Differential Cellular Mechanisms Underlie Language and Executive Decline in Amyotrophic Lateral Sclerosis

Joana Petrescu[1,2]*, Cláudio Gouveia Roque[1]*, Christopher A Jackson[1]*, Aidan Daly[1]*, Kristy Kang[1], Obadele Casel[1], Matthew Leung[1], Luke Reilly[1], Jacqueline Eschbach[2], Karina McDade[4,5], Jenna M. Gregory[6,7], Colin Smith[4,5], Hemali Phatnani[1,2,3] †

[1]Center for Genomics of Neurodegenerative Disease, New York Genome Center, 101 Avenue of the Americas, New York, NY 10013.

[2]Department of Neurology, Center for Translational and Computational Neuroimmunology, Columbia University, New York, NY 10032, USA.

[3]Center for Motor Neuron Biology and Disease, Columbia University Irving Medical Center, New York, NY, USA.

[4]Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, EH16 4TG, UK.

[5]UK Dementia Research Institute, University of Edinburgh, Edinburgh EH16 4SB, UK

[6]Institute of Medical Sciences, University of Aberdeen, Aberdeen, UK.

[7]Department of Pathology, NHS Grampian Tissue Biorepository, Aberdeen, UK.

*These authors contributed equally to this work.

† Correspondence: hphatnani@nygenome.org (H.P.)

Abstract

Half of all amyotrophic lateral sclerosis (ALS) patients demonstrate a spectrum of cognitive and behavioral changes over the course of the disease, but the mechanisms underlying this heterogeneity remain unclear. We assemble a high-resolution cellular map of prefrontal cortex regions of the ALS brain by integrating spatial and single-nucleus transcriptomic profiles of a cognitively stratified ALS patient cohort that includes non-neuropathological controls. We find cellular programs characteristic of ALS, including a frequent gliotic response. We also find that executive and language cognitive impairments differ from ALS without cognitive impairment, and from each other, in the extent and pattern of neuronal dysregulation and neuron-glial interactions across different brain regions. These findings reveal a relationship between cognitive phenotype and prefrontal cortex dysfunction in ALS.

Introduction

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disorder underpinned by a multicellular etiology and characterized by progressive paralysis, cognitive impairment, and behavioral changes (*1*, *2*). While up to half of all patients exhibit some degree of extra-motor involvement, the clinical presentation of ALS is highly heterogeneous, existing on a spectrum ranging from cases with normal cognition and behavior to ALS with frontotemporal dementia (ALS-FTD) (*1*, *2*). A mid-spectrum range, which encompasses about 35% of ALS patients, is defined by varying patterns and degrees of cognitive and/or behavioral manifestations, including deficits in language (e.g., spelling and naming), a decline in executive functions (e.g., decision-making), and apathy (*1*, *2*). Brain imaging and postmortem analyses have shown that these cognitive and behavioral profiles are predictive of underlying dysfunction in specific areas of the brain, highlighting potential regional susceptibilities to ALS pathophysiology (*1*, *3*–*7*). It remains unclear, however, whether the clinical phenotypes observed in ALS stem from a shared root mechanism or instead emerge from brain region-dependent interactions of distinct cellular, molecular, neuropathological, and genetic factors.

To address this question, an integrative approach across disciplines is required—one that bridges clinical data and autopsy findings with contextualized cell-level responses. In practice, however, several factors make its implementation challenging, including a lack of standardized cognitive testing in ALS clinical care (*1*). This limitation, in turn, has hindered efforts to disentangle biological differences behind the cognitive heterogeneity observed in ALS. Additionally, while technological advances have recently enabled a cellular view of ALS (*8*–*12*), the interactions and dependencies within each cellular community remain poorly understood. Here, we combined spatially resolved and single-nucleus transcriptomic profiling of a deeply

phenotyped ALS patient cohort to interrogate how differences in cell types and intercellular relationships relate to cognitive defects. We performed these analyses across two non-motor cortical regions implicated in ALS cognitive impairment, the dorsolateral prefrontal cortex and the inferior frontal gyrus, and prioritized patients with predominant language or executive dysfunction in our sampling.

Globally, our experimental strategy highlighted clear cellular and molecular differences between cognitive subgroups, particularly in neuronal cells, but also common signatures shared by every ALS patient. Pan-cellular programs broadly affected within each brain region were also identified, with the synapse emerging as a focal point of gene dysfunction associated with cognitive deficits. While some of these differences were already present in cognitively normal patients, the patterns and scope of dysregulation we observed were generally more pronounced in cognitively impaired ALS individuals. Overall, our combined patient stratification approach with spatially-resolved, transcriptome-wide molecular profiling uncovered measurable early signatures associated with cognitive decline in ALS. The analyses we provide further show that language and executive deficits in ALS have distinct molecular and cellular profiles, informing future investigations and the pursuit of tailored therapeutic strategies.

Results

Study Overview and Cohort Demographics

We investigate the biological bases of cognitive impairment in ALS in a group of ALS patients and sudden-death controls (Figure 1A). ALS patients underwent neuropsychological testing using the Edinburgh Cognitive and Behavioral ALS Screen (ECAS), a multidomain assessment tool that includes quantitative evaluations of executive function, language, and verbal fluency—three cognitive domains commonly affected in ALS patients (*1*, *2*, *13*). The ECAS evaluation has individual cognitive domain scores that capture subtleties in cognitive presentation and an aggregate ALS score that is diagnostic for cognitive impairment in general (*13*). Cognitive deficits in ALS are associated with localized cerebral and subcortical dysfunction, and the underlying regional abnormalities are predicted by the ECAS profile of cognitive change in each patient (*1*, *3*–*7*). We leveraged this structure-function relationship to investigate molecular and cellular signatures linked to domain-specific deficits in ALS. For patients with predominant executive dysfunction profiles, we sampled from the dorsolateral prefrontal cortex (DLPFC), specifically Brodmann Area (BA) 46, which supports decision-making and other high-level executive processes (*4*, *14*–*16*). For those predominantly exhibiting language impairments, we prioritized BA44/45 in the inferior frontal gyrus of the left hemisphere, commonly known as Broca's area (*17*), given its strong links to speech production and articulation (*4*, *16*, *18*). Verbal fluency, which requires activation of BA44/45 and BA46 networks, was investigated in connection to both anatomical regions in patients exhibiting impairments in this domain (*4*, *6*, *16*). Building on this sampling strategy, we combined capture-based spatial transcriptomics (ST) and single-nucleus RNA sequencing (snRNA-seq) to obtain an unbiased and comprehensive view of the ALS brain.

All ALS individuals in our study (*n* = 36) were participants in the CARE-MND initiative (*19*), a Scotland-wide patient register unmatched in its depth, detail, and systematization of clinical information. This patient cohort was complemented by a control group consisting of non-neurological, sudden death donors (*n* = 17) (*20*). Postmortem brain tissue from both ALS patients and control individuals was sourced from a single biorepository—the Edinburgh Brain Bank—to ensure procedural reproducibility across samples, including consistent preservation methods, standardized neuropathological assessments, and controlled shipping conditions. Based on ECAS test scores, each donor was stratified into one of three analysis groups: (1) ALS patients with cognition in the normal range (hereafter referred to as 'ALS'), representing a pure motor form of the disease; (2) ALS patients with cognitive impairment (ALSci); and (3) neurotypical controls. We note that ALSci patients are non-demented and do not meet the diagnostic criteria for ALS-FTD, falling instead within a mid-spectrum range and exhibiting milder, more focal impairments than ALS-FTD cases (*1*, *2*).

We provide clinical phenotyping data and disease metrics for this unique cohort in Figures 1B-H. Age at death was balanced between females and males across the three groups (Figure 1B), and the median survival of ALS patients from disease onset was comparable between ALS and ALSci patients, although the ALSci group included a single long-term survivor (Figure 1C). The number of donors was evenly distributed across clinical groups (Figure 1D), and we provide in Figure 1D a breakdown of ECAS subtest scores for each ALS patient. Consistent with published literature (*13*, *21*), language was the most frequently affected domain in ALSci donors, but the final cohort had also good coverage of executive and fluency impairments (Figures 1D, E). Figure 1F summarizes key cohort demographics alongside details of our sampling approach, we observed a weak correlation (R = -0.11) between patient age and

ECAS test scores (Figure 1G), suggesting that something intrinsic to the pathophysiology of ALS—and not the age of the participant—primarily drives the decline in cognitive performance seen in some patients. Finally, although the cohort is heterogeneous in relation to both TDP-43 pathology and *C9ORF72* genotype, these variables were overall poor predictors of cognitive status (Figure 1H).

A high-resolution spatial cell atlas of BA44/45 and BA46 from the ALS Prefrontal Cortex

From this donor cohort, we sequenced and analyzed 246 Visium arrays (BA44/45: 137, BA46: 109) which passed quality control, and we annotate 953,073 individual array spots from these arrays into cortex layers L1-L6 and white matter (WM) (Supplemental Figure S1). We preprocess and project each array spot into a Uniform Manifold Approximation and Projection (UMAP) plot, where cortex layer L1, layers L2-6, and WM are clearly separated (Figure 1I). The proportion of PFC layers, with large L3 and L5 layers, is expected and does not change between Control, ALS, and ALSci donors (Figure 1J).

Layer-specific markers localize normally (Figure 1K, Supplemental Figure S2A), and deconvoluted cell types localize to the expected layers, with no change in abundance between our Control, ALS, and ALSci donors (Supplemental Figure S2B-E), with the possible exception of a small decrease in inhibitory neurons (INs) in ALS and ALSci donors. We conclude from this that there are no gross structural changes present in ALS and ALSci, as compared to the controls. We also have a notable outlier; ALS donor SD0015/16 (N=7 arrays) is entirely separated from the rest of the cohort in the UMAP plot, and has extremely high interferon and antiviral transcriptomic expression (Supplemental Figure S2F). We exclude this individual outlying donor from our analyses.

A paired snRNA-seq and snATAC-seq cell atlas of BA44/45 and BA46 from the ALS Prefrontal Cortex

In parallel from this donor cohort, we processed and sequenced 48 10x paired snRNA-seq and snATAC-seq libraries (BA44/45: 24, BA46: 24). To facilitate analysis and comparison, we also obtained 2.6m snRNA-seq PFC cells from published literature, and used it to build a machine learning classifier that predicts annotations from prefrontal cortex gene expression (Supplemental Figure S3). After preprocessing and annotating cells with our classifier (Supplemental Figure S4), we have 144,260 cells from 48 samples, annotated into broad cell types (Figure 1L). We find that low-count ATAC cells contribute little meaningful information, and so from the 144k annotated gene expression cells, we filter to only cells with 2500 ATAC counts, which allows for preprocessing and projection without requiring aggregation, imputation, or denoising. This yields 119,153 high-quality paired RNA and ATAC cells that largely separate by cell type when projected using only ATAC count data (Figure 1M). The snRNA-seq cells are split evenly between neurons and glial cells, although we have poor representation of endothelial cells that is most likely due to the sample processing method (Figure 1N). All of the broad cell types are present in almost all libraries, and excitatory neurons (ENs) and inhibitory neurons (INs) are present in a 3:1 ratio, consistent with published single cell prefrontal cortex datasets (*22*). Finally, as with the spatial data, we do not observe changes in the abundance of the broad cell types between control, ALS, and ALSci donors (Figure 1O).

Glial cells from ALS and ALSci donors have common transcriptional changes

In our preliminary data exploration, we noted that astrocytes, microglia, oligodendrocytes, and oligodendrocyte precursor cells (OPCs) show some ALS phenotype-specific separation in UMAP plots (Figure 2A). Some of this patterning is likely attributable to donor-specific transcriptional differences (Supplemental Figure S5A), but there are no significant changes to the number of glial cells present in controls vs ALS or ALSci donors (Supplemental Figure S5B-D). Instead, we find that there is differential gene expression between control and ALS/ALSci donors within each glial cell type, and when analyzed for gene ontology enrichment, these expression changes suggest alteration in core glial cell functions (Supplemental Figure S5E).

ALS and ALSci donors have glial cells with transcriptional signatures of gliosis

We categorize astrocytes into homeostatic or reactive subtypes; microglia into homeostatic, reactive, phagocytic, other immune cell, or a donor-specific stress/heat shock subtype; oligodendrocytes into high myelinating and low myelinating subtypes; and oligodendrocyte progenitor cells into homeostatic, maturing, and stressed subtypes (Figure 2B). We do this by separately clustering each glial cell type into small, high-resolution cell clusters (see Methods for details), and then functionally annotating every cluster into one of several broad functional categories based on the expression of highly-expressed functional marker genes (Figure 2C, Supplemental Figure S6A-D).

The proportion of reactive microglia, and stressed OPCs significantly increases in ALS and ALSci donors relative to controls, with larger increases in ALSci than ALS, while the proportion of reactive (*GFAP*$^{+}$) astrocytes significantly increases in ALSci patients only (Figure

2D, Supplemental Figure S6E). Furthermore, we find that the proportion of glial cells with homeostatic transcriptomes are correlated within samples, as are the proportion of glial cells with reactive transcriptomes, and the proportion of homeostatic glial cells is anticorrelated with reactive glial cells (Figure 2E). Relationships between these functional proportions are not always linear, as exemplified by the high correlation between reacting astrocytes and stressed OPCs (Figure 2F, Supplemental Figure S6F).

We identify differentially expressed transcripts between each functional group, and filter for differentially expressed transcripts which are specific to the glial cell type (Supplemental Figure S6G). We use these transcripts to define functional modules which we can quantify from our spatial transcriptomic data, and then infer spatial changes in glial cell state between control and ALS donors. With this approach, we determine that the reactive astrocytes, reactive microglia, and stressed OPCs are increasing uniformly in L1-L6 and the white matter of ALS/ALSci donors, but that changes to oligodendrocytes appear limited mainly to neuron-rich L2-L6 (Figure 2G). We conclude that oligodendrocyte functional changes most likely have a different cause than the other reactive glial changes, and that this cause may be linked to neuronal changes.

Spatial gene expression is modeled with Splotch and clustered into spatial co-expression modules

Based on differences in the spatial distribution of glial cells, we suspect that ALS donors have more than one spatially correlated transcriptomic signature. In order to identify these transcriptomic signatures in our BA44/45 and BA46 ST cohort, we fit Splotch (*12*, *23*, *24*), a Bayesian hierarchical model designed for spatial transcriptomic data (Figure 3A). This model fits

a zero-inflated poisson distribution to the observed count data, where the Poisson lambda parameter is the expected mean transcript abundance at an array spot; this parameter is in turn modeled as a linear combination of variables accounting for the effects of spot- and donor-level metadata (the cortical layer annotation and ALS diagnosis/presentation), spatial autocorrelation with nearby spots, and spot-specific noise (Supplemental Figure S7A-C). We use the fit model to determine $\log_2$ fold change between control and ALS/ALSci donors for every gene within annotated cortex layers, and calculate bayes factors between posterior distributions over relevant model parameters for significance testing (Supplemental Figure S7D-E; Methods).

