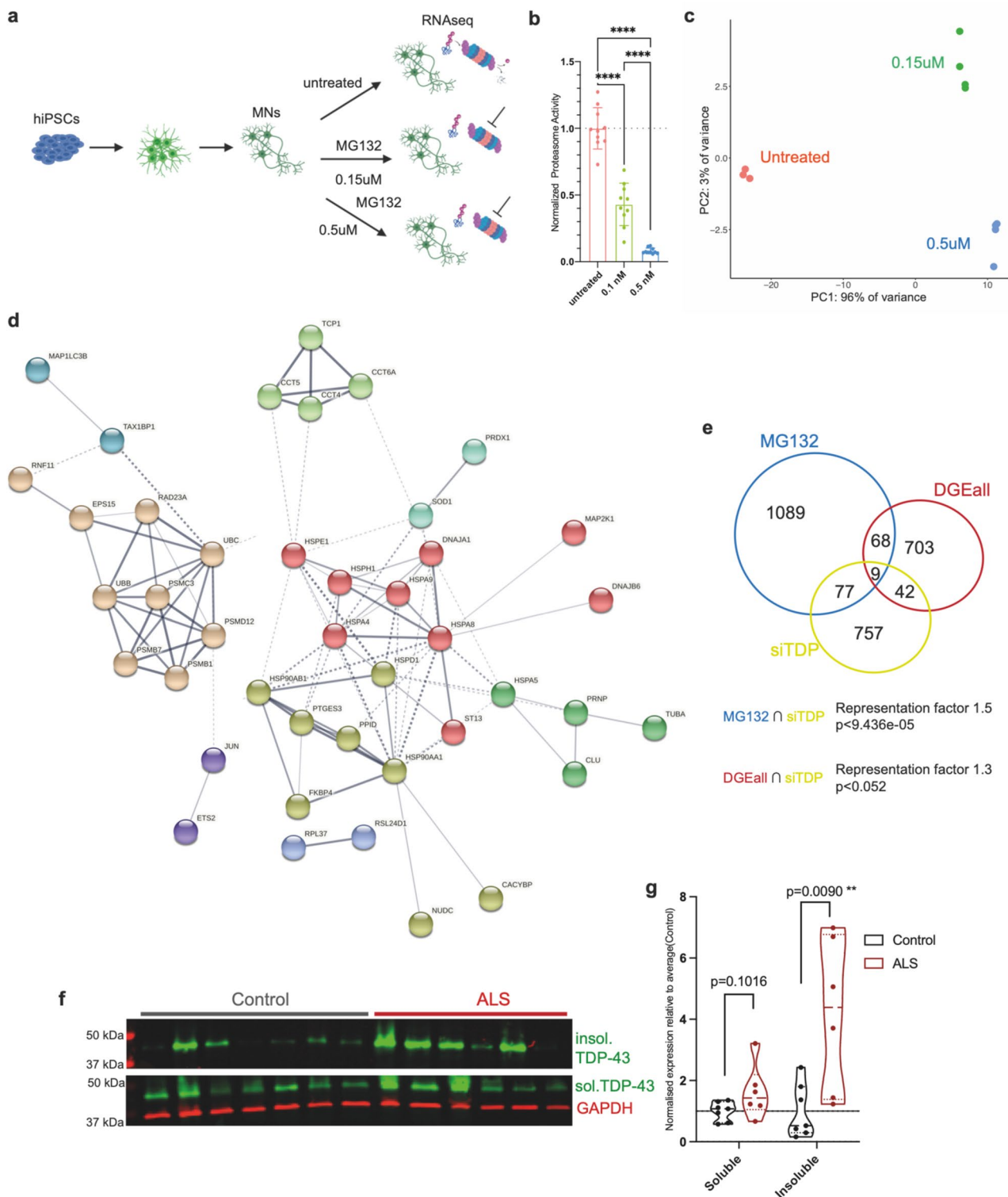
Extended Data Fig. 6 | Proteostatic stress in hPSC-derived neurons resembles changes in excitatory neurons from brain of ALS patients.
a. Diagram of neuronal differentiation from Pluripotent Stem Cells and treatment with proteasome inhibitors for bulk RNA-sequencing. **b**. Quantification of proteasome inhibition (mean ± SD, n = 3 biological replicates, n = 9 technical total, (t-test, ****p < 0.001). **c**. Principle Component Analysis plot showing strong effect of treatments compared to untreated controls (n = 3 biological replicates).

