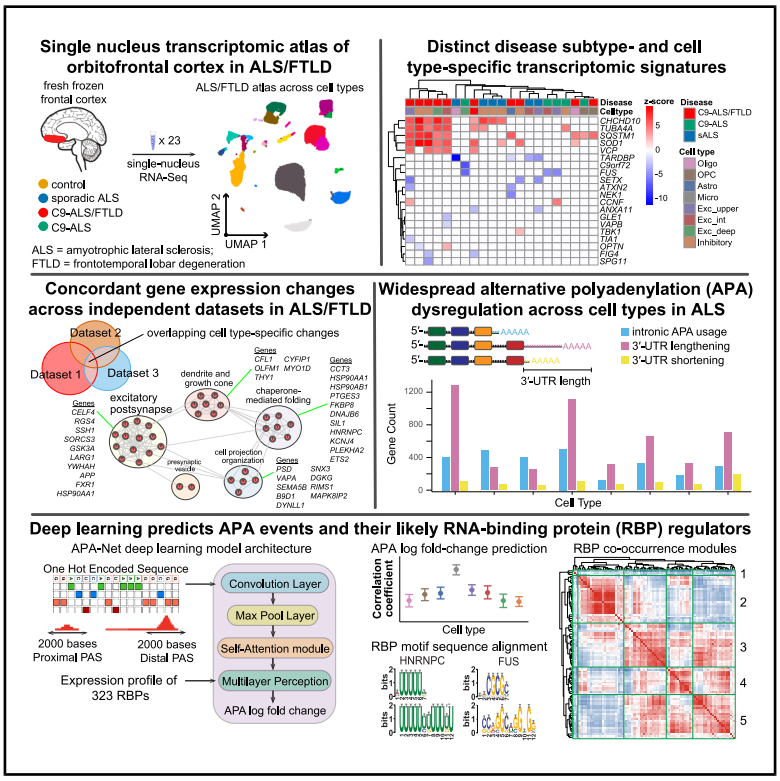
Cell Genomics

Article

Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms

Graphical abstract



Highlights

- snRNA-seq atlas of orbitofrontal cortex in ALS identified cell-type-specific changes
- Comparative analyses with existing ALS datasets uncover shared molecular signatures
- Alternative polyadenylation is dysregulated across neuronal and glial cell types in ALS
- Deep learning model, APA-Net, decodes APA regulation from snRNA-seq data

Authors

Paul M. McKeever, Aiden M. Sababi, Raghav Sharma, ..., Hani Goodarzi, Gary D. Bader, Janice Robertson

Correspondence

gary.bader@utoronto.ca (G.D.B.), jan.robertson@utoronto.ca (J.R.)

In brief

McKeever et al. use single-nucleus RNA sequencing to map molecular changes in the orbitofrontal cortex across ALS with and without FTLD. They identify cell-type-specific transcriptional alterations and widespread dysregulation of alternative polyadenylation and apply a deep learning model to uncover the regulatory landscape of APA.

McKeever et al., 2025, Cell Genomics *5*, 101007
December 10, 2025
https://doi.org/10.1016/j.xgen.2025.101007

CellPress

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

Cell Genomics

CellPress
OPEN ACCESS

Article

Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms

Paul M. McKeever,[1,12] Aiden M. Sababi,[2,3,12] Raghav Sharma,[1,4] Zhiyu Xu,[5] Shangxi Xiao,[1] Philip McGoldrick,[1] Troy Ketela,[5] Christine Sato,[1] Danielle Moreno,[1] Naomi Visanji,[1] Gabor G. Kovacs,[1,4] Julia Keith,[6] Lorne Zinman,[7] Ekaterina Rogaeva,[1] Hani Goodarzi,[8,9] Gary D. Bader,[2,3,5,10,11,*] and Janice Robertson[1,4,13,*]

[1]Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, ON, Canada
[2]Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada
[3]The Donnelly Centre, University of Toronto, Toronto, ON, Canada
[4]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada
[5]Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada
[6]Laboratory Medicine and Molecular Diagnostics, Division of Anatomical Pathology, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, ON, Canada
[7]Division of Neurology, Department of Medicine, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, ON, Canada
[8]Department of Biophysics and Biochemistry, University of California, San Francisco, San Francisco, CA, USA
[9]Arc Institute, Palo Alto, CA, USA
[10]Lunenfeld-Tanenbaum Research Institute, Sinai Health System, University of Toronto, Toronto, ON, Canada
[11]Department of Computer Science, University of Toronto, Toronto, ON, Canada
[12]These authors contributed equally
[13]Lead contact
*Correspondence: gary.bader@utoronto.ca (G.D.B.), jan.robertson@utoronto.ca (J.R.)
https://doi.org/10.1016/j.xgen.2025.101007

SUMMARY

Amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) are fatal neurodegenerative diseases sharing clinical and pathological features. Both involve complex neuron-glia interactions, but cell-type-specific alterations remain poorly defined. We performed single-nucleus RNA sequencing of the frontal cortex from *C9orf72*-related ALS (with and without FTLD) and sporadic ALS (sALS). Neurons showed prominent changes in mitochondrial function, protein homeostasis, and chromatin remodeling. Comparison with independent datasets from other cortical regions revealed consistent pathway alterations, including upregulation of *STMN2* and *NEFL* across brain regions and subtypes. We further examined dysregulation of alternative polyadenylation (APA), an understudied post-transcriptional mechanism, uncovering cell-type-specific APA patterns. To investigate its regulation, we developed the alternative polyadenylation network (APA-Net), a multi-modal deep learning model integrating transcript sequences and RNA-binding protein (RBP) expression profiles to predict APA. This atlas advances our understanding of ALS/FTLD molecular pathology and provides a valuable resource for future mechanistic studies.

INTRODUCTION

The advent of single-cell sequencing technology facilitates a deeper interrogation of the mechanisms that underlie disease, especially in complex tissues such as the brain. Amyotrophic lateral sclerosis (ALS) is an adult-onset neurodegenerative disease caused by motor neuron degeneration in the brain, brain stem, and spinal cord. Extending beyond motor pathways, over 50% of ALS patients also exhibit progressive impairments in cognition, behavior, or language caused by frontotemporal lobar degeneration (FTLD), with 10%–15% of patients fulfilling the diagnostic criteria for frontotemporal dementia.[1] Neuroimaging studies further support early frontal cortical involvement in ALS, highlighting structural and functional disruptions specifically in regions such as the prefrontal and orbitofrontal cortices,[2–4] areas critical for higher-order cognitive processing and behavioral regulation. Although the vast majority of ALS cases occur sporadically (sporadic ALS [sALS]), without clear genetic causes, familial studies have identified mutations in several genes, including *SOD1*, *TARDBP* (TDP-43), and *FUS*, as important contributors to disease pathogenesis.[5] Among these, a hexanucleotide (G4C2) repeat expansion within intron 1 of *C9orf72* (C9) represents the most common genetic cause shared by ALS and FTLD,[6,7] highlighting a clear molecular continuum between motor and cortical dysfunction in ALS. This mutation triggers multiple pathological mechanisms, including dipeptide repeat protein toxicity, RNA-binding protein (RBP) sequestration by repeat-containing RNA, and reduced *C9orf72*

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

expression.[6–8] Additionally, C9orf72 itself plays diverse roles in neuronal health through modulation of small guanine triphosphatases (GTPases) and regulation of essential cellular pathways, such as autophagy, lysosomal processing, immune response, synaptic plasticity, and nucleocytoplasmic transport.[9–12]

The molecular mechanisms underlying cortical dysfunction associated with FTLD in ALS remain elusive but have been linked to changes in expression and/or subcellular localizations of various RBPs. Nuclear depletion and cytoplasmic aggregation of TDP-43 represents the hallmark pathology in motor neurons of ALS and cortical neurons of FTLD.[13,14] Other RBPs implicated in ALS/FTLD include FUS,[15] SFPQ,[16,17] TIA1,[18] and heterogeneous nuclear ribonucleoproteins (hnRNPs).[19,20] RBPs regulate diverse functions across cell types, including transcription,[21–24] auto-regulation,[25,26] alternative splicing (AS),[22] and alternative polyadenylation (APA).[26–30] Disruption of these post-transcriptional regulatory processes has been documented in ALS,[22,31] including the observation of brain region-specific APA patterns in postmortem bulk RNA sequencing (RNA-seq) from ALS patients[32] and in induced pluripotent stem cell-derived motor neurons from patients harboring *TARDBP* and *VCP* mutations.[33] APA is a critical post-transcriptional mechanism that contributes to transcriptomic diversity by generating mRNA isoforms with distinct 3′ untranslated region (3′ UTR) regions or coding sequences. Beyond its role in diversifying mRNA, APA plays essential functions during development, differentiation, and the cellular response to environmental cues.[34,35] Aberrant APA has been reported in several neurodegenerative diseases, including ALS, Parkinson's disease, and Alzheimer's disease.[36–38] Such alterations in 3′ end processing can influence mRNA stability, localization, and translation, potentially contributing to disease pathogenesis. Despite these insights, the cell-type-specific regulation and dysregulation of APA in ALS remain poorly characterized. Recent advances in methods for APA quantification in single-cell RNA-seq datasets[39–43] now enable systematic investigation of APA at cell type resolution, providing an opportunity to explore this understudied regulatory dimension in ALS.

Here, we generated a single-nucleus RNA-seq (snRNA-seq) atlas comprising 103,076 nuclei to uncover transcriptomic changes in postmortem orbitofrontal cortices from ALS cases caused by mutations in *C9orf72* (C9-ALS) and sALS cases. We also analyzed C9-ALS cases with FTLD, identified by TDP-43 pathology in the orbitofrontal cortex (C9-ALS/FTLD). Our snRNA-seq analysis of the orbitofrontal cortex revealed distinct molecular signatures in C9-ALS/FTLD, C9-ALS, and sALS cases. Each disease subtype showed unique patterns of pathway dysregulation across cell types while also sharing common alterations in processes such as protein homeostasis and mitochondrial dysfunction. By integrating our findings with published datasets from the dorsolateral prefrontal and primary motor cortex,[44–47] we discovered concordant cell-type-specific transcriptional changes across brain regions, suggesting that these represent fundamental disease mechanisms rather than region-specific responses. These molecular signatures were particularly striking in neuronal populations, where we observed consistent patterns of gene expression changes across datasets.

We also investigated cell-type-specific dysregulation of APA in ALS, identifying thousands of significant APA events in 3′ UTRs and upstream internal regions in C9-ALS and sALS compared to controls. To further understand this APA dysregulation, we developed an interpretable deep learning method called alternative polyadenylation network (APA-Net). APA-Net enabled us to identify a range of *cis* and *trans* regulators correlated with the observed APA events in the disease. By interpreting APA-Net, we reveal potential RBP interactions in the dysregulation of APA in ALS, shedding light on the regulatory programs likely to induce APA in ALS. These findings enhance our understanding of cell-type-specific transcriptomic changes occurring in the frontal cortex in ALS and provide mechanistic insights into the potential coordinated interaction of RBPs in regulating APA in disease.

RESULTS

Orbitofrontal cortex cell types in ALS patients with and without FTLD show converging and diverging transcriptomic changes

We leveraged snRNA-seq to characterize both gene expression and APA modalities in ALS (Figures 1A and 1B). Our single-cell atlas comprised 103,076 nuclei derived from 23 snRNA-seq samples from the orbitofrontal cortex (Brodmann area 11) of C9-ALS/FTLD (n = 6 individuals), C9-ALS without FTLD (n = 3), and sALS without FTLD (n = 8) cases compared to controls (n = 6) (Figure S1A). Individual- and group-level demographic information is summarized in Tables S1 and S2, respectively. Standard quality control steps were applied to the snRNA-seq data.[48] We used an established snRNA-seq data integration method,[49] ensuring consistent alignment of cell clusters across technologies for sex, condition, and samples (Figures S1B–S1D).

To delineate cell types, we applied an established clustering algorithm[50] and cell annotation approach[51,52] (Figures S1E and S1F; STAR Methods). We uncovered 23 orbitofrontal cortex cell subtypes, including oligodendrocytes, oligodendrocyte precursor cells (OPCs), astrocytes, endothelial cells, vascular leptomeningeal cells (VLMCs), microglia and perivascular macrophages (PVMs), 10 excitatory neuron subtypes, and seven inhibitory neuron subtypes (Figure S1G). A catalog of specific gene expression markers for each cell type is presented in Figure S1H. To facilitate downstream analysis, we grouped cell subtypes into nine major cell types: oligodendrocytes, OPCs, astrocytes, endothelial VLMCs, microglia-PVMs, upper-layer excitatory neurons, intermediate-layer excitatory neurons, deep-layer excitatory neurons, and inhibitory neurons, based on canonical markers (Figures S2A and S2B). The nuclei from endothelial VLMCs were excluded from all downstream analyses due to poor yield across samples. All other cell types were detected across samples, showing similar distributions of cell types when categorized by disease subtype (Figures S2C–S2E).

To assess whether transcriptional differences between controls and ALS cases reflect true biological separation rather than random variation, we performed principal-component analysis on pseudobulk profiles across all cell types (Figure S3A). We then calculated pairwise Euclidean distances between samples, comparing within-group (i.e., control vs. control, ALS vs. ALS) and between-group (control vs. ALS) variation. Bootstrapped confidence intervals reveal a trend toward

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

CellPress
OPEN ACCESS

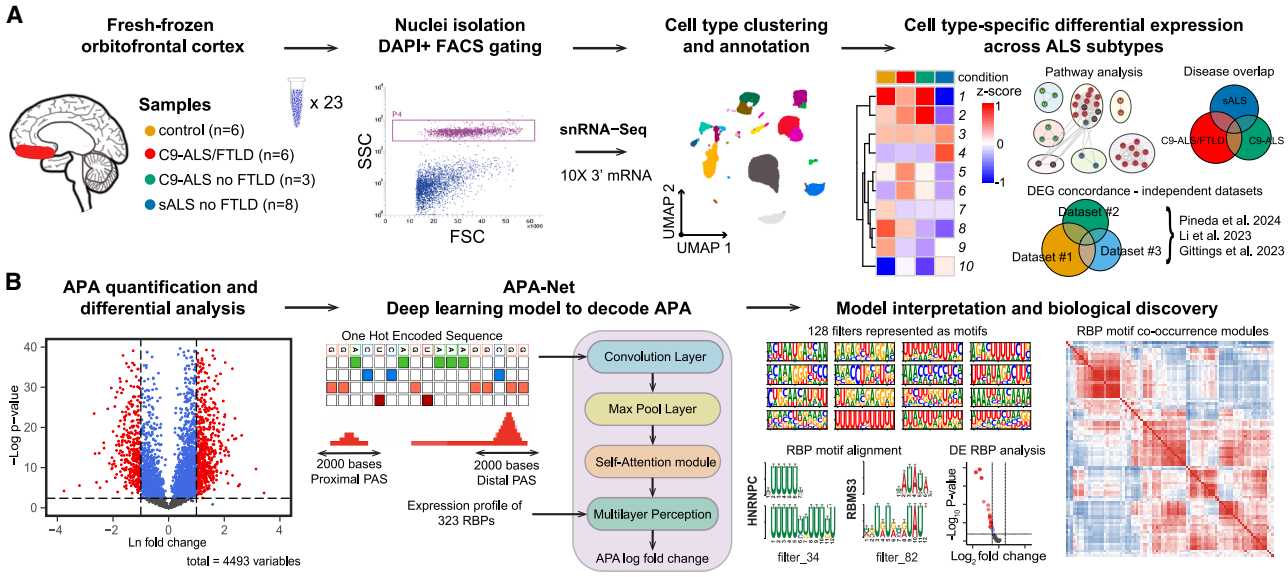


Figure 1. Single-nucleus atlas for the study of transcriptomic changes in the ALS frontal cortex

(A) Workflow for the generation of a single-nucleus atlas of the frontal cortex from controls ($n = 6$), C9-ALS/FTLD ($n = 6$), C9-ALS without FTLD ($n = 3$), and sALS without FTLD ($n = 8$). All n values represent distinct biological replicates (individuals). Nuclei were isolated using fluorescence-activated cell sorting, and nuclei were labeled and gated with DAPI. Analyses include cell type clustering, machine learning and reference-based annotation, differential expression analysis, pathway analysis, overlap across ALS subtypes, and concordance with independent single-nucleus datasets. DEG = differentially expressed genes.

(B) APA analysis across ALS subtypes and cell types. The APAlog software facilitated quantification and differential expression analysis of APA events from assigned reads. APA grammar was decoded using a multi-input deep learning model called APA-Net, which consists of a CNN with an MAT mechanism. APA-Net uses the sequence information around polyadenylation site (PAS) and RBP expression profiles to predict APA LFC.

greater between-group distances across multiple cell types (Figure S3B). To test whether this separation exceeded chance expectations, we conducted permutation testing. This confirmed statistically significant group separation in all excitatory neuron subtypes, astrocytes, and microglia ($p < 0.05$, permutation test; Figure S3B). Together, these results suggest that the transcriptional differences distinguishing ALS from control samples are unlikely to be driven by random noise.

We next explored these differences further by performing a differential expression analysis,[53] comparing C9-ALS/FTLD, C9-ALS, and sALS cases with controls (STAR Methods). This analysis uncovered widespread cell-type-specific transcriptomic changes across all disease subtypes (false discovery rate (FDR) < 0.01; $|\log_2 \text{fold change}| > 0.5$; Figures 2A–2G and S4A–S4R; Table S3). The majority of gene expression changes are found in excitatory and inhibitory neurons, with the most significant variances observed in C9-ALS/FTLD compared to C9-ALS and sALS (Figure 2A). Venn diagrams show the convergent and divergent changes among the three disease subtypes compared with controls for cortical-layer excitatory neurons, inhibitory neurons, and glial cells: oligodendrocytes, OPCs, astrocytes, and microglia (Figures 2H–2O).

Neurons in ALS are perturbed across all subtypes and cortical layers

To gain insights into the biological underpinnings of the identified differential gene expression changes across disease subtypes, we conducted a pathway analysis[54] across cell types (STAR Methods; Table S4). Upper-layer excitatory neurons across disease subtypes show upregulation of ribosomal complexes and glutamatergic synapse genes with downregulation of processes related to chromatin remodeling, presynaptic signaling, and dendritic spines and postsynapses[55,56] (Figure 3A). Major pathway changes in excitatory neurons across layers in C9-ALS/FTLD include upregulation of the electron transport and respiratory chains, indicative of mitochondrial dysfunction, as well as alterations in protein homeostasis pathways, including autophagy, proteosome function, and endoplasmic reticulum stress (Figures 3A, S5A, and S5B).