The Poisson lambda parameters from the fit splotch model are denoised estimates for transcript expression at every array spot. These denoised estimates have very similar spatial correlation in both BA44/45 and BA46 samples (Figure 3B), reflecting the underlying cortex structure and cell composition. We embed each transcript into a k-Nearest Neighbors (k-NN) graph, using pearson correlation distance, and cluster them with the leiden graph clustering algorithm into co-expressed spatial modules (Figure 3C). This yields 23 spatial modules, each of which represents a set of genes that are expressed in the same spatial locations within the cortex, although one is composed only of all chromosome Y (chrY) genes, and one is composed of all mitochondrially-encoded (chrM) genes (Figure 3Di).

There are 21 co-expressed spatial modules that are not chromosome-specific, for which we can calculate a module score for each cell in the snRNA-seq data and for each Visium array spot. Module scores represent the aggregate expression of all transcripts in the co-expressed module within a cell or spatial location. By examining the highest scored module within each snRNA-seq cell (Figure 3Dii), and by taking the average expression for each module in each cell type (Figure 3E), we determine the cell type specificity for each spatially co-expressed

module. Most glial cell types have one or two characteristic spatial modules, but neuronal cells are represented by many spatial co-expression modules, indicating more complex spatial and transcriptional structure.

ALS donors have spatially localized transcriptomic changes in co-expression modules

Aggregating spatial co-expression module scores from the spatial arrays shows cortex layer-specific expression, while module scores from the snRNA-seq data further show cellular association. Modules M15, M6, M5, and M18 represent L2, L3-L4, L4-5, and L5-L6 neuronal expression respectively, in contrast to modules M0, M2 and M3 which represent L2-L6 neuronal expression (Figure 3F). 7 of the 21 gene modules are significantly upregulated in ALS patients, including M1 (endothelial/microglia), M2 (neurons), M4 (oligodendrocytes), M9 (endothelial/neurons/oligodendrocytes), M11 (astrocytes), M12 (endothelial/neurons), and M19 (endothelial/astrocytes) (Figure 3E-G). Overlaying scores for M11 and M12 onto hematoxylin and eosin (H&E) stained slides confirms the layer specificity of these modules (to L1 and WM, respectively), as well as a shift in expression between control donors and ALS donors (Figure 3H).

Because this patient cohort has been scored for cognitive function, we can correlate spatial co-expression module scores to distinct cognitive dysfunctions as measured by ECAS scores. Notably, we see distinct and different spatial correlations to executive function and to language/verbal fluency function (Figure 3I, Supplemental Figure S8). Loss of executive function correlates with a downregulation of M7 and M8 in BA46, which are modules associated with oxidative metabolism and neurons. This contrasts with loss of language/verbal fluency,

which correlates with an upregulation of M1, M12, and M19 in both BA44/45 and BA46, which are modules that are associated with glial (endothelial, astrocyte, and microglia) cells.

Module scores themselves correlate as expected based on their individual cellular and spatial associations. For example, modules associated with neurons tend to correlate and be highly expressed in L2-L6 (with a single inhibitory neuron module expressed in L1), and modules associated with glial cells tend to correlate and either be highly expressed in white matter or without layer specificity (Figure 3J). As spatial co-expression modules can be associated with multiple distinct cell types, we further separate these modules into submodules with more cell type specificity. For each module, module transcripts are embedded into a k-NN graph using snRNA-seq transcript correlation distance as a metric, and reclustered into submodules (Supplemental Figure S9).

We select spatial module M1 for further examination, as it is upregulated in ALS/ALSci donors and linked to language/verbal fluency dysfunction (Figure 3F, Supplemental Figure S8C). After separating into submodules, we see M1 has spatially overlapping immune response, vascular development, and protein folding ontology enrichment (Figure 3K), and each submodule is associated with endothelial, microglial or neuronal gene expression (Figure 3L). These spatial submodules contain a number of transcripts that are differentially expressed in snRNA-seq between controls, ALS, and ALSci donors (Figure 3M). Comparing submodule scores in the spatial arrays shows increases in expression for all M1 submodules in ALS donors (Figure 3N). The M1 submodules additionally become more spatially correlated across endothelial, microglial, and neuronal cells in ALS donors, despite demonstrating very little cross-cell type correlation in control donors (Figure 3O). This may indicate a disease-associated response across different cell types occurring in the same physical locations in the cortex. We

interpret this as evidence for spatially coordinated glial-neuronal interactions associated with ALS.

Neuronal dysfunction in ALS correlates with clinical phenotype

Due to the predicted changes in glial functionality in L2-L6 of ALS/ALSci donors and the evidence for glial-neuronal functional correlation in our ST data, we examined the neuronal cell population within our snRNA-seq data for disease-associated changes in abundance/function. We first cluster neurons, and then assign each cluster a neuronal subtype based on expression of key marker genes (Supplemental Figure S10), comparing our subtype clusters to published data (*25*). We identify 8 well-characterized EN subtypes, including L2/3, L4, L5, L6, and L6 Car3 intratelencephalic (IT) neurons, L5/L6 near-projecting (NP) neurons, L6 corticothalamic (CT) projection neurons, and L6b neurons, but we are unable to isolate the relatively rare L5 extratelencephalic (ET) projecting neurons (Figure 4A). We also find a ninth cluster specific to a few donors that likely includes multiple neuronal cell subtypes, and is characterized by abnormally high neurogranin (*NRGN*) production. The proportion of neurons in each of these neuronal subtypes does not significantly change between control and ALS donors.

We identify 9 well-characterized IN subtypes, including *VIP*$^+$, basket and chandelier *PVALB*$^+$, three *LAMP5*$^+$ subtypes, and three *SST*$^+$ interneuron subtypes (Figure 4B). We annotate *SST*$^+$ clusters as Martinotti, non-Martinotti, and long-range projection interneurons based on expression of comparable mouse marker transcripts(*26*). We note that there may be neuronal transcriptomic changes that correlate with donor age in most neuronal subtypes (Supplemental Figure S11). Comparing the proportion of IN subtypes, there is no significant change in quantity between control and ALS donors. There may be a modest, non-significant decrease in *SST*$^+$

cells, which is consistent with the spatial deconvolution into interneuron subtypes (Supplemental Figure S2E). However, *SST* immunohistochemistry quantified algorithmically from 25 of our donor cohort indicates no change between control and ALS donors (Supplemental Figure S12).

To build on these findings, we examined in greater detail how distinct patterns of neuronal dysfunction are connected to cognitive heterogeneity in ALS. The ECAS employs specific score thresholds to indicate impairment, but a borderline range above those cutoffs has been proposed to optimize usage of the test (*13*). We reasoned that these extended ECAS guidelines could help identify early features of cognitive decline. Specifically, we used this extended range to define an intermediate group of ALS patients (hereafter referred to as 'borderline') with cognitive scores 5-points above the established abnormality cut-off of the ECAS, as previously reported (Figure 4C) (*13*). Differential gene expression analyses across neuronal subtypes revealed a graded transcriptional response across each group in both BA44/45 and BA46 regions (Supplemental Figure S13), with ALSci patients generally exhibiting the greatest number of differentially expressed genes (DEGs), followed by borderline cases, and pure-motor ALS individuals showing the most muted responses (Figure 4D). The largest number of DEGs was detected in excitatory neurons, with BA46 nuclei generally displaying a more robust transcriptional response than those isolated from BA44/45 (Figure 4D). This observation cannot be attributed to sampling biases, as we profiled a comparable number of excitatory nuclei from both brain regions across experimental conditions (Supplemental Figure S10C,E). Interestingly, the degree of transcriptional dysregulation was also variable between BA46 excitatory subtypes, ranging from less than 10 DEGs in layer 5/layer 6 near-projecting (L5/L6 NP) neurons to nearly 3000 DEGs in layer 6 intratelencephalic projecting neurons (L6 IT). In contrast, inhibitory neurons exhibited mixed and more nuanced transcriptional behaviors. For

instance, in ALSci cases, we detected *n* = 371 DEGs in *VIP*$^+$ BA44/45 interneurons compared to control donors, whereas this subpopulation remained relatively unchanged at the transcriptional level in BA46 (*n* = 3 DEGs). LAMP5 interneurons further exemplify this regionally divergent molecular response: whereas PAX6$^+$/SNCG$^+$ LAMP5 cells derived from BA44/45 exhibited small but measurable transcriptional shifts in ALSci individuals, no changes were observed in the BA46 region; conversely, the subtype annotated as LAMP5$^+$ showed the most pronounced dysregulation profile in BA46-derived nuclei (*n* = 387 DEGs), differently from the same subtype in BA44/45, where fewer than 10 DEGs were detected in ALSci individuals. Taken together, the findings highlight differences in how neurons across regions are affected by disease and implicate distinct neuronal subpopulations in the cognitive heterogeneity that characterizes ALS.

Divergent intra- and cross-region signatures of neuronal dysfunction in ALS

To deepen our understanding of what distinguishes ALS individuals with and without cognitive impairment, we next conducted a detailed exploration of disease-associated transcriptional changes. Using the three-group stratification strategy of ALS patients described above—pure-motor ALS, borderline, and ALSci—we carried out an integrative functional enrichment analysis covering neuronal subtypes, clinical diagnoses, and brain regions. Across all conditions, both excitatory and inhibitory neurons converged on the same top 5 gene ontology (GO) terms: cytoplasmic translation (GO:0002181), the postsynapse (GO:0098794), mitochondrial respiration (GO:0031966 and GO:0098803), the axon (GO:0030424), and the presynapse (GO:0098793) (Figure 4E). Notably, enrichment scores were generally highest in ALSci, or were at least comparable between ALSci and borderline cases, though this varied between neuronal subtypes and also depended on brain region. For example, L6 IT neurons

exhibited strong and comparable responses across translation, mitochondria, and neuronal connectivity functions in BA46-derived nuclei from borderline and ALSci patients; while in BA44/45, only translation-associated genes showed pronounced enrichment in the same neuronal subtype. By contrast, in pure-motor ALS, we found no functional enrichment in BA46 L6 IT neurons, and only subtle enrichment signals related to pre- and postsynaptic gene dysregulation in BA44/45. This limited functional enrichment in pure-motor ALS L6 IT neurons was in fact a common feature observed across all neuronal subtypes, suggesting that neuronal transcriptional changes in non-motor regions of the ALS brain mostly co-occur with declining cognitive performance.

While GO analyses suggest the same pathways are recurrently affected across neuronal subtypes, gene-level differential expression analyses reveal the involvement of distinct gene subsets and variable directions of transcriptional regulation across conditions (Figure 4E). Indeed, while a similar set of processes was enriched across both regions, the directionality of these changes was not always coincident and we additionally uncovered some level of donor-to-donor variability (Figure 4G). These opposing responses were most evident for excitatory neurons in translation-related genes —whereas BA44/45 showed strong upregulation of ribosomal and translation factors in ALSci patients, neurons in BA46 exhibited a similarly robust but opposite shift in expression (i.e., downregulation). As already indicated by our previous analyses, neurons from pure-motor ALS cases displayed a limited transcriptional response. Borderlines cases, on the other hand, showed a sizeable transcriptional signature in a subset of neuronal BA46 subpopulations, including L5 ITs, L6 ITs, and Parvalbumin-expressing (PVALB^{+}) basket cells, consistent with a early role for these subpopulations in deterioration of executive function. In line with this idea, the strongest

transcriptional responses were observed in the lower layers of BA46, specifically in L5 IT and L6 IT neurons of ALSci cases, further indicating that these subpopulations contribute to the decline of cognitive function in ALS. Interestingly, a transcriptional signature in VIP$^+$ interneurons centered exclusively on neuronal connectivity dysfunction and enriched in postsynapse-associated genes was detected only in the BA44 region of ALSci patients, highlighting yet another pattern of regional specificity. Taken together, these analyses show that molecular dysfunction in ALS co-occurs with clinical phenotype and pinpoint vulnerable neuronal subtypes associated with distinct forms of cognitive impairment.

A gene-level view of the data provided an additional layer of insight into these relationships. Intersecting the top five functional categories enriched in neurons with genes linked to TDP43—whose dysfunction is a pathological hallmark of ALS (*27*)—including mRNA targets detected by eCLIP-seq (*28*), direct protein-protein interaction partners (*29*), and transcriptional changes associated with TDP43 nuclear depletion identified in postmortem FTD-ALS brains (*30*), revealed notable subtleties (Figure 4E). While TDP43 protein interactors predominantly converged on ribosomal and other translation-related genes, eCLIP-defined TDP43 target genes differentially expressed in ALSci were primarily associated with the postsynapse and, to a lesser extent, with axonal and presynaptic genes. Similarly, gene expression changes linked to the loss of nuclear TDP43 closely mirrored the pattern identified for TDP43 mRNA targets, aligning with postsynapse-, presynapse-, and axon-associated transcriptional dysregulation in ALSci. BA44/45 VIP$^+$ interneurons in ALSci are particularly noteworthy in this regard, as 40.6% of the differential transcriptional response observed in these cells overlapped with TDP43 mRNA targets. Nonetheless, an overview of the analysis presented in Figure 4E indicates that a substantial portion of the molecular dysregulation

detected in the ALSci brain may be unrelated to loss TDP43 function. Indeed, only 129 out of 929 genes associated with the top five GO terms highlighted in this analysis overlap with TDP43 targets. In line with these observations, TDP43 pathology in our cohort was generally not predictive of neuronal transcriptional dysregulation, except for VIP^{+} interneurons (Figure 4F), overall recapitulating previous findings that TDP43 aggregation correlates poorly with cognitive symptom manifestation (*31*).

Neuron-glia correlates of cognitive impairment

Our analyses so far reveal both shared and region-specific patterns of neuronal and glial dysregulation, which, when considered together, may provide a better understanding of the cellular mechanisms underlying cognitive impairment in ALS. Within-donor correlation analyses revealed a consistent pattern of gliosis across BA44/45 and BA46 primarily driven by astrocytes and microglia (Figure 4H and Supplemental Figure S14A). Notably, this gliosis signature appeared to be regionally coordinated—when present in BA44/45, a corresponding response was consistently observed in BA46 . This suggests that certain glial activation patterns may be pan-regional within the ALS frontal cortex, potentially reflecting a common gliotic response to ALS pathology irrespective of cognitive phenotype. While astrocytic and microglial responses were generally consistent between BA44/45 and BA46, myelinating oligodendrocytes (*PLP1*^{High}) and maturing OPCs displayed region-specific behaviors, as revealed by the weak correlation in their proportions across these regions. Interestingly, the abundance of each subpopulation closely tracked with neuronal transcriptional dysregulation, an effect that was particularly evident in relation to translation- and mitochondria-associated genes (Figure 4I and Supplemental Figure S14B-D). Overall, this oligodendrocyte/maturing OPC/neuron correlation emerged as

one the strongest within-sample associations, and was largely independent of neuronal subtype. While the basis for this regional heterogeneity remains unclear, these results uncover a complex landscape of neuron-glia interactions that correlate with cognitive impairment in ALS.