d. Protein-protein interaction network of shared genes. **e**. Venn Diagram depicting shared upregulated genes in treated hPSC-derived neurons (MG132), excitatory neurons from ALS patients (DGEall) and genes misregulated in human neurons after TDP-43 siRNA from Klim et al. (proportional test) **f-g**. Western Blot quantification of soluble and insoluble TDP-43 from motor cortex of ALS patients and age-matched controls (t-test).
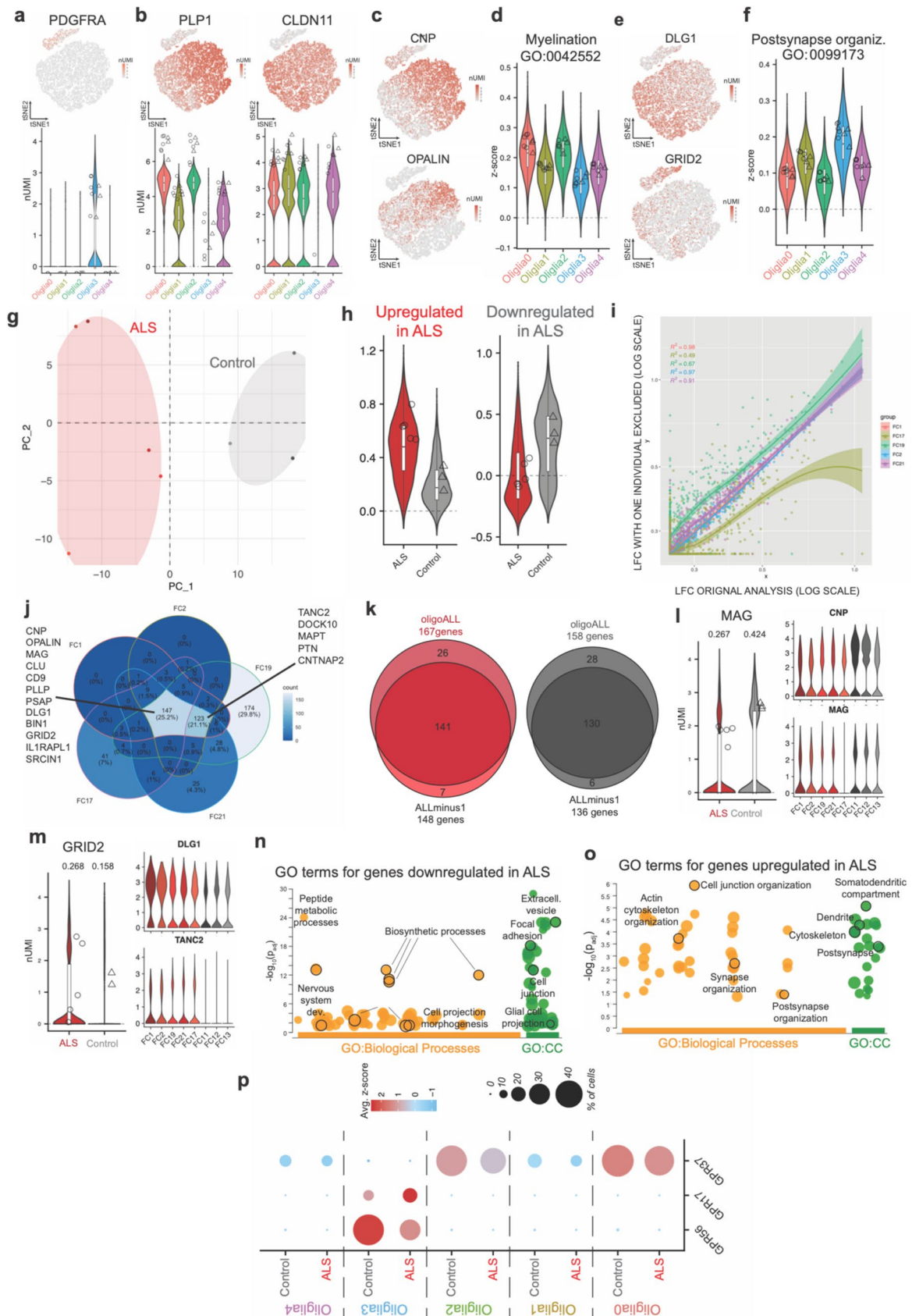
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Oligodendrocytes polarize between myelinating and neuro-engaged states. a,b. *t*-SNE projection and Violin plot of markers of Oligodendrocyte Progenitor Cells (OPCs) and mature oligodendrocytes (geometric boxplots represent median and interquantile ranges – symbols: log2(AverageExpression) per individual) (fraction of cell expressing). **c,d**. *t*-SNE projection of markers of actively myelinating oligodendrocytes and violin plots representing z-score for selected GO terms by cluster (geometric boxplots represent median and interquantile ranges – symbols: average score per individual). **e,f**. *t*-SNE projection of markers of neuronally-engaged oligodendrocytes and violin plots representing z-score for selected GO terms by cluster (geometric boxplots represent median and interquantile ranges – symbols: average score per individual). **g**. DGE-based PCA plot showing separation of individuals by diagnosis. **h**. Violin plot for z-score for genes upregulated and downre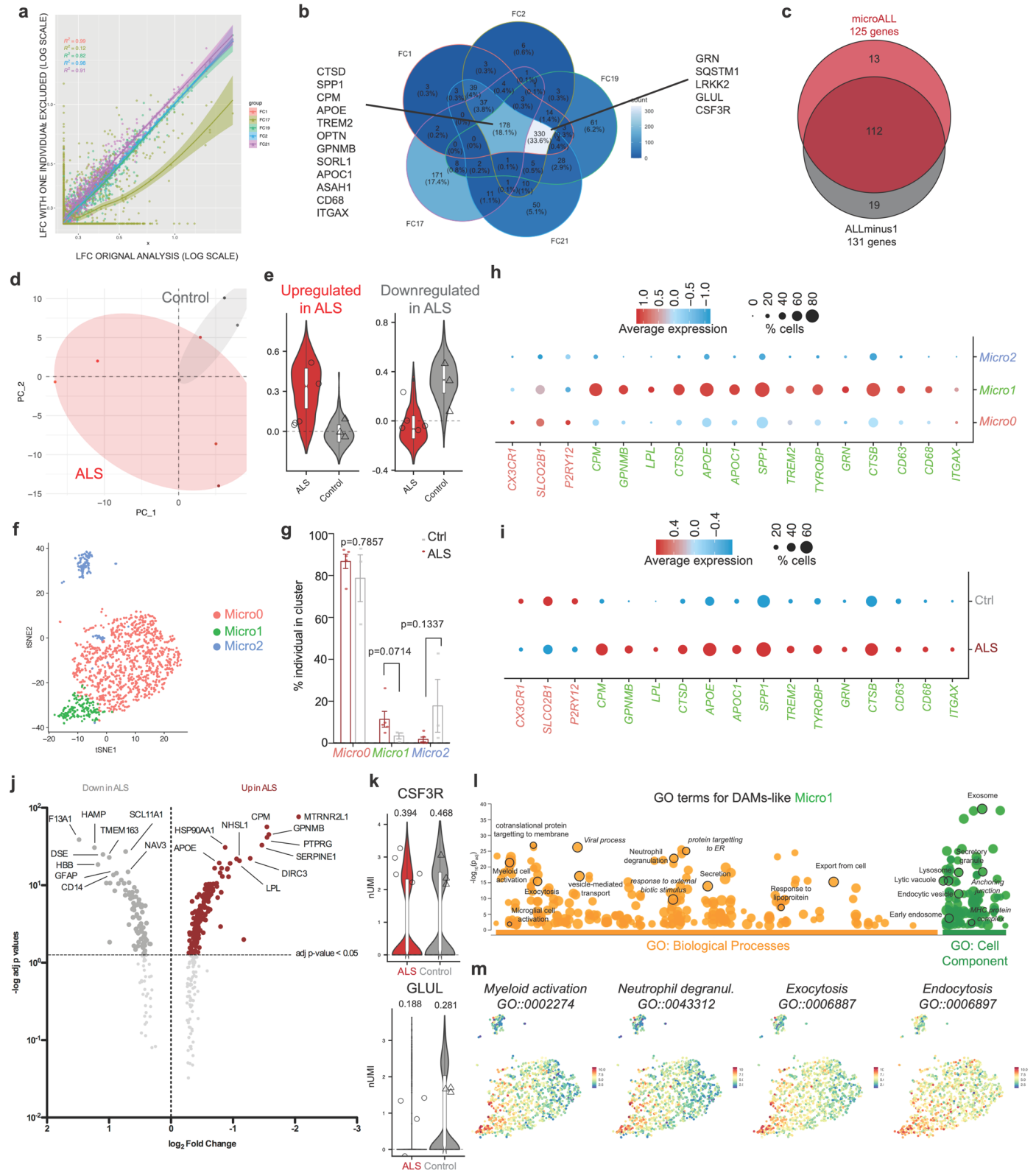gulated in ALS patients (geometric boxplots represent median and interquantile ranges – symbols: average score per individual). **i**. Correlation plot of Log Fold Changes in ALLminus1 comparisons. **j**. Venn diagram comparing genes shared between ALLminus1 analyses. **k**. Venn diagram comparing genes in common between ALLminus1 analysis and OligoAll. **l**. Violin plot of markers of actively myelinating oligodendrocytes (geometric boxplots represent median and