Intermediate and deep-layer excitatory neurons in C9-ALS specifically show upregulation of postsynaptic organization and downregulation of cerebral cortex migration, respectively (Figures S5A and S5B). In sALS, excitatory neurons across cortical layers show downregulation of mRNA metabolism, with DNA damage pathways altered specifically in upper-layer neurons (Figures 3A, S5A, and S5B). Deep-layer excitatory neurons in sALS uniquely show upregulation of glutamatergic and GABAergic synapses, axonogenesis, and forebrain development (Figure S5B), highlighting previously documented dysregulation of axonal and potassium channel receptor genes in ALS.[57,58] Convergent gene expression changes across ALS subtypes involve pathways related to autophagy, glutamatergic neuronal function, mRNA metabolism, and the respiratory chain (Figures 3B–3D and S5C–S5G), with consistent upregulation of ALS-related genes such as *SQSTM1*, *VCP*, *CHCHD10*, *UCHL1*, and *NEFL*.[5,59]

We next evaluated the intersections of differentially expressed genes from upper-layer excitatory neurons between our dataset

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

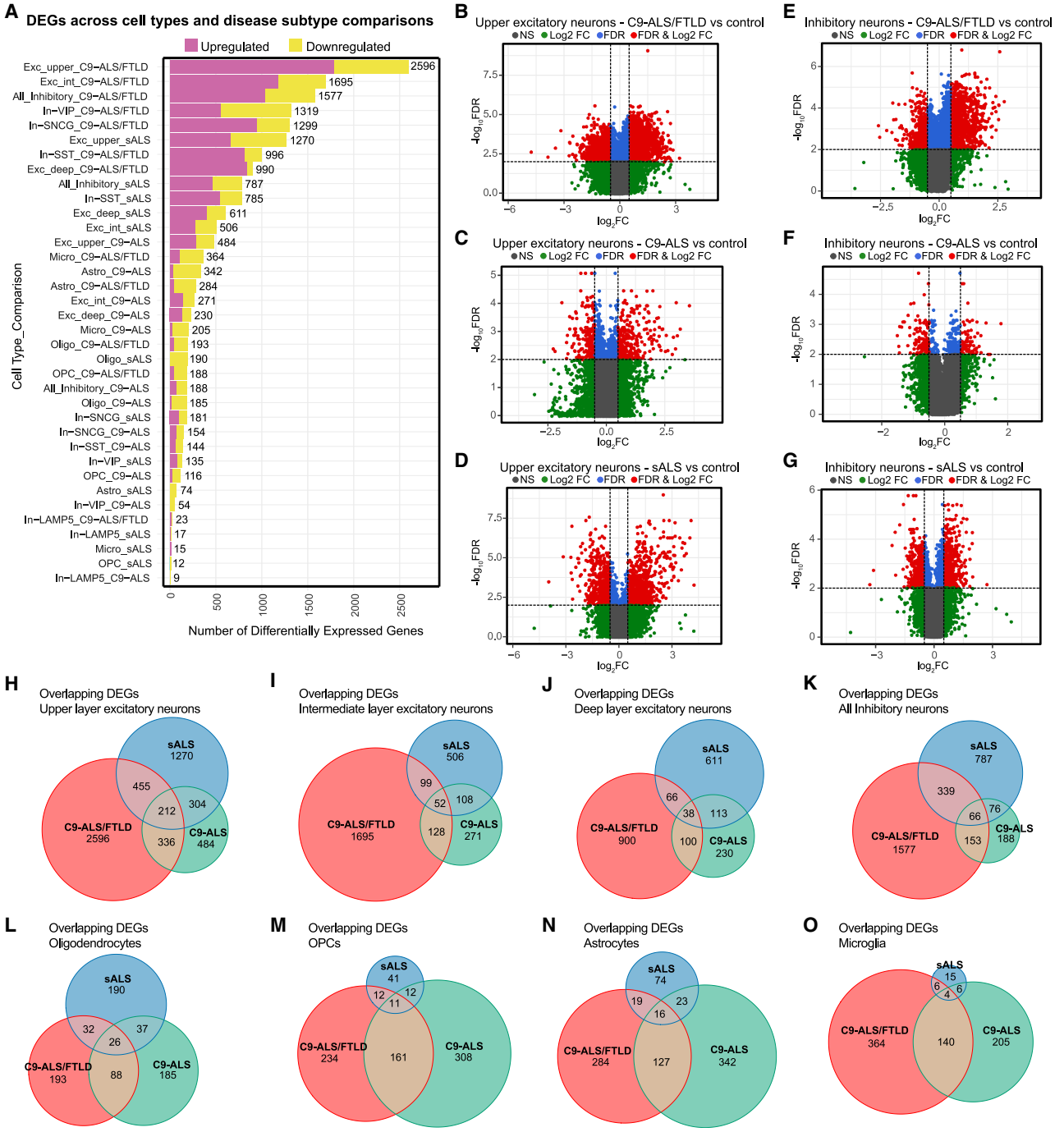


Figure 2. Distribution of differentially expressed genes across cell types in ALS

(A) Stacked bar plot indicating the number of differentially expressed genes on the *y* axis across ALS cell subtypes. Purple indicates upregulation and yellow downregulation.

(B–G) Volcano plots of differentially expressed genes in upper-layer excitatory neurons in (B) C9-ALS/FTLD cases vs. controls, (C) C9-ALS vs. controls,; and (D) sALS vs. controls and inhibitory neurons in (E) C9-ALS/FTLD cases vs. controls, (F) C9-ALS vs. controls, and (G) sALS vs. controls. Only differentially expressed genes that passed the FDR < 0.01 and $|LFC| \geq 0.5$ cutoffs were considered significant (shown in red). Differential expression was performed using the model-based analysis of single-cell transcriptomics (MAST) method[53] with a mixed-effect hurdle model including diagnosis, number of detected genes, unique molecular index count, sex, and age as fixed effects and donor ID as a random effect. All tests were two sided and adjusted for multiple comparisons with the Benjamini-Hochberg method. Effect sizes are expressed as model-derived $\log_2$ fold changes.

(H–O) Venn diagrams indicating overlapping differentially expressed genes between C9-ALS/FTLD, C9-ALS, and sALS across cell types.

All data are based on 4 individuals with C9-ALS/FTLD, 2 with C9-ALS without FTLD, 4 with sALS, and 6 controls.

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

CellPress
OPEN ACCESS

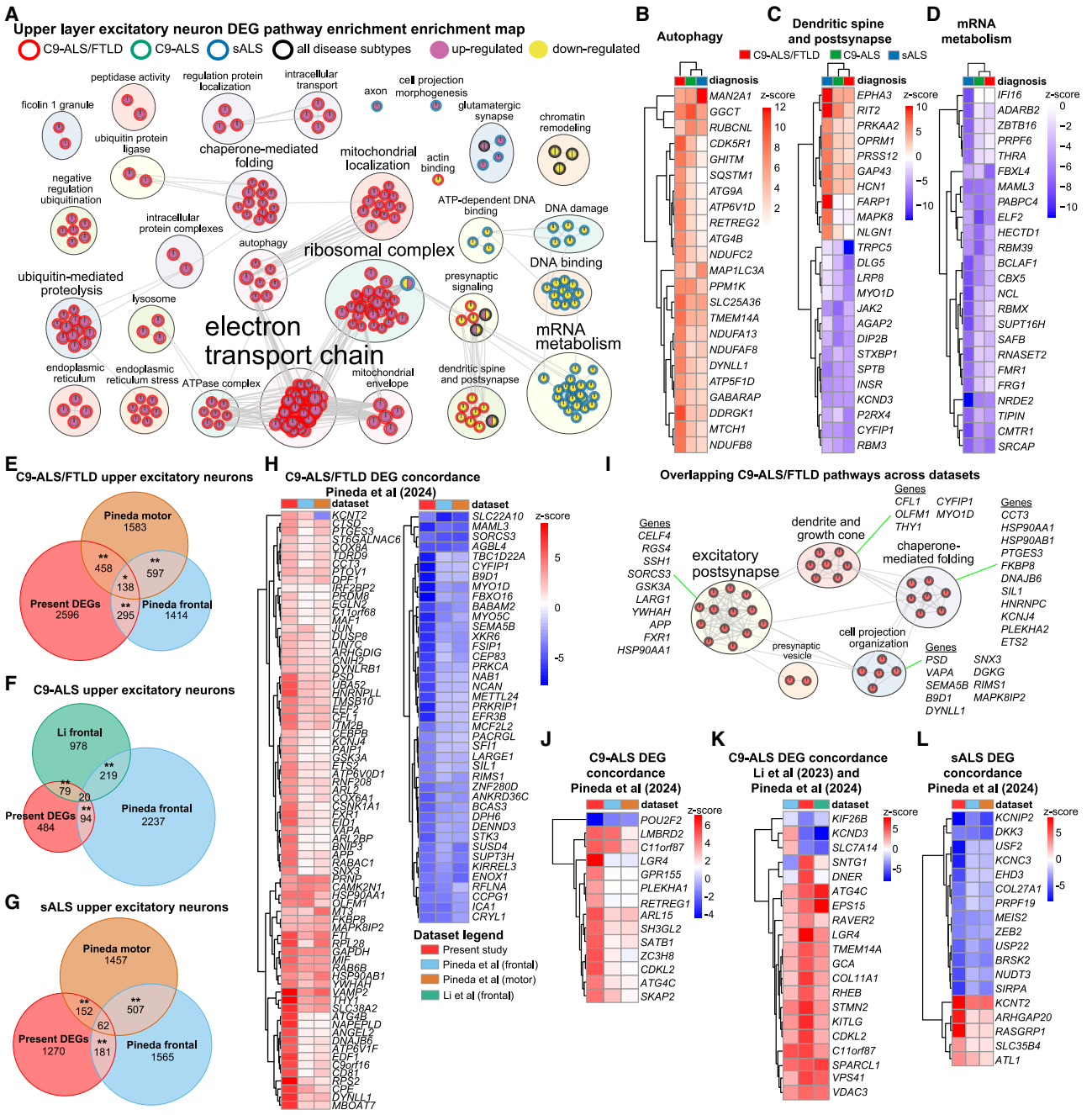


Figure 3. Upper-layer excitatory neurons demonstrate converging and diverging changes across ALS disease subtypes

(A) Pathway analysis enrichment map for differentially expressed genes in upper-layer excitatory neurons across ALS subtypes. Circles represent pathways grouped by similarity (bubbles) and lines indicate shared genes between pathways. Circle center and border colors indicate disease subtype and differential expression status.

(B–D) Clustered heatmap for top-ranked genes identified in annotated enrichment map clusters showing convergence across ALS subtypes for (B) autophagy, (C) dendritic spine and postsynapse, and (D) mRNA metabolism.

(E) Venn diagram of C9-ALS/FTLD differentially expressed genes in upper excitatory neurons that overlap between the present study and the motor and frontal cortices.[45] The triple intersection is tested for significance using a log-linear model (χ^2, two-sided) with degrees of freedom = 2; $^*p = 0.0356$. Pairwise overlaps are tested by separate hypergeometric analyses (one sided); $^{**}p < 0.001$ after Bonferroni correction.

(F) Venn diagram of C9-ALS differentially expressed genes in upper excitatory neurons which overlap between the present study and the frontal cortex from two independent studies.[44,45]

(G) Venn diagram of sALS differentially expressed genes in upper excitatory neurons that overlap between the present study and the motor and frontal cortices.[45]

(legend continued on next page)

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

and two independent single-nucleus datasets in dorsolateral prefrontal and primary motor cortices from C9-ALS, C9-FTLD, and sALS cases (Figures 3E–3L and S6A–S6D; Table S5).[44,45] We identified 138 significantly overlapping genes between our C9-ALS/FTLD cases and C9-FTLD cases described by Pineda et al. (Figures 3E, 3H, and S6A; chi-square $p < 0.05$), with 115 showing concordant expression changes (74 up, 41 down) (Figures 3, 3H, and S6A). Concordant genes include *PACRGL* and *ICA1*, known to contain cryptic exons repressed by TDP-43,[60] and pathway changes related to excitatory postsynapses, dendrites and axonal growth cones, chaperone-mediated folding, and cell projection organization (Figure 3I). These results identify a core group of overlapping genes and pathways perturbed in C9-ALS/FTLD and C9-FTLD.

For C9-ALS, we observe 20 overlapping differentially expressed genes between our dataset and those from the frontal cortex in Li et al. and Pineda et al., with 16 of these showing concordance (Figures 3F, 3K, S6B, and S6C). *STMN2*, known to undergo cryptic splicing and premature polyadenylation under TDP-43 deficiency in ALS,[28,60–62] shows consistent upregulation across datasets in upper-layer excitatory neurons. Additionally, *ATG4C*, *C11orf87*, and *CDKL12* are upregulated across all regions and datasets in C9-ALS (Figure S6C), though their specific role in ALS remains unclear. In sALS, we observe 62 overlapping genes with Pineda et al.'s motor and frontal cortex data (Figure 3G). Concordantly dysregulated genes involve potassium channels (*KCNT2*, *KCNIP2*, and *KCNC3*), neuronal differentiation (*BRSK2*, *ATL1*, and *ARHGAP20*), and neuronal stress/structural markers (*USF2*, *EHD3*, *PRPF19*, *ZEB2*, *USP22*, and *SIRPA*) and are predominantly downregulated (Figure 3L). These results underscore shared and distinct molecular vulnerabilities in excitatory neurons between sALS and C9-linked ALS.

We next sought to determine whether similar patterns of pathway dysregulation observed in excitatory neurons were also evident in inhibitory neurons across ALS subtypes. Pathway analysis of inhibitory neurons revealed both unique and overlapping alterations across ALS subtypes (Figure S7A). The electron transport chain is consistently upregulated across ALS subtypes, strongly implicating mitochondrial dysfunction in disease. Chromatin remodeling is downregulated similarly across inhibitory and excitatory neurons in ALS subtypes. C9-ALS/FTLD inhibitory neurons show upregulation of ribosomal subunits, RNA localization, mitochondrial function, and proteostasis (Figure S7A). In C9-ALS, we observe increased axonal growth cone and downregulation in double-stranded DNA break repair pathways (Figure S7A). In sALS, we note downregulation in presynaptic signaling, dendritic structure, mRNA processing, histone binding, and DNA damage response pathways (Figure S7A). Analysis of top-ranked genes in RNA localization, DNA break repair, and chromatin remodeling shows convergence across ALS subtypes, including upregulation of the disease-relevant genes *SOD1* and *SQSTM1* (Figures S7B–S7D).

Comparisons with independent datasets from inhibitory neurons further support these findings (Figures S7E–S7K; Table S5).[44–46] We find 84 differentially expressed genes significantly overlapping between our frontal and motor cortex data from Pineda et al., of which 66 genes showed concordance (57 up, 9 down) (Figures S6E, S7E, and S7H; chi-square $p < 0.01$). Concordant genes include the ALS-related *VCP*[63] with pathways involving synaptic regulation, dendritic structures, ribonucleoprotein granules, and translation (Figure S7I).[45,47] For C9-ALS, we observe 21 overlapping genes with datasets from Li et al. and Pineda et al. (Figures S7F and S6F), including upregulation of *HSP90AA1*, *STMN2*, *NEFL*, *MDH1*, *RGS4*, and *VDAC3* (Figure S7J).[44,45] In sALS inhibitory neurons, we uncover 63 overlapping differentially expressed genes when comparing our study with frontal and motor cortex data from Pineda et al. (Figures S6G and S7G), including concordant upregulation of ribosomal subunits (*RPL7A*, *RPS7*, and *RPS26*) and *STMN2* (Figure S7K). This consistent dysregulation underscores shared molecular vulnerabilities in inhibitory neurons across sALS and C9-ALS subtypes.

Overall, these data highlight both converging and distinct molecular signatures in inhibitory neurons, closely overlapping with changes observed in excitatory neurons across cortical regions in ALS (summarized in Table S6).

Protein homeostasis pathways are dysregulated in glial cells across ALS subtypes

Oligodendrocytes may drive neuronal dysfunction in ALS via impaired myelination, metabolic deficits, and pro-inflammatory signaling.[47,64] Across ALS subtypes, oligodendrocytes show upregulation of chaperone-mediated folding genes, and downregulation of genes involved in mRNA metabolism and chromatin remodeling (Figures S8A–S8C). C9-ALS/FTLD oligodendrocytes show elevated apoptosis and impaired neuronal development pathways, whereas C9-ALS shows reduced DNA break repair. sALS oligodendrocytes uniquely exhibit downregulation of *TARDBP* without specific pathway enrichment (Figures S8A and S8C; Table S3).[65] Comparing oligodendrocyte gene expression changes with data from the frontal and motor cortices from Pineda et al., both C9-ALS/FTLD and sALS demonstrate downregulation of myelin-associated scaffold gene *SEPTIN4*[66] (Figure S8D–S8F). C9-ALS/FTLD oligodendrocytes also show decreased expression of *NAALADL2* (Figure S8D), which is associated with myelin deficiency,[67] and downregulation of the

(H) Clustered heatmap of concordant C9-ALS/FTLD differentially expressed genes (positive correlation between *Z* scores), separated by upregulated (left) and downregulated (right) genes.
(I) Enrichment map for overlapping biological processes enriched in C9-ALS/FTLD frontal and motor areas (based on the three-way intersection in E). Select genes from each cluster are highlighted.
(J) Clustered heatmap of concordant C9-ALS differentially expressed genes in frontal regions alone.
(K) Clustered heatmap of concordant C9-ALS differentially expressed genes in frontal and motor regions.
(L) Clustered heatmap of concordant sALS differentially expressed genes in frontal and motor regions.
Across independent dataset comparisons, significance was set at FDR < 0.01 and $|LFC| \geq 0.5$. All data are based on 4 individuals with C9-ALS/FTLD, 2 with C9-ALS without FTLD, 4 with sALS, and 6 controls.

chromatin remodeling and splicing genes *BAZ2B* and *SRRM1*, respectively (Figure S8D). Concordant with independent datasets, C9-ALS oligodendrocytes show increased expression of the autophagy-related genes *ATG4B* and *HSP90AB1* (Figure S8E), both RNA targets of TDP-43.[68,69] Oligodendrocytes from sALS (Figure S8F) show downregulation of the myelin sheath component *PLLP* and TDP-43 kinase *CSNK1E*.[70]

In OPCs, ATP metabolism is upregulated across ALS subtypes (Figures S8G and S8H), suggesting a stressed cellular state. Chaperone-mediated folding is upregulated in C9-ALS/FTLD and C9-ALS, while both C9-ALS and sALS show downregulation of glycosylphosphatidylinositol (GPI) anchor processes and chromatin remodeling (Figures S8G and S8I). Specifically, C9-ALS/FTLD OPCs show elevated ubiquitin protein ligase activity but reduced pathways related to cell morphogenesis and tubulin binding (Figure S8G). C9-ALS OPCs demonstrate upregulation of endogenous hormone response genes and downregulation in ATP-dependent DNA binding pathways (Figure S8G). In sALS, OPCs show downregulation of synapse organization pathways (Figure S8G), suggesting impaired neuronal connectivity. Despite limited overlap with independent datasets from Li et al. and Pineda et al., *ANKRD10* shows concordant upregulation specifically in C9-ALS, consistent with its elevation in skeletal muscle from ALS patients.[71] In sALS, *APOD*, critical for remyelination,[72,73] is consistently downregulated, aligning with proteomic alterations observed in cerebrospinal fluid from ALS patients.[74]

Astrocytes are linked with the pathophysiology of ALS through loss of homeostasis and reactive toxicity.[75,76] In C9-ALS/FTLD and C9-ALS (Figures 2A and 2N), astrocytes feature upregulation of the unfolded protein response and actin filament pathways alongside downregulation of chromatin organization (Figures S8J–S8L). C9-ALS/FTLD astrocytes specifically show elevated cholesterol and microRNA (miRNA) metabolism pathways, reduced filopodium tip gene expression, and increased reactive markers *GFAP* and *CHI3L1*[77,78] (Figure S8J; Table S3), common markers of reactive astrocytes. Astrocytes in C9-ALS uniquely demonstrate increased protein localization and decreased cilium assembly formation[79] (Figure S7J). In sALS, astrocytes show downregulation of nuclear protein complex genes and *TARDBP* (Figures S8J and S8L), consistent with reactive astrocyte phenotypes linked to TDP-43 loss.[80] Analysis of independent frontal and motor cortex datasets reveals concordant downregulation of the synapse-associated genes *SLC4A4* and *GRIA2* in C9-ALS/FTLD astrocytes (Figure S8M; Table S5), while C9-ALS astrocytes show reduced *NTRK2*, essential for astrocyte morphology and brain-derived neurotrophic factor signaling[75,76,81,82] (Figure S8N; Table S5). In sALS astrocytes, there is concordant downregulation of *SNRNP70*, a splicing-related RBP[83] linked to ALS through interactions with FUS.[84] Together, these findings highlight distinct and overlapping astrocytic alterations across ALS and FTLD subtypes and brain regions.