Discussion

ALS is traditionally framed as a disease of motor neurons (*32*, *33*). Together with other recent reports (*8*), our study provides a broader understanding of this neurological disorder. By combining careful patient stratification and transcriptome-wide profiling of two prefrontal cortex regions, we identify spatially defined signatures associated with cognitive alterations and explore cellular and molecular features correlated with language and executive impairment. These changes occur in the absence of overt neuronal degeneration, suggesting that cognitive symptoms in ALS precede neuronal loss. Instead, the findings we report support a model of the disease in which abnormal neuron–glia interactions, rather than the simple degeneration of vulnerable neurons, are mechanistically involved in the decline of cognitive function. These interactions exhibit phenotype-specific characteristics, indicative of a heterogeneous disease etiology driven by distinct mechanisms across affected regions rather than a uniform pathological process.

TDP43 proteinopathy is a pathognomonic feature of the broader ALS clinical spectrum (*27*). This RNA-binding protein is ubiquitously expressed in the brain and typically localizes to the nucleus. However, in disease states, TDP43 is redistributed to the cytoplasm, where it forms insoluble aggregates. The resulting loss of nuclear TDP43 is linked to perturbations in gene expression, which are hypothesized to contribute to ALS pathogenesis (*27*). Notably, our analysis revealed that the relationship between TDP43 neuropathology and clinical outcomes is

more nuanced than initially expected. Indeed, we found an overall weak association between the transcriptional programs characterized in the ALS brain and canonical TDP43–associated signatures described in the literature (*28–30*), raising questions about the interplay between regional TDP43 pathology and clinically significant dysfunction. Our studies highlight the need to more closely examine how TDP-43 dysfunction interacts with additional disease processes, such as the coordinated responses we observed between neurons and glia.

In summary, here we have uncovered measurable early signatures associated with cognitive decline in ALS. Our study supports a view of the disease wherein distinct molecular and cellular mechanisms contribute to its cognitive heterogeneity. Whether some of these changes are themselves active drivers of disease or a secondary response to the disease process remains an open question.

METHODS

Tissue sourcing and donor information

Postmortem tissue specimens were obtained from the Medical Research Council (MRC) Edinburgh Brain Bank (Research Ethics Committee approval: 21/ES/0087), following review by the Bank's internal ethics committee, the Academic and Clinical Central Office for Research and Development Medical Research Ethics Committee, and the Biomedical Research Alliance of New York. Tissue specimens originated from patients recruited through the Scottish Motor Neurone Disease Register (a prospective, population-based ALS study in Scotland) who had consented to the collection and use of their health-related information for research purposes. Clinical metadata associated with each donor, including ECAS test scores, were extracted from the Clinical Audit Research Evaluation (CARE)-MND database. Tissue from neurotypical control cases – defined in general terms as individuals without neurological or psychiatric conditions and no substantial neuropathology postmortem – was provided by the Edinburgh Sudden Death Brain Bank(*20*). All applicable federal, state, and local regulations were followed when transferring materials to the New York Genome Center.

The diagnosis of patients fulfilled the revised El Escorial criteria for clinically definite or probable ALS. Comorbid neurological disorders, including frontotemporal dementia, or psychiatric history were exclusion criteria for the selection of study participants. We note also that dementia was not documented in the clinical files of any recruited patients. Clinical metadata is provided in Supplemental Table 1.

Cognitive profiling

Cognitive performance was assessed in life using the Edinburgh Cognitive and Behavioural ALS Screen (ECAS) following standard guidelines (https://ecas.psy.ed.ac.uk). Briefly, the ECAS

incorporates 16 subtests assessing a range of cognitive domains, including executive functions (reverse digit span, alternation, inhibitory sentence completion, and social cognition), verbal fluency (free and fixed), language (naming, comprehension, and spelling), memory (immediate recall, delayed percentage retention, and delayed recognition), and visuospatial functions (dot counting, cube counting, and number location), complemented by a separate semi-structured interview with a relative or carer designed to detect behavioral changes(*13*, *34*). Performance across all tasks generates the ECAS Total Score, which adds up to an aggregated maximum of 136 points. The ECAS Total Score is composed of two fractions, the ALS-Specific and ALS Non-Specific Scores. Subtest evaluations of executive functions, language, and verbal fluency are combined to generate the ALS-Specific Score, which can reach a maximum of 100 points in cognitively healthy individuals. Memory and visuospatial functions tasks contribute to the ALS Non-Specific Score. Abnormal scores are defined as ≤77 for the ALS-Specific Score, ≤24 for the ALS Non-Specific Score (out of a maximum of 36 points), and ≤105 for the ECAS Total Score. Further subdomain-specific thresholds include ≤33 out of 48 for executive function, ≤14 out of24 for fluency, ≤26 out of 28 for language, ≤13 out of 24 for memory, and ≤10 out of 12 for visuospatial functions. As discussed by Niven et al. (*13*), ECAS Total Scores and ALS-Specific Scores in the range of 105–110 and 77–82, respectively, were used to define ALS cases with ‘borderline’ cognitive impairment.

Neuropathology Assessments

Brain regions of interest, including BA44/45 and BA46, were fixed in 10% neutral-buffered formalin for a minimum of 24 hours to ensure tissue integrity and morphology. Post-fixation, tissue processing involved a dehydration step through an ascending ethanol series

(70%–100%), followed by clearing with three rounds of xylene treatment. Specimens were then embedded in molten paraffin wax over three consecutive 5-hour stages and, after solidifying at room temperature, sectioned onto Superfrost Plus slides on a Leica microtome at 4 μm thickness with overnight drying at 40°C. Aggregated TDP-43 detection employed Proteintech's antibody against phosphorylated TDP-43 (Ser409/410) at a 1:1,000 dilution, combined with 3,3′-diamino-benzidine (DAB) visualization. TDP-43 pathology was staged using a semiquantitative four-tiered scoring system, where a score of 0 indicates the absence of pathology, and scores 1 through 3 represent progressively increasing levels of TDP-43 pathology burden (mild, moderate, and severe, respectively). Scores were defined as follows: 1 (mild) for up to 5 affected cells, 2 (moderate) for 6–15 affected cells, and 3 (severe) for more than 15 affected cells, all within at least a single 20× field of view per tissue section. This scoring scheme was applied independently to both neurons and glia using haematoxylin counterstaining for morphology-based cell classification (*4*).

Nuclei isolation from frozen post-mortem brain tissue and library preparation

Nuclei isolation protocol was adapted from previously described methods (PMID: 31042697; PMID: 36385529). All procedures were carried out on ice unless otherwise stated. Frozen post-mortem brain tissue was retrieved from storage at -80 °C and equilibrated on dry ice for 20–30 minutes before being transferred to a cryostat at -18 °C. Tissue samples were left for 15–20 minutes inside the cryostat before sectioning. We collected 20-50 mg of tissue per preparation, taking care to avoid excessive sampling of white matter. Tissue cuts were placed in a pre-chilled Dounce homogenizer containing 700 μL of ice-cold homogenization buffer (320 mM sucrose, 5 mM $CaCl_2$ [#21115; Millipore Sigma], 3 mM magnesium acetate [#63052; Millipore

Sigma], 10 mM Tris HCl [pH 7.8; prepared from Thermo Fisher's #AM9850G and #AM9855G], 0.1 mM EDTA [pH 8.0; Thermo Fisher #AM9260G], 0.1% IGEPAL CA-630 [#I8896; Millipore Sigma], 1 mM DTT [#646563, Millipore Sigma], and 0.4 U/µl Protector RNase inhibitor [#03335402001, Millipore Sigma]). Homogenization was performed with 10 strokes of the loose pestle followed by 10 strokes of the tight pestle. After a 5-minute incubation on ice, the homogenized tissue was passed through a 70-µm Flowmi Cell Strainer (Bel-Art; H13680-0070) and combined with an equal volume of a solution made of 50% OptiPrep (#07820, STEMCELL Technologies), 5 mM $CaCl_2$, 3 mM magnesium acetate, 10 mM Tris HCl (pH 7.8), 0.1 mM EDTA (pH 8.0), and 1 mM DTT. Using a 15-mL conical tube, this mixture was then layered onto a density gradient made up of 750 µL of a 30% OptiPrep solution supplemented with 134 mM sucrose, 5 mM $CaCl_2$, 3 mM magnesium acetate, 10 mM Tris HCl (pH 7.8), 0.1 mM EDTA (pH 8.0), 1 mM DTT, 0.04% IGEPAL CA-630, and 0.17 U μl^{-1} RNase inhibitor) over 300 µL of 40% OptiPrep solution supplemented with 96 mM sucrose, 5 mM $CaCl_2$, 3 mM magnesium acetate, 10 mM Tris HCl (pH 7.8), 0.1 mM EDTA (pH 8.0), 1 mM DTT, 0.03% IGEPAL CA-630, and 0.12 U µl−1 RNase inhibitor. Nuclei were separated by centrifugation at 3,000 × *g* for 20 minutes at 4 °C using a pre-chilled swinging-bucket rotor with the brake off. Approximately 150 µL of nuclei were collected from the 30%/40% gradient interphase and subsequently washed in PBS containing 1% BSA and 1 U μl^{-1} RNase inhibitor. Following a 5-minute incubation on ice, the nuclei were centrifuged at 450 × *g* for 3 minutes at 4 °C, and the pellet resuspended in 15 µL of diluted 10X Genomics Nuclei Buffer. Nuclei counting was performed using a Countess II FL instrument (Thermo Fisher) by staining with SYTOX Green (ThermoFisher; S7020). Additional quality control measures included Trypan Blue staining (#EBT-001, NanoEntek) to evaluate adequate permeabilization (cutoff: >95% of 'dead cells') and the assessment of nuclear morphology before proceeding with transposition.

Matched snRNA-seq and snATAC-seq libraries were prepared using the Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Kit (#PN-1000283) according to the protocol provided by 10X Genomics. Nuclei (16,100 per reaction; targeted recovery: 10,000 nuclei) from a single donor were used for each library preparation. Partitioning of nuclei was performed on a Chromium X instrument (10X Genomics). Quality control of cDNAs and libraries was conducted using Agilent High Sensitivity DNA Kits and a BioAnalyzer instrument. Final library quantification prior to pooling utilized KAPA Biosystems' Library Quantification Kit for Illumina Platforms (KK4824). Libraries were sequenced on a NovaSeq X system using a 2 × 100 bp configuration run on a 10B flow cell. BA44/45 and BA46 libraries (n = 48) were prepared and sequenced in independent batches.

Prefrontal Cortex Atlas and Classifier

Published 10x snRNA-seq sequencing data was acquired from NCBI GEO or Synapse repositories: GSE157827 (*35*), GSE167494 (*36*), GSE174332 (*8*), GSE174367 (*37*), GSE207334 (*38*), GSE219280 (*11*), PRJNA434002 (*39*), syn45351388 (*40*), and syn51123517 (*22*),

Sequencing data was aligned to a standard reference genome (hg38; GENCODE v32/Ensembl98) and counted (cellranger v7.1.0). Individual cells were filtered so that the number of transcripts (UMIs) was between 500 and 50000, and the maximum fraction of mitochondrially-encoded transcripts was 0.2. Prefrontal cortex atlas was restricted to donor samples from patients between 50 and 90 years of age at death taken from the prefrontal cortex. The resulting dataset consists of 2,603,756 cells from 407 donors.

Each cell was scaled to a total of 10,000 counts, the count data was transformed by log(x + 1), and each gene was scaled by dividing by the interquartile range (RobustScaler). Genes which would have a variance greater than 1 after scaling were further scaled so that their variance was equal to 1. Cells were embedded into a k-nearest neighbors (k-NN) graph from 50 principal components and 25 neighbors, then clustered into 114 separate clusters using the leiden algorithm (resolution set to 2.5). Labels for cell type and cell subtype were hand annotated for each cluster based on known marker genes (Supplemental Table 4).

This standardized, transformed, scaled data was used as input to a neural network (NN) autoencoder/classifier that predicts cell type, subtype, and denoised transcript expression data. The NN encoder module takes 36601 transcript counts as input, and is fully connected to a 1000 node hidden layer, a 250 node hidden layer, then a 100 node hidden layer, each of which has a hyperbolic tangent (tanh) activation function. The 100 node hidden layer is connected to a 50 node (h_{width}) hidden layer with a tanh activation function that represents a shared latent space embedding. This hidden layer is shared between three decoder modules. The transcript expression predictor connects the 50 node shared latent space embedding to a 50 and then 250 node hidden layers with a tanh activation function, then a 36601 node output layer with a rectified linear (ReLU) activation function. The cell type classifier connects the 50 node shared latent space embedding to a 50 node hidden layer with a tanh activation function, and connects that hidden layer to a 9 node output softmax layer that predicts cell type label. The cell subtype classifier connects the 50 node shared latent space embedding to a 50 node hidden layer with a tanh activation function, and connects that hidden layer to a 14 node output softmax layer that predicts cell subtype label.

The NN model was constructed in pyTorch and trained with the Adam optimizer. Model hyperparameters h_{width}, weight decay, and learning rate were tuned by training on randomly sampled 465,000 cell training sets for 200 epochs and scoring on 155,000 cell validation sets. Transcript expression was trained to minimize mean squared error, and cell type and subtype classifiers were trained to minimize cross-entropy. Full training was performed using 1.95m cells for 1500 epochs, and scored on 650k cells held out. Held out test cells were not used during cross-validation. Models were evaluated based on coefficient of determination (R^2) for transcript expression, and on F1 score for cell type and subtype classifiers.

10X Genomics Multiome Data Processing

Sequenced 10x gene expression and ATAC libraries were aligned to a standard reference genome (hg38; GENCODE v32/Ensembl98) and counted (Cellranger-ARC v2.0.2; Supplemental Table 2). ATAC peaks were aligned between samples for quantification (Cellranger-ARC aggr v2.0.2). Individual cells were filtered so that the number of gene expression transcripts (UMIs) was between 500 and 50000, the maximum fraction of mitochondrially-encoded transcripts was 0.2, and the number of quantified ATAC fragments in peaks was between 250 and 100,000. A total of 165,884 cells passed this initial filtering. Cell types and subtypes were determined by the prefrontal cortex cell type classifier, and validated by comparing to published marker genes (Supplemental Table 4). Doublets and low-quality cells were removed, leaving 144,260 usable cells from 48 libraries for analysis.