interquantile ranges – symbols: log2(AverageExpression) per individual) (fraction of cell expressing). **m**. Violin plot of markers of neuronally-engaged oligodendrocytes (geometric boxplots represent median and interquantile ranges – symbols: log2(AverageExpression) per individual) (fraction of cell expressing). **n,o**. GO analysis for genes downregulated and upregulated in ALS oligodendrocytes, highlighted terms involved in myelination (CC=Cellular Component). **p**. Dotplot representing genes characteristic of maturation and development of OPCs in myelinating oligodendrocytes in each subcluster split by diagnosis.
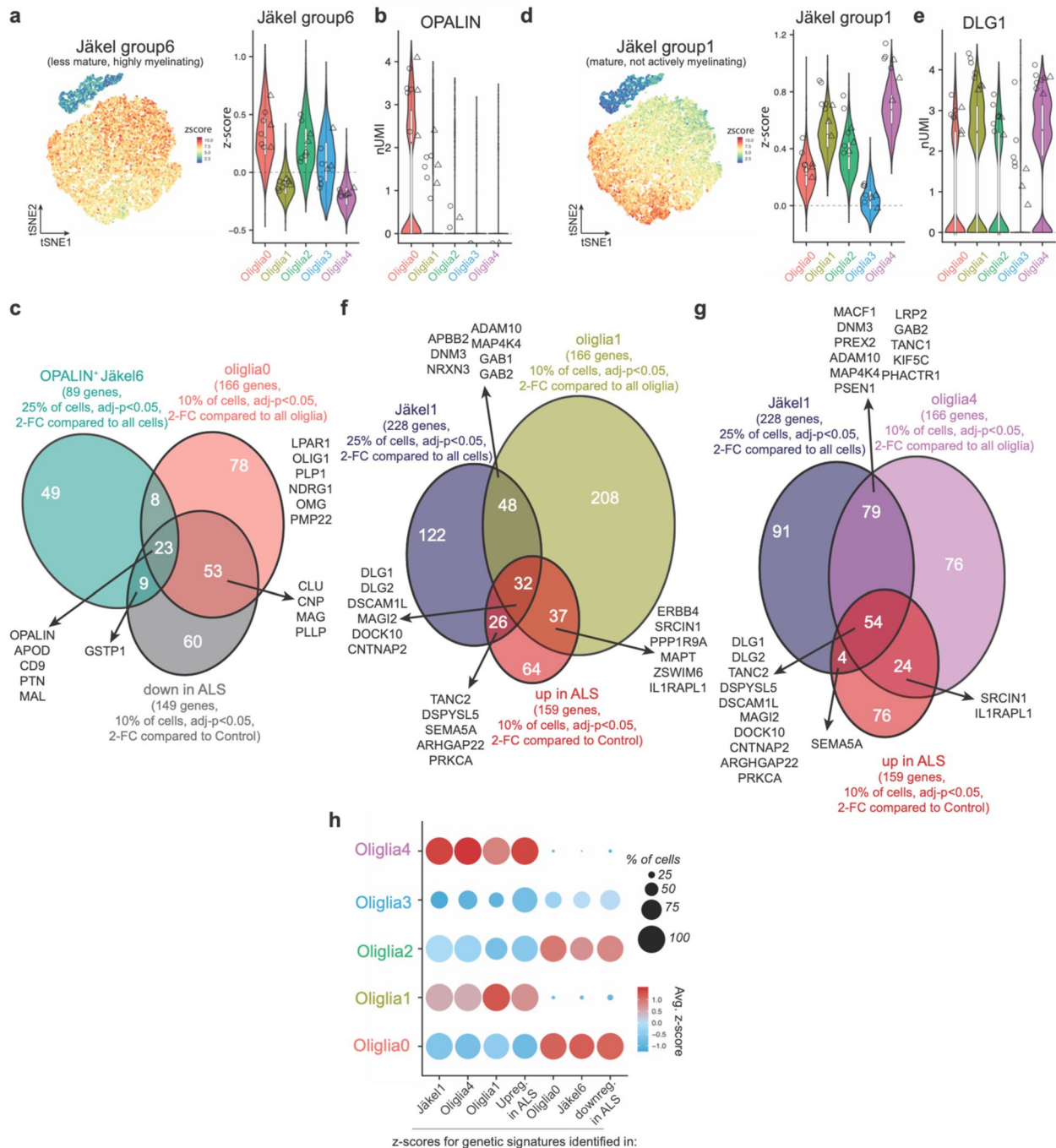
Extended Data Fig. 8 | Comparison of ALS-driven changes with study with similar signature disrupted in disease (MS). a,b. *t*-SNE projection and violin plot representing z-score for highly myelinating, OPALIN+ oligodendrocytes in Jäkel et al (geometric boxplots represent median and interquantile ranges – symbols: average per individual). **c.** Comparison of genes downregulated in oligodendroglia from ALS patients with genes characteristic of highly myelinating, OPALIN+ subtypes identified by this study (oliglia0) and by Jäkel et al (Jäkel6), highlighted genes are shared with GO terms shown in figures. **d,e.** *t*-SNE projection and violin plot representing z-score for genes of mature, not-actively myelinating oligodendrocytes in Jäkel et al (geometric boxplots represent median and interquantile ranges – symbols: average per individual). **f,g.** Comparison of genes upregulated in oligodendroglia from ALS patients with genes characteristic of mature, lowly myelinating groups in this study (oliglia1 and 4) and by Jäkel et al (Jäkel1), highlighted genes are shared with GO terms shown in figures. **h.** Dotplot representing z-scores for the genetic signatures identified in the actively myelinating cells, the mature lowly myelinating cells and DEGs identified in this study. *(for whole panel n = 3 Control individuals, n = 5 sALS patients)*

Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Shared features between ALS-driven changes and reactive subcluster of microglia. a. Correlation plot of Log Fold Changes in ALLminus1 comparisons. **b**. Venn diagram comparing genes shared between ALLminus1 analyses. **c**. Venn diagram comparing genes in common between ALLminus1 analysis and MicroAll. **d**. DGE-based PCA plot showing separation of individuals by diagnosis. **e**. Violin plot for z-score for genes upregulated and downregulated in ALS patients (geometric boxplots represent median and interquantile ranges – symbols: average score per individual). **f**. t-SNE projection of subclusters identified within microglia (Micro0 = Homeo = homeostatic, Micro1 = DAMs = Disease-associated microglia, Micro2 = Cycling cells)). **g**. Distribution of microglia within clusters by diagnosis. **h**. Dotplot representing genes identified as characteristic of Homeostatic microglia and DAMs by subcluster. **i**. Dotplot representing genes identified as characteristic of Homeostatic microglia and DAMs by diagnosis. **j**. Volcano plot of statistically significant differentially expressed genes between Control and ALS microglia (top ten upregulated and top ten downregulated genes highlighted). **k**. Violin plots of representative DEGs downregulated in ALS patients of genes associated with homeostatic microglia (geometric boxplots represent median and interquantile ranges – symbols: log2(AverageExpression) per individual) (fraction of cell expressing). **l**. Gene Ontology analysis of terms associated with genes characteristic of DAMs microglia, highlighted terms playing important role in microglial biology and/or pathogenesis of the disease. **m**. t-SNE projections representing z-score for selected, statistically significant GO terms.
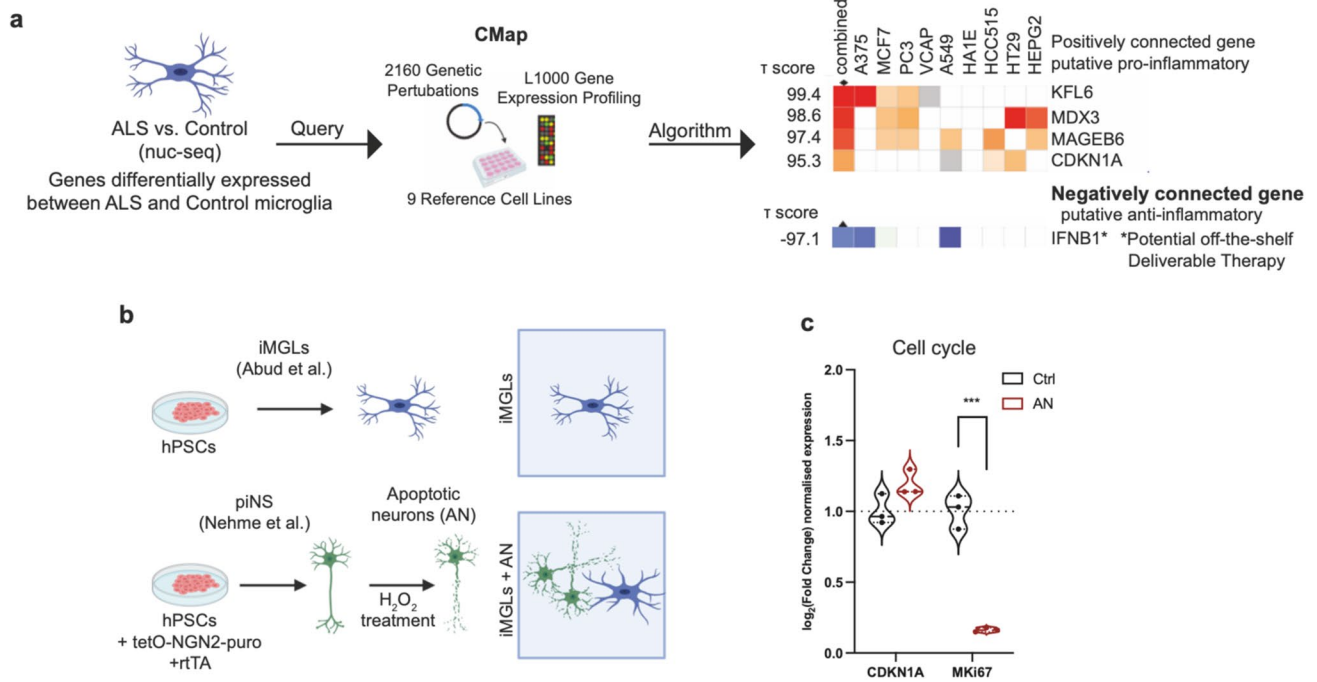
Extended Data Fig. 10 | Apoptotic neurons upregulate lysosomal genes in microglia. a. Schematic of workflow and results from the Connectivity Map project for the genes upregulated in ALS microglia. Heatmap shows what cellular signature is most closely related to the query. **b**. Diagram of microglia and neuronal differentiation from Pluripotent Stem Cells, induction of apoptosis neurons and feeding to iMGLs. **c**. RT-qPCR quantification of cell cycle-associated genes after feeding (t-test, *p < 0.05, **p < 0.01, ***p < 0.001; n = 3 biological replicates).