Altered cell states in microglia, the primary immune cells of the central nervous system, are a hallmark of ALS and neurodegeneration.[85–87] Compared to sALS, microglia from C9-ALS (with and without FTLD) demonstrate more pronounced gene expression changes (Figures 2A and S9A), consistent with immune-related function related to *C9orf72*.[88,89] Across ALS subtypes, microglia show increased glucose metabolism, cytokine signaling, and chaperone-mediated refolding alongside decreased small GTPase signaling, neuron projection morphogenesis, and interleukin-6 pathways (Figures S9A–S9D. C9-ALS/FTLD microglia specifically upregulate the innate immune response, antigen presentation, phagocytosis, and iron ion homeostasis pathways (Figure S9A) but downregulate chemotaxis and glutamatergic synapse function genes (Figures S9A and S9B). C9-ALS microglia shift toward pro-inflammatory signaling, programmed cell death, and growth pathways (Figure S9A). Notably, *JAK1* and *JAK2* are consistently downregulated across all ALS subtypes (Figure S9D), consistent with data demonstrating the importance of JAK/STAT signaling in ALS pathogenesis.[90]

Microglia from our orbitofrontal data show gene expression changes concordant with independent frontal and motor cortex datasets,[44–46] including upregulation of the microglial-enriched marker *SYTL3*[91] and downregulation of *ATP8B4* (Figure S9E), a gene implicated in Alzheimer's disease risk.[92] We examined microglial states across ALS subtypes by scoring gene expression[93] across four independent datasets[44–46] using established neurodegenerative disease markers.[86] Microglia from C9-ALS and C9-FTLD subtypes consistently show reduced homeostatic microglial markers *CSF1R*, *CX3CR1*, and *P2RY12* and increased expression of disease-associated microglia (DAM) stage 1 (*TYROBP*, *APOE*, and *B2M*) and stage 2 (*SPP1*) markers (Figures S9F and S10A–S10C). This disease-associated cell state shift mirrors patterns observed in other neurodegenerative diseases.[85–87]

ALS-related risk gene expression across frontal cortex cell types

To investigate the cell-type-specific expression of ALS-related genes, we examined the expression of 30 genes associated with familial ALS or sporadic risk.[94,95] Hierarchical clustering identified microglia-enriched (modules 1–3), oligodendrocyte/OPC/astrocyte-enriched (modules 4–5), and neuron-enriched (modules 6–7) gene modules (Figure 4A). Cell-type-specific differential expression revealed neuron-enriched *CHCHD10*, *TUBA4A*, and *SOD1* upregulated in C9-ALS/FTLD and sALS. *SQSTM1* is upregulated in neurons and glia in C9-ALS/FTLD and C9-ALS, while *VCP* is specifically upregulated in neurons from C9-ALS/FTLD. Upregulation of the mitochondrion-associated genes *CHCHD10*, *TUBA4A*, *SQSTM1*, *SOD1*, and *VCP* in excitatory and inhibitory neurons from C9-ALS/FTLD emphasizes mitochondrial dysfunction (Figure 4B). *CCNF* upregulation is specific to upper-layer excitatory neurons from C9-ALS. *TARDBP* is downregulated in oligodendrocytes, astrocytes, and excitatory neurons in sALS. *C9orf72* is downregulated in oligodendrocytes from C9 cases, while *FUS* shows downregulation across all excitatory neuron subtypes. *SETX* is downregulated in upper-layer neurons in C9-ALS/FTLD. In sALS, *ANXA11* is downregulated in upper-layer excitatory neurons and vasoactive intestinal peptide-expressing inhibitory neurons. These findings highlight cell-type-specific alterations in ALS-related gene expression, providing insights into the molecular mechanisms driving ALS pathogenesis.

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

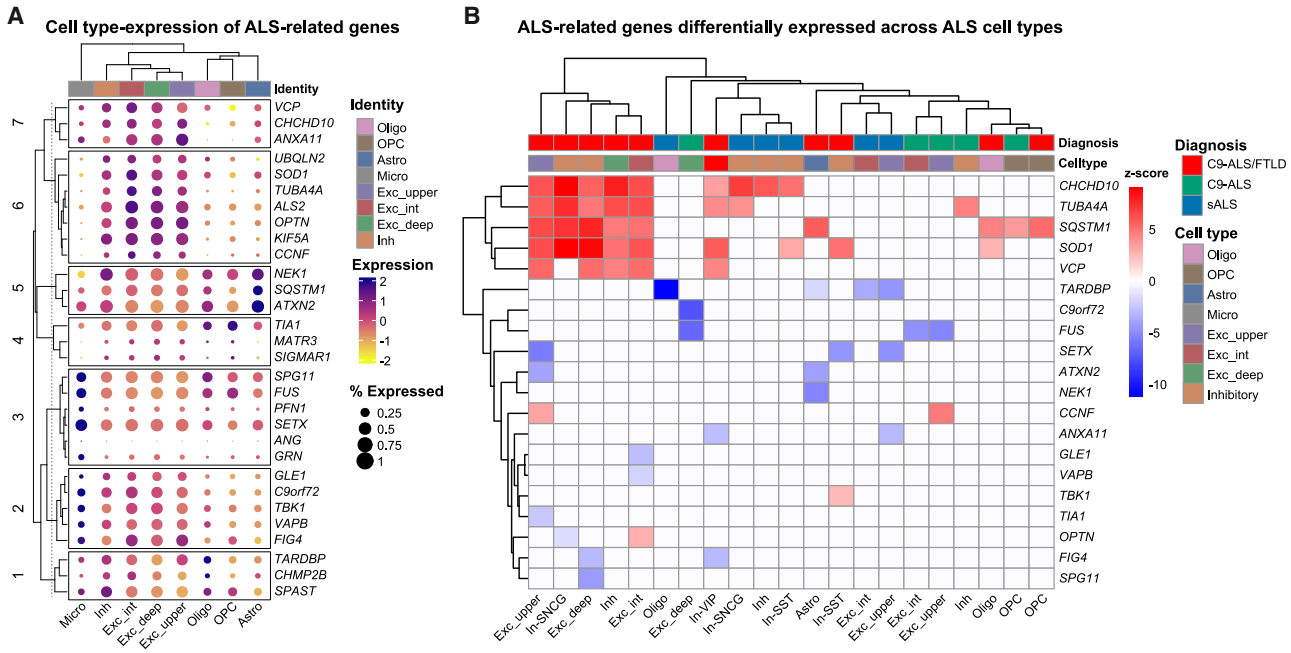


Figure 4. Distribution of differentially expressed ALS-related genes across orbitofrontal cell types in ALS

(A) Clustered heatmap dot plot analysis comparing average ALS-related gene expression levels (*y* axis) in cell types. The size of each dot corresponds to the percentage of cells of a given cell type expressing the corresponding gene of interest. Cell type identity is indicated on the upper *x* axis by color coding and lower *x* axis by label. Columns are clustered hierarchically, whereas rows are clustered and partitioned with k-means clustering (7 clusters).

(B) Clustered heatmap of ALS-related genes intersecting with differentially expressed genes across cell types and ALS subtypes. Scale is in *Z* scores, where the red scale indicates upregulation, the blue scale shows downregulation, and white shows that the gene was unchanged in a particular cell type for a given disease condition. Differential expression was performed using MAST (two sided) with the covariates and random effects described in STAR Methods, and multiple comparisons were corrected using the Benjamini-Hochberg method. Only genes meeting FDR < 0.05 and $|LFC| \geq 0.5$ are shown.

All data are based on 4 individuals with C9-ALS/FTLD, 2 with C9-ALS without FTLD, 4 with sALS, and 6 controls.

Taken together, these findings highlight converging and distinct alterations across ALS subtypes and cortical cell types, underscoring widespread disruptions in cellular pathways linked to mitochondrial function, protein homeostasis, and RNA metabolism.

Widespread and cell-type-specific dysregulation of APA across ALS subtypes

Given the widespread dysregulation of RNA metabolism genes across neuronal and glial populations in our results, we next sought to investigate whether APA, which has previously been shown to be altered in bulk tissue from ALS patients,[21,32,36] also exhibits cell-type-specific changes. We identified known polyadenylation (PA) sites in transcripts across our ALS subtypes and controls and performed differential PA analysis to compare ALS with controls (Table S7)[43,96] (STAR Methods). To mitigate false positives, we conducted an empirical FDR analysis using the control samples to establish cell-type-specific FDR thresholds (Table S8). These stringent criteria enabled us to filter out potential false positives across different cell types, ensuring an FDR of 10% or less for significant APA events. To further validate the APA findings and rule out potential confounding factors, we investigated the relationship between APA events and differentially expressed genes. We examined whether genes exhibiting APA, either lengthening, shortening, or intronic, also show significant changes in overall expression levels across the cell types analyzed. Our analysis reveals no significant correlation between differentially expressed genes and their APA events (Figures S11A–S11E), consistent with a recent study showing that mRNA abundance and 3′ UTR length are two independent measures of gene output.[97] This also suggests that APA events represent a distinct regulatory layer, potentially influencing disease-specific cellular functions without altering overall gene expression levels.

Our differential analysis revealed numerous significant PA differences between ALS subtypes and controls across all major cell types (Figures 5A and 5B). In both C9-ALS (with and without FTLD) and sALS, we observe a global trend toward distal PA site usage (Figures 5A and 5B). However, we also identified a considerable number of intronic and shortened genes, indicating that APA occurs in both directions. This pattern is evident across major cell types in both C9-ALS and sALS, but a higher proportion of intronic APA usage is detected in all neuron subtypes, oligodendrocytes, microglia, and astrocytes in sALS compared to C9-ALS (Figures 5A and 5B). Intronic APA usage is particularly prevalent in intermediate- and deep-layer excitatory neurons of sALS cases. The abundance of 3′ UTR shortened and intronic APAs in ALS compared to controls suggests a potential increase in truncated gene transcripts, which could significantly impact gene function. Conversely, the shift toward distal PA sites and longer 3′ UTR could affect function by altering mRNA metabolism, transport, and stability.[98,99]

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

CellPress
OPEN ACCESS

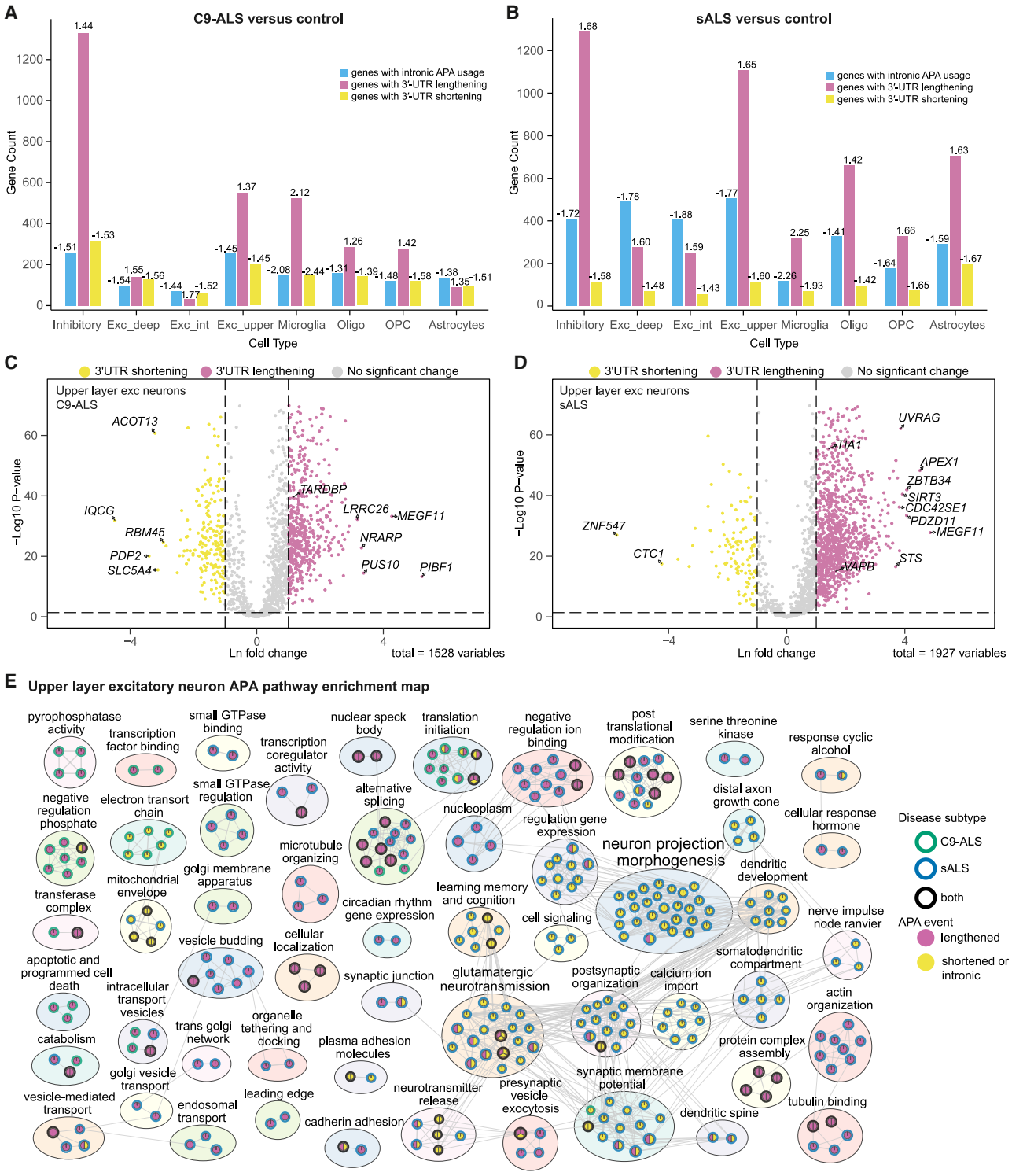


Figure 5. Dysregulation of the APA landscape in ALS compared to a control

(A) Barplot depicting the number of genes undergoing 3′ UTR lengthening, shortening, and intronic PA site usage in C9-ALS compared to a control across major cell types. n = 6 individuals (4 C9-ALS/FTLD and 2 C9-ALS without FTLD) vs. control (n = 6 individuals). Measurements were taken from distinct donor samples, with APA events quantified from pseudobulked nuclei per donor to ensure statistical independence. APA differences were tested using the model-based analysis of alternative polyadenylation using 3′ end-linked reads (MAAPER) method, with p values adjusted by the Benjamini-Hochberg procedure and further filtered

(legend continued on next page)

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

We observe cell-type-specific APA dysregulation involving key ALS genes in C9-ALS and sALS (Figures 5C, 5D, S12A–S12F, and S13A–S13H). Among excitatory neurons, upper-layer cells in C9-ALS display lengthening of *TARDBP*, whereas sALS cells show lengthening of *TIA1* and *VAPB*. In the intermediate layer, sALS is associated with lengthening of *TIA1*, *C9orf72*, and *TARDBP*. Deep-layer excitatory neurons in C9-ALS exhibit lengthening of *ALS2* along with shorter 3′ UTRs for *ANXA11* and *TARDBP*, while in sALS, deep-layer neurons show shorter 3′ UTRs for *UNC13A*, *TIA1*, and *ANXA11*. In inhibitory neurons, C9-ALS is marked by lengthening of *SPAST*, *ALS2*, *C9orf72*, and *TBK1*, accompanied by shorter 3′ UTRs for *ANXA11* and *VCP*. By contrast, sALS inhibitory neurons display lengthening of *SPAST*, *NEK1*, *SETX*, *ALS2*, *C9orf72*, and *VAPB* together with shortened *ANXA11*.

In glial populations, across oligodendrocytes, C9-ALS shows lengthening of *ALS2*, whereas sALS shows lengthening of *VCP*. In OPCs, sALS exhibits lengthening of *TIA1* and shorter *ATXN2*. Astrocytes in sALS displays lengthening of *TIA1*, *TARDBP*, and *C9orf72* but shows a shorter 3′ UTR for *ANXA11*. In C9-ALS, astrocytes express shorter *TARDBP*. Microglia in C9-ALS are characterized by lengthening of *ALS2*, *TIA1*, *TBK1*, and *SETX*, whereas sALS microglia show lengthening of *SQSTM1* and *TARDBP* alongside shorter *ANXA11*. Taken together, these findings highlight distinct APA profiles and emphasize layer- and cell type-dependent shifts, particularly for *TARDBP*.

To assess the biological significance of APA dysregulation in C9-ALS and sALS, we analyzed pathway enrichment across neuronal and glial cell types (Figures 5E, S14A–S14C, and S15A–S15D; Table S9). Our analysis reveals distinct APA signatures between C9-ALS and sALS subtypes, where APA events across cell types primarily affect a greater number of pathways in sALS than in C9-ALS. Both conditions share disruptions in mRNA metabolism, splicing, and stress response pathways through 3′ UTR lengthening in neurons, oligodendrocytes, and microglia.

Excitatory neurons show pronounced layer-specific effects, with upper-layer excitatory neurons showing intronic APAs in glutamatergic signaling pathways (Figure 5E). Strikingly, sALS, but not C9-ALS, shows a widespread switch to intronic APA usage in intermediate- and deep-layer neurons, affecting dendritic structure, synaptic organization, and calcium transport (Figures S14A and S14B). These intronic APA changes encompass known disease-associated genes such as *UNC13A*.[20,60,100] In addition, sALS shows lengthening in AS pathways, while C9-ALS exhibits this pattern in deep-layer neurons.

Specific disruptions in inhibitory neurons from C9-ALS include 3′ UTR shortening in mitochondrial pathways, capturing changes also identified from the differential expression analysis (Figure S14C). In microglia, specific changes in C9-ALS include 3′ UTR lengthening in myeloid differentiation and endosomal membrane response, whereas transcriptional machinery and chromatin organization are impacted in sALS (Figure S15D). Moreover, sALS exhibits extensive intronic APA events related to neuron projection and synapses in OPCs and astrocytes (Figures S15B and S15C), revealing pathway changes not detected by differential gene expression analysis.

Collectively, this analysis reveals distinct APA signatures between two ALS subtypes, most notably the prevalence of intronic APAs in sALS neurons and glia when compared to C9-ALS. While certain pathways are affected across multiple cell types, the cell-type- and disease-specific patterns of APA regulation suggest diverse post-transcriptional mechanisms contributing to ALS pathogenesis. These APA-mediated changes often converge with pathways identified through differential gene expression analysis, highlighting the multi-modal nature of gene regulation in ALS.

APA-Net reveals mechanistic insights into APA dysregulation in ALS

To decode the grammar of APA dysregulation in our ALS cohorts, we developed a deep learning model called APA-Net. To comprehensively examine APA changes, we performed differential analysis on all PA pairs of transcripts that showed significant transcript-wise APA compared to controls (see STAR Methods, Table S7). We observe thousands of significant APA changes across cell types (Figure S16). Comparison of APA events between C9-ALS and sALS identified that 12.6%–30.8% of events are shared between both disease subtypes (Figure S17A). These shared events demonstrate remarkably high Pearson correlation coefficients, ranging from 0.65 ($p < 5.3 \times 10^{-159}$) to 0.87 ($p < 7.2 \times 10^{-176}$) across all cell types, with particularly strong correlations observed in neurons (Figure S17B). While most APA events are unique to either C9-ALS or sALS, suggesting distinct regulatory mechanisms in each subtype, these events maintain a positive, though weaker, correlation between the subtypes.

APA-Net is trained to predict the APA log-fold change (LFC) (Figure S16) using the RNA sequences surrounding proximal and distal PA sites for each APA event, along with the RBP expression profiles per cell type (Figure 6A) as input. To assess APA-Net performance, we compared the predicted to the observed APA LFC values using Pearson correlation across

using an empirical FDR $\leq$ 10% from control-control comparisons. Effect sizes are shown as median relative expression difference (RED) for 3′-UTR size difference (REDu, considering 3′-most exon PASs only) and RED for internal PASs (REDi) on the bars.

(B) Same as (A) for sALS (n = 4 individuals) vs. control (n = 6 individuals).

(C) Volcano plot for upper-layer excitatory neurons in C9-ALS for genes undergoing 3′ UTR lengthening and shortening. Only genes that passed the 10% empirical FDR threshold for transcript-wise APA usage are shown (STAR Methods). Vertical dashed lines indicate $\log_2$ fold change thresholds of -1 and 1; a horizontal dashed line represents an adjusted $p < 0.05$. Tests are two sided; multiple comparisons were corrected (Benjamini-Yekutieli where applicable). Effect sizes are APA $\log_2$ fold changes.

(D) Same as (C) for sALS vs. control.

(E) Enrichment map showcasing the pathways enriched for genes undergoing lengthening and shortening in C9-ALS and sALS for upper-layer excitatory neurons. The node border indicates ALS subtype; fill shows whether the pathway is enriched for transcripts with a lengthened 3′ UTR, shortened 3′ UTR, and/or intronic APA. Pathway enrichment was tested using g:Profiler2 (right tailed, one-sided hypergeometric) with $q < 0.1$ after Benjamini-Hochberg correction; the effect size is the proportion of pathway genes meeting the APA criterion. Annotation text size for clusters is scaled by the number of nodes within each cluster.