To maintain the difference in count depth between cell types, individual cells were standardized for count depth by dividing all transcript counts for the cell by a size factor, to reach the median count depth for the assigned cell type. Standardized counts were log(x+1)

transformed, and each gene *g* was then scaled by dividing by the interquartile range (IQR). Genes which would have a variance (Var) greater than 1 after scaling were further scaled so that their variance was equal to 1, giving log standardized counts.

$$\mathrm{Scaled}(\mathbf{X_g}) = \begin{cases} \mathbf{X_g}/\sqrt{\mathrm{Var}(\mathbf{X_g})} & \mathrm{Var}(\mathbf{X_g}) \geq \mathrm{IQR}(\mathbf{X_g})^2 \\ \mathbf{X_g}/\mathrm{IQR}(\mathbf{X_g}) & \mathrm{Var}(\mathbf{X_g}) < \mathrm{IQR}(\mathbf{X_g})^2 \\ \mathbf{X_g} & \mathrm{IQR}(\mathbf{X_g}) = 0 \end{cases}$$

Principal Component Analysis (PCA) was performed using a truncated singular value decomposition with 25 components on the log standardized counts. Cells were embedded into a 25-neighbor k-Nearest Neighbors (k-NN) graph using these 25 components. Cells were projected into 2 dimensions with Uniform Manifold Approximation and Projection (UMAP) using the k-NN embedding (Minimum distance of 0.5; initialized from spectral graph embedding).

The 144,260 usable cells were further filtered for cells with at least 2500 ATAC peak counts, yielding 119,153 cells. Individual cells were standardized for ATAC count depth by dividing all peak counts for the cell by a size factor, to reach the median ATAC count depth for the assigned cell type. Standardized ATAC counts were then log(x+1) transformed to give log standardized ATAC counts. PCA, k-NN embedding, and UMAP projection were performed identically to gene expression data from log standardized ATAC counts.

Glial Cell Functional Annotation

snRNA-seq log standardized counts for each glial cell type were separated for PCA (25 components), k-NN graph embedding (25 neighbors), and UMAP as for the full dataset. Cells from each cell type were separately clustered into a large number of groups with the leiden

algorithm (resolution of 2), and then each cluster was assigned to one of a few functional groups based on relative expression of key functional marker genes (Supplemental Table 5).

Neuronal Cell Subtype Annotation

snRNA-seq log standardized counts for excitatory (EN) and inhibitory neurons (IN) were further filtered by transcript count to neurons with a minimum of 2500 UMI counts, yielding 41,933 ENs and 15,553 INs. ENs and INs were separately processed for PCA (25 components), k-NN graph embedding (25 neighbors), and UMAP as for the full dataset. ENs and INs were then clustered into a few transcriptionally distinct clusters with the leiden algorithm (resolution of 0.15 for EN, 0.2 for IN), and assigned to a neuronal subtype based on marker expression (Supplemental Table 5).

Cell Proportions Per Sample

For each single nucleus sample (n=48), the number of cells in each functional group for a glial cell type are divided by the total number of cells of that type (e.g. 60 homeostatic astrocytes and 40 reactive astrocytes have proportions of 0.6 and 0.4). The correlation between these ratios is calculated as the within-sample correlation of glial cell function proportions.

Single-cell Differential Expression Analysis

Transcript differential expression analysis was performed on raw integer counts with pyDeseq2 (*41*, *42*), standardizing using only non-zero counts ('poscounts') and excluding mitochondrially

encoded genes and *MALAT1* from size factor calculations. Statistical testing was against a null hypothesis of log2(Fold Change) ≤ ±0.263, rejecting the null for adjusted p-values ≤ 0.05.

Gene Ontology Enrichment Analysis

Gene ontology enrichment was performed with Metascape using the default complete proteome as the background of each analysis (*43*), and g:Profiler(*44*), which tests gene lists for enrichment against background a multiple hypothesis test corrected hypergeometric test. Here, gene lists are differentially expressed genes against a background of genes which are not all-zero in counts, or genes in a module against a background of genes which were included in the Splotch spatial model.

Tissue processing, RNA Capture, and Library Preparation for Spatial Transcriptomics

Frozen tissue blocks, kept at -80°C prior to processing, were equilibrated to -16°C inside a Leica CM3050S cryostat, embedded in Tissue Plus OCT Compound (Fisher Healthcare, #4585), and sectioned in a perpendicular orientation with regard to the pial surface. Tissue sections (10 µm-thick) from quality controlled blocks (hematoxylin and eosin [H&E] staining and imaging) were collected onto prechilled Visium Spatial Gene Expression Slides (10X Genomics, #1000185) while briefly warming the back of the slide with a gloved finger to improve adherence and minimize artifacts. We sought to capture cortical layers 1 through 6 and white matter within each spatially barcoded array. For each subject, four technical replicates, sampled from tissue cuts approximately 20 µm apart, were collected across two Visium slides. A deliberate effort was made to section control and ALS subjects onto the same Visium slide when possible, with sample processing proceeding in parallel from that point onward. Where necessary due to

tissue sectioning artifacts or poor yield, additional replicates were collected and processed on a later date.

Generation of spatially resolved transcriptomes was performed in accordance with 10X Genomics guidelines (CG000239, Revisions F and G). Briefly, tissue sections were fixed in chilled methanol, stained using an abbreviated H&E protocol, and imaged on a Zeiss Axio Observer Z1 fitted with an EC Plan-Neofluar 10x/0.3 M27 objective and a Zeiss Axiocam 506 Mono camera. Tissue sections were permeabilized for 12 minutes following optimization (10X Genomics, CG000238, Revision E). The number of cDNA amplification cycles for each Visium array was calculated using qPCR, as instructed by 10X Genomics.

Visium libraries were prepared using unique Illumina-compatible PCR primers with Dual Index Kit TT Set A (10X Genomics, #1000215). Size distribution profiles of the final libraries were assessed by running Agilent High Sensitivity DNA chips (#5067-4626) on a 2100 Bioanalyzer system (Agilent, #G2939BA). Libraries outside the expected range (450 ± 50 bp) or that contained adapter dimer contaminants were discarded and reprocessed. Total yields were initially measured on a Qubit 3.0 fluorometer before final qPCR quantification (KAPA Library Quantification Kit for Illumina Platforms, KAPA Biosystems, #KK4824). Per batch, libraries from 48-52 Visium arrays were pooled at a concentration of 10 nM prior to sequencing. We processed 246 arrays from 46 unique donors.

10x Genomics Visium Data Processing and Annotation

Each sequenced 10x Visium library was aligned to a standard reference genome (hg38; GENCODE v32/Ensembl98) and counted with spaceranger (v3.0.0). We initially filtered any spot with fewer than 100 UMI transcript counts. Visium spots were manually assigned to one of the

six cortical layers or the white matter using 10X Genomics Loupe Browser (version 8.0). Layer annotations were based on cortical architecture and cellular morphology. The white matter was identified by the absence of neuronal cell bodies and the linear arrangement of glial nuclei along neuronal processes. Cortical layer 6 was defined as a heterogeneous neuronal layer spanning from the white matter to the beginning of cortical layer 5, characterized by its large pyramidal neurons oriented perpendicularly to the pial surface. Cortical layer 4 was identified by the presence of small, granular neurons adjacent to the large pyramidal neurons of cortical layer 5. Cortical layer 3 was labeled as the pyramidal cell layer located between granular cortical layers 2 and 4. Cortical layer 1 was distinguished by its paucity of cells and lack of large neuronal somas. Spots containing meninges contiguous to cortical layer 1 were excluded from downstream analysis. Additionally, spots overlying tissue artifacts, such as freeze fractures, folds, or tears, were excluded. After annotation, we have 953,073 individual spots across 246 arrays.

To maintain the difference in count depth between cell types, individual spots were standardized for count depth by dividing all transcript counts for the cell by a size factor, to reach the median count depth for the annotated cortex layer. Size factors were clipped to a minimum of 0.25 to mitigate the risk of overinflating low-count spots. Standardized spatial counts were then scaled by dividing by the interquartile range (RobustScaler). Genes which would have a variance greater than 1 after scaling were further scaled so that their variance was equal to 1, giving sandardized spatial counts.

PCA was performed using a truncated singular value decomposition with 25 components on the standardized counts. Cells were embedded into a 25-neighbor k-NN graph using these

25 components. Cells were projected into 2 dimensions with UMAP using the k-NN embedding (Minimum distance of 0.5; initialized from spectral graph embedding).

Splotch modeling of spatial gene expression across clinical groups

Raw sequencing data was filtered to remove mitochondrially encoded genes, lncRNAs, pseudogenes, and genes expressed in less than 1% of spots (this filtering was applied separately to BA44/45 and BA46 datasets). Spots with fewer than 100 UMIs across remaining genes were discarded prior to Splotch analysis. Finally, spots without a valid annotation were discarded, and any spots without at least one immediate neighbor in the Visium grid were discarded to prevent singularities in the spatial autocorrelation component of the Splotch model. This yielded expression data for 13,188 genes measured across 401,976 spots (286,581 ALS, 115,386 control) with an average depth of 3,679 UMIs in BA46, and 13,116 genes measured across 551,106 spots (347,548 ALS, 203,558 control) with an average depth of 3,447 UMIs in BA44.

Splotch analysis was conducted as described previously(*12*, *23*, *24*). Briefly, Splotch employs a zero-inflated Poisson likelihood function (where the zero-inflation accounts for technical dropouts) to model the expression (λ) of each gene in each spot. This expression level is in turn modeled using a generalized linear model (GLM) with the following three components: the characteristic expression of the spot's donor ID and ALS phenotype (β), autocorrelation with the spot's spatial neighbors (ψ), and spot-specific variation (ε). Furthermore, the characteristic expression rate (β) is hierarchically formulated to account for the experimental design, enabling the investigation of the effect of sample covariates on differential gene expression. By performing posterior inference on the model given the observed spatial expression data, we

identified clinical group level trends in spatial gene expression (β) that best explain our observations.

Compared to other computational methods for analysis of spatial transcriptomics data, Splotch i) enables quantification of expression differences between conditions and anatomical regions, ii) is designed to rigorously account for the uncertainty of low counts, and iii) analyzes multiple tissue sections simultaneously, stratifying patients based on clinical phenotype in order to quantify biological and technical variation.

Parameter inference

Splotch has been implemented in the probabilistic programming language Stan, and is available at https://github.com/adaly/cSplotch. For all analyses, Bayesian inference was performed over the parameters using Stan's adaptive Hamiltonian Monte-Carlo (HMC) sampler with default parameters. Four independent chains were sampled, each with 175 warm-up iterations and 175 sampling iterations (700 total), and convergence was monitored using the R-hat statistic.

Splotch differential expression analysis

To quantify difference in expression between two conditions using Splotch, the Bayes factor between posterior distributions over characteristic expression coefficients β estimated by the model was calculated. Without loss of generality, difference in expression was quantified between conditions represented by $\beta^{(1)}$ and $\beta^{(2)}$, which may differ across any combination of genes, donor covariates, or cortical layers. A random variable $\Delta_\beta = \beta^{(1)} - \beta^{(2)}$ was defined, which captures the difference between $\beta^{(1)}$ and $\beta^{(2)}$. If Δ_β is tightly centered around zero,

then the two distributions are very similar to each other, and the null hypothesis of identical expression cannot be rejected. To quantify this similarity, the posterior distribution $\Delta_\beta|\mathcal{D}$ (where $\mathcal{D}$ represents the data used to train the model) was compared to the prior distribution Δ_β using the Savage-Dickey density ratio(*45*) that estimates the Bayes factor between the conditions:

$$\mathrm{BF} \approx \frac{p(\Delta_\beta = 0)}{p(\Delta_\beta = 0|\mathcal{D})},$$

where the probability density functions were evaluated at zero. If expression is different between the two conditions, then the posterior $\Delta_\beta|\mathcal{D}$ will have very little mass at 0, and the estimated Bayes factor will be large (by convention, $\mathrm{BF} > 5$ indicates substantial support). Conversely, for similar expression regimes, the posterior will place a mass equal to or greater than that of the prior at zero, and the Bayes factor will be ≤ 1. $p(\Delta_\beta|\mathcal{D})$ is calculated using the posterior samples obtained in the previous section.

Assigning Genes to Spatial Modules

Spatial modules are inferred from the mean posterior estimates (λ) of the depth-normalized expression of each gene in each spatial spot (a matrix of dimension 13,188 genes x 401,976 spots in BA46; 13,116 genes x 551,106 spots in BA44/45). Mean posterior gene expression estimates λ for each gene were clipped to the 99th percentile to remove the effect of outlier spots. Gene expression values were then scaled by dividing by the gene interquartile range, and genes which would have a variance greater than 1 after scaling were further scaled so that their variance was equal to 1. Pearson correlation coefficients were then calculated across all spots for all pairs of genes, resulting in a 13,188 x 13,188 correlation matrix in BA46 and 13,116 x 13,116 correlation matrix in BA44/45.

Correlation distance between 13,253 genes was calculated by taking the mean Pearson correlation distance (1 - pearson correlation coefficient) for a gene-gene pair from both brain regions. This correlation distance matrix was used as the distance matrix for a gene-gene k-NN graph embedding (k=10). Genes were projected into 2 dimensions with the UMAP algorithm, and clustered into 23 clusters using the leiden algorithm (resolution of 2). Each cluster is considered a spatially-correlated gene module. Clusters 21 and 22 contain all chrM and chrY-encoded transcripts, respectively. Python functions for module assignment are available in the scself (single-cell self-supervised) package.

Module Scores

For each gene, standardized transcript count (snRNA-seq) or modeled expression lambda (spatial) was standardized to a 0-1 range for each gene, where 0 is fixed to the 1st percentile, 1 is fixed to the 99th percentile, and values are linearly scaled between 0 and 1. Values below the 1st percentile are clipped to 0, and values above the 99th percentile are clipped to 1. For each module, a module score is calculated for each cell (snRNA-seq) or array spot (spatial) by taking the mean of the 0-1 scaled gene scores for every gene in the module.

$$\mathrm{MinMaxScaled}(x) = \begin{cases} \frac{x - \mathbf{X}_{q=0.01}}{\mathbf{X}_{q=0.99} - \mathbf{X}_{q=0.01}} & \mathbf{X}_{q=0.01} \leq x \leq \mathbf{X}_{q=0.01} \\ 0 & x < \mathbf{X}_{q=0.01} \\ 1 & x > \mathbf{X}_{q=0.99} \end{cases}$$

Spatial spot-level module scores are aggregated by taking the median module score for all spots in a layer (L1-L6 and WM) for each array, and rescaling the median layer-array module scores to a 0-1 range. These scaled layer-array module scores are used for statistical testing, and are further aggregated into mean module scores per ALS phenotype for plotting.