Corresponding author(s): Francesco Limone & Kevin Eggan

Last updated by author(s): Jan 17, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No custom software was used to collect data. Other programs used to collect and analyze data include: Licor (Image Studios version 2.1), ImageJ (FIJI v 2.14) with Nikon NIS Elements (Version 4.0).
Data analysis	snRNAseq dataset was analysed as previously described in methods, no new code was generated for this study. Softwares and packages used: Seurat v3.0.2 and V4.0.1 ; GSEA v4.0.3 ; GraphPad Prism 7 ; ImageJ (FIJI v2.14) ENSEMBL human reference genome (build GRCh37/hg19) using STAR (v.24.0), Cufflinks (v.2.2.1)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Matrices counts are available at GEO accession number GSE226753. The authors will make any other data available to readers upon reasonable request. All other data analyzed from previously published sources will be available at publications references in the manuscript (for Schirmer et al. Sequence Read Archive (SRA) under accession number PRJNA544731 and NCBI Bioproject ID: 544731 ; for Velmeshev et al. Sequence Read Archive, accession number PRJNA434002 ; for Maynard et al. through GitHub at https://github.com/LieberInstitute/HumanPilot and https://github.com/LieberInstitute/spatialLIBD).

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	No sex- nor gender-specific analyses were conducted because cohort size was not big enough to draw strong sex-based conclusions. Individuals included in the study were picked trying to maintain a sex balance (n= 4 F and 4 M). No gender information was available to researchers because not relevant for this particular study.
Population characteristics	Relevant information provided in Extended Data Table 1. average age 61.125, n=4 males, n=4 females. average disease duration 3.875 years
Recruitment	Individuals were selected by Target ALS Nueropathology Core and MGH Biobank and partners following IRB and approved protocols.
Ethics oversight	Informed consent for motor cortices from ALS patients and controls were obtained at MGH using a Partners IRB approved protocol and stored at –80°C. Study protocol was further approved by Harvard Stem Cell and Regenerative Biology Department, Harvard University. Informed consent and study protocol for human Stem Cell work was provided by Stanley Center for Psychiatric Research at Broad Institute of MIT and Harvard, United States, and the Harvard Stem Cell and Regenerative Biology Department at Harvard University, United States.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

Sample size	Samples were selected based on availability from biobanks. Samples were selected to balance sex and diagnosis. Number of nuclei sampled and sample size are sufficient to draw preliminary conclusions given the protein validation is carried out on additional cohort.
Data exclusions	No data was excluded
Replication	All findings were replicated in technical replicates. V alidations in stem cell -based studies were carried out in several technical replicate and biological replicates. Protein validations for RNA changes was carried out in replication cohort of patients.
Randomization	Experimental groups were based on diagnosis and/or treatments
Blinding	Investigators were not blinded for sequencing studies, patients IDs were available with samples from the beginning of the study . Investigators were blinded to treatments for stem cell based experiments

All studies must disclose on these points even when the disclosure is negative.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	TDP-43 (Peprotech 10782-2-AP), TUJ1 (R&D NL1195G), GAPDH (Millipore Cat# MAB374; CST 2118 (14C10)) MBP (ThermoFisher PA-1-10008), CNP (Abcam ab6319 (11-5b)), 20S (Enzo BML-PW8195-0025), Ubiquitin (CST 3936T (P4D1)). All used in 1:1000 dilutions. IRDye provided by Licor used as secondary antibodies (1:10,000 dilution).