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

CellPress
OPEN ACCESS

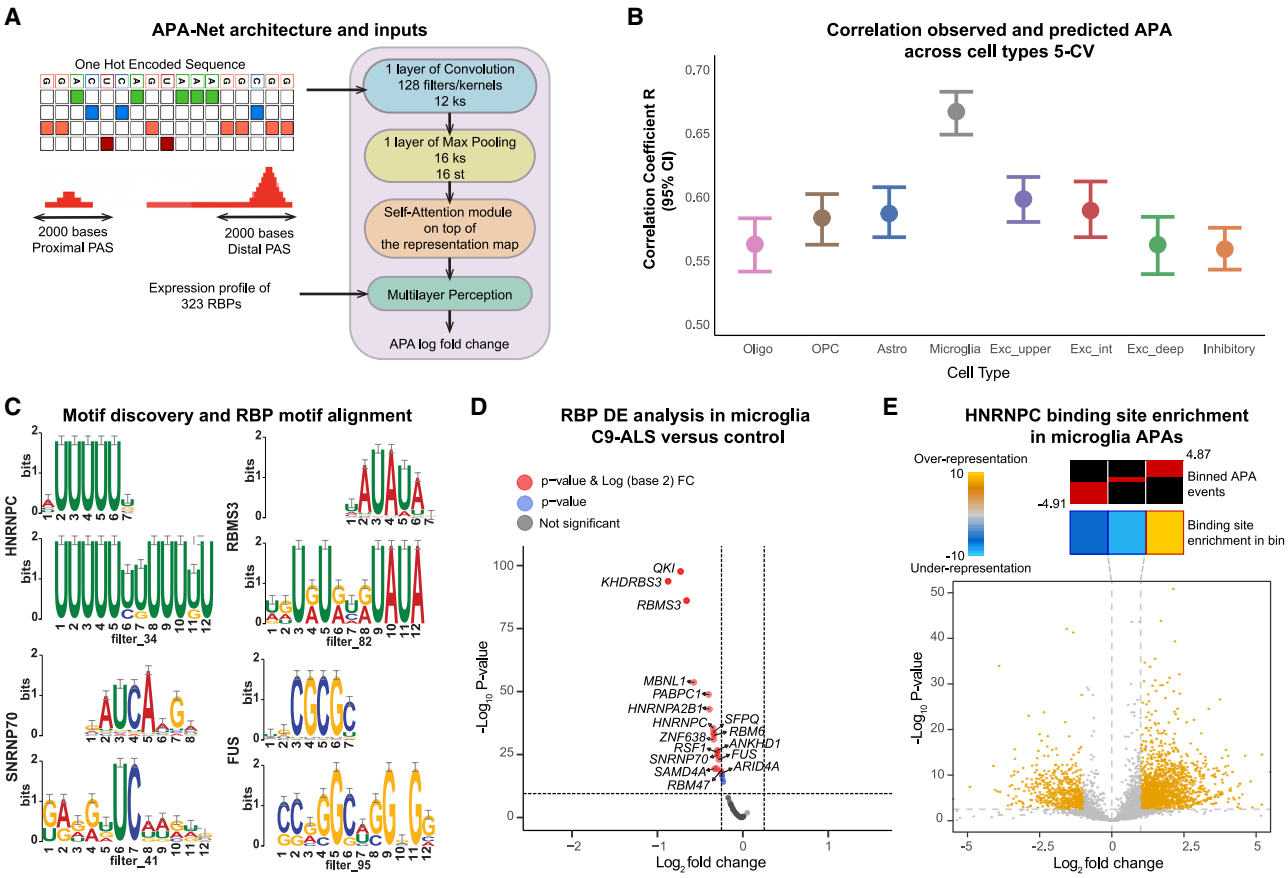


Figure 6. Development and analysis of a deep learning model to unravel the grammar of APA regulation in C9-ALS and sALS cases

(A) Schematic of the APA-Net architecture and inputs. The model uses two inputs: sequences surrounding the PA sites and the expression profiles of 323 RBPs.

(B) Performance of APA-Net in predicting APA events across cell types using 5-fold cross-validation. The plot shows Pearson's correlation coefficient (R) between observed and predicted APA LFCs with 95% confidence intervals.

(C) APA-Net filters interpreted as motifs, which were subsequently aligned to an RBP database to identify corresponding RBPs.

(D) Differential expression analysis performed on the identified RBPs in microglia from C9-ALS vs. control. n = 6 vs. n = 6 individuals. Measurements are from distinct donor samples. Analysis used the MAST mixed-effect hurdle model (two sided) with Benjamini-Hochberg correction. Effect sizes are model $\log_2$ fold changes.

(E) Enrichment analysis of HNRNPC binding sites for APA events in microglia. Top: APA sequences are divided into equally populated bins based on their LFC values. Red shows the proportion of APA events in each bin. Middle: Enrichment score indicating under-representation and over-representation of the binding sites. Hypergeometric tests (right tailed) were used, with multiple-comparison correction (Benjamini-Hochberg). Bins with significant enrichment (corrected $p <$ 0.05, red) or depletion (blue) of poly(U) motifs are denoted with a bolded border. Bottom: volcano plot showing the distribution of changes in APA LFC in microglia from C9-ALS compared to controls. Effect sizes are APA $\log_2$ fold changes. Significant observations are highlighted in orange.

5-fold cross-validation. APA-Net achieves a strong Pearson correlation coefficient across the entire dataset (Figure 6B). Thus, APA-Net robustly learns cell-type-specific APA profiles across disease subtypes, demonstrating its potential as a powerful tool for studying APA in complex diseases like ALS.

Using a convolutional neural network (CNN) architecture augmented with multi-head attention (MAT), our model is designed to identify *cis*-regulatory elements that influence PA site selection across different cell types. We optimized the model architecture,[101,102] including kernel size and max pooling steps, to capture relevant genomic motifs (Figure 6A). This hybrid CNN-MAT architecture enhances the model's ability to capture both local and long-range dependencies in sequence data. This is important for accurately identifying *cis*-regulatory elements influencing APA, particularly those involved in binding of RBPs other than the core PA (CPA) machinery.

To identify potential *cis*-regulatory elements involved in APA, we used the filter weights from the CNN module of APA-Net, which represent learned sequence motifs (Figure S18; STAR Methods). We scanned every sequence within the test dataset using the model's filters to identify regional subsequences where the filters showed high activation or response. Next, we aligned RBPs from the compendium of RNA-binding motifs[103] with the motifs found by APA-Net (Figure 6C; Table S10). We observed several ALS-related genes, such as *FUS*,[104] *TARDBP*,[105] *FXR1*,[106] and *G3BP2*,[107] as well as numerous APA and AS factors, such as *HNRNPC*, *SNRNP70*, *SFPQ*, *MBNL1*, and *SRSF7*,[108–112] among the aligned RBPs (Table S10).

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

Differential expression analysis[113] reveals that many of these RBPs, along with others, are significantly dysregulated across major cell types in both C9-ALS and sALS compared to controls (Figures S19 and S20). *HNRNPC* shows significant downregulation in C9-ALS microglia (Figure 6D), where APA-Net has its strongest predictive performance, though it also remains significantly predictive across other cell types. To investigate the functional implications of this finding, we examined HNRNPC binding patterns using in-house enhanced crosslinking and immunoprecipitation sequencing data generated from HEK293T cells (STAR Methods). HNRNPC is known to bind regions surrounding proximal PA sites, and its knockdown can promote distal PA site usage.[114–116] Based on its downregulation in C9-ALS microglia (Figure 6D), we hypothesized that HNRNPC binding sites would be enriched in proximal sites of the APA events showing lengthening. We analyzed whether HNRNPC binding sites are enriched in microglia APA events. We observe a significant enrichment of binding sites for the lengthened APA events ($p < 1 \times 10^{-4}$) and a lower enrichment for the shortened APA events (Figure 6E), which supports a regulatory role of HNRNPC in APA in ALS.

Overall, our findings highlight the value of APA-Net's interpretability, enabling us to generate mechanistic hypotheses about APA dysregulation in ALS.

RBP interactions reveal cell-type-specific mechanisms and dysregulation of APA in ALS

Leveraging APA-Net's interpretability, we examined RBPs potentially involved in APA dysregulation in ALS. To this end, we analyzed the model's convolutional filter activation patterns by scanning input sequences to identify co-occurring sequence motifs, which were subsequently mapped to specific RBPs. This allowed us to generate RBP activation profiles across sequences, which we used to calculate pairwise correlations between RBPs based on their co-activation patterns. These patterns provide an interpretable proxy for motif-level co-occurrence learned by the model. These correlation matrices were then clustered to identify regulatory modules, defined as groups of RBPs with coordinated activation patterns, within microglia, neurons, and other glial cell types (Figures 7 and S21–S27). These modules suggest how combinations of RBPs may cooperatively regulate APA in a cell type-specific manner. Furthermore, these profiles revealed distinct patterns of RBP interactions across cell types (Figures S28A and S28B). Notably, the dissimilarity heatmap (Figure S28A) highlights cell-type-specific regulatory networks that are not captured by RBP expression patterns alone. While many identified modules contain RBPs from the same family or with similar binding preferences, a substantial subset of interacting RBP pairs shows little to no sequence motif similarity Figure S29, suggesting that functionally coordinated RBPs may regulate APA through distinct sequence motifs.

In microglia, module 2 (Figure 7) highlights a diverse array of RBPs involved in various aspects of RNA processing and regulation. This module includes RBPs that regulate AS, such as ESRP2,[117] U2AF, and several members of the CUGBP Elav-like family (CELF3, CELF4, CELF5, and CELF6).[118] Additionally, module 2 features members of the CPEB family, specifically CPEB2 and CPEB4, which bind to specific sites within the 3′ UTR region and regulate mRNA transport and metabolism in the cytoplasm.[119,120] Dysregulation of CPEBs has been documented in various neuronal pathologies[120] and is correlated with APA changes in Huntington's disease.[121] This module also encompasses single-stranded RBPs, including A1CF, RBMS1, and MSI1, as well as HNRNPC, which is involved in maintaining short 3′ UTRs, and SFPQ, which plays a role in transcript shortening by activating cryptic last exons in ALS.[122] This clustering of diverse RBPs with roles in splicing, PA, and 3′ UTR processing suggests a coordinated regulation of post-transcriptional events in microglia that are relevant to ALS.

Module 3 includes RBPs involved in RNA processing and stress response, such as QKI,[123] RBOFX1,[124] and YBX1/YBX2,[125] which are important in stress granule formation. It also features RBPs with K-homologous (KH) domains, including PCBP1, PCBP2, KHDRBS3, and HNRNPK, which bind to poly(C) regions and regulate AS, stability, and translation of transcripts.[126,127]

Module 4 is characterized by RBPs involved in mRNA processing and export, such as RBM8A and RBM4,[128,129] which are also involved in AS and stress granule localization. This module includes hnRNPs such as HNRNPA2B1, HNRNPH2, and FUS, which play roles in pre-mRNA splicing and mRNA export.[15,20,130,131] SAMD4A, which is associated with the coordinated regulation of transcription, AS, and APA,[132] is also part of this module. The composition of module 4 suggests an interplay between splicing regulation, mRNA export, and post-transcriptional processing, which may contribute to microglial alterations in ALS.

Module 5 includes a diverse array of RBPs involved in RNA processing, splicing, and stress response. This module features KHDRBS1 and KHDRBS2, which are involved in AS and 3′ UTR formation, and members of the PABC family (PABPC1, PABPC4, and PABPC5) known for their role in binding to the 3′ poly(A) tails of eukaryotic mRNAs. The splicing regulators SART3, RBM5, and SNRNP70, as well as several members of the SR-rich family (SRSF1, SRSF7, SRSF9, and SRSF10) are also included, with SRSF7 implicated in transcript shortening.[111] Module 5 includes FXR1 and FXR2, RNA-processing proteins from the FXP family that regulate stress responses and have shown potential links to ALS pathogenesis.[106] The co-occurrence of ALS-linked RBP motifs alongside additional splicing and PA factors in this module suggests a complex interplay between these processes in ALS.

We next investigated the expression of RBPs known to be the CPA factors regulating APA[133] alongside the RBPs identified by APA-Net. Across cell types, we observe eight distinct expression patterns for RBPs identified by APA-Net (Figure S28B) and five expression patterns for CPA factors (Figure S30A). Additionally, we detect cell-type-specific downregulation of several CPA factors in ALS subtypes compared to a control (Figure S30B). This cell type specificity in RBP interactions and expression profiles hints at the presence of distinct *cis*- and *trans*-regulatory elements that may influence APA in ALS. Collectively, our findings begin to map cell-type-specific APA regulation in ALS and elucidate the complex molecular landscape underlying disease.

DISCUSSION

In this study, we present the first snRNA-seq atlas of the orbitofrontal cortex in ALS, a region associated with behavioral

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

CellPress
OPEN ACCESS

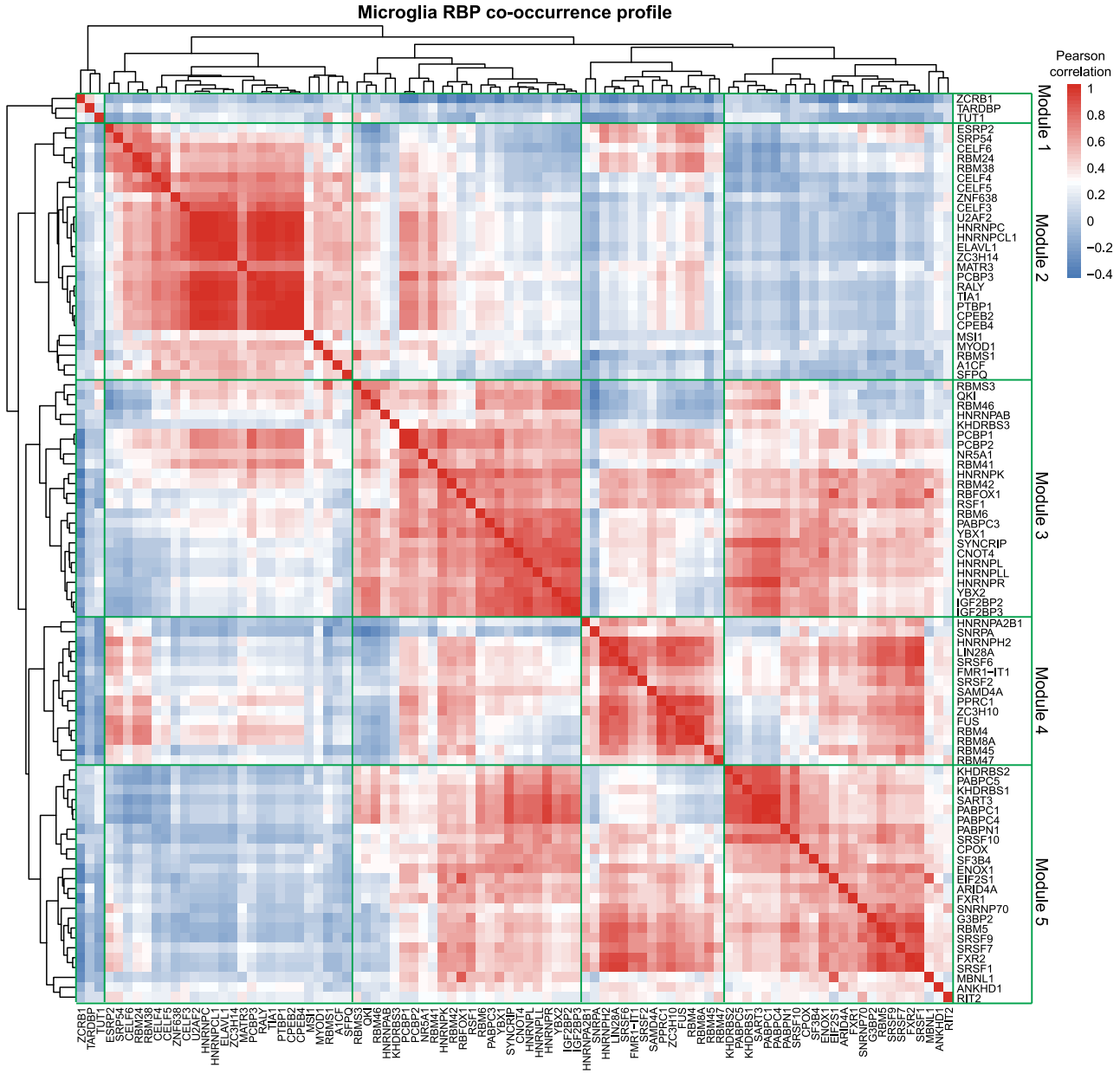


Figure 7. RBP interactions in microglia reveal cell-type-specific mechanisms and dysregulation of APA in C9-ALS and sALS

Clustered RBP motif co-occurrence profile for microglia. The heatmap is computed using Pearson correlation of co-occurrence of RBP-aligned filters in microglia APA events. Five RBP modules are defined based on hierarchical clustering (labeled on the *y* and *x* axes). Effect size is the Pearson correlation coefficient. Data are based on 4 individuals with C9-ALS/FTLD, 2 with C9-ALS without FTLD, 4 with sALS, and 6 controls.

impairments along the ALS-FTLD spectrum.[134–140] We compared the orbitofrontal cortex with published dorsolateral prefrontal and primary motor cortex datasets.[44–47] This analysis provides a resource of concordant cell-type-specific gene expression changes across ALS subtypes and affected brain regions. We characterized cell-type-specific APA dysregulation in ALS subtypes compared to neurologically healthy controls. To further understand this dysregulation, we developed APA-Net, an interpretable deep learning model that integrates transcript sequences and RBP expression profiles from snRNA-seq data to predict APA changes. Our findings underscore that C9-ALS (with and without FTLD) and sALS drive distinct cellular states marked by unique signatures in gene expression and APA profiles.

Molecular dysregulation in the orbitofrontal cortex reveals consistent patterns across the ALS-FTLD spectrum. Protein homeostasis pathways, particularly chaperone-mediated folding, show enrichment across cell types in C9-ALS/FTLD, while chromatin remodeling pathways show widespread downregulation across multiple cell populations, including upper-layer excitatory

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

neurons, inhibitory neurons, oligodendrocytes, and astrocytes. Cell-type-specific changes in energy metabolism emerge through upregulation of electron transport chain components in neurons, glucose metabolism in microglia, and ATP metabolism in OPCs, aligning with established patterns of mitochondrial dysfunction in ALS.[37,141–143] Our analysis identified widespread dysregulation of ribosomal complexes across upper-layer, intermediate-layer, and inhibitory neuronal populations, most prominently in C9-ALS/FTLD, with a subset of these changes also present in sALS.

Independent datasets strongly corroborate our molecular findings, particularly in C9-ALS/FTLD upper-layer excitatory neurons, where 115 of 138 genes maintain their expression patterns, highlighting disruptions in protein homeostasis, mitochondrial function, and ribosomal processes.[44,45,47] Both excitatory and inhibitory neurons demonstrate consistent upregulation of ribosomal subunits, *VCP*, and *SQSTM1* across cortical regions, with inhibitory neurons specifically showing increased expression of *RGS4*, *MDH1*, and *SLC38A2*, along with similarly corroborated changes in *HSP90AA1*, *STMN2*, and *NEFL*.[45] Our analysis reveals increased transcription of both *STMN2* and *NEFL* across neuronal populations. The elevated *STMN2* expression, in conjunction with its premature PA and cryptic splicing, and loss of function with TDP-43 deficiency,[28,60–62] potentially reflects cellular compensation for compromised post-transcriptional processing. Similarly, *NEFL* upregulation occurs while its protein product, NfL, serves as a validated marker of axonal degeneration in patient biofluids,[144] suggesting a coordinated neuronal response to maintain axonal integrity. Further corroborating changes include C9-ALS neurons exhibiting changes in actin dynamics (*CFL1*, *RAC1*, and *YWHAH*)[12] and autophagy (*HSP90AA1*, *HSP90AB1*, and *DNAJB6*),[45] while sALS displays hallmark potassium channel dysregulation.[57,58,145] The consistency of these alterations across brain regions indicates that they represent fundamental disease processes rather than region-specific responses.

In glial populations, astrocytes demonstrate pronounced changes in C9-ALS/FTLD, including upregulation of cholesterol and miRNA metabolism pathways alongside reactive markers (*GFAP* and *CHI3L1*),[44,45] while sALS astrocytes show *TARDBP* downregulation.[80] Common to astrocytes across ALS subtypes is the dysregulation of protein homeostasis and actin filament organization.[146] In oligodendrocytes and OPCs, we observed consistent alterations across ALS subtypes. In sALS across independent datasets,[45] oligodendrocytes show downregulation of *PLLP*, while OPCs specifically demonstrate reduced *APOD* expression. Given the role of APOD in remyelination,[73] its transcript reduction in OPCs may significantly impair the regenerative capacity of myelin in sALS. This finding aligns with broader evidence of compromised myelin maintenance across ALS subtypes, as demonstrated by the downregulation of the myelination scaffold protein *SEPTIN4* across sALS and C9-ALS/FTLD datasets. In C9-ALS oligodendrocytes specifically, we found concordant upregulation of autophagy-related genes (*ATG4B* and *HSP90AB1*), suggesting that protein homeostasis disruption could exacerbate myelin maintenance deficits in this C9 cases.