Assigning Genes to Cell Type Submodules

A gene-gene correlation matrix for all genes in a spatial module was constructed from standardized snRNA-seq transcript expression. This was subtracted from 1 to get a correlation distance matrix, which was used as the distance matrix for a gene-gene k-NN graph embedding (k=10). Genes were projected into 2 dimensions with the UMAP algorithm, and clustered into 2 to 5 clusters using the leiden algorithm (resolution of 0.75). Each cluster was interpreted as a gene submodule, and this process was repeated for all 23 spatial modules. Modules 21 and 22 (chrM and chrY) did not yield more than one submodule. Submodule scores were calculated for individual cells and spatial spots as above. Spatial modules and submodules are listed in Supplemental Table 6.

Spatial Co-occurrence Analysis

Pearson correlation was calculated for spatial module and spatial submodule scores across all spatial spots. Modules and submodules are assigned to cortex layers by aggregating into layer-array module scores as above, and then taking the average module score across all layers. The module is assigned to the layer with the highest module score, and any other layer within 95% of that value. If all scores are within 10%, the module is assigned to the pan-layer label ‘ALL’.

Spatial Cell Type Deconvolution

Per-spot estimates of cell type abundances were generated using Cell2location(*46*). The prefrontal cortex atlas assembled in *Prefrontal Cortex Atlas and Classifier* was used to train a negative binomial (NB) regression model of gene expression for each of the 5 major glial cell types, for the 4 inhibitory neuron subtypes, and for the 3 excitatory neuron layer subtypes (Supplementary Table 3), which was used for all subsequent deconvolution experiments. Deconvolution with Cell2location was performed separately for each donor, treating individual ST arrays as batches and assuming an average of five cells per Visium spot with all other hyperparameters set to default values. Mean estimates of the abundance of each cell type in each spot were extracted and used for differential abundance testing.

SST^+ Cell Counting

Raw hematoxylin and eosin (H&E) images were first analyzed to differentiate between white matter and gray matter regions. Each field of view (FOV) was classified based on its most dominant colors, excluding the lightest and darkest 20% of pixels to avoid artifacts from holes in tissue, as well as the cells and SST stainings themselves. The 3 dominant colors were determined using k-means clustering with 3 clusters, and the cluster centers were used as representative colors. A FOV was classified as gray matter if the blue channel intensity in RGB space exceeded the red channel by at least 25 intensity units for any of the 3 dominant colors.

For FOVs identified as gray matter, color deconvolution was performed to separate the cellular component from the SST signal. Cell segmentation was conducted using a Cellpose machine learning model that was trained from scratch with human-in-the-loop training on several manually-annotated FOVs (*47*). The resultant cell masks were dilated by 50% of the cell's radius to include surrounding cytoplasmic regions. The SST image was converted to

grayscale, and pixel values below 90% of the median FOV pixel intensity were removed to eliminate background noise. The remaining SST signal was thresholded based on size, retaining only regions larger than 20 pixels in area. Finally, colocalization of the thresholded SST signal and dilated cell masks was performed, and any masks containing SST signal were classified as SST-positive.

General procedures

Unless otherwise noted, data analysis was performed in Python, using the scanpy (*48*) ecosystem, plots were generated with matplotlib (*49*), and figures were assembled using Adobe Illustrator.

ACKNOWLEDGEMENTS

With thanks to the CARE MND Register, hosted by the Euan Macdonald Centre for MND Research. This work was supported by awards from the NIH: R01NS118183, R01NS116350, R01NS127186, and R01NS118570. This work was also supported by Target ALS grant FD-2023-GEN-S1. The text of this manuscript was generated and proofread in part using a large language model (OpenAI; GPT-4o). This work utilized computational resources and assistance provided by the New York Genome Center. The computations in this work were, in part, run at facilities supported by the Scientific Computing Core at the Flatiron Institute, a division of the Simons Foundation.

AUTHOR CONTRIBUTIONS

Conceptualization: HP, JP, CS, CGR, JG; Methodology: JP, CGR, CAJ, AD, KK, JE; Software: AD, CAJ, JE; Validation: AD, CAJ , CGR, JE, JP, KK; Formal analysis: AD, CAJ, CGR, JE; Investigation: CGR, JP, KK, OC, ML, LR; Resources: AD, CAJ, CGR, CM, JE, JG, JP, KK, KM; Data curation: AD, CAJ, CGR, JP; Visualization: AD, CAJ, CGR, JP, JE; Writing – original draft preparation: CAJ, CGR, HP; Writing – review and editing: N/A; Supervision: CS, HP; Funding acquisition: HP; Project administration: HP, CS.

DATA AVAILABILITY

Single-nucleus multiomic data has been deposited in NCBI GEO (GSE290359) . Spatial transcriptomic data is available on the Target ALS portal. Whole genome sequencing data is available on the Target ALS portal.

COMPETING INTEREST

The authors declare no competing interests

REFERENCES

1. S. Abrahams, Neuropsychological impairment in amyotrophic lateral sclerosis-frontotemporal spectrum disorder. *Nat. Rev. Neurol.* **19**, 655–667 (2023).

2. M. J. Strong, S. Abrahams, L. H. Goldstein, S. Woolley, P. Mclaughlin, J. Snowden, E. Mioshi, A. Roberts-South, M. Benatar, T. HortobáGyi, J. Rosenfeld, V. Silani, P. G. Ince, M. R. Turner, Amyotrophic lateral sclerosis - frontotemporal spectrum disorder (ALS-FTSD): Revised diagnostic criteria. *Amyotroph. Lateral Scler. Frontotemporal Degener.* **18**, 153–174 (2017).

3. S. Abrahams, L. H. Goldstein, J. J. Kew, D. J. Brooks, C. M. Lloyd, C. D. Frith, P. N. Leigh, Frontal lobe dysfunction in amyotrophic lateral sclerosis. A PET study. *Brain* **119 (Pt 6)**, 2105–2120 (1996).

4. J. M. Gregory, K. McDade, T. H. Bak, S. Pal, S. Chandran, C. Smith, S. Abrahams, Executive, language and fluency dysfunction are markers of localised TDP-43 cerebral pathology in non-demented ALS. *J. Neurol. Neurosurg. Psychiatry* **91**, 149–157 (2020).

5. J. J. Kew, L. H. Goldstein, P. N. Leigh, S. Abrahams, N. Cosgrave, R. E. Passingham, R. S. Frackowiak, D. J. Brooks, The relationship between abnormalities of cognitive function and cerebral activation in amyotrophic lateral sclerosis. A neuropsychological and positron emission tomography study. *Brain* **116 (Pt 6)**, 1399–1423 (1993).

6. S. Abrahams, L. H. Goldstein, A. Simmons, M. J. Brammer, S. C. R. Williams, V. P. Giampietro, C. M. Andrew, P. N. Leigh, Functional magnetic resonance imaging of verbal fluency and confrontation naming using compressed image acquisition to permit overt responses. *Hum. Brain Mapp.* **20**, 29–40 (2003).

7. S. Abrahams, L. H. Goldstein, A. Simmons, M. Brammer, S. C. R. Williams, V. Giampietro, P. N. Leigh, Word retrieval in amyotrophic lateral sclerosis: a functional magnetic resonance imaging study. *Brain* **127**, 1507–1517 (2004).

8. S. S. Pineda, H. Lee, M. J. Ulloa-Navas, R. M. Linville, F. J. Garcia, K. Galani, E. Engelberg-Cook, M. C. Castanedes, B. E. Fitzwalter, L. J. Pregent, M. E. Gardashli, M. DeTure, D. V. Vera-Garcia, A. T. S. Hucke, B. E. Oskarsson, M. E. Murray, D. W. Dickson, M. Heiman, V. V. Belzil, M. Kellis, Single-cell dissection of the human motor and prefrontal cortices in ALS and FTLD. *Cell* **187**, 1971–1989.e16 (2024).

9. A. Yadav, K. J. E. Matson, L. Li, I. Hua, J. Petrescu, K. Kang, M. R. Alkaslasi, D. I. Lee, S. Hasan, A. Galuta, A. Dedek, S. Ameri, J. Parnell, M. M. Alshardan, F. A. Qumqumji, S. M. Alhamad, A. P. Wang, G. Poulen, N. Lonjon, F. Vachiery-Lahaye, P. Gaur, M. A. Nalls, Y. A. Qi, D. Maric, M. E. Ward, M. E. Hildebrand, P.-F. Mery, E. Bourinet, L. Bauchet, E. C. Tsai, H. Phatnani, C. E. Le Pichon, V. Menon, A. J. Levine, A cellular taxonomy of the adult human spinal cord. *Neuron* **111**, 328–344.e7 (2023).

10. F. Limone, D. A. Mordes, A. Couto, B. J. Joseph, J. M. Mitchell, M. Therrien, S. D. Ghosh, D. Meyer, Y. Zhang, M. Goldman, L. Bortolin, I. Cobos, B. Stevens, S. A. McCarroll, I. Kadiu, A. Burberry, O. Pietiläinen, K. Eggan, Single-nucleus sequencing reveals enriched expression of genetic risk factors in extratelencephalic neurons sensitive to degeneration in ALS. *Nat. Aging* **4**, 984–997 (2024).

11. J. Li, M. K. Jaiswal, J.-F. Chien, A. Kozlenkov, J. Jung, P. Zhou, M. Gardashli, L. J. Pregent, E. Engelberg-Cook, D. W. Dickson, V. V. Belzil, E. A. Mukamel, S. Dracheva, Divergent single cell transcriptome and epigenome alterations in ALS and FTD patients with C9orf72 mutation. *Nat. Commun.* **14**, 5714 (2023).

12. S. Maniatis, T. Äijö, S. Vickovic, C. Braine, K. Kang, A. Mollbrink, D. Fagegaltier, Ž. Andrusivová, S. Saarenpää, G. Saiz-Castro, M. Cuevas, A. Watters, J. Lundeberg, R. Bonneau, H. Phatnani, Spatiotemporal dynamics of molecular pathology in amyotrophic lateral sclerosis. *Science* **364**, 89–93 (2019).

13. E. Niven, J. Newton, J. Foley, S. Colville, R. Swingler, S. Chandran, T. H. Bak, S. Abrahams, Validation of the Edinburgh Cognitive and Behavioural Amyotrophic Lateral Sclerosis Screen (ECAS): A cognitive tool for motor disorders. *Amyotroph. Lateral Scler. Frontotemporal Degener.* **16**, 172–179 (2015).

14. A. F. T. Arnsten, Stress weakens prefrontal networks: molecular insults to higher cognition. *Nat. Neurosci.* **18**, 1376–1385 (2015).

15. S. J. Gilbert, P. W. Burgess, Executive function. *Curr. Biol.* **18**, R110–4 (2008).

16. J. M. Gregory, E. Elliott, K. McDade, T. Bak, S. Pal, S. Chandran, S. Abrahams, C. Smith, Neuronal clusterin expression is associated with cognitive protection in amyotrophic lateral sclerosis. *Neuropathol. Appl. Neurobiol.* **46**, 255–263 (2020).

17. S. S. Keller, T. Crow, A. Foundas, K. Amunts, N. Roberts, Broca's area: nomenclature, anatomy, typology and asymmetry. *Brain Lang.* **109**, 29–48 (2009).

18. E. Fedorenko, A. A. Ivanova, T. I. Regev, The language network as a natural kind within the broader landscape of the human brain. *Nat. Rev. Neurosci.* **25**, 289–312 (2024).

19. D. Leighton, J. Newton, S. Colville, A. Bethell, G. Craig, L. Cunningham, M. Flett, D. Fraser, J. Hatrick, H. Lennox, L. Marshall, D. McAleer, A. McEleney, K. Millar, A. Silver, L. Stephenson, S. Stewart, D. Storey, G. Stott, C. Thornton, C. Webber, H. Gordon, G. Melchiorre, L. Sherlock, E. Beswick, D. Buchanan, S. Abrahams, A. Bateman, J. Preston, C. Duncan, R. Davenport, G. Gorrie, I. Morrison, R. Swingler, S. Chandran, S. Pal, Clinical audit research and evaluation of motor neuron disease (CARE-MND): a national electronic platform for prospective, longitudinal monitoring of MND in Scotland. *Amyotroph. Lateral Scler. Frontotemporal Degener.* **20**, 242–250 (2019).

20. C. Smith, T. Millar, Brain donation procedures in the Sudden Death Brain Bank in Edinburgh. *Handbook of clinical neurology* **150** (2018).

21. L. Taylor, R. G. Brown, S. Tsermentseli, S. Tsermentseli, A. Al-Chalabi, C. Shaw, C. Ellis, P. Leigh, L. Goldstein, Is language impairment more common than executive dysfunction in amyotrophic lateral sclerosis? *Journal of Neurology, Neurosurgery, and Psychiatry* **84**, 494–498 (2012).

22. H. Mathys, Z. Peng, C. A. Boix, M. B. Victor, N. Leary, S. Babu, G. Abdelhady, X. Jiang, A. P. Ng, K. Ghafari, A. K. Kunisky, J. Mantero, K. Galani, V. N. Lohia, G. E. Fortier, Y. Lotfi, J. Ivey, H. P. Brown, P. R. Patel, N. Chakraborty, J. I. Beaudway, E. J. Imhoff, C. F. Keeler, M. M. McChesney, H. H. Patel, S. P. Patel, M. T. Thai, D. A. Bennett, M. Kellis, L.-H. Tsai, Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to

Alzheimer's disease pathology. *Cell* **186**, 4365–4385.e27 (2023).

23. T. Äijö, S. Maniatis, S. Vickovic, K. Kang, M. Cuevas, C. Braine, H. Phatnani, J. Lundeberg, R. Bonneau, Splotch: Robust estimation of aligned spatial temporal gene expression data, *bioRxiv* (2019)p. 757096.

24. A. C. Daly, F. Cambuli, T. Äijö, B. Lötstedt, N. Marjanovic, O. Kuksenko, M. Smith-Erb, S. Fernandez, D. Domovic, N. Van Wittenberghe, E. Drokhlyansky, G. K. Griffin, H. Phatnani, R. Bonneau, A. Regev, S. Vickovic, Tissue and cellular spatiotemporal dynamics in colon aging, *bioRxiv* (2024)p. 2024.04.22.590125.

