



pTDP-43 levels correlate with cell type–specific molecular alterations in the prefrontal cortex of *C9orf72* ALS/FTD patients

Hsiao-Lin V. Wang^{a,b}, Jian-Feng Xiang^a, Chenyang Yuan^{a,c}, Austin M. Veire^d, Tania F. Gendron^d, Melissa E. Murray^d , Malú G. Tansey^{e,f} , Jian Hu^{a,c}, Marla Gearing^{b,g,h}, Jonathan D. Glass^{b,g} , Peng Jin^{a,b} , Victor G. Corces^{a,b,1} , and Zachary T. McEachin^{a,b}

Affiliations are included on p. 11.

Contributed by Victor G. Corces; received October 1, 2024; accepted January 24, 2025; reviewed by Udal Pandey and Yin Shen

Repeat expansions in the *C9orf72* gene are the most common genetic cause of amyotrophic lateral sclerosis and familial frontotemporal dementia (ALS/FTD). To identify molecular defects that take place in the dorsolateral frontal cortex of patients with *C9orf72* ALS/FTD, we compared healthy controls with *C9orf72* ALS/FTD donor samples staged based on the levels of cortical phosphorylated TAR DNA binding protein (pTDP-43), a neuropathological hallmark of disease progression. We identified distinct molecular changes in different cell types that take place during FTD development. Loss of neurosurveillance microglia and activation of the complement cascade take place early, when pTDP-43 aggregates are absent or very low, and become more pronounced in late stages, suggesting an initial involvement of microglia in disease progression. Reduction of layer 2-3 cortical projection neurons with high expression of CUX2/LAMP5 also occurs early, and the reduction becomes more pronounced as pTDP-43 accumulates. Several unique features were observed only in samples with high levels of pTDP-43, including global alteration of chromatin accessibility in oligodendrocytes, microglia, and astrocytes; higher ratios of premature oligodendrocytes; increased levels of the non-coding RNA NEAT1 in astrocytes and neurons, and higher amount of phosphorylated ribosomal protein S6. Our findings reveal progressive functional changes in major cell types found in the prefrontal cortex of *C9orf72* ALS/FTD patients that shed light on the mechanisms underlying the pathology of this disease.

neurodegeneration | amyotrophic lateral sclerosis | frontotemporal dementia | chromatin | epigenetics

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are progressive neurodegenerative disorders characterized by the loss of neuronal cell populations in the central nervous system (CNS). In ALS, upper motor neurons in the primary motor cortex and lower motor neurons in the spinal cord degenerate, leading to paralysis and respiratory failure (1). FTD is a heterogenous disorder characterized by degeneration of the frontal and temporal cortex leading to progressive cognitive impairments (2). Despite being symptomatically distinct, ALS and FTD have clinical, genetic, and neuropathological overlap, suggesting that these two disorders lie on a disease continuum. In support of this, a G₄C₂ hexanucleotide repeat expansion in the gene *C9orf72* was identified as the most common cause of both ALS and FTD (3).

A neuropathological hallmark of ALS and FTD is the phosphorylation, mislocalization, and aggregation of TAR DNA binding protein 43 (TDP-43) (4). TDP-43 is a ubiquitously expressed, nuclear RNA/DNA-binding protein that performs important functions associated with RNA metabolism, including alternative splicing and mRNA stability. Present in approximately 95% of ALS cases and 50% of FTD cases, pTDP-43 burden has been shown to correlate with degeneration of affected cell populations in both ALS and FTD (5, 6). In *C9orf72* carriers specifically, semiquantitative analyses suggest that the extent of TDP-43 pathology in an affected CNS region correlates with clinical phenotypes (7). Given the variability in symptom onset in ALS vs FTD and discordant timing of clinical progression despite a shared genetic etiology, postmortem samples from *C9orf72* ALS/FTD donors with quantitative measurements of pTDP-43 abundance provide a unique opportunity to identify molecular cascades that promote and/or result from TDP-43 dysfunction. Here, we use multiome single-nucleus analysis of postmortem human brain cortex tissue from *C9orf72* ALS/FTD patients and cognitively healthy age-matched controls to gain insights into the cellular and molecular events altered in different cell types and stages of disease progression based on pTDP-43 abundance. The results suggest a sequential cascade of alterations in the regulatory landscape of *C9orf72* ALS/FTD, highlighting the contribution of pTDP-43

Significance

Changes in the brain of patients with neurodegenerative disorders during disease progression are challenging to study because of difficulties in obtaining brain samples. Here, we analyze changes in the frontotemporal cortex in the brain of patients with ALS/FTD who die of ALS but present with cognitive symptoms of progressive severity that correlate with the levels of the protein pTDP-43. We find that microglia are affected early in the disease, accompanied by changes in the neurons of specific layers of the cortex. Changes in oligodendrocytes and astrocytes take place later. We identify alterations in the expression of specific genes that accompany changes in cellular function. These findings suggest a cascade of events responsible for neurodegeneration.