The analysis of microglia reveals shared and unique cell-type-specific changes across ALS subtypes. These include increased cytokine signaling pathways and elevated glucose metabolism, suggesting a microglial state with heightened energy demands. We find consistent patterns of depleted homeostatic markers and increased DAM stage markers (*APOE*, *TYROBP*, and *SPP1*) across both C9-ALS and sALS.[44–47] C9-ALS microglia specifically show impaired neuronal surveillance, marked by loss of *CX3CR1* and *P2RY12*,[86,147] along with increased inflammatory and stress responses. Both sALS and C9-ALS exhibit altered genes related to JAK/STAT signaling with depleted *IL6R*, *JAK1*, and *JAK2* but elevated *JAK3* and *STAT3*, supporting the therapeutic potential of JAK/STAT targeting.[90,148,149] These microglial signatures reveal both unique and shared features with other neurodegenerative conditions while maintaining strong concordance across existing ALS datasets.

Our APA analysis revealed both shared and subtype-specific patterns across cell types. Differential gene expression and APA changes were largely uncorrelated across all cell types and ALS subtypes. In microglia, C9-ALS showed lengthening of transcripts involved in myeloid differentiation, mitotic, and focal adhesion pathways,[86,150] whereas sALS displayed similar changes in pathways related to chromatin organization and DNA-templated transcription. Both subtypes exhibited lengthening in transcripts associated with nucleocytoplasmic transport,[9] transcription factor binding, and stress response pathways.

Neurons display subtype-specific dysregulation of both gene expression and APA in key pathways, including mitochondrial function, stress response, synaptic signaling, and chromatin remodeling. sALS excitatory neurons show distinctive APA dysregulation in postsynaptic glutamatergic neurotransmission and calcium signaling. All neuronal subtypes exhibit downregulation of ALS-relevant RBPs (*TARDBP*, *FUS*, *ATXN2*, *SETX*, and *ANXA11*) and lengthening of AS transcripts, highlighting mRNA processing as a central mechanism in neuronal dysfunction.[8,148]

Microglia highlight the distinct yet intersecting roles of APA and gene expression in regulatory mechanisms. In C9-ALS, both processes converge on pathways such as GTPase signaling and mRNA processing but through distinct sets of genes. Conversely, sALS microglia exhibit minimal gene expression changes beyond RBP dysregulation while showing substantial APA alterations, uncovering an unrecognized layer of transcriptional complexity in sporadic disease.[97] This integrated analysis of gene expression and APA reveals regulatory intricacies in ALS that traditional approaches would overlook.

To further investigate cell-type-specific APA, we developed APA-Net. While significant progress has been made in deep learning approaches for APA prediction,[151–155] APA-Net is uniquely trained on cell-type-resolved APA profiles derived from snRNA-seq data. Using APA-Net, we identified *cis*-regulatory elements and cell-type-specific RBP interactions directly from model predictions. This analysis revealed RBPs and their potential cooperative interactions involved in APA regulation across ALS subtypes. It is important to note that, while these RBP interaction profiles are based on pairwise correlations of filter co-occurrence, they do not imply direct physical interactions between RBPs. Instead, they reflect shared sequence-level features learned by APA-Net and may capture higher-order or indirect associations, including context-specific or adapter-mediated interactions. This approach not only enhances our

understanding of APA regulation but also uncovers the complex interplay of RBPs in disease contexts, providing a valuable framework for exploring post-transcriptional regulation in ALS and other neurodegenerative diseases. Looking ahead, incorporating additional data modalities, such as transcription factor activity and chromatin accessibility profiles, should further boost the predictive power of models like APA-Net and refine our grasp of APA regulation in disease.

Our single-nucleus transcriptomic profiling of the orbitofrontal cortex provides a granular understanding of cellular heterogeneity in ALS and FTLD, complementing existing single-cell ALS mapping efforts.[44–47] Through evaluation of the orbitofrontal cortex, we reveal concordant and divergent ALS-associated changes across cell types, brain regions, and disease subtypes. Moreover, APA-Net's ability to predict cell-type-specific APA events can be applied to other single-cell transcriptomic atlases. Overall, our findings illuminate the complex, cell-type-specific pathomechanisms of ALS, offering a valuable resource for advancing therapeutic research in ALS and other neurodegenerative diseases.

Limitations of the study

Our study provides a foundational molecular characterization of ALS/FTLD, though certain limitations require mention. Control and ALS/FTLD tissues were sourced from different brain banks (Douglas-Bell Canada Brain Bank/University Health Network Neurodegenerative Brain Collection and Tanz Centre for Research in Neurodegenerative Diseases, respectively). Although all samples were flash frozen and stored at −80°C to minimize postmortem variability, premortem clinical factors such as hypoxia, nutritional status, or respiratory support may have influenced gene expression profiles. Richer clinical annotations and prospectively collected tissue will be essential in future work to disentangle these confounding effects, including cell-type-specific variability related to age in the orbitofrontal cortex.[156]

Although neuronal loss is a hallmark of ALS and FTLD, our study was not designed to reliably quantify changes in cell type abundance. Variability in nucleus recovery and the size of the disease cohorts, together with disease stage-specific differences in cellular composition,[157,158] and known snRNA-seq biases against rare or fragile populations[159] precluded formal analysis of compositional changes between disease conditions.

Independent validation is critical, with spatial transcriptomics and protein-level studies as prime approaches. Moreover, our pseudobulk 3′ end analysis could be refined by long-read sequencing and QuantSeq REV on sorted nuclei from ALS patients, with or without FTLD, compared with controls.[160]

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Janice Robertson (jan.robertson@utoronto.ca).

Materials availability

No new, unique reagents were generated in this study.

Data and code availability

Raw FASTQ snRNA-seq files have been deposited in the NCBI Sequence Read Archive Database: PRJNA918304. The exact version of the analysis code used in this study is archived on Zenodo at https://doi.org/10.5281/zenodo.16879041 (release v.1.0.0). The actively maintained codebase is available at https://github.com/BaderLab/APA-Net.

ACKNOWLEDGMENTS

The authors are thankful to the ALS patients and families for their tissue donations in this study, as well as the Douglas-Bell Canada Brain Bank (McGill University). This research was enabled in part by access to high-performance computing clusters provided through the Digital Research Alliance of Canada (https://alliancecan.ca). This work was funded by the James Hunter Family Initiative in ALS Research, a project grant from the ALS Society of Canada, and an ERA-LEARN (E-Rare-3/Canadian Institute for Health Research) grant (063-REPETOMICS). P.M.M. was supported by the ALS Association Milton Safenowitz Postdoctoral Fellowship and currently holds the Christopher Chiu Fellowship from ALS Double Play. J.R. is the James Hunter Family Chair in ALS Research. E.R. is supported in part by the Canadian Consortium on Neurodegeneration in Aging. This work was supported by NRNB (US National Institutes of Health, National Center for Research Resources grant P41 GM103504).

AUTHOR CONTRIBUTIONS

The project was conceived and planned by P.M.M., A.M.S., G.D.B., and J.R. Autopsies on ALS/FTLD tissue were performed by P.M.M., P.M., and S.X. Clinical care and diagnosis of ALS/FTLD patients were provided by L.Z. Neuropathological diagnosis was performed by J.K., and screening for phospho-TDP-43 was conducted by P.M.M. Genetic analysis was conducted by C.S., D.M., and E.R. Autopsies on two control tissues were performed by N.V. and G.G.K. Single-nucleus preparations and 10× library creation were performed by Z.X. under the supervision of T.K. All analyses were conducted by P.M.M., A.M.S., and R.S. in consultation with H.G. and G.D.B. P.M.M. and A.M.S. generated figures and wrote the manuscript. G.D.B. and J.R. provided primary comments and edits.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
  - Human brain samples and QC
  - Single nucleus RNA-seq by fluorescent activated cell sorting (FACS)
  - snRNA-seq sample processing, QC, and clustering
  - Cluster annotation by machine learning and reference atlas
  - PCA and sample-level distance analysis
  - Differential gene expression analysis
  - Comparison with existing ALS datasets
  - APA quantification and profiling
  - Empirical FDR analysis for APA
  - Pathway enrichment analysis
  - Deep learning model architecture
  - Deep learning model interpretation
  - HNRNPC eCLIP-seq processing and enrichment analysis
  - RBP interaction dissimilarity across cell types

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xgen.2025.101007.

Received: January 13, 2025
Revised: April 28, 2025
Accepted: August 15, 2025

REFERENCES

1. Strong, M.J., Abrahams, S., Goldstein, L.H., Woolley, S., Mclaughlin, P., Snowden, J., Mioshi, E., Roberts-South, A., Benatar, M., HortobáGyi, T., et al. (2017). Amyotrophic lateral sclerosis-frontotemporal spectrum disorder (als-ftsd): Revised diagnostic criteria. Amyotroph. Lateral Scler. Frontotemporal Degener. *18*, 153–174.
2. Benbrika, S., Doidy, F., Carluer, L., Mondou, A., Pélerin, A., Eustache, F., Viader, F., and Desgranges, B. (2021). Longitudinal study of cognitive and emotional alterations in amyotrophic lateral sclerosis: clinical and imaging data. Front. Neurol. *12*, 620198.
3. Zhang, Y., Qiu, T., Yuan, X., Zhang, J., Wang, Y., Zhang, N., Zhou, C., Luo, C., and Zhang, J. (2019). Abnormal topological organization of structural covariance networks in amyotrophic lateral sclerosis. Neuroimage Clin. *21*, 101619.
4. Prado, L.d.G.R., Bicalho, I.C.S., Magalhães, D., Caramelli, P., Teixeira, A. L., and Souza, L.C.d. (2015). C9orf72 and the ftd-als spectrum: a systematic review of neuroimaging studies. Dement. Neuropsychol. *9*, 413–421.
5. Akçimen, F., Lopez, E.R., Landers, J.E., Nath, A., Chiò, A., Chia, R., and Traynor, B.J. (2023). Amyotrophic lateral sclerosis: translating genetic discoveries into therapies. Nat. Rev. Genet. *24*, 642–658.
6. Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Laaksovirta, H., van Swieten, J.C., Myllykangas, L., et al. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron *72*, 257–268.
7. DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron *72*, 245–256.
8. Ling, S.-C., Polymenidou, M., and Cleveland, D.W. (2013). Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. Neuron *79*, 416–438.
9. McGoldrick, P., and Robertson, J. (2023). Unraveling the impact of disrupted nucleocytoplasmic transport systems in c9orf72-associated ALS. Front. Cell. Neurosci. *17*, 1247297.
10. Chua, J.P., De Calbiac, H., Kabashi, E., and Barmada, S.J. (2022). Autophagy and als: mechanistic insights and therapeutic implications. Autophagy *18*, 254–282.
11. Pang, W., and Hu, F. (2021). Cellular and physiological functions of c9orf72 and implications for als/ftd. J. Neurochem. *157*, 334–350.
12. Sivadasan, R., Hornburg, D., Drepper, C., Frank, N., Jablonka, S., Hansel, A., Lojewski, X., Sterneckert, J., Hermann, A., Shaw, P.J., et al. (2016). C9orf72 interaction with cofilin modulates actin dynamics in motor neurons. Nat. Neurosci. *19*, 1610–1618.
13. Neumann, M., Sampathu, D.M., Kwong, L.K., Truax, A.C., Micsenyi, M. C., Chou, T.T., Bruce, J., Schuck, T., Grossman, M., Clark, C.M., et al. (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science *314*, 130–133.
14. Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., Mann, D., Tsuchiya, K., Yoshida, M., Hashizume, Y., and Oda, T. (2006). Tdp-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem. Biophys. Res. Commun. *351*, 602–611.
15. Lagier-Tourenne, C., Polymenidou, M., Hutt, K.R., Vu, A.Q., Baughn, M., Huelga, S.C., Clutario, K.M., Ling, S.-C., Liang, T.Y., Mazur, C., et al. (2012). Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. Nat. Neurosci. *15*, 1488–1497.
16. Luisier, R., Tyzack, G.E., Hall, C.E., Mitchell, J.S., Devine, H., Taha, D.M., Malik, B., Meyer, I., Greensmith, L., Newcombe, J., et al. (2018). Intron retention and nuclear loss of SFPQ are molecular hallmarks of ALS. Nat. Commun. *9*, 2010.
17. Taylor, R., Hamid, F., Fielding, T., Gordon, P.M., Maloney, M., Makeyev, E.V., and Houart, C. (2022). Prematurely terminated intron-retaining mRNAs invade axons in SFPQ null-driven neurodegeneration and are a hallmark of ALS. Nat. Commun. *13*, 6994.
18. Mackenzie, I.R., Nicholson, A.M., Sarkar, M., Messing, J., Purice, M.D., Pottier, C., Annu, K., Baker, M., Perkerson, R.B., Kurti, A., et al. (2017). TIA1 mutations in amyotrophic lateral sclerosis and frontotemporal dementia promote phase separation and alter stress granule dynamics. Neuron *95*, 808–816.e9.
19. Humphrey, J., Emmett, W., Fratta, P., Isaacs, A.M., and Plagnol, V. (2017). Quantitative analysis of cryptic splicing associated with TDP-43 depletion. BMC Med. Genomics *10*, 38.
20. Koike, Y., Pickles, S., Estades Ayuso, V., Jansen-West, K., Qi, Y.A., Li, Z., Daughrity, L.M., Yue, M., Zhang, Y.-J., Cook, C.N., et al. (2023). Tdp-43 and other hnrnps regulate cryptic exon inclusion of a key als/ftd risk gene, unc13a. PLoS Biol. *21*, e3002028.
21. Liu, E.Y., Russ, J., Cali, C.P., Phan, J.M., Amlie-Wolf, A., and Lee, E.B. (2019). Loss of nuclear TDP-43 is associated with decondensation of LINE retrotransposons. Cell Rep. *27*, 1409–1421.e6.
22. Tollervey, J.R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A.L., Župunski, V., et al. (2011). Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nat. Neurosci. *14*, 452–458.
23. Amlie-Wolf, A., Ryvkin, P., Tong, R., Dragomir, I., Suh, E., Xu, Y., Van Deerlin, V.M., Gregory, B.D., Kwong, L.K., Trojanowski, J.Q., et al. (2015). Transcriptomic changes due to cytoplasmic TDP-43 expression reveal dysregulation of histone transcripts and nuclear chromatin. PLoS One *10*, e0141836.
24. Xiao, S., Sanelli, T., Dib, S., Sheps, D., Findlater, J., Bilbao, J., Keith, J., Zinman, L., Rogaeva, E., and Robertson, J. (2011). RNA targets of TDP-43 identified by UV-CLIP are deregulated in ALS. Mol. Cell. Neurosci. *47*, 167–180.
25. Avendaño-Vázquez, S.E., Dhir, A., Bembich, S., Buratti, E., Proudfoot, N., and Baralle, F.E. (2012). Autoregulation of TDP-43 mRNA levels involves interplay between transcription, splicing, and alternative polyA site selection. Genes Dev. *26*, 1679–1684.
26. Ayala, Y.M., De Conti, L., Avendaño-Vázquez, S.E., Dhir, A., Romano, M., D'Ambrogio, A., Tollervey, J., Ule, J., Baralle, M., Buratti, E., and Baralle, F.E. (2011). TDP-43 regulates its mRNA levels through a negative feedback loop. EMBO J. *30*, 277–288.
27. Imaizumi, K., Ideno, H., Sato, T., Morimoto, S., and Okano, H. (2022). Pathogenic mutation of TDP-43 impairs RNA processing in a cell type-specific manner: Implications for the pathogenesis of ALS/FTLD. eNeuro *9*, ENEURO.0061-22.2022.
28. Melamed, Z., López-Erauskin, J., Baughn, M.W., Zhang, O., Drenner, K., Sun, Y., Freyermuth, F., McMahon, M.A., Beccari, M.S., Artates, J.W., et al. (2019). Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. Nat. Neurosci. *22*, 180–190.
29. Modic, M., Grosch, M., Rot, G., Schirge, S., Lepko, T., Yamazaki, T., Lee, F.C.Y., Rusha, E., Shaposhnikov, D., Palo, M., et al. (2019). Cross-regulation between TDP-43 and paraspeckles promotes pluripotency-differentiation transition. Mol. Cell *74*, 951–965.e13.
30. Rot, G., Wang, Z., Huppertz, I., Modic, M., Lenče, T., Hallegger, M., Haberman, N., Curk, T., von Mering, C., and Ule, J. (2017). High-resolution RNA maps suggest common principles of splicing and polyadenylation regulation by TDP-43. Cell Rep. *19*, 1056–1067.
31. Butti, Z., and Patten, S.A. (2019). RNA dysregulation in amyotrophic lateral sclerosis. Front. Genet. *9*, 712.

32. Prudencio, M., Belzil, V.V., Batra, R., Ross, C.A., Gendron, T.F., Pregent, L.J., Murray, M.E., Overstreet, K.K., Piazza-Johnston, A.E., Desaro, P., et al. (2015). Distinct brain transcriptome profiles in c9orf72-associated and sporadic als. Nat. Neurosci. *18*, 1175–1182.

33. Ziff, O.J., Harley, J., Wang, Y., Neeves, J., Tyzack, G., Ibrahim, F., Skehel, M., Chakrabarti, A.M., Kelly, G., and Patani, R. (2023). Nucleocytoplasmic mRNA redistribution accompanies RNA binding protein mislocalization in ALS motor neurons and is restored by VCP ATPase inhibition. Neuron *111*, 3011–3027.e7.

34. Tian, B., and Manley, J.L. (2017). Alternative polyadenylation of mRNA precursors. Nat. Rev. Mol. Cell Biol. *18*, 18–30.

35. Guvenek, A., and Tian, B. (2018). Analysis of alternative cleavage and polyadenylation in mature and differentiating neurons using rna-seq data. Quant. Biol. *6*, 253–266.

36. Cui, Y., Arnold, F.J., Peng, F., Wang, D., Li, J.S., Michels, S., Wagner, E. J., La Spada, A.R., and Li, W. (2023). Alternative polyadenylation transcriptome-wide association study identifies apa-linked susceptibility genes in brain disorders. Nat. Commun. *14*, 583.

37. Patel, R., Brophy, C., Hickling, M., Neve, J., and Furger, A. (2019). Alternative cleavage and polyadenylation of genes associated with protein turnover and mitochondrial function are deregulated in parkinson's, alzheimer's and als disease. BMC Med. Genomics *12*, 60.

38. Rhinn, H., Qiang, L., Yamashita, T., Rhee, D., Zolin, A., Vanti, W., and Abeliovich, A. (2012). Alternative α-synuclein transcript usage as a convergent mechanism in parkinson's disease pathology. Nat. Commun. *3*, 1084.

39. Shulman, E.D., and Elkon, R. (2019). Cell-type-specific analysis of alternative polyadenylation using single-cell transcriptomics data. Nucleic Acids Res. *47*, 10027–10039.

40. Patrick, R., Humphreys, D.T., Janbandhu, V., Oshlack, A., Ho, J.W.K., Harvey, R.P., and Lo, K.K. (2020). Sierra: discovery of differential transcript usage from polya-captured single-cell rna-seq data. Genome Biol. *21*, 167.

41. Gao, Y., Li, L., Amos, C.I., and Li, W. (2021). Analysis of alternative polyadenylation from single-cell rna-seq using scdapars reveals cell subpopulations invisible to gene expression. Genome Res. *31*, 1856–1866.

42. Li, G.-W., Nan, F., Yuan, G.-H., Liu, C.-X., Liu, X., Chen, L.-L., Tian, B., and Yang, L. (2021). Scapture: a deep learning-embedded pipeline that captures polyadenylation information from 3 tag-based rna-seq of single cells. Genome Biol. *22*, 221.

43. Li, W.V., Zheng, D., Wang, R., and Tian, B. (2021). Maaper: model-based analysis of alternative polyadenylation using 3 end-linked reads. Genome Biol. *22*, 222.

44. Li, J., Jaiswal, M.K., Chien, J.-F., Kozlenkov, A., Jung, J., Zhou, P., Gardashli, M., Pregent, L.J., Engelberg-Cook, E., Dickson, D.W., et al. (2023). Divergent single cell transcriptome and epigenome alterations in ALS and FTD patients with c9orf72 mutation. Nat. Commun. *14*, 5714.

45. Pineda, S.S., Lee, H., Ulloa-Navas, M.J., Linville, R.M., Garcia, F.J., Galani, K., Engelberg-Cook, E., Castanedes, M.C., Fitzwalter, B.E., Pregent, L.J., et al. (2024). Single-cell dissection of the human motor and prefrontal cortices in als and ftld. Cell *187*, 1971–1989.e16.

46. Gittings, L.M., Alsop, E.B., Antone, J., Singer, M., Whitsett, T.G., Sattler, R., and Van Keuren-Jensen, K. (2023). 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.