25. M. I. Gabitto, K. J. Travaglini, V. M. Rachleff, E. S. Kaplan, B. Long, J. Ariza, Y. Ding, J. T. Mahoney, N. Dee, J. Goldy, E. J. Melief, A. Agrawal, O. Kana, X. Zhen, S. T. Barlow, K. Brouner, J. Campos, J. Campos, A. J. Carr, T. Casper, R. Chakrabarty, M. Clark, J. Cool, R. Dalley, M. Darvas, S.-L. Ding, T. Dolbeare, T. Egdorf, L. Esposito, R. Ferrer, L. E. Fleckenstein, R. Gala, A. Gary, E. Gelfand, J. Gloe, N. Guilford, J. Guzman, D. Hirschstein, W. Ho, M. Hupp, T. Jarsky, N. Johansen, B. E. Kalmbach, L. M. Keene, S. Khawand, M. D. Kilgore, A. Kirkland, M. Kunst, B. R. Lee, M. Leytze, C. L. Mac Donald, J. Malone, Z. Maltzer, N. Martin, R. McCue, D. McMillen, G. Mena, E. Meyerdierks, K. P. Meyers, T. Mollenkopf, M. Montine, A. L. Nolan, J. K. Nyhus, P. A. Olsen, M. Pacleb, C. M. Pagan, N. Peña, T. Pham, C. A. Pom, N. Postupna, C. Rimorin, A. Ruiz, G. A. Saldi, A. M. Schantz, N. V. Shapovalova, S. A. Sorensen, B. Staats, M. Sullivan, S. M. Sunkin, C. Thompson, M. Tieu, J. T. Ting, A. Torkelson, T. Tran, N. J. Valera Cuevas, S. Walling-Bell, M.-Q. Wang, J. Waters, A. M. Wilson, M. Xiao, D. Haynor, N. M. Gatto, S. Jayadev, S. Mufti, L. Ng, S. Mukherjee, P. K. Crane, C. S. Latimer, B. P. Levi, K. A. Smith, J. L. Close, J. A. Miller, R. D. Hodge, E. B. Larson, T. J. Grabowski, M. Hawrylycz, C. D. Keene, E. S. Lein, Integrated multimodal cell atlas of Alzheimer's disease. *Nat. Neurosci.* **27**, 2366–2383 (2024).

26. J. Fisher, M. Verhagen, Z. Long, M. Moissidis, Y. Yan, C. He, J. Wang, E. Micoli, C. M. Alastruey, R. Moors, O. Marín, D. Mi, L. Lim, Cortical somatostatin long-range projection neurons and interneurons exhibit divergent developmental trajectories. *Neuron* **112**, 558–573.e8 (2024).

27. T. R. Suk, M. W. C. Rousseaux, The role of TDP-43 mislocalization in amyotrophic lateral sclerosis. *Mol. Neurodegener.* **15**, 45 (2020).

28. O. H. Tam, N. V. Rozhkov, R. Shaw, D. Kim, I. Hubbard, S. Fennessey, N. Propp, The NYGC ALS Consortium, H. Phatnani, J. Kwan, D. Sareen, J. R. Broach, Z. Simmons, X. Arcila-Londono, E. B. Lee, V. M. Van Deerlin, N. A. Shneider, E. Fraenkel, L. W. Ostrow, F. Baas, N. Zaitlen, J. D. Berry, A. Malaspina, P. Fratta, G. A. Cox, L. M. Thompson, S. Finkbeiner, E. Dardiotis, T. M. Miller, S. Chandran, S. Pal, E. Hornstein, D. J. MacGowan, T. Heiman-Patterson, M. G. Hammell, N. A. Patsopoulos, O. Butovsky, J. Dubnau, A. Nath, R. Bowser, M. Harms, E. Aronica, M. Poss, J. Phillips-Cremins, J. Crary, N. Atassi, D. J. Lange, D. J. Adams, L. Stefanis, M. Gotkine, R. Baloh, S. Babu, T. Raj, S. Paganoni, O. Shalem, C. Smith, B. Zhang, B. T. Harris, D. Fagegaltier, B. T. Harris, L. W. Ostrow, H. Phatnani, J. Ravits, J. Dubnau, M. G. Hammell, Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia. *CellReports* **29**, 1164–1177.e5 (2019).

29. B. R. Helmold, K. E. Pauss, P. H. Ozdinler, TDP-43 protein interactome informs about perturbed canonical pathways and may help develop personalized medicine approaches for patients with TDP-43 pathology. *Drug Discov. Today* **28**, 103769 (2023).

30. E. Y. Liu, J. Russ, C. P. Cali, J. M. Phan, A. Amlie-Wolf, E. B. Lee, Loss of nuclear TDP-43 is associated with decondensation of LINE retrotransposons. *Cell Rep.* **27**, 1409–1421.e6 (2019).

31. H. Spence, F. M. Waldron, R. S. Saleeb, A.-L. Brown, O. M. Rifai, M. Gilodi, F. Read, K. Roberts, G. Milne, D. Wilkinson, J. O'Shaughnessy, A. Pastore, P. Fratta, N. Shneider, G. G. Tartaglia, E. Zacco, M. H. Horrocks, J. M. Gregory, RNA aptamer reveals nuclear TDP-43 pathology is an early aggregation event that coincides with STMN-2 cryptic splicing and precedes clinical manifestation in ALS. *Acta Neuropathol.* **147**, 50 (2024).

32. Kurosawa K., Amyotrophic lateral sclerosis. *Ryoikibetsu Shokogun Shirizu* **3**, 178–180 (2001).

33. R. J. Mead, N. Shan, H. J. Reiser, F. Marshall, P. J. Shaw, Amyotrophic lateral sclerosis: a neurodegenerative disorder poised for successful therapeutic translation. *Nat. Rev. Drug Discov.* **22**, 185–212 (2023).

34. S. Abrahams, J. Newton, E. Niven, J. Foley, T. H. Bak, Screening for cognition and behaviour changes in ALS. *Amyotroph. Lateral Scler. Frontotemporal Degener.* **15**, 9–14 (2014).

35. S.-F. Lau, H. Cao, A. K. Y. Fu, N. Y. Ip, Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and neuroprotective glia in Alzheimer's disease. *Proceedings of the National Academy of Sciences* **117**, 25800–25809 (2020).

36. J. S. Sadick, M. R. O'Dea, P. Hasel, T. Dykstra, A. Faustin, S. A. Liddelow, Astrocytes and oligodendrocytes undergo subtype-specific transcriptional changes in Alzheimer's disease. *Neuron* **110**, 1788–1805.e10 (2022).

37. S. Morabito, E. Miyoshi, N. Michael, S. Shahin, A. C. Martini, E. Head, J. Silva, K. Leavy, M. Perez-Rosendahl, V. Swarup, Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nat. Genet.* **53**, 1143–1155 (2021).

38. S. Ma, M. Skarica, Q. Li, C. Xu, R. D. Risgaard, A. T. N. Tebbenkamp, X. Mato-Blanco, R. Kovner, Ž. Krsnik, X. de Martin, V. Luria, X. Martí-Pérez, D. Liang, A. Karger, D. K. Schmidt, Z. Gomez-Sanchez, C. Qi, K. T. Gobeske, S. Pochareddy, A. Debnath, C. J. Hottman, J. Spurrier, L. Teo, A. G. Boghdadi, J. Homman-Ludiye, J. J. Ely, E. W. Daadi, D. Mi, M. Daadi, O. Marín, P. R. Hof, M.-R. Rasin, J. Bourne, C. C. Sherwood, G. Santpere, M. J. Girgenti, S. M. Strittmatter, A. M. M. Sousa, N. Sestan, Molecular and cellular evolution of the primate dorsolateral prefrontal cortex. *Science* **377**, eabo7257 (2022).

39. D. Velmeshev, L. Schirmer, D. Jung, M. Haeussler, Y. Perez, S. Mayer, A. Bhaduri, N. Goyal, D. H. Rowitch, A. R. Kriegstein, Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* **364**, 685–689 (2019).

40. L. M. Gittings, E. B. Alsop, J. Antone, M. Singer, T. G. Whitsett, R. Sattler, K. Van Keuren-Jensen, Cryptic exon detection and transcriptomic changes revealed in single-nuclei RNA sequencing of C9ORF72 patients spanning the ALS-FTD spectrum. *Acta Neuropathol.* **146**, 433–450 (2023).

41. B. Muzellec, M. Teleńczuk, V. Cabeli, M. Andreux, PyDESeq2: a python package for bulk RNA-seq differential expression analysis. *Bioinformatics* **39**, btad547 (2023).

42. S. Anders, W. Huber, Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

43. Y. Zhou, B. Zhou, L. Pache, M. Chang, A. H. Khodabakhshi, O. Tanaseichuk, C. Benner, S. K. Chanda, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).

44. L. Kolberg, U. Raudvere, I. Kuzmin, P. Adler, J. Vilo, H. Peterson, g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res.* **51**, W207–W212 (2023).

45. E.-J. Wagenmakers, T. Lodewyckx, H. Kuriyal, R. Grasman, Bayesian hypothesis testing for psychologists: a tutorial on the Savage-Dickey method. *Cogn. Psychol.* **60**, 158–189 (2010).

46. V. Kleshchevnikov, A. Shmatko, E. Dann, A. Aivazidis, H. W. King, T. Li, R. Elmentaite, A. Lomakin, V. Kedlian, A. Gayoso, M. S. Jain, J. S. Park, L. Ramona, E. Tuck, A. Arutyunyan, R. Vento-Tormo, M. Gerstung, L. James, O. Stegle, O. A. Bayraktar, Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.* **40**, 661–671 (2022).

47. M. Pachitariu, C. Stringer, Cellpose 2.0: how to train your own model. *Nat. Methods* **19**, 1634–1641 (2022).

48. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).

49. J. D. Hunter, Matplotlib: A 2D Graphics Environment. *Computing in Science Engineering* **9**, 90–95 (2007).

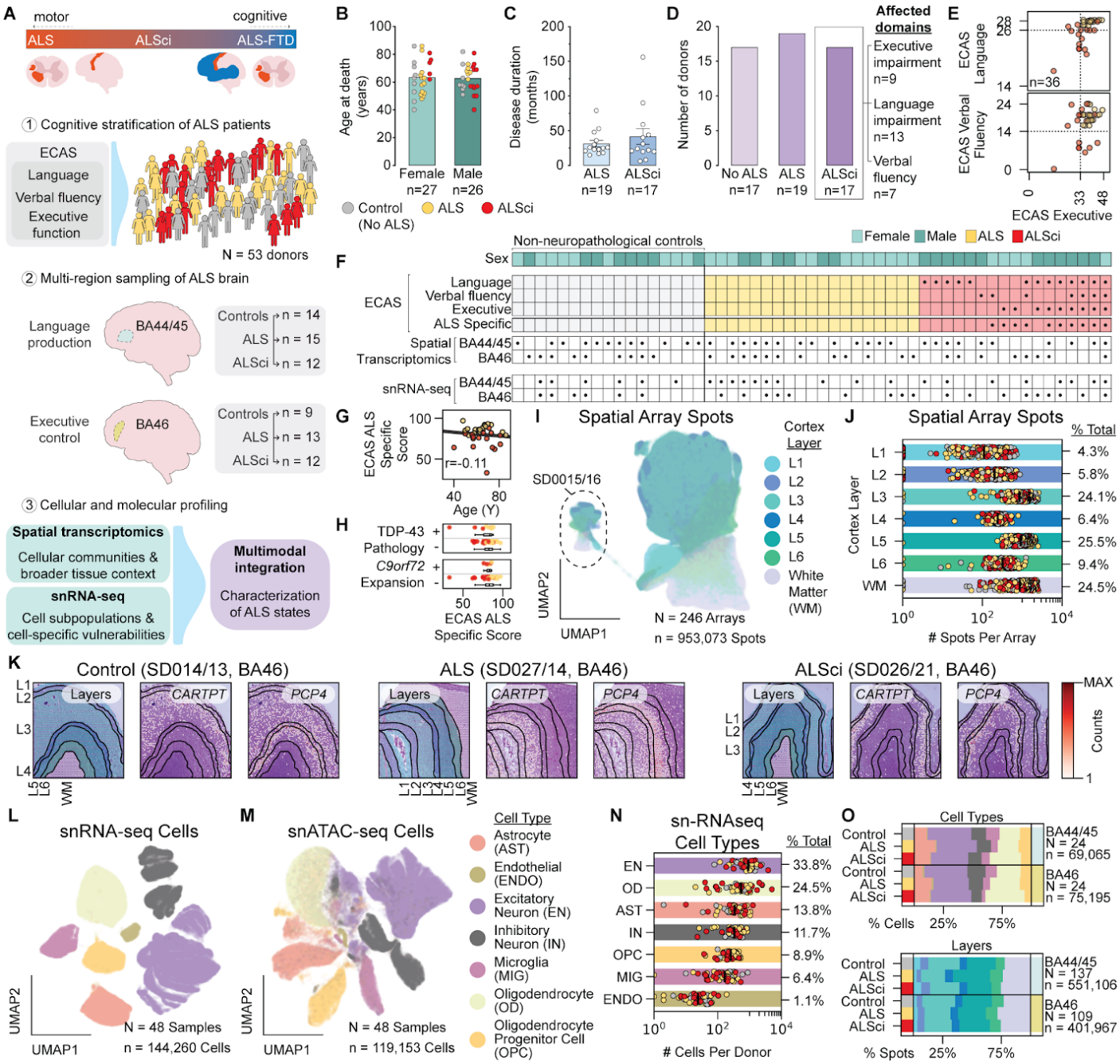
A
motor
cognitive
ALS
ALSci
ALS-FTD
① Cognitive stratification of ALS patients
ECAS
Language
Verbal fluency
Executive function
N = 53 donors
② Multi-region sampling of ALS brain
Language production
BA44/45
Controls n = 14
ALS n = 15
ALSci n = 12
Executive control
BA46
Controls n = 9
ALS n = 13
ALSci n = 12
③ Cellular and molecular profiling
Spatial transcriptomics
Cellular communities & broader tissue context
snRNA-seq
Cell subpopulations & cell-specific vulnerabilities
Multimodal integration
Characterization of ALS states
B
Age at death (years)
Female n=27
Male n=26
Control (No ALS)
ALS
ALSci
C
Disease duration (months)
ALS n=19
ALSci n=17
D
Number of donors
No ALS n=17
ALS n=19
ALSci n=17
Affected domains
Executive impairment n=9
Language impairment n=13
Verbal fluency n=7
E
ECAS Language
ECAS Verbal Fluency
n=36
ECAS Executive
F
Non-neuropathological controls
Female
Male
ALS
ALSci
Sex
ECAS
Language
Verbal fluency
Executive
ALS Specific
Spatial Transcriptomics
BA44/45
BA46
snRNA-seq
G
ECAS ALS Specific Score
r=-0.11
Age (Y)
H
TDP-43 Pathology
C9orf72 Expansion
ECAS ALS Specific Score
I
Spatial Array Spots
SD0015/16
UMAP2
UMAP1
N = 246 Arrays
n = 953,073 Spots
Cortex Layer
L1
L2
L3
L4
L5
L6
White Matter (WM)
J
Spatial Array Spots
% Total
Cortex Layer
L1 4.3%
L2 5.8%
L3 24.1%
L4 6.4%
L5 25.5%
L6 9.4%
WM 24.5%
Spots Per Array
K
Control (SD014/13, BA46)
ALS (SD027/14, BA46)
ALSci (SD026/21, BA46)
Layers
CARTPT
PCP4
MAX
Counts
1
L
snRNA-seq Cells
UMAP2
UMAP1
N = 48 Samples
n = 144,260 Cells
M
snATAC-seq Cells
N = 48 Samples
n = 119,153 Cells
Cell Type
Astrocyte (AST)
Endothelial (ENDO)
Excitatory Neuron (EN)
Inhibitory Neuron (IN)
Microglia (MIG)
Oligodendrocyte (OD)
Oligodendrocyte Progenitor Cell (OPC)
N
sn-RNAseq Cell Types
% Total
EN 33.8%
OD 24.5%
AST 13.8%
IN 11.7%
OPC 8.9%
MIG 6.4%
ENDO 1.1%
Cells Per Donor
O
Cell Types
Control
ALS
ALSci
BA44/45 N = 24 n = 69,065
BA46 N = 24 n = 75,195
% Cells
25%
75%
Layers
BA44/45 N = 137 n = 551,106
BA46 N = 109 n = 401,967
% Spots