Author contributions: H.-L.V.W., V.G.C., and Z.T.M. designed research; H.-L.V.W., J.-F.X., A.M.V., T.F.G., M.E.M., and Z.T.M. performed research; H.-L.V.W., J.-F.X., C.Y., M.G.T., M.G., J.D.G., and P.J. contributed new reagents/analytic tools; H.-L.V.W., C.Y., J.H., and Z.T.M. analyzed data; and H.-L.V.W., V.G.C., and Z.T.M. wrote the paper.

Reviewers: U.P., University of Pittsburgh; and Y.S., Institute for Human Genetics and Department of Neurology - University of California, San Francisco.

The authors declare no competing interest.

Copyright © 2025 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹To whom correspondence may be addressed. Email: vgcorges@gmail.com or zmceach@emory.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2419818122/-DCSupplemental>.

Published February 25, 2025.

accumulation during the progression of neurodegeneration in cell types affected in FTD.

Results

Multiome Analysis of the Human Dorsolateral Prefrontal Cortex from *C9orf72* ALS/FTD Donors. To investigate molecular changes in specific cell types of the dorsolateral prefrontal cortex of *C9orf72* ALS/FTD patients with different burden of pTDP-43, we utilized a multiome approach to simultaneously analyze the transcriptome and epigenome of single nuclei from Brodmann area 9 of postmortem human brain tissue from *C9orf72* ALS/FTD donors ($n = 19$) and age/sex-matched healthy controls with normal cognitive and motor function ($n = 7$) (SI Appendix, Fig. S1A and Dataset S1). The average ages are 71 and 69 for control and *C9orf72* ALS/FTD donors, respectively. Fourteen samples were obtained from the Goizueta Emory Alzheimer's Disease Center Brain Bank (Emory cohort) and 12 samples were obtained from the Mayo Clinic Brain Bank (Mayo cohort). All *C9orf72* ALS/FTD donors have clinical diagnoses of ALS and/or FTD, with neuromuscular abnormalities and different degrees of cognitive impairment. Additional information for each case including age, sex, and copathologies is listed in Dataset S1. Quantitative measurements of pTDP-43 abundance were performed on all samples using Meso Scale Discovery immunoassay in lysates from the same cortical tissue used for multiome analysis. *C9orf72* ALS/FTD donors were then grouped into terciles, referred to as TDPneg, TDPMed, and TDPhigh, based on pTDP-43 levels, which were confirmed using immunohistochemistry for each TDP donor group (SI Appendix, Fig. S1B and C and Dataset S1). We assigned samples with relative pTDP-43 levels of less than 100 to the pTDPneg group, samples with levels between 100 and 2,000 to the pTDPMed group, and samples with levels above 2,000 to the pTDPhigh group. On average, the pTDPhigh samples had pTDP-43 levels at least 100-fold higher than the pTDPneg samples and at least 4-fold higher levels than the pTDPMed samples. The presence of cytoplasmic pTDP-43 aggregates is the defining neuropathological hallmark of FTLD-TDP and it has been reported to associate with more rapid cognitive decline and often found in patients with dementia (8). We also measured levels of the dipeptide repeat proteins poly(GA) and poly(GR) (Dataset S1), which are known to be produced from *C9orf72* repeats by non-AUG translation. We did not observe correlations between the levels of poly(GA) and poly(GR) and the levels of pTDP-43.

We analyzed the multiome datasets from all 26 samples but observed a strong batch effect linked to the brain bank of origin. Therefore, we separately analyzed the samples from each cohort and performed parallel analyses on each set. We obtained a total of 34,874 and 53,331 single-nucleus multiomes (snRNA-seq and snATAC-seq) from the Emory and Mayo cohorts, respectively, after quality control filtration using the ArchR multiome pipeline (9) and Seurat snRNA-seq guidelines (10) (Methods and SI Appendix, Figs. S1D, S2 and S3 and Dataset S2). Dimensionality reduction was performed for each snRNA-seq and snATAC-seq dataset using the ArchR optimized iterative LSI method (9) and batch effect correction was performed using Harmony (11). Uniform manifold approximation and projection (UMAP) and unsupervised clustering (12) of the combined snATAC-seq and snRNA-seq datasets identified 31 cell clusters for Emory samples and 20 for Mayo samples, excluding unassigned clusters (SI Appendix, Figs. S1D and S4). We used SoupX (13) to detect ambient RNA contamination and found it to be negligible (SI Appendix, Fig. S5A). Gene activity scores derived from snATAC-seq chromatin accessibility at proximal promoter regions were used to identify marker genes