47. Limone, F., Mordes, D.A., Couto, A., Joseph, B.J., Mitchell, J.M., Therrien, M., Ghosh, S.D., Meyer, D., Zhang, Y., Goldman, M., et al. (2024). Single-nucleus sequencing reveals enriched expression of genetic risk factors in extratelencephalic neurons sensitive to degeneration in als. Nat. Aging *4*, 984–997.

48. Andrews, T.S., Kiselev, V.Y., McCarthy, D., and Hemberg, M. (2020). Tutorial: guidelines for the computational analysis of single-cell RNA sequencing data. Nat. Protoc. *16*, 1–9.

49. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with harmony. Nat. Methods *16*, 1289–1296.

50. Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From louvain to leiden: guaranteeing well-connected communities. Sci. Rep. *9*, 5233.

51. Aevermann, B., Zhang, Y., Novotny, M., Keshk, M., Bakken, T., Miller, J., Hodge, R., Lelieveldt, B., Lein, E., and Scheuermann, R.H. (2021). A machine learning method for the discovery of minimum marker gene combinations for cell type identification from single-cell RNA sequencing. Genome Res. *31*, 1767–1780.

52. Clarke, Z.A., Andrews, T.S., Atif, J., Pouyabahar, D., Innes, B.T., MacParland, S.A., and Bader, G.D. (2021). Tutorial: guidelines for annotating single-cell transcriptomic maps using automated and manual methods. Nat. Protoc. *16*, 2749–2764.

53. Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A.K., Slichter, C.K., Miller, H.W., McElrath, M.J., Prlic, M., et al. (2015). Mast: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell rna sequencing data. Genome Biol. *16*, 278.

54. Kolberg, L., Raudvere, U., Kuzmin, I., Vilo, J., and Peterson, H. (2020). gprofiler2–an r package for gene list functional enrichment analysis and namespace conversion toolset g: Profiler. F1000Res. *9*, 709.

55. Reale, L.A., Dyer, M.S., Perry, S.E., Young, K.M., Dickson, T.C., Woodhouse, A., and Blizzard, C.A. (2023). Pathologically mislocalised tdp-43 in upper motor neurons causes a die-forward spread of als-like pathogenic changes throughout the mouse corticomotor system. Prog. Neurobiol. *226*, 102449.

56. Gelon, P.A., Dutchak, P.A., and Sephton, C.F. (2022). Synaptic dysfunction in als and ftd: anatomical and molecular changes provide insights into mechanisms of disease. Front. Mol. Neurosci. *15*, 1000183.

57. Kanai, K., Kuwabara, S., Misawa, S., Tamura, N., Ogawara, K., Nakata, M., Sawai, S., Hattori, T., and Bostock, H. (2006). Altered axonal excitability properties in amyotrophic lateral sclerosis: impaired potassium channel function related to disease stage. Brain *129*, 953–962.

58. Shibuya, K., Misawa, S., Arai, K., Nakata, M., Kanai, K., Yoshiyama, Y., Ito, K., Isose, S., Noto, Y.-i., Nasu, S., et al. (2011). Markedly reduced axonal potassium channel expression in human sporadic amyotrophic lateral sclerosis: an immunohistochemical study. Exp. Neurol. *232*, 149–153.

59. Oeckl, P., Weydt, P., Thal, D.R., Weishaupt, J.H., Ludolph, A.C., and Otto, M. (2020). Proteomics in cerebrospinal fluid and spinal cord suggests uchl1, map2 and gpnmb as biomarkers and underpins importance of transcriptional pathways in amyotrophic lateral sclerosis. Acta Neuropathol. *139*, 119–134.

60. Ma, X.R., Prudencio, M., Koike, Y., Vatsavayai, S.C., Kim, G., Harbinski, F., Briner, A., Rodriguez, C.M., Guo, C., Akiyama, T., et al. (2022). TDP-43 represses cryptic exon inclusion in the FTD-ALS gene UNC13A. Nature *603*, 124–130.

61. Klim, J.R., Williams, L.A., Limone, F., Guerra San Juan, I., Davis-Dusenbery, B.N., Mordes, D.A., Burberry, A., Steinbaugh, M.J., Gamage, K.K., Kirchner, R., et al. (2019). ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. Nat. Neurosci. *22*, 167–179.

62. Baughn, M.W., Melamed, Z., López-Erauskin, J., Beccari, M.S., Ling, K., Zuberi, A., Presa, M., Gonzalo-Gil, E., Maimon, R., Vazquez-Sanchez, S., et al. (2023). Mechanism of stmn2 cryptic splice-polyadenylation and its correction for tdp-43 proteinopathies. Science *379*, 1140–1149.

63. Johnson, J.O., Mandrioli, J., Benatar, M., Abramzon, Y., Van Deerlin, V. M., Trojanowski, J.Q., Gibbs, J.R., Brunetti, M., Gronka, S., Wuu, J.,

et al. (2010). Exome sequencing reveals vcp mutations as a cause of familial als. Neuron *68*, 857–864.

64. Kang, S.H., Li, Y., Fukaya, M., Lorenzini, I., Cleveland, D.W., Ostrow, L. W., Rothstein, J.D., and Bergles, D.E. (2013). Degeneration and impaired regeneration of gray matter oligodendrocytes in amyotrophic lateral sclerosis. Nat. Neurosci. *16*, 571–579.
65. Wang, J., Ho, W.Y., Lim, K., Feng, J., Tucker-Kellogg, G., Nave, K.-A., and Ling, S.-C. (2018). Cell-autonomous requirement of tdp-43, an als/ftd signature protein, for oligodendrocyte survival and myelination. Proc. Natl. Acad. Sci. *115*, E10941–E10950.
66. Patzig, J., Erwig, M.S., Tenzer, S., Kusch, K., Dibaj, P., Möbius, W., Goebbels, S., Schaeren-Wiemers, N., Nave, K.-A., and Werner, H.B. (2016). Septin/anillin filaments scaffold central nervous system myelin to accelerate nerve conduction. eLife *5*, e17119.
67. Lim, R.G., Al-Dalahmah, O., Wu, J., Gold, M.P., Reidling, J.C., Tang, G., Adam, M., Dansu, D.K., Park, H.-J., Casaccia, P., et al. (2022). Huntington disease oligodendrocyte maturation deficits revealed by single-nucleus rnaseq are rescued by thiamine-biotin supplementation. Nat. Commun. *13*, 7791.
68. Ling, J.P., Pletnikova, O., Troncoso, J.C., and Wong, P.C. (2015). Tdp-43 repression of nonconserved cryptic exons is compromised in als-ftd. Science *349*, 650–655.
69. San Gil, R., Pascovici, D., Venturato, J., Brown-Wright, H., Mehta, P., Madrid San Martin, L., Wu, J., Luan, W., Chui, Y.K., Bademosi, A.T., et al. (2024). A transient protein folding response targets aggregation in the early phase of tdp-43-mediated neurodegeneration. Nat. Commun. *15*, 1508.
70. Krach, F., Batra, R., Wheeler, E.C., Vu, A.Q., Wang, R., Hutt, K., Rabin, S. J., Baughn, M.W., Libby, R.T., Diaz-Garcia, S., et al. (2018). Transcriptome–pathology correlation identifies interplay between tdp-43 and the expression of its kinase ck1e in sporadic als. Acta Neuropathol. *136*, 405–423.
71. Pradat, P.-F., Dubourg, O., De Tapia, M., Di Scala, F., Dupuis, L., Lenglet, T., Bruneteau, G., Salachas, F., Lacomblez, L., Corvol, J.-C., et al. (2011). Muscle gene expression is a marker of amyotrophic lateral sclerosis severity. Neurodegener. Dis. *9*, 38–52.
72. Diez-Hermano, S., Mejias, A., Sanchez, D., Gutierrez, G., and Ganfornina, M.D. (2020). Control of the neuroprotective lipocalin apolipoprotein d expression by alternative promoter regions and differentially expressed mrna 5' utr variants. PLoS One *15*, e0234857.
73. Xiao, Y., and Czopka, T. (2023). Myelination-independent functions of oligodendrocyte precursor cells in health and disease. Nat. Neurosci. *26*, 1663–1669.
74. Venkatraman, V., Filiano, A.J., Xu, L., Collins, L., Luo, E., Ripple, K.M., de Castro, G.C., Boua, J.-V.K., Marius, C., Giamberardino, C., et al. (2022). Filtered cerebrospinal fluid from patients with amyotrophic lateral sclerosis displays an altered proteome and affects motor phenotype in a mouse model. Cureus *14*, e32980.
75. Yamanaka, K., and Komine, O. (2018). The multi-dimensional roles of astrocytes in als. Neurosci. Res. *126*, 31–38.
76. Humphrey, J., Venkatesh, S., Hasan, R., Herb, J.T., de Paiva Lopes, K., Küçükali, F., Byrska-Bishop, M., Evani, U.S., Narzisi, G., Fagegaltier, D., et al. (2023). Integrative transcriptomic analysis of the amyotrophic lateral sclerosis spinal cord implicates glial activation and suggests new risk genes. Nat. Neurosci. *26*, 150–162.
77. Pekny, M., and Nilsson, M. (2005). Astrocyte activation and reactive gliosis. Glia *50*, 427–434.
78. Li, F., Liu, A., Zhao, M., and Luo, L. (2023). Astrocytic chitinase-3-like protein 1 in neurological diseases: Potential roles and future perspectives. J. Neurochem. *165*, 772–790.
79. Ki, S.M., Jeong, H.S., and Lee, J.E. (2021). Primary cilia in glial cells: An oasis in the journey to overcoming neurodegenerative diseases. Front. Neurosci. *15*, 736888.
80. Peng, A.Y.T., Agrawal, I., Ho, W.Y., Yen, Y.-C., Pinter, A.J., Liu, J., Phua, Q.X.C., Koh, K.B., Chang, J.-C., Sanford, E., et al. (2020). Loss of tdp-43 in astrocytes leads to motor deficits by triggering a1-like reactive phenotype and triglial dysfunction. Proc. Natl. Acad. Sci. USA *117*, 29101–29112.
81. Holt, L.M., Hernandez, R.D., Pacheco, N.L., Torres Ceja, B., Hossain, M., and Olsen, M.L. (2019). Astrocyte morphogenesis is dependent on bdnf signaling via astrocytic trkb.t1. eLife *8*, e44667.
82. Gupta, V., You, Y., Gupta, V., Klistorner, A., and Graham, S. (2013). Trkb receptor signalling: implications in neurodegenerative, psychiatric and proliferative disorders. Int. J. Mol. Sci. *14*, 10122–10142.
83. Nakaya, T. (2022). A specific gene-splicing alteration in the snrnp70 gene as a hallmark of an als subtype. Gene *818*, 146203.
84. Chi, B., O'Connell, J.D., Yamazaki, T., Gangopadhyay, J., Gygi, S.P., and Reed, R. (2018). Interactome analyses revealed that the u1 snrnp machinery overlaps extensively with the rnap ii machinery and contains multiple als/sma-causative proteins. Sci. Rep. *8*, 8755.
85. Olah, M., Menon, V., Habib, N., Taga, M.F., Ma, Y., Yung, C.J., Cimpean, M., Khairallah, A., Coronas-Samano, G., Sankowski, R., et al. (2020). Single cell RNA sequencing of human microglia uncovers a subset associated with alzheimer's disease. Nat. Commun. *11*, 6129.
86. Sun, N., Victor, M.B., Park, Y.P., Xiong, X., Scannail, A.N., Leary, N., Prosper, S., Viswanathan, S., Luna, X., Boix, C.A., et al. (2023). Human microglial state dynamics in alzheimer's disease progression. Cell *186*, 4386–4403.e29.
87. Keren-Shaul, H., Spinrad, A., Weiner, A., Matcovitch-Natan, O., Dvir-Szternfeld, R., Ulland, T.K., David, E., Baruch, K., Lara-Astaiso, D., Toth, B., et al. (2017). A unique microglia type associated with restricting development of alzheimer's disease. Cell *169*, 1276–1290.e17.
88. McCauley, M.E., O'Rourke, J.G., Yáñez, A., Markman, J.L., Ho, R., Wang, X., Chen, S., Lall, D., Jin, M., Muhammad, A.K.M.G., et al. (2020). C9orf72 in myeloid cells suppresses sting-induced inflammation. Nature *585*, 96–101.
89. Chen, H., Kankel, M.W., Su, S.C., Han, S.W.S., and Ofengeim, D. (2018). Exploring the genetics and non-cell autonomous mechanisms underlying ALS/FTLD. Cell Death Differ. *25*, 648–662.
90. Pang, W., and Hu, F. (2023). C9ORF72 suppresses JAK-STAT mediated inflammation. iScience *26*, 106579.
91. Sjöstedt, E., Zhong, W., Fagerberg, L., Karlsson, M., Mitsios, N., Adori, C., Oksvold, P., Edfors, F., Limiszewska, A., Hikmet, F., et al. (2020). An atlas of the protein-coding genes in the human, pig, and mouse brain. Science *367*, eaay5947.
92. Holstege, H., Hulsman, M., Charbonnier, C., Grenier-Boley, B., Quenez, O., Grozeva, D., van Rooij, J.G.J., Sims, R., Ahmad, S., Amin, N., et al. (2022). Exome sequencing identifies rare damaging variants in atp8b4 and abca1 as risk factors for alzheimer's disease. Nat. Genet. *54*, 1786–1794.
93. Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat. Methods *14*, 1083–1086.
94. Goutman, S.A., Hardiman, O., Al-Chalabi, A., Chió, A., Savelieff, M.G., Kiernan, M.C., and Feldman, E.L. (2022). Emerging insights into the complex genetics and pathophysiology of amyotrophic lateral sclerosis. Lancet Neurol. *21*, 465–479.
95. van Rheenen, W., van der Spek, R.A.A., Bakker, M.K., van Vugt, J.J.F.A., Hop, P.J., Zwamborn, R.A.J., de Klein, N., Westra, H.-J., Bakker, O.B., Deelen, P., et al. (2021). Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology. Nat. Genet. *53*, 1636–1648.
96. Wang, R., Zheng, D., Yehia, G., and Tian, B. (2018). A compendium of conserved cleavage and polyadenylation events in mammalian genes. Genome Res. *28*, 1427–1441.

97. Fansler, M.M., Mitschka, S., and Mayr, C. (2024). Quantifying 3'UTR length from scRNA-seq data reveals changes independent of gene expression. Nat. Commun. *15*, 4050.

98. Chen, C.Y.A., and Shyu, A.B. (1995). AU-rich elements: characterization and importance in mRNA degradation. Trends Biochem. Sci. *20*, 465–470.

99. O'Brien, J., Hayder, H., Zayed, Y., and Peng, C. (2018). Overview of microrna biogenesis, mechanisms of actions, and circulation. Front. Endocrinol. *9*, 402.

100. Brown, A.-L., Wilkins, O.G., Keuss, M.J., Kargbo-Hill, S.E., Zanovello, M., Lee, W.C., Bampton, A., Lee, F.C.Y., Masino, L., Qi, Y.A., et al. (2022). TDP-43 loss and ALS-risk SNPs drive mis-splicing and depletion of UNC13A. Nature *603*, 131–137.

101. Ghotra, R., Lee, N.K., Tripathy, R., and Koo, P.K. (2021). Designing interpretable convolution-based hybrid networks for genomics. Preprint at bioRxiv. https://doi.org/10.1101/2021.07.13.452181.

102. Koo, P.K., and Eddy, S.R. (2019). Representation learning of genomic sequence motifs with convolutional neural networks. PLoS Comput. Biol. *15*, e1007560.

103. Ray, D., Kazan, H., Cook, K.B., Weirauch, M.T., Najafabadi, H.S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., et al. (2013). A compendium of RNA-binding motifs for decoding gene regulation. Nature *499*, 172–177.

104. Masuda, A., Takeda, J.-i., Okuno, T., Okamoto, T., Ohkawara, B., Ito, M., Ishigaki, S., Sobue, G., and Ohno, K. (2015). Position-specific binding of FUS to nascent RNA regulates mRNA length. Genes Dev. *29*, 1045–1057.

105. Shenouda, M., Zhang, A.B., Weichert, A., and Robertson, J. (2018). Mechanisms associated with TDP-43 neurotoxicity in ALS/FTLD. Adv. Neurobiol. *20*, 239–263.

106. Mueller, S., Decker, L., Menge, S., Ludolph, A.C., and Freischmidt, A. (2023). The fragile x protein family in amyotrophic lateral sclerosis. Mol. Neurobiol. *60*, 3898–3910.

107. Yang, P., Mathieu, C., Kolaitis, R.-M., Zhang, P., Messing, J., Yurtsever, U., Yang, Z., Wu, J., Li, Y., Pan, Q., et al. (2020). G3bp1 is a tunable switch that triggers phase separation to assemble stress granules. Cell *181*, 325–345.e28.

108. Feng, H., Moakley, D.F., Chen, S., McKenzie, M.G., Menon, V., and Zhang, C. (2021). Complexity and graded regulation of neuronal cell-type-specific alternative splicing revealed by single-cell RNA sequencing. Proc. Natl. Acad. Sci. USA *118*, e2013056118.

109. Hu, Z., Li, M., Huo, Z., Chen, L., Liu, S., Deng, K., Lu, X., Chen, S., Fu, Y., and Xu, A. (2022). U1 snrnp proteins promote proximal alternative polyadenylation sites by directly interacting with 3 end processing core factors. J. Mol. Cell Biol. *14*, mjac054.

110. Nasiri-Aghdam, M., Garcia-Garduño, T.C., and Jave-Suárez, L.F. (2021). CELF family proteins in cancer: Highlights on the RNA-binding protein/noncoding RNA regulatory axis. Int. J. Mol. Sci. *22*, 11056.

111. Schwich, O.D., Blümel, N., Keller, M., Wegener, M., Setty, S.T., Brunstein, M.E., Poser, I., Mozos, I.R.D.L., Suess, B., Münch, C., et al. (2021). SRSF3 and SRSF7 modulate 3'UTR length through suppression or activation of proximal polyadenylation sites and regulation of CFIm levels. Genome Biol. *22*, 82.

112. Shkurin, A., Pour, S.E., and Hughes, T.R. (2023). Known sequence features explain half of all human gene ends. NAR Genomics Bioinform. *5*, lqad031.

113. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

114. Fischl, H., Neve, J., Wang, Z., Patel, R., Louey, A., Tian, B., and Furger, A. (2019). hnRNPC regulates cancer-specific alternative cleavage and polyadenylation profiles. Nucleic Acids Res. *47*, 7580–7591.

115. Goodarzi, H., Najafabadi, H.S., Oikonomou, P., Greco, T.M., Fish, L., Salavati, R., Cristea, I.M., and Tavazoie, S. (2012). Systematic discovery of structural elements governing stability of mammalian messenger RNAs. Nature *485*, 264–268.

116. Navickas, A., Asgharian, H., Winkler, J., Fish, L., Garcia, K., Markett, D., Dodel, M., Culbertson, B., Miglani, S., Joshi, T., et al. (2023). An mrna processing pathway suppresses metastasis by governing translational control from the nucleus. Nat. Cell Biol. *25*, 892–903.

117. Bhate, A., Parker, D.J., Bebee, T.W., Ahn, J., Arif, W., Rashan, E.H., Chorghade, S., Chau, A., Lee, J.-H., Anakk, S., et al. (2015). Esrp2 controls an adult splicing programme in hepatocytes to support postnatal liver maturation. Nat. Commun. *6*, 8768.

118. Ladd, A.N., Charlet-B, N., and Cooper, T.A. (2001). The celf family of rna binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. Mol. Cell Biol. *21*, 1285–1296.

119. Huang, Y.-S., Kan, M.-C., Lin, C.-L., and Richter, J.D. (2006). Cpeb3 and cpeb4 in neurons: analysis of rna-binding specificity and translational control of ampa receptor glur2 mrna. EMBO J. *25*, 4865–4876.

120. Kozlov, E., Shidlovskii, Y.V., Gilmutdinov, R., Schedl, P., and Zhukova, M. (2021). The role of cpeb family proteins in the nervous system function in the norm and pathology. Cell Biosci. *11*, 64.

121. Picó, S., Parras, A., Santos-Galindo, M., Pose-Utrilla, J., Castro, M., Fraga, E., Hernández, I.H., Elorza, A., Anta, H., Wang, N., et al. (2021). Cpeb alteration and aberrant transcriptome-polyadenylation lead to a treatable slc19a3 deficiency in huntington's disease. Sci. Transl. Med. *13*, eabe7104.