| Validation | All antibodies have their respective source company and clone number and are validated for the applications used within this manuscript. This information is available on the manufacturers publicly available datasheets.
TDP-43 Positive WB detected in SH-SY5Y cells, HeLa cells, C2C12 cells, Neuro-2a cells. TUJ1 Detects mammalian and chicken neuron-specific beta -III tubulin but not other beta -tubulin isotypes in Western blots. GAPDH GAPDH (14C10) Rabbit mAb detects endogenous levels of total GAPDH protein. GAPDH enzyme is detected in many non-muscle cells lines including HeLa, HCT-116 cells, U937 and THP-1 cells among others. MBP rat spinal cord whole tissue homogenates. CNP Human and Mouse Spinal Cord and Brain tissue lysates. 20S Recognizes the α1, 2, 3, 5, 6 & 7 subunits of the 20S proteasome. Ubiquitin Ubiquitin (P4D1) Mouse mAb detects ubiquitin, polyubiquitin and ubiquitinated proteins |

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	HUES3 Hb9::GFP is a human embryonic stem cell line derived and available at Harvard University study. WA01 is a human embryonic stem cell line derived and available at the University of Wisconsin Madison. The iPS cell lines (11a, 15b, 17a, 18a, 20b) were generated in our lab with fibroblasts under IRB approved protocols of collaborative study with Dr. Chris Henderson and are available at the ALS Clinic at Columbia University.
Authentication	Cell lines were tested for karyotipic abnormalities.
Mycoplasma contamination	All cell lines were Mycoplasma negative, tested twice a month
Commonly misidentified lines (See ICLAC register)	No commonly misidentifed lines were used in the study.

Palaeontology and Archaeology

Specimen provenance	*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.*
Specimen deposition	*Indicate where the specimens have been deposited to permit free access by other researchers.*
Dating methods	*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

| Ethics oversight | *Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.* |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	*For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Reporting on sex	*Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies
All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

| Clinical trial registration | *Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.* |

Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No | Yes
⊠ ☐ Public health
⊠ ☐ National security
⊠ ☐ Crops and/or livestock
⊠ ☐ Ecosystems
⊠ ☐ Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No | Yes
⊠ ☐ Demonstrate how to render a vaccine ineffective
⊠ ☐ Confer resistance to therapeutically useful antibiotics or antiviral agents
⊠ ☐ Enhance the virulence of a pathogen or render a nonpathogen virulent
⊠ ☐ Increase transmissibility of a pathogen
⊠ ☐ Alter the host range of a pathogen
⊠ ☐ Enable evasion of diagnostic/detection modalities
⊠ ☐ Enable the weaponization of a biological agent or toxin
⊠ ☐ Any other potentially harmful combination of experiments and agents

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

| Data access links
May remain private before publication.	*For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*
Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session
(e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	*Describe the experimental replicates, specifying number, type and replicate agreement.*
Sequencing depth	*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*
Antibodies	*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Peak calling parameters	*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*

Data quality	*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*
Software	*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*
Instrument	*Identify the instrument used for data collection, specifying make and model number.*
Software	*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*
Cell population abundance	*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*
Gating strategy	*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI   ☐ Used   ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g.*

| Normalization template | *original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.* |

| Noise and artifact removal | *Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).* |

| Volume censoring | *Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.* |

Statistical modeling & inference

| Model type and settings | *Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).* |

| Effect(s) tested | *Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.* |

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

| Statistic type for inference (See Eklund et al. 2016) | *Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.* |

| Correction | *Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).* |

Models & analysis

n/a | Involved in the study
☐ ☐ Functional and/or effective connectivity
☐ ☐ Graph analysis
☐ ☐ Multivariate modeling or predictive analysis

| Functional and/or effective connectivity | *Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).* |

| Graph analysis | *Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).* |

| Multivariate modeling and predictive analysis | *Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.* |