Figure 1: A cognitively stratified ALS cohort: study design, donor clinical data, and initial QC
A) An age- and sex-matched cohort of ALS donors (n=36) that have been screened for cognitive impairment and non-ALS controls (sudden death; n=17). Donors with abnormal scores on the ECAS test are classified as cognitively impaired (ALSci). Left hemisphere dorsolateral prefrontal cortex (BA46) and inferior frontal gyrus (BA44/45; Broca's area) are profiled with spatial transcriptomics and snRNA-seq. **B**) Age at death separated by sex and cognitive status. **C**) The interval between diagnosis and date of death (disease duration) for ALS donors. **D**) Cohort distribution among control, ALS, and ALSci groups, alongside the distinct impairment profiles comprising the ALSci group. **E**) ECAS test scores (Executive function, Verbal fluency, and Language) for ALS donors in the cohort. **F**) Cohort summary, including clinical and cognitive variables, and which donor regions were profiled by spatial transcriptomics or snRNA-seq. **G**) Donor age plotted against ECAS ALS-specific aggregate cognitive test score (sum of the language, verbal fluency, and executive scores). **H**) The presence of TDP-43 aggregate pathology and the presence of C9orf72 genomic expansions in donors plotted against ECAS ALS-specific aggregate cognitive test score. **I**) Uniform manifold approximation and projection (UMAP) of spatial transcriptomics data. Individual data points are visium array spots, colored by cortical layer annotation: layer 1 (L1) through layer 6 (L6) and white matter (WM). Seven arrays from a single donor (SD0015/16) are separate from the rest of the data and are highlighted with a dashed box. **J**) The number of spots annotated as a cortical layer in a single Visium array (n = 246 arrays), colored by donor phenotype. The proportion of each layer's spots to all spots is in percent, and the median spot count for each layer is annotated with a black bar. **K**) Annotated cortical layers, *CARTPT*, and *PCP4* transcript expression overlaid on a Visium spatial array with a H&E-stained tissue section. Examples are shown for Control, ALS, and ALSci donors. Max expression is standardized to each array for each transcript. **L**) UMAP plot projecting gene expression from 48 10x multiomic snRNA-seq libraries (24 BA44/45, 24 BA46; 144,260 cells), colored by broad cell type. **M**) UMAP plot projecting ATAC peak counts from the same nuclei from snATAC-seq libraries (24 BA44/45, 24 BA46; 119,153 cells after filtering for >2500 ATAC counts) **N**) The number of nuclei in each snRNA-seq library for the 7 broad cell types, after QC and removing doublets, colored by donor phenotype. The proportion of each cell type to the total number of cells is in percent. **O**) Proportion of broad cell types in snRNA-seq libraries for Control, ALS, and ALSci donors, and proportion of annotated cortex layers in spatial arrays for Control, ALS, and ALSci donors.

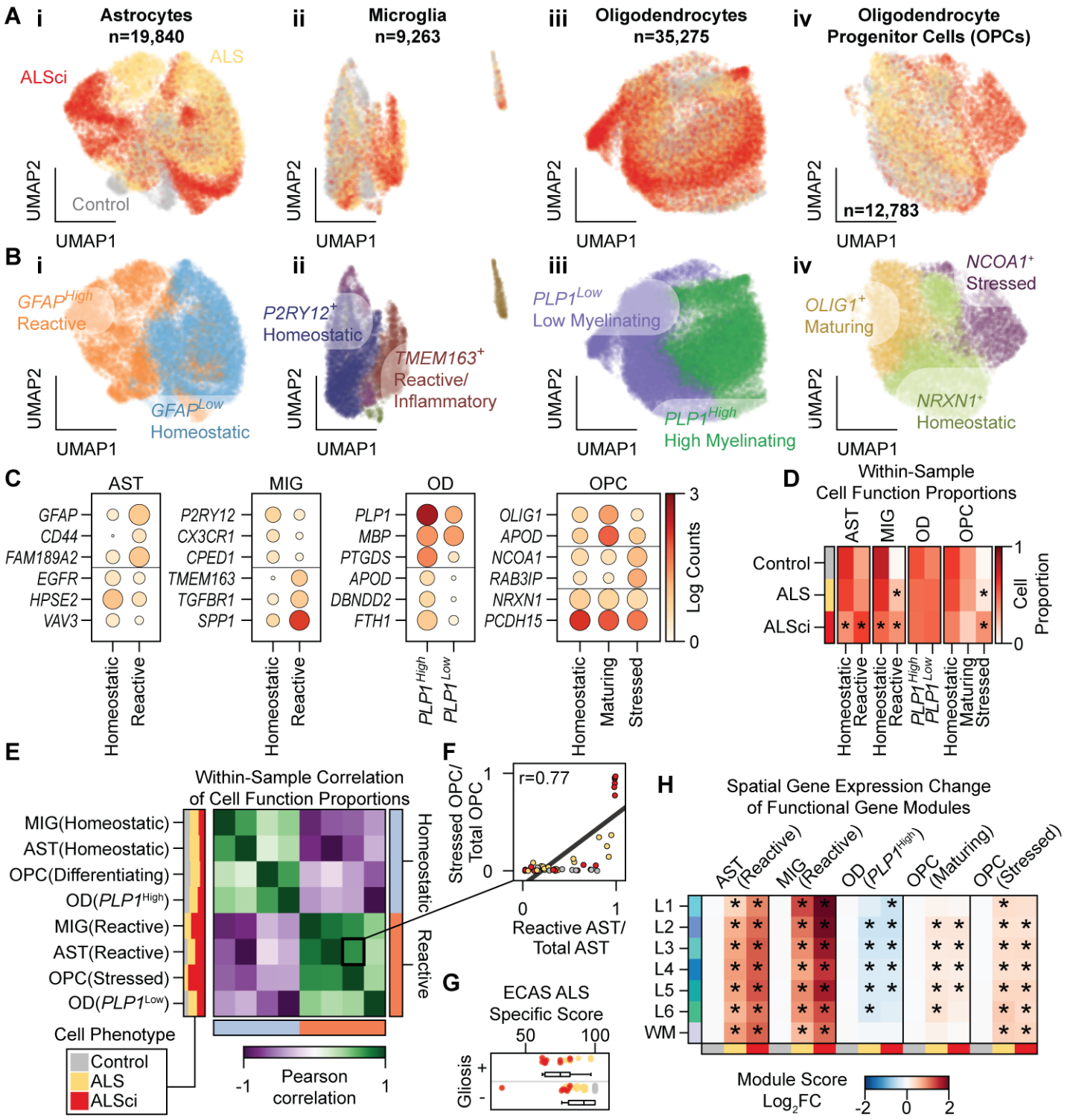
A
i
Astrocytes
n=19,840
ALSci
ALS
Control
UMAP2
UMAP1
ii
Microglia
n=9,263
iii
Oligodendrocytes
n=35,275
iv
Oligodendrocyte Progenitor Cells (OPCs)
n=12,783
B
GFAP High Reactive
GFAP Low Homeostatic
P2RY12+ Homeostatic
TMEM163+ Reactive/ Inflammatory
PLP1 Low Low Myelinating
PLP1 High High Myelinating
OLIG1+ Maturing
NCOA1+ Stressed
NRXN1+ Homeostatic
C
AST
MIG
OD
OPC
GFAP
CD44
FAM189A2
EGFR
HPSE2
VAV3
P2RY12
CX3CR1
CPED1
TMEM163
TGFBR1
SPP1
PLP1
MBP
PTGDS
APOD
DBNDD2
FTH1
OLIG1
APOD
NCOA1
RAB3IP
NRXN1
PCDH15
Homeostatic
Reactive
PLP1 High
PLP1 Low
Maturing
Stressed
Log Counts
D
Within-Sample Cell Function Proportions
Control
ALS
ALSci
Cell Proportion
E
Within-Sample Correlation of Cell Function Proportions
MIG(Homeostatic)
AST(Homeostatic)
OPC(Differentiating)
OD(PLP1 High)
MIG(Reactive)
AST(Reactive)
OPC(Stressed)
OD(PLP1 Low)
Cell Phenotype
Control
ALS
ALSci
Pearson correlation
F
r=0.77
Stressed OPC/ Total OPC
Reactive AST/ Total AST
G
ECAS ALS Specific Score
Gliosis
H
Spatial Gene Expression Change of Functional Gene Modules
AST (Reactive)
MIG (Reactive)
OD (PLP1 High)
OPC (Maturing)
OPC (Stressed)
L1
L2
L3
L4
L5
L6
WM
Module Score Log2FC

Figure 2: ALS donors with and without cognitive impairment have transcriptomic signatures of gliosis **A**) Broad glial cell populations are separated and plotted by UMAP, colored by the disease phenotype of the donor **B**) Glial cell populations are plotted based on functional characterization. (i) Astrocytes are categorized as homeostatic (*GFAP*^{Low}) or reactive (*GFAP*^{High}). (ii) Microglia are categorized as homeostatic (*NRXN1*$^{+}$), reacting/inflammatory (*TMEM163*$^{+}$), phagocytic, stress-response, and non-microglial immune cells. (iii) Oligodendrocytes are categorized as high myelinating *PLP1*^{High} or low myelinating *PLP1*^{Low}. (iv) Oligodendrocyte progenitor cells are categorized as homeostatic (*NRXN1*$^{+}$), maturing (*OLIG1*$^{+}$), or stressed (*NCOA1*$^{+}$). **C**) Highly-expressed transcriptomic markers of each glial cell functional category, plotted as a dot plot colored by log(standardized count) and sized based on the proportion of nonzero counts. **D**) Glial cell functional category proportions for each donor plotted as the average for non-ALS, ALS, and ALSci. Asterisks denote differences from non-ALS Controls determined by Welch's t-test ($p < 0.05$). **E**) Pearson correlation between glial cell functional category proportions for each donor. The proportion of cells from non-ALS, ALS, and ALSci donors are plotted on the left for each category. **F**) Proportion of stressed to total OPCs plotted against proportion of reactive to total astrocytes for each donor, annotated with least squares regression line and Pearson correlation coefficient (r) **G**) ALS specific ECAS score plotted for donors with and without transcriptional signatures of gliosis (>80% reactive astrocytes or >20% reactive microglia). Control donors are plotted with ALS specific scores of 100, although those donors do not have ECAS scores. **H**) Spatial distribution of aggregate module scores across annotated Visium arrays, plotted as $\log_2$ fold change compared to controls. Asterisks denote significance by Welch's t-test against non-ALS Controls ($p < 0.05$), corrected for false discovery rate (0.05).

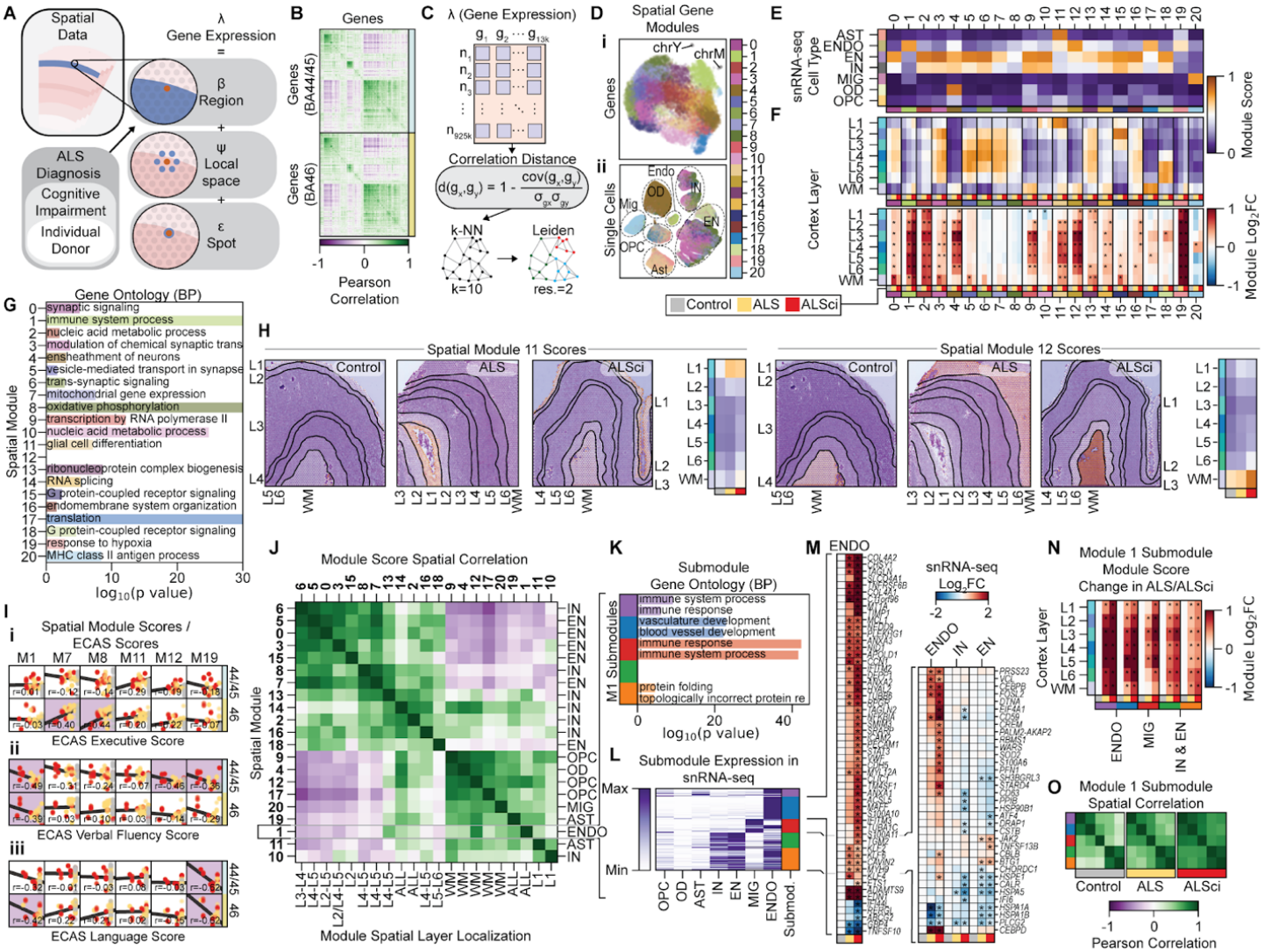
A
Spatial Data
λ
Gene Expression
=
β
Region
+
ψ
Local space
+
ε
Spot
ALS Diagnosis
Cognitive Impairment
Individual Donor
B
Genes
Genes (BA44/45)
Genes (BA46)
Pearson Correlation
C
λ (Gene Expression)
Correlation Distance
k-NN
Leiden
k=10
res.=2
D
Spatial Gene Modules
Genes
Single Cells
chrY
chrM
Endo
OD
IN
Mig
EN
OPC
Ast
E
snRNA-seq Cell Type
AST
ENDO
EN
IN
MIG
OD
OPC
Module Score
F
Cortex Layer
L1
L2
L3
L4
L5
L6
WM
Module Log2FC
Control
ALS
ALSci
G
Gene Ontology (BP)
Spatial Module
synaptic signaling
immune system process
nucleic acid metabolic process
modulation of chemical synaptic trans
ensheathment of neurons
vesicle-mediated transport in synapse
trans-synaptic signaling
mitochondrial gene expression
oxidative phosphorylation
transcription by RNA polymerase II
nucleic acid metabolic process
glial cell differentiation
ribonucleoprotein complex biogenesis
RNA splicing
G protein-coupled receptor signaling
endomembrane system organization
translation
G protein-coupled receptor signaling
response to hypoxia
MHC class II antigen process
log10(p value)
H
Spatial Module 11 Scores
Spatial Module 12 Scores
Control
ALS
ALSci
I
Spatial Module Scores / ECAS Scores
M1
M7
M8
M11
M12
M19
ECAS Executive Score
ECAS Verbal Fluency Score
ECAS Language Score
J
Module Score Spatial Correlation
Spatial Module
Module Spatial Layer Localization
K
Submodule Gene Ontology (BP)
M1 Submodules
immune system process
immune response
vasculature development
blood vessel development
immune response
immune system process
protein folding
topologically incorrect protein re
log10(p value)
L
Submodule Expression in snRNA-seq
Max
Min
OPC
OD
AST
IN
EN
MIG
ENDO
Submod.
M
ENDO
snRNA-seq Log2FC
ENDO
IN
EN
N
Module 1 Submodule Module Score Change in ALS/ALSci
Cortex Layer
Module Log2FC
ENDO
MIG
IN & EN
O
Module 1 Submodule Spatial Correlation
Control
ALS
ALSci
Pearson Correlation