122. Gordon, P.M., Hamid, F., Makeyev, E.V., and Houart, C. (2021). A conserved role for the ALS-linked splicing factor SFPQ in repression of pathogenic cryptic last exons. Nat. Commun. *12*, 1918.

123. Zhao, Z., Qing, Y., Dong, L., Han, L., Wu, D., Li, Y., Li, W., Xue, J., Zhou, K., Sun, M., et al. (2023). Qki shuttles internal m7g-modified transcripts into stress granules and modulates mrna metabolism. Cell *186*, 3208–3226.e27.

124. Park, C., Choi, S., Kim, Y.-E., Lee, S., Park, S.-H., Adelstein, R.S., Kawamoto, S., and Kim, K.K. (2017). Stress granules contain rbfox2 with cell cycle-related mrnas. Sci. Rep. *7*, 11211.

125. Somasekharan, S.P., El-Naggar, A., Leprivier, G., Cheng, H., Hajee, S., Grunewald, T.G., Zhang, F., Ng, T., Delattre, O., and Evdokimova, V. (2015). Yb-1 regulates stress granule formation and tumor progression by translationally activating g3bp1. J. Cell Biol. *208*, 913–929.

126. Huang, S., Luo, K., Jiang, L., Zhang, X.-D., Lv, Y.-H., and Li, R.-F. (2021). Pcbp1 regulates the transcription and alternative splicing of metastasis-related genes and pathways in hepatocellular carcinoma. Sci. Rep. *11*, 23356.

127. Zhang, T., Huang, X.-H., Dong, L., Hu, D., Ge, C., Zhan, Y.-Q., Xu, W.-X., Yu, M., Li, W., Wang, X., et al. (2010). Pcbp-1 regulates alternative splicing of the cd44 gene and inhibits invasion in human hepatoma cell line hepg2 cells. Mol. Cancer *9*, 72.

128. Zou, D., McSweeney, C., Sebastian, A., Reynolds, D.J., Dong, F., Zhou, Y., Deng, D., Wang, Y., Liu, L., Zhu, J., et al. (2015). A critical role of rbm8a in proliferation and differentiation of embryonic neural progenitors. Neural Dev. *10*, 18.

129. Lin, J.-C., Hsu, M., and Tarn, W.-Y. (2007). Cell stress modulates the function of splicing regulatory protein rbm4 in translation control. Proc. Natl. Acad. Sci. *104*, 2235–2240.

130. Jia, Q., Nie, H., Yu, P., Xie, B., Wang, C., Yang, F., Wei, G., and Ni, T. (2019). HNRNPA1-mediated 3' UTR length changes of HN1 contributes to cancer- and senescence-associated phenotypes. Aging *11*, 4407–4437.

131. Bampton, A., Gittings, L.M., Fratta, P., Lashley, T., and Gatt, A. (2020). The role of hnrnps in frontotemporal dementia and amyotrophic lateral sclerosis. Acta Neuropathol. *140*, 599–623.

132. Anvar, S.Y., Allard, G., Tseng, E., Sheynkman, G.M., de Klerk, E., Vermaat, M., Yin, R.H., Johansson, H.E., Ariyurek, Y., den Dunnen, J.T., et al. (2018). Full-length mRNA sequencing uncovers a widespread

coupling between transcription initiation and mRNA processing. Genome Biol. *19*, 46.

133. Kowalski, M.H., Wessels, H.-H., Linder, J., Dalgarno, C., Mascio, I., Choudhary, S., Hartman, A., Hao, Y., Kundaje, A., and Satija, R. (2024). Multiplexed single-cell characterization of alternative polyadenylation regulators. Cell *187*, 4408–4425.e23.
134. Consonni, M., Cappa, S.F., Dalla Bella, E., Contarino, V.E., and Lauria, G. (2019). Cortical correlates of behavioural change in amyotrophic lateral sclerosis. J. Neurol. Neurosurg. Psychiatry *90*, 380–386.
135. Crockford, C., Newton, J., Lonergan, K., Chiwera, T., Booth, T., Chandran, S., Colville, S., Heverin, M., Mays, I., Pal, S., et al. (2018). ALS-specific cognitive and behavior changes associated with advancing disease stage in ALS. Neurology *91*, e1370–e1380.
136. Lillo, P., Garcin, B., Hornberger, M., Bak, T.H., and Hodges, J.R. (2010). Neurobehavioral features in frontotemporal dementia with amyotrophic lateral sclerosis. Arch. Neurol. *67*, 826–830.
137. Lillo, P., Mioshi, E., Zoing, M.C., Kiernan, M.C., and Hodges, J.R. (2011). How common are behavioural changes in amyotrophic lateral sclerosis? Amyotroph Lateral Scler. *12*, 45–51.
138. Lomen-Hoerth, C. (2004). Characterization of amyotrophic lateral sclerosis and frontotemporal dementia. Dement. Geriatr. Cogn. Disord *17*, 337–341.
139. Meier, S.L., Charleston, A.J., and Tippett, L.J. (2010). Cognitive and behavioural deficits associated with the orbitomedial prefrontal cortex in amyotrophic lateral sclerosis. Brain: J. Neurol. *133*, 3444–3457.
140. Tsujimoto, M., Senda, J., Ishihara, T., Niimi, Y., Kawai, Y., Atsuta, N., Watanabe, H., Tanaka, F., Naganawa, S., and Sobue, G. (2011). Behavioral changes in early ALS correlate with voxel-based morphometry and diffusion tensor imaging. J. Neurol. Sci. *307*, 34–40.
141. Bannwarth, S., Ait-El-Mkadem, S., Chaussenot, A., Genin, E.C., Lacas-Gervais, S., Fragaki, K., Berg-Alonso, L., Kageyama, Y., Serre, V., Moore, D.G., et al. (2014). A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. Brain: J. Neurol. *137*, 2329–2345.
142. Smith, E.F., Shaw, P.J., and De Vos, K.J. (2019). The role of mitochondria in amyotrophic lateral sclerosis. Neurosci. Lett. *710*, 132933.
143. Tsai, Y.-L., and Manley, J.L. (2021). Multiple ways to a dead end: diverse mechanisms by which ALS mutant genes induce cell death. Cell Cycle *20*, 631–646.
144. Poesen, K., and Van Damme, P. (2019). Diagnostic and prognostic performance of neurofilaments in als. Front. Neurol. *9*, 1167.
145. Tam, O.H., Rozhkov, N.V., Shaw, R., Kim, D., Hubbard, I., Fennessey, S., Propp, N., Fagegaltier, D., Harris, B.T., Ostrow, L.W., et al. (2019). Postmortem cortex samples identify distinct molecular subtypes of als: retrotransposon activation, oxidative stress, and activated glia. Cell Rep. *29*, 1164–1177.e5.
146. Valori, C.F., Sulmona, C., Brambilla, L., and Rossi, D. (2023). Astrocytes: dissecting their diverse roles in amyotrophic lateral sclerosis and frontotemporal dementia. Cells *12*, 1450.
147. Guedes, J.R., Ferreira, P.A., Costa, J.M., Cardoso, A.L., and Peça, J. (2022). Microglia-dependent remodeling of neuronal circuits. J. Neurochem. *163*, 74–93.
148. Mead, R.J., Shan, N., Reiser, H.J., Marshall, F., and Shaw, P.J. (2023). Amyotrophic lateral sclerosis: a neurodegenerative disorder poised for successful therapeutic translation. Nat. Rev. Drug Discov. *22*, 185–212.
149. Richardson, P.J., Smith, D.P., de Giorgio, A., Snetkov, X., Almond-Thynne, J., Cronin, S., Mead, R.J., McDermott, C.J., and Shaw, P.J. (2023). Janus kinase inhibitors are potential therapeutics for amyotrophic lateral sclerosis. Transl. Neurodegener. *12*, 47.
150. Quek, H., Cuní-López, C., Stewart, R., Colletti, T., Notaro, A., Nguyen, T. H., Sun, Y., Guo, C.C., Lupton, M.K., Roberts, T.L., et al. (2022). Als monocyte-derived microglia-like cells reveal cytoplasmic tdp-43 accumulation, dna damage, and cell-specific impairment of phagocytosis associated with disease progression. J. Neuroinflammation *19*, 58.
151. Arefeen, A., Xiao, X., and Jiang, T. (2019). DeepPASTA: deep neural network based polyadenylation site analysis. Bioinformatics *35*, 4577–4585.
152. Leung, M.K.K., Delong, A., and Frey, B.J. (2018). Inference of the human polyadenylation code. Bioinformatics *34*, 2889–2898.
153. Linder, J., Koplik, S.E., Kundaje, A., and Seelig, G. (2022). Deciphering the impact of genetic variation on human polyadenylation using aparent2. Genome Biol. *23*, 232.
154. Li, Z., Li, Y., Zhang, B., Li, Y., Long, Y., Zhou, J., Zou, X., Zhang, M., Hu, Y., Chen, W., and Gao, X. (2022). Deerect-apa: prediction of alternative polyadenylation site usage through deep learning. Genom. Proteom. Bioinform. *20*, 483–495.
155. Bak, M., van Nimwegen, E., Kouzel, I.U., Gur, T., Schmidt, R., Zavolan, M., and Gruber, A.J. (2024). Mapp unravels frequent co-regulation of splicing and polyadenylation by rna-binding proteins and their dysregulation in cancer. Nat. Commun. *15*, 4110.
156. Fröhlich, A.S., Gerstner, N., Gagliardi, M., Ködel, M., Yusupov, N., Matosin, N., Czamara, D., Sauer, S., Roeh, S., Murek, V., and Chatzinakos, C. (2024). Single-nucleus transcriptomic profiling of human orbitofrontal cortex reveals convergent effects of aging and psychiatric disease. Nat. Neurosci. *27*, 2021–2032.
157. Wang, H.-L.V., Xiang, J.-F., Yuan, C., Veire, A.M., Gendron, T.F., Murray, M.E., Tansey, M.G., Hu, J., Gearing, M., Glass, J.D., and Jin, P. (2025). ptdp-43 levels correlate with cell type–specific molecular alterations in the prefrontal cortex of c9orf72 als/ftd patients. Proc. Natl. Acad. Sci. *122*, e2419818122.
158. Neumann, M., and Mackenzie, I.R.A. (2019). Review: Neuropathology of non-tau frontotemporal lobar degeneration. Neuropathol. Appl. Neurobiol. *45*, 19–40.
159. Andrews, T.S., Atif, J., Liu, J.C., Perciani, C.T., Ma, X.-Z., Thoeni, C., Slyper, M., Eraslan, G., Segerstolpe, A., Manuel, J., et al. (2022). Single-cell, single-nucleus, and spatial rna sequencing of the human liver identifies cholangiocyte and mesenchymal heterogeneity. Hepatol. Commun. *6*, 821–840.
160. Pressl, C., Mätlik, K., Kus, L., Darnell, P., Luo, J.-D., Paul, M.R., Weiss, A. R., Liguore, W., Carroll, T.S., Davis, D.A., et al. (2024). Selective vulnerability of layer 5a corticostriatal neurons in huntington's disease. Neuron *112*, 924–941.e10.
161. Siletti, K., Hodge, R., Mossi Albiach, A., Lee, K.W., Ding, S.-L., Hu, L., Lönnerberg, P., Bakken, T., Casper, T., Clark, M., et al. (2023). Transcriptomic diversity of cell types across the adult human brain. Science *382*, eadd7046.
162. Gabitto, M.I., Travaglini, K.J., Rachleff, V.M., Kaplan, E.S., Long, B., Ariza, J., Ding, Y., Mahoney, J.T., Dee, N., Goldy, J., et al. (2024). Integrated multimodal cell atlas of alzheimer disease. Nat. Neurosci. *27*, 2366–2383. https://doi.org/10.1038/s41593-024-01774-5.
163. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., and Satija, R. (2023). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat. Biotechnol. *42*, 293–304.
164. Hao, Y., Hao, S., Andersen-Nissen, E., III, W.M.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zagar, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29.
165. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., III, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21.
166. Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

167. Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502.

168. Lun, A.T.L., McCarthy, D.J., and Marioni, J.C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data. F1000Res. *5*, 2122.

169. Germain, P.-L., Lun, A., Macnair, W., and Robinson, M.D. (2021). Doublet identification in single-cell sequencing data using scDblFinder. F1000Res. *10*, 979.

170. Marsh, S., Salmon, M., and Hoffman, P. (2023). samuel-marsh/scCustomize: Version 1.1.3. Zenodo.

171. Bunis, D.G., Andrews, J., Fragiadakis, G.K., Burt, T.D., and Sirota, M. (2021). dittoSeq: universal user-friendly single-cell and bulk RNA sequencing visualization toolkit. Bioinformatics *36*, 5535–5536.

172. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for rna-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

173. Andreatta, M., and Carmona, S.J. (2021). UCell: Robust and scalable single-cell gene signature scoring. Comput. Struct. Biotechnol. J. *19*, 3796–3798.

174. Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., Wadi, L., Meyer, M., Wong, J., Xu, C., et al. (2019). Pathway enrichment analysis and visualization of omics data using g:profiler, GSEA, cytoscape and EnrichmentMap. Nat. Protoc. *14*, 482–517.

175. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. *17*, 10–12.

176. Guo, L., and Huo, H. (2024). An efficient burrows–wheeler transform-based aligner for short read mapping. Comput. Biol. Chem. *110*, 108050.

177. Shah, A., Qian, Y., Weyn-Vanhentenryck, S.M., and Zhang, C. (2017). Clip tool kit (ctk): a flexible and robust pipeline to analyze clip sequencing data. Bioinformatics *33*, 566–567.

178. Elemento, O., Slonim, N., and Tavazoie, S. (2007). A universal framework for regulatory element discovery across all genomes and data types. Mol. Cell *28*, 337–350.

179. Brooks, B.R., Miller, R.G., Swash, M., and Munsat, T.L.; Neurology Research Group on Motor Neuron Diseases (2000). El escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph. Lateral Scler. Other Motor Neuron Disord. *1*, 293–299.

180. McGoldrick, P., Zhang, M., van Blitterswijk, M., Sato, C., Moreno, D., Xiao, S., Zhang, A.B., McKeever, P.M., Weichert, A., Schneider, R., et al. (2018). Unaffected mosaic c9orf72 case: RNA foci, dipeptide proteins, but upregulated c9orf72 expression. Neurology *90*, e323–e331.

181. Xi, Z., van Blitterswijk, M., Zhang, M., McGoldrick, P., McLean, J.R., Yunusova, Y., Knock, E., Moreno, D., Sato, C., McKeever, P.M., et al. (2015). Jump from pre-mutation to pathologic expansion in c9orf72. Am. J. Hum. Genet. *96*, 962–970.

182. Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. *37*, 38–44.

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

184. Ye, C., Zhou, Q., Wu, X., Yu, C., Ji, G., Saban, D.R., and Li, Q.Q. (2020). scdapa: detection and visualization of dynamic alternative polyadenylation from single cell rna-seq data. Bioinformatics *36*, 1262–1264.

185. Yang, X., Tong, Y., Liu, G., Yuan, J., and Yang, Y. (2022). scapaatlas: an atlas of alternative polyadenylation across cell types in human and mouse. Nucleic Acids Res. *50*, D356–D364.

186. Quang, D., and Xie, X. (2016). Danq: a hybrid convolutional and recurrent deep neural network for quantifying the function of dna sequences. Nucleic Acids Res. *44*, e107.

187. Zimmerman, K.D., Espeland, M.A., and Langefeld, C.D. (2021). A practical solution to pseudoreplication bias in single-cell studies. Nat. Commun. *12*, 738.

188. Soneson, C., and Robinson, M.D. (2018). Bias, robustness and scalability in single-cell differential expression analysis. Nat. Methods *15*, 255–261.

189. Squair, J.W., Gautier, M., Kathe, C., Anderson, M.A., James, N.D., Hutson, T.H., Hudelle, R., Qaiser, T., Matson, K.J.E., Barraud, Q., et al. (2021). Confronting false discoveries in single-cell differential expression. Nat. Commun. *12*, 5692.

190. Goodarzi, H., Elemento, O., and Tavazoie, S. (2009). Revealing global regulatory perturbations across human cancers. Mol. Cell *36*, 900–911.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-phospho TDP-43 (Ser409/Ser410) Antibody, clone 1D3	CiteAb	Cat# MABN14; RRID: AB_11212279
Chemicals, peptides, and recombinant proteins		
PBS, pH 7.4 1X	Thermo Fisher Scientific	Cat# 10010023
SYBR™ Green II RNA Gel Stain, 10,000X concentrate in DMSO	Thermo Fisher Scientific	Cat# S7564
MACS BSA Stock Solution	MACS Miltenyi Biotec	Cat# 130-091-376
Sucrose (>99.5%)	Sigma-Aldrich	Cat# S0389
Calcium chloride solution ~ 1.00M	Sigma-Aldrich	Cat# 21115
Magnesium acetate solution	Sigma-Aldrich	Cat# 63052
Invitrogen™ UltraPure™ 1 M Tris-HCl Buffer, pH 7.5	Fisher Scientific	Cat# 15567027
Triton X-100	Sigma-Aldrich	Cat# X-100
0.5M EDTA pH 8.0	Sigma-Aldrich	Cat# 324506
UltraPure™ DNase/RNase-Free Distilled Water	Fisher Scientific	Cat# 10977015
Flowmi® Cell Strainers, 40 μm, for 1000 μL Pipette Tips	Sigma-Aldrich	Cat# BAH136800040
DAPI	Sigma-Aldrich	Cat# 10236276001
Protector RNase inhibitor	Sigma-Aldrich	Cat# 3335402001
C-Chip™ Disposable Hemacytometers	Fisher Scientific	Cat# 22-600-100
Fisherbrand Razor Blades	Fisher Scientific	Cat# 12–640
KIMBLE Dounce tissue grinder set	Sigma-Aldrich	Cat# D9063
Critical commercial assays		
10X Chromium Single Cell 3′ v2	10X Genomics	Cat# PN-120267
10X Chromium Single Cell 3′ v3	10X Genomics	Cat# PN-1000268
Illumina NovaSeq 6000 SP flow cell	Illumina	Cat# 20028402
Illumina NovaSeq 6000 S2 flow cell	Illumina	Cat# 20028314
Deposited data		
snRNA-Seq data from human orbitofrontal cortex	Current study	Sequencing Read Archive Database (SRA): https://www.ncbi.nlm.nih.gov/bioproject/PRJNA918304
Processed human motor (M1) cortex data	Siletti et al.[161]	https://portal.brain-map.org/atlases-and-data/rnaseq/human-m1-10x
Seattle Alzheimer Disease Cell Atlas	Gabitto et al.[162]	https://portal.brain-map.org/atlases-and-data/rnaseq/human-mtg-10x_sea-ad
Raw and processed snRNA-Seq data from prefrontal cortex	Gittings et al.[46]	Synapse: https://www.synapse.org/Synapse:syn45351388/files/
Raw and processed snRNA-Seq data from prefrontal and motor cortex	Li et al.[44]	Gene Expression Omnibus: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE219281
Raw and processed snRNA-Seq data from human prefrontal and motor cortex	Pineda et al.[45]	SRA & Synapse: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1073234/, https://www.synapse.org/Synapse:syn51105515/files/

(*Continued on next page*)

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Cell Ranger v5.0.1	10X Genomics	https://www.10xgenomics.com/support/software/cell-ranger/downloads/previous-versions
Seurat (v4)	Satija lab[163–167]	https://satijalab.org/seurat/articles/install_v5#install-seurat-v4
scran	Lun et al.[168]	https://bioconductor.org/packages/devel/bioc/html/scran.html
scDblFinder	Germain et al.[169]	https://github.com/plger/scDblFinder
Harmony	Korsunsky et al.[49]	https://github.com/immunogenomics/harmony
Leiden	Traag et al.[50]	https://arxiv.org/abs/1810.08473
ScCustomize (v2.1.2)	Marsh et al.[170]	https://samuel-marsh.github.io/scCustomize/
DittoSeq (v1.2.0)	Bunis et al.[171]	https://www.bioconductor.org/packages/release/bioc/html/dittoSeq.html
NSForest (v3.9.1)	Aevermann et al.[51]	https://github.com/JCVenterInstitute/NSForest
DESeq2	Love et al.[113]	https://github.com/thelovelab/DESeq2
limma	Ritchie et al.[172]	https://bioconductor.org/packages/release/bioc/html/limma.html
MAST	Finak et al.[53]	https://github.com/RGLab/MAST
UCell (v2.4)	Andreatta et al.[173]	https://carmonalab.github.io/UCell_demo/UCell_Seurat_vignette.html
MAAPER	Li et al.[43]	https://github.com/Vivianstats/MAAPER
PolyA DB (v3)	Wang et al.[96]	https://exon.apps.wistar.org/polya_db/v3/
APAlog	Navickas et al.[116]	https://github.com/goodarzilab/APAlog
gProfiler2	Kolberg et al.[54]	https://biit.cs.ut.ee/gprofiler/gost
Gene Ontology gene sets (2024 release)	Bader lab Lab[174]	http://baderlab.org/GeneSets/
Cytoscape (v3.9.1) and EnrichmentMap Plugin (v3.4.0)	Reimand et al.[174]	https://apps.cytoscape.org/apps/enrichmentmap
APA–Net (v1.0.0)	Current study	Zenodo: https://doi.org/10.5281/zenodo.16879041; GitHub: https://github.com/BaderLab/APA-Net
RBP compendium	Ray et al.[103]	https://hugheslab.ccbr.utoronto.ca/supplementary-data/RNAcompete_eukarya/
cutadapt	Martin[175]	https://ojsupgrade.com/index.php/embnetjournal/article/view/200
BWA (v0.7.17)	Guo and Ho[176]	https://github.com/lh3/bwa
CLIP Tool Kit (CTK v1.1.13)	Shah et al.[177]	https://github.com/chaolinzhanglab/ctk
FIRE	Elemento et al.[178]	https://github.com/hanig/FIRE and https://tavazoielab.c2b2.columbia.edu/FIRE/

METHOD DETAILS

Human brain samples and QC

Informed consent was obtained from all participants in accordance with the Ethics Review Boards at Sunnybrook Health Sciences Center and University of Toronto. ALS clinical diagnosis was determined based on the El Escorial revisited clinical criteria.[179] Fresh frozen orbitofrontal cortex tissues were collected from ALS cases with pathologically confirmed FTLD or no FTLD, and six non-neurological disease control cases. Demographic information for individual- and group-level (Tables S1 and S2). C9-genotypes were determined as described previously.[180,181] Postmortem patient samples comprised C9-ALS/FTLD (n = 6), C9-ALS with no FTLD (n = 3), and sALS with no FTLD (n = 8). Postmortem non-neurological control samples (n = 6) were obtained from the Douglas-Bell Canada Brain Bank (DBCBB, Montreal, Canada) (n = 4) or University Health Network–Neurodegenerative Brain Collection (UHN–NBC,

Toronto, Canada) ($n = 2$). To ensure consistency across tissue sources, all samples were collected with postmortem intervals < 24 h, dissected by experts in neuroanatomy to preserve laminar structure (layers I–VI and white matter), flash-frozen in liquid nitrogen immediately after extraction, and stored at −80°C until processing. These standardized protocols minimized variability in RNA quality despite procurement from different brain banks.