in each cell cluster (SI Appendix, Figs. S1 D–F and S3 B–D and Dataset S3). A total of seven major cortical cell types were identified for both cohorts (SI Appendix, Fig. S1 D–F and S4). We subclustered each major cell type and identified excitatory neurons (EX; 6,576 nuclei and 9 clusters for the Emory cohort; 2,570 nuclei and 5 clusters for the Mayo cohort), inhibitory neurons (IN; 4,501 nuclei and 7 clusters for the Emory cohort; 1,407 nuclei and 5 clusters for the Mayo cohort), astrocytes (ASC; 4,437 nuclei and 6 clusters for the Emory cohort; 4,165 nuclei and 3 clusters for the Mayo cohort), microglia (MG; 3,255 nuclei and 4 clusters for the Emory cohort; 2,580 nuclei and 2 clusters for the Mayo cohort), oligodendrocytes (ODC; 1,2746 nuclei and 4 clusters for the Emory cohort; 29,789 nuclei and 4 clusters for the Mayo cohort), oligodendrocyte progenitor cells (OPC; 2,595 nuclei; 3 clusters for the Emory cohort; 1,270 nuclei and 2 clusters for the Mayo cohort), and endothelial cells (ENDO; 337 nuclei; 1 cluster for the Emory cohort; 282 nuclei and 1 cluster for the Mayo cohort) (SI Appendix, Fig. S4 and Dataset S2B). The cell type identification was verified by a module score composed of known cell-type specific marker genes (SI Appendix, Fig. S4D). Notably, the Mayo samples have fewer neuronal nuclei (SI Appendix, Fig. S4E), which could explain the strong batch effect observed when attempting to integrate samples from both cohorts. Interestingly, oligodendrocytes exhibited the highest similarity in marker genes between the two datasets. Using pseudobulk and PCA analysis for each cell type, we found that astrocytes, microglia, and oligodendrocyte lineage cells exhibited high reproducibility between the two cohorts, whereas neurons did not (SI Appendix, Fig. S5F). We then performed pseudobulk per sample correlation analysis and PCA analysis of pseudo-snRNA-seq data for both cohorts (SI Appendix, Fig. S5 B and C) and confirmed a strong batch effect between Emory and Mayo samples. However, all samples show strong correlations within each sample group. We also conducted additional analyses to correlate marker genes across both datasets, revealing a high degree of similarity (SI Appendix, Fig. S5D). In the rest of the manuscript, we treat the Emory cohort as the primary dataset and use results from the Mayo cohort to test the validity of the major findings, except for neurons, which are present in low numbers in the Mayo samples. Subclusters within each cell type were named separately for each cohort.

Chromatin Accessibility Is Altered in Oligodendrocyte Lineage Cells, Microglia, and Astrocytes from *C9orf72* ALS/FTD Donors with High Levels of pTDP-43. To address whether the transcriptome and chromatin accessibility are altered progressively and correlate with pTDP-43 levels in each cell type of *C9orf72* ALS/FTD cortex tissue, we performed systematic pair-wise comparisons of gene expression and chromatin accessibility between controls and different groups of *C9orf72* ALS/FTD donors in each of the seven major cell types observed in the Emory cohort (Fig. 1 and Dataset S4). A total of 404,124 reproducible 501 bp peaks of chromatin-accessible regions were identified in the snATAC-seq dataset. Utilizing the pseudobulk method, we aggregated fragment counts per sample-cell type combination and analyzed them using DESeq2 (14) with multifactor design to assess the significance of differentially accessible regions between control and *C9orf72* ALS/FTD samples with varying levels of pTDP-43. We identified 3500 differentially accessible regions (DARs) (Fig. 1A and Dataset S4A). The majority of DARs were observed when comparing control samples with pTDPhigh samples (Fig. 1A), suggesting that alteration of transcription factor binding might be a hallmark of late disease stages. Surprisingly, DARs are more frequently found in nonneuronal cells, primarily in oligodendrocytes (Fig. 1B). The frequent observation of oligodendroglial cytoplasmic