Figure 3: ALS donors have spatially correlated transcriptional changes with variable impact on ECAS-scored clinical cognitive phenotype

A) Schematic of the splotch model fit to each gene captured in the spatial transcriptomic data set. Transcript expression at each spatial spot on each array is modeled as a combination of ALS disease state, cognitive impairment state, and donor ID (β), a spatial autoregressive component (φ), and a spot noise (ε) component. **B**) Pearson correlation between genes in modeled gene expression in BA44/45 (top) and BA46 (bottom). **C**) Schematic showing the gene-gene k-NN Pearson correlation distance graph embedding and spatial module graph clustering. **D**) (**i**) UMAP plot of genes, colored by spatial module. Modules 21 and 22 correspond to all chrM and chrY genes respectively and are excluded from additional analysis (**ii**) UMAP plot of snRNA-seq cells, colored by the spatial module for which they have the highest relative expression. **E**) Mean spatial module expression for each gene module in snRNA-seq cells by broad cell type. **F**) (**i**) Mean spatial module expression for each gene module in cortex layers L1-L6 and WM for Control, ALS, and ALSci donors. (**ii**) Change in spatial module expression, comparing ALS and ALSci to Control donors, expressed as Log_2 Fold Change. Asterisks denote significant differences to controls by Welch's t-test ($p < 0.05$), corrected for false discovery rate (0.05). **G**) Most significant gene ontology bioprocess terms associated with each gene module. Length of the bar corresponds to log_{10}(p-value). **H**) Spatial module scores for modules 11 and 12 overlaid on a Visium spatial array with a H&E-stained tissue section. Examples are shown for Control, ALS, and ALSci donors, and mean module scores for each cortex layer is plotted for the example donor arrays. **I**) Plots of individual ECAS scores (Executive function, Verbal Fluency, and Language) plotted against module scores that have been averaged over each Visium array for tissue sections from BA44/45 and from BA46. Panels highlighted in lilac have a statistically significant non-zero linear relationship by Wald test ($p < 0.01$, controlled for false discovery rate of 0.01). Pearson correlation (r) is shown; ECAS scores decrease with cognitive impairment, so negative correlation is higher module expression in cognitively impaired donors, and positive correlation is higher module expression in cognitively normal donors. **J**) Pearson correlation between module scores at each spatial array spot. Labels on the left and top are the gene module. Label on the bottom is the spatial region (L1-L6 & WM) with the highest module expression. Modules which do not have enrichment for any region are annotated as 'ALL'. Label on the right is the cell type which has the largest proportion of single cells with the highest expression of the module. **K**) Most significant two gene ontology bioprocess terms associated with each submodule of module 1. Length of the bar corresponds to log_{10}(p-value). **L**) Average expression of each gene in module 1 across cell types in snRNA-seq. Each gene is scaled min to max for plotting. **M**) snRNA-seq differential gene expression for selected cell types as log_2 fold change for genes from module 1 submodules, comparing ALS and ALSci to controls. Significant changes (DESeq2 padj < 0.05) labeled with asterisks. **O**) Pearson correlation between submodule scores for module 1, stratified by Control, ALS, and ALSci donors.

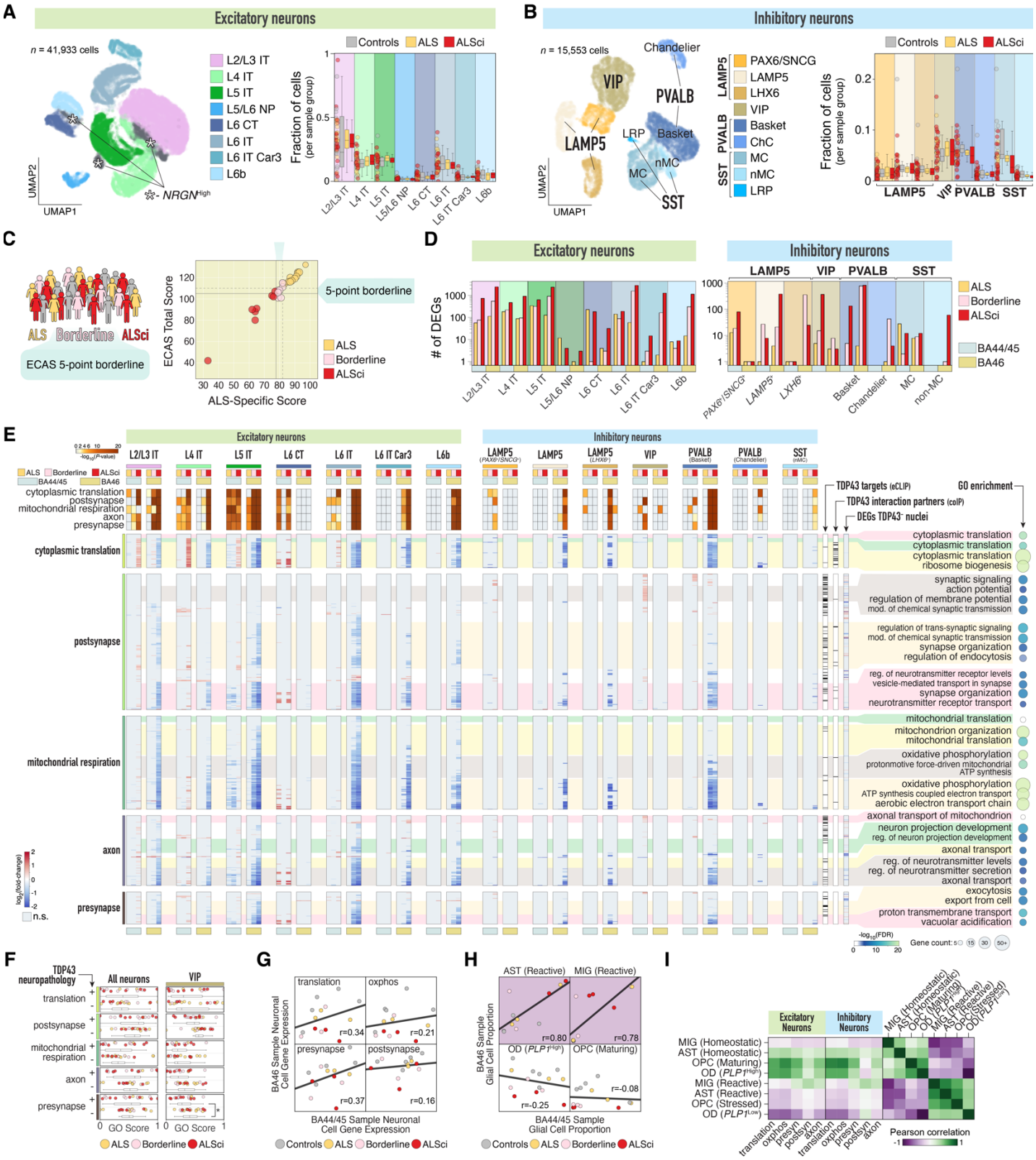
A
Excitatory neurons
n = 41,933 cells
L2/L3 IT
L4 IT
L5 IT
L5/L6 NP
L6 CT
L6 IT
L6 IT Car3
L6b
NRGN^High
UMAP1
UMAP2
Controls ALS ALSci
Fraction of cells (per sample group)
B
Inhibitory neurons
n = 15,553 cells
VIP
LAMP5
PVALB
Chandelier
Basket
LRP
nMC
MC
SST
PAX6/SNCG
LAMP5
LHX6
VIP
Basket
ChC
MC
nMC
LRP
C
ALS Borderline ALSci
ECAS 5-point borderline
ECAS Total Score
ALS-Specific Score
5-point borderline
D
Excitatory neurons
Inhibitory neurons
# of DEGs
BA44/45
BA46
E
Excitatory neurons
Inhibitory neurons
TDP43 targets (eCLIP)
TDP43 interaction partners (coIP)
DEGs TDP43⁻ nuclei
GO enrichment
cytoplasmic translation
postsynapse
mitochondrial respiration
axon
presynapse
n.s.
Gene count
F
TDP43 neuropathology
All neurons
VIP
GO Score
G
BA46 Sample Neuronal Cell Gene Expression
BA44/45 Sample Neuronal Cell Gene Expression
translation
oxphos
presynapse
postsynapse
r=0.34
r=0.21
r=0.37
r=0.16
H
AST (Reactive)
MIG (Reactive)
OD (PLP1^High)
OPC (Maturing)
BA46 Sample Glial Cell Proportion
BA44/45 Sample Glial Cell Proportion
r=0.80
r=0.78
r=-0.25
r=-0.08
I
Excitatory Neurons
Inhibitory Neurons
MIG (Homeostatic)
AST (Homeostatic)
OPC (Maturing)
OD (PLP1^High)
MIG (Reactive)
AST (Reactive)
OPC (Stressed)
OD (PLP1^Low)
Pearson correlation

Figure 4: Divergent intra- and cross-region signatures of neuronal dysfunction in ALS.
A) Excitatory neurons from snRNA-seq are plotted by UMAP and colored by neuronal subtype (left). The proportion of excitatory neurons in each subtype from each sample are plotted and summarized for each ALS phenotype with boxplots (right); whiskers extend to the farthest data point within 1.5 times the interquartile range (IQR). **B**) Inhibitory neurons from snRNA-seq, plotted and summarized as in **A**. (ChC, Chandelier Cells; MC, Martinotti cells; nMC, non-Martinotti cells) **C**) Additional cohort stratification into borderline cognitive impairment in an ECAS 5-point borderline range. **D**) The number of differentially expressed transcripts (DEseq2; Multiple hypothesis test corrected Wald test $p \leq 0.05$) compared to control donors across neuronal subtypes, clinical phenotypes, and brain regions. **E**) Gene ontology (GO) analysis of differentially expressed genes across neuronal subtypes, clinical phenotypes, and brain regions. Enrichment is plotted as $-\log_{10}$(p-value) (multiple hypothesis test corrected hypergeometric test) for the 5 functional terms with the smallest p-value, including biological process (GO:BP) and cellular component (GO:CC) terms (top). Differentially expressed transcripts compared to controls, for neuronal subtypes in different brain regions, are plotted as a $\log_2$(fold change) heatmap for genes annotated with translation, mitochondrial respiration, and neuronal connectivity ontology terms (bottom). Transcript $\log_2$(fold changes) which are not significantly changing are masked. TDP-43 binding transcripts, protein-coding transcripts for proteins that physically interact with TDP-43, and transcripts that are differentially expressed in nuclei when TDP-43 is genetically ablated are annotated (right). Gene ontology enrichment is shown for subclusters of the differentially expressed transcripts (far right). **F**) All transcripts annotated with gene ontology terms for translation, mitochondrial respiration, and neuronal connectivity ontology terms are aggregated into per-sample ontology module scores from all neurons (left) or from only VIP interneurons (right). These scores are stratified and plotted for donors positive and negative for TDP-43 proteinopathy, with dots colored by ALS phenotype and summary boxplots. Statistically significant differences between TDP-43$^+$ and TDP-43$^-$ ontology module scores are determined by Welch's t-test ($p < 0.05$) **G**) Correlation between BA44/45 and BA46 neuronal ontology module scores taken from the same donor. Black line shows best fit from ordinary least squares regression, and slope of the best-fit line is tested by Wald test (all are non-significant). **H**) Correlation between the glial cell proportion in BA44/45 and BA46 of homeostatic and reactive astrocytes, homeostatic and reactive microglia, high- and low-myelinating oligodendrocytes, and maturing and stressed OPCs. Black line shows best fit from ordinary least squares regression, and plots highlighted in lilac have best-fit lines with slopes significantly different from 0 by Wald test ($p < 0.01$ corrected for false discovery rate). **I**) Correlation between glial cell functional proportions and the aggregate snRNA-seq expression of transcripts annotated with specific ontologies from neuronal cells in BA44/45 (top) and BA46 (bottom), with correlation between glial cell functional proportions (right).