Single nucleus RNA-seq by fluorescent activated cell sorting (FACS)

Frozen orbitofrontal cortex (approximately 50 mg per sample) was dounce homogenized on ice in lysis buffer (0.32 mM sucrose, 5 mM $CaCl_2$, 3 mM $Mg(Ac)_2$, 20 mM Tris-HCl [pH 7.5], 0.1% Triton X-100, 0.5 M EDTA [pH 8.0], 40U/mL RNase inhibitor in H_2O), centrifuged at 800 × *g* for 10 min at 4°C. The supernatant was removed and the pellet was washed twice and resuspended in a resuspension buffer (1x PBS, 1% BSA, 0.2U/μL RNase inhibitor). Resuspended nuclei were sorted by FACS with DAPI (Roche) labeling, removing any debris and nuclei aggregates within DAPI-positive gating, aiming to capture approximately 6000 nuclei per sample. Library preparation was performed using either the 10X Chromium Single Cell 3′ v2 or v3 platform following the manufacturer protocols. The QC of cDNA libraries was conducted on a 2100 Bioanalyzer (Agilent). The cDNA libraries were 100-bp paired end sequenced on either an Illumina NovaSeq 6000 SP or NovaSeq6000 S2 standard flow cell at the Princess Margaret Genomic Center (Toronto, Ontario). Raw Illumina base call files from each sample were demultiplexed to produce FASTQ files with the cellranger mkfastq pipeline (10X Genomics). Reads were aligned to the pre-mRNA GRCh38-2020-A genome and quantified using cellranger count command on Cell Ranger v5.0.0 (10X Genomics).

snRNA-seq sample processing, QC, and clustering

All processing and QC of snRNA-seq samples was performed using Seurat (v4),[163–167] and custom R and Python scripts. Seurat objects were created for each sample using the filtered feature-barcode matrices obtained from Cell Ranger (v5.0.0). For each sample, nuclei containing mitochondrial reads with a threshold greater than three mean absolute deviations from the median number of mitochondrial reads with a maximum cut-off of 5% were removed. Next, nuclei with fewer than 200 and greater than 12000 detected genes were removed. Reads pertaining to the cell cycle were scored using the scran R package.[168] Potential doublets were estimated and removed using scDblFinder with default parameters.[169] The remaining singlet transcriptomes were merged and batch corrected with Harmony[49] on log1p normalized counts. Dimensionality reduction was performed using PCA on 50 principal components and then visualized using uniform manifold projection and approximation (UMAP).[182] Clustering was performed using the Leiden algorithm[50] at a resolution of 0.6. Clusters with fewer than 200 cells were filtered out as background[183] based on poor representation across samples and disease subtypes. Visualization was performed using the scCustomize[170] or dittoSeq[171] R packages.

Cluster annotation by machine learning and reference atlas

To uncover orbitofrontal cell types from the identified clusters, we employed marker discovery by machine learning (NSForest v3.9.1)[51] and reference-based annotation[52] using the Allen Brain Atlas https://portal.brain-map.org/atlases-and-data/rnaseq.[161] Default parameters in NSForest were applied to identify binary markers for each cluster. The machine learning-based marker classification approach identified binary markers for each cluster using random forest feature selection and expression scoring. These binary markers for clusters were then used to confirm cortical cell type identity using reference-based annotation. Markers from each dataset were then compared with ref. 10X Genomics experiments from two cortical regions in humans, including the primary motor cortex, dorsolateral prefrontal cortex, and medial temporal gyrus from the Seattle Alzheimer Disease Cell Atlas.[161,162] Final confirmation of cell type identity was confirmed for each cell subtype by manual annotation[52] using canonical cortical cell markers.

PCA and sample-level distance analysis

To assess transcriptional variation between diagnostic groups at the sample level, we computed pseudobulk gene expression matrices for each cell type by aggregating raw counts across cells from the same sample. After variance-stabilizing transformation using DESeq2,[113] expression matrices were adjusted to regress out sample identity as a batch variable using *removeBatchEffect* from the limma package.[172] PCA was performed on the top 10 components of the residualized expression matrix.

Pairwise Euclidean distances between samples in PCA space were calculated for each cell type, and median within-group (control vs. control, disease vs. disease) and between-group (control vs. disease) distances were summarized. To estimate confidence intervals, we performed non-parametric bootstrapping with 1,000 resamples. To evaluate whether observed between-group separation exceeded chance expectations, we conducted a permutation test by shuffling diagnosis labels 1,000 times and recalculating the median difference within- and between-groups. One-sided empirical *p*-values were calculated as the proportion of permutations where the difference in median distances (between-group minus within-group) was greater than or equal to the observed value, with a +1 continuity correction applied.

Differential gene expression analysis

To minimize technical variability and ensure robust downstream comparisons, differential expression and APA analyses were restricted to samples processed with 10X v3 chemistry. The subset of samples selected for analysis is described below. A subset of samples were processed using the 10X v2 chemistry for C9-ALS/FTLD ($n = 3$), C9-ALS no FTLD ($n = 1$), and sALS no FTLD ($n = 4$). These v2 samples were used exclusively for cell type annotation due to their lower gene detection rates, while all downstream

analyses (differential expression, APA) were restricted to v3 samples to avoid chemistry-specific biases. For these analyses, the C9-ALS/FTLD ($n = 4$) and C9-ALS without FTLD ($n = 2$) were combined to create the C9-ALS cohort ($n = 6$) to improve statistical power, as both share C9orf72 mutations causing ALS. The sALS without FTLD ($n = 4$) samples constituted the sALS cohort. These disease subtypes (C9-ALS and sALS) were then compared to the non-neurological disease control group ($n = 6$). We also compared C9-ALS/FTLD ($n = 4$) cases with all ALS no FTLD ($n = 4$ sALS no FTLD, $n = 2$ C9-ALS no FTLD) in addition to all ALS no FTLD ($n = 6$) with controls ($n = 6$). These findings are available in Table S3.

Differentially expressed genes across ALS subtypes relative to control in each cell type were uncovered using the MAST approach.[53] A mixed-effect hurdle model was employed with the counts data, analyzing log2-normalized counts as described.[53] The model was designed to include both fixed effects and the individual sample as a random effect as follows:

$$y \sim D + G + U + S + A + (1|I)$$

where fixed effects are denoted as y for the log2-normalized count of the gene, D for the diagnosis (C9-ALS/FTLD vs. control, C9-ALS vs. control, and sALS vs. control), G is the number of genes detected, U is the unique molecular index (UMI) distribution, S is the sex of the donor, A is the age of the donor, and $1|I$ is the random effect for donor ID, accounting for donor-specific variability.

The hurdle component modeled the probability of zero versus non-zero counts using logistic regression, while the count component modeled non-zero counts using a Poisson distribution. This approach enabled us to address the excess zeros in our data and assess the influence of each predictor on the gene expression counts.

Next, we performed a likelihood ratio test to identify differentially expressed genes by comparing models with and without the diagnosis term, as described.[44] Briefly, hurdle p-values were reported using the MAST method, and p-values were adjusted for multiple comparisons using the Benjamini and Hochberg FDR method. The MAST model provided LFC due to the disease effect, derived from both the continuous component (nonzero expression) and the discrete component (expressed or not) of the hurdle model. Additionally, we computed average LFC by subtracting the mean log2 counts per million of control nuclei from that of disease nuclei, ensuring consistency between model LFC and average fold change (FC), and excluding genes with significant discrepancies. We considered significance set at FDR < 0.01 and model $|\text{LFC}| \geq 0.5$. Further analysis of transcriptional signatures within cell types and across disease subtypes was performed using the R package UCell version 2.4.[173] UCell uses the Mann-Whitney U statistic to assign and rank signature scores from a set of pre-determined genes based on their relative expression in cells.

Comparison with existing ALS datasets

We retrieved pre-processed snRNA-seq data and relevant metadata from three independent ALS/FTLD studies, including Pineda et al.[45] from https://www.synapse.org/Synapse:syn51105515/files/ and https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1073234/; Li et al.[44] from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE219281, and Gittings et al.[46] from https://www.synapse.org/Synapse:syn45351388/files/. To allow for the comparative analysis and visualization of gene expression changes, we created Seurat objects for each dataset as described above using the meta-data provided by each study. For differentially expressed genes, we retrieved differential expression data from the published supplementary tables where available.[44,45,47] For each cell type, we considered significance at either an adjusted p-value or FDR < 0.01 and $|\text{LFC}| \geq 0.5$. Gene intersections were visualized using Venn diagrams.

We assessed whether gene intersections between each pair of differential gene sets was greater than expected by chance using hypergeometric tests. For a pair of differentially expressed gene sets (A and B) out of the total universe of N genes tested (all unique, expressed genes included across differential expression analyses), the background size was set to N, with $|A| = K$, $|B| = n$, and $|A \cap B| = k$. We computed one-tailed hypergeometric p-values ($p(X \geq k)$) to evaluate the enrichment of shared genes beyond random expectation. A $p < 0.05$ was considered significant. To evaluate whether the triple intersection was significantly larger than predicted by each set's size and their pairwise overlaps, we fit a log-linear model (chi-squared framework) to a $2 \times 2 \times 2$ contingency table reflecting membership or non-membership in each dataset. We compared the model containing main effects and pairwise interactions against a model that additionally included the three-way interaction term. A significant likelihood-ratio test ($p < 0.05$) indicated that the triple intersection deviated from random chance given the marginal and pairwise distributions.

APA quantification and profiling

To address the inherent sparsity of single-nucleus 3′-end data, we aggregated barcodes from each cell type into pseudo-bulk alignment files, thereby boosting coverage at PA sites and enhancing statistical power for detecting APA events. Although pseudo-bulk approaches can obscure heterogeneity in APA usage across individual cells, they remain a well-established strategy for capturing robust 3′-end usage patterns in single cell RNA-Seq data.[39,40,184,185] This is particularly appropriate in our case, as cell types were already defined through transcriptomic analysis.

We employed the MAAPER software[43] to assign sequencing reads to known PA, as defined in the PolyA DB v3 database.[96] PA sites were considered only if there were at least 25 reads aligning to the sites. To identify genes with significant changes in the length of their 3′-most exon, we used the REDu metric provided by MAAPER. REDu measures the relative expression levels between the two most differentially expressed isoforms in the 3′-most exon. A positive REDu value indicates transcript lengthening events, while a negative value points to shortening events. For pinpointing genes exhibiting intronic APA usage, we used the REDi metric. REDi

compares the relative expression levels of the top differentially expressed isoform in the 3′-most exon and the top differentially expressed isoform in an intron or internal exon. The RED score, comparing conditions 1 and 2, is computed using the formula:

$$RED = \log_2\left(\frac{\alpha_{2d}}{\alpha_{2p}}\right) - \log_2\left(\frac{\alpha_{1d}}{\alpha_{1p}}\right)$$

where the proportions of the distal and proximal PAs in conditions 1 and 2 are denoted as α_{2d}, α_{2p}, α_{1d}, and α_{1p} respectively.[43,96]

More than 80% of the genes in PolyA DB v3 database[96] have 3 or more annotated PAs. Hence, to investigate differential poly(A) site usage patterns in more detail, we used the APAlog package. APAlog operates on the normalized counts of reads mapped to each PA site to assess the extent and nature of differential usage. For a comprehensive comparison, APAlog was run in Pairwise Test mode, which enables the comparison of all possible pairs of PA sites per transcript.[116] Similarly, the APA LFC values were calculated to identify distal and proximal PA site usage:

$$APA_{LFC} = \log_2\left(\frac{\beta_{2d}}{\beta_{2p}}\right) - \log_2\left(\frac{\beta_{1d}}{\beta_{1p}}\right)$$

where the proportions of the distal and proximal PAs in conditions 1 and 2 are denoted as β_{2d}, β_{2p}, β_{1d}, and β_{1p} respectively.

Empirical FDR analysis for APA

We conducted an empirical FDR analysis using the controls from individuals without neurological diseases to mitigate false positives. This enabled us to establish cell type-specific FDR thresholds where only 10% of the *p*-values fell below each threshold (Table S8). For downstream analysis, we included only transcripts that passed this FDR analysis, ensuring a reliable FDR of 10% or less in our results.

APA analysis was performed on all possible pairs of control vs. control samples, and transcript-level APA *p*-values were examined. The threshold was selected based on the 10% quantile. For further analysis, we retained only significant APA events after applying the Benjamini-Hochberg FDR correction to the RED values from MAAPER. Similarly, APAlog was used to identify transcripts with significant APA, and the Benjamini-Yekutieli method was applied to correct the *p*-values of APA LFC values due to the mutual non-independence of *p*-values from testing pairs of PA sites in transcripts with more than two sites.

Pathway enrichment analysis

Pathway enrichment analyses for snRNA-seq data, including from differentially expressed genes and APA, were performed using gProfiler2[54] with curated Gene ontology (GO) biological process and cellular compartment gene sets with no inferred electronic annotation were downloaded from http://baderlab.org/GeneSets/(April 2024 release). We used a minimum gene set size of 15 and the maximum gene set size of 200. Visualization of GO results focused on minimizing GO term redundancy by plotting enrichment maps created with the Enrichment Map plugin for Cytoscape (v3.9.1) in either Linux or Windows.[174] A *q*-value cutoff < 0.1 and edge cutoff (similarity) of 0.375 was used for plotting either the differentially expressed genes or APA enrichment map results.

Deep learning model architecture

APA-Net architecture is designed to map the complex regulatory mechanisms underlying APA from 3′ single-cell transcriptomics data. Hybrid models that combine convolutional neural networks (CNNs) with architectures traditionally developed for natural language processing have emerged as powerful tools in genomic sequence analysis.[101,186] APA-Net employs a CNN architecture supplemented with a multi-head attention module, specifically optimized to identify *cis*-regulatory elements impacting PA site selection across varied cell types. The input region for the CNN and MAT modules encompasses 2 kb surrounding both the proximal and distal polyadenylation site (PAS). The CNN module contains a single convolutional layer, comprising 128 kernels, each with a size of 12 and a stride of 1. This is followed by a max pooling layer with a kernel size of 16 and stride of 16. The output from this stage feeds into a multi-head attention module, in which each position in the representation map functions as a distinct token. In this context, a 'token' refers to a discrete unit of information, which is essential for the attention mechanism to effectively process and interpret the complex patterns within the data. A residual connection links the CNN and multi-head attention modules. This residual connection, a key component in deep learning architectures, helps in mitigating the vanishing gradient problem by allowing the flow of information and gradients directly across layers. The attention module's output, along with the RBP expression profile specific to the cell type, is forwarded to a multi-layer perceptron for final APA effect prediction through a regression task.

The objective of APA-Net therefore is to predict the APA LFC values and the loss for regression can be represented as:

$$L_{\hat{x}} = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mathbf{x}}_i - \mathbf{x}_i)^2$$

To train APA-Net, we utilized the pairwise APA events identified by APAlog across different cell types. For each APA event, we extracted genomic sequences from the hg38 reference genome surrounding both the proximal and distal PAS, transcribed them to RNA sequences, and oriented them in the sense direction to preserve the proximal/distal usage information. These sequences served as input features, along with the cell type specific RBP expression profiles. The corresponding APA log fold change (LFC)

Please cite this article in press as: McKeever et al., Single-nucleus transcriptome atlas of orbitofrontal cortex in ALS with a deep learning-based decoding of alternative polyadenylation mechanisms, Cell Genomics (2025), https://doi.org/10.1016/j.xgen.2025.101007

values were used as regression labels. We employed 5-fold cross-validation, where for each fold, the data was split into 80% training and 20% testing sets while maintaining the representation of APA events across different cell types. This cross-validation approach ensured robust evaluation of the model's performance across different subsets of the data.

Deep learning model interpretation

To identify potential *cis*-regulatory elements, we used the filter weights from the CNN module of APA-Net, which represent learned sequence motifs. Each sequence within the test dataset was scanned using the model's filters to identify regions with maximal filter activation, hypothesized to correspond to biologically relevant RNA sequence motifs. These motifs were modeled as Position-Weight Matrices (PWMs) to capture the patterns each filter had learned for APA sequences in ALS.

Next, we aligned RBPs from the compendium of RNA-binding motifs[103] with the motifs identified by APA-Net. This alignment enabled us to associate the learned motifs with known RBPs. To identify differentially expressed RBPs, we employed DESeq2.[113] This pseudo-bulk approach enabled the preservation of biological relevance and expected effect sizes of altered RBPs while mitigating cell type variability.[187–189] We used the *FindMarkers* function from Seurat with the list of identified RBPs to perform the analysis. We then applied stringent statistical cutoffs of adjusted $p < 0.01$ and used a $|\text{LFC}| < 0.25$ to capture a large net of dysregulated RBPs.

HNRNPC eCLIP-seq processing and enrichment analysis

For eCLIP-seq data processing, reads were initially processed using UMI-tools to extract unique molecular identifiers, followed by quality filtering ($Q > 15$) and adapter removal using cutadapt.[175] The processed reads were aligned to the human genome (hg38) using BWA (v0.7.17). PCR duplicates were removed and peaks were called using the CLIP Tool Kit (CTK v1.1.13)[177] implementing a valley-seeking algorithm with multiple testing correction. Peak boundaries were determined by combining replicates. We used FIRE[190] for binding site enrichment analysis.

RBP interaction dissimilarity across cell types

We used the Frobenius norm, a measure of the difference between two matrices (similar to Euclidean distance for vectors), defined as the square root of the sum of the absolute squares of their element-wise differences, to measure the dissimilarity of the RBPs interaction profiles across cell types. The Frobenius distance between two matrices X and Y is given by:

$$d(X,Y) = \|X - Y\|_F = \sqrt{\text{tr}((X - Y)^\top (X - Y))}$$

where.

(1) $\text{tr}(A)$ denotes the trace of matrix A, which is the sum of its diagonal elements.
(2) $A^\top$ denotes the transpose of matrix A.
(3) $\|\cdot\|_F$ denotes the Frobenius norm.