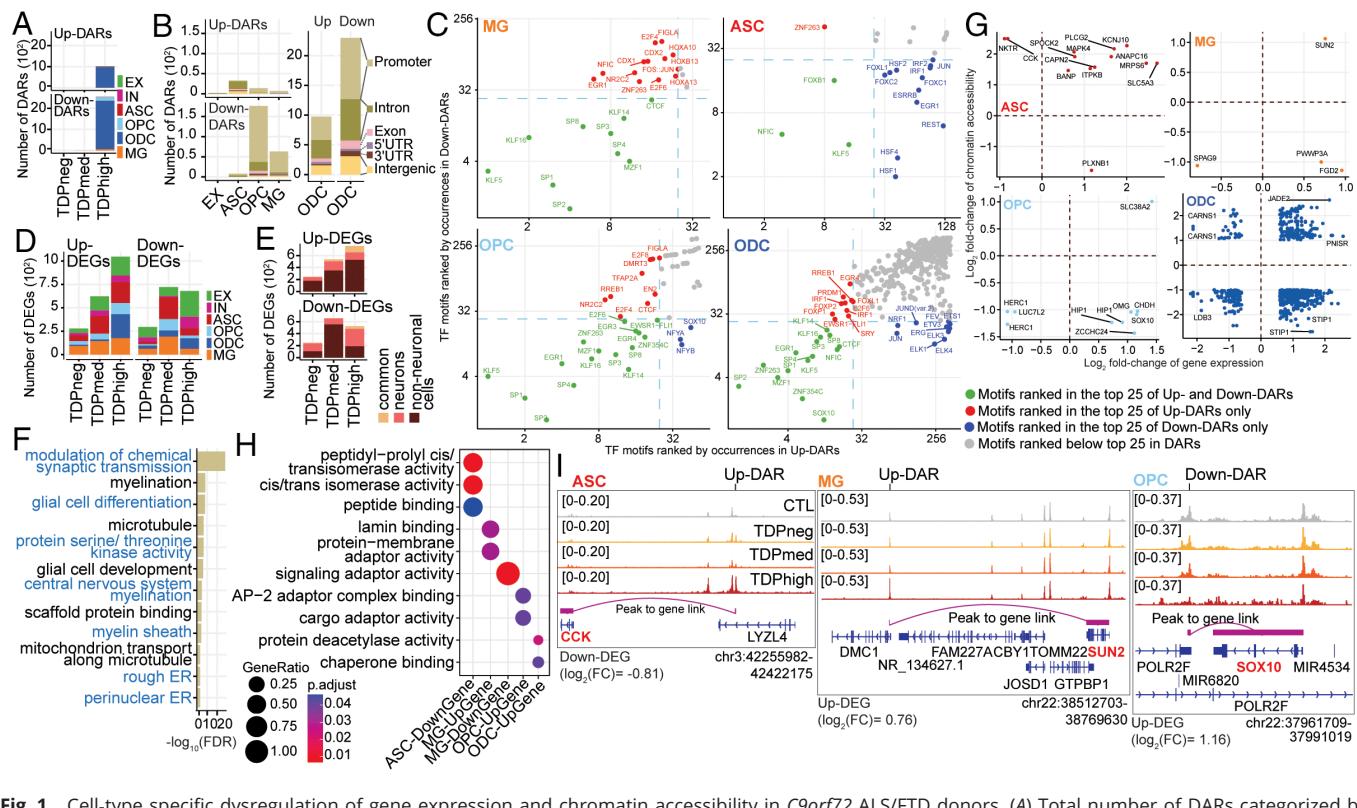


Fig. 1. Cell-type specific dysregulation of gene expression and chromatin accessibility in *C9orf72* ALS/FTD donors. (A) Total number of DARs categorized by cell type and pTDP-43 donor group. (B) Total number of DARs found in pTDPhigh donor group categorized by cell type (Left) and distribution of DARs related to functional annotations in oligodendrocyte lineage cells (Right). (C) Occurrence of Transcription Factor (TF) motifs in DARs in nonneuronal cells found in the pTDPhigh donor group ranked by the frequency of occurrence. The TF with motif occurrences ranked highest in both *Up-* and *Down-*DARs are highlighted in green in the Bottom Left quadrant of each panel. TFs with motif occurrences ranked highest only in Up-DARs are marked in red and displayed in the Top Left quadrant, while TFs with motif occurrences ranked highest only in Down-DARs are marked in blue and displayed in the Bottom Right quadrant (D) Number of DEGs categorized by cell type and pTDP-43 donor group. (E) Number of differentially expressed genes common among neuronal and nonneuronal cells for each pTDP-43 donor group. (F) Gene Ontology enrichment analysis of genes located near DARs found in the pTDPhigh donor group. (G) DEGs with a linked DAR in astrocytes, microglia, and oligodendrocyte lineage cells. (H) Gene Ontology enrichment analysis of DEGs with linked DARs as shown in panel (G). (I) Genome track visualization of the *CCK* (chr3:42,255,982-42,422,175), *SUN2* (chr22:38,512,703-38,769,630), and *SOX10* (chr22:37,961,709-37,991,019) loci in astrocytes, microglia, and oligodendrocyte precursor cells (OPCs), respectively.

pTDP-43 inclusions in *C9orf72* brain tissues with FTLD-TDP (15) suggests the possibility that changes in chromatin accessibility observed in oligodendrocytes with high levels of cortical pTDP-43 could be due to direct or indirect effects of pTDP-43 accumulation and/or reduction of nuclear TDP-43. Motifs associated with transcription factors involved in cell differentiation were prevalent at these differentially accessible sites (Fig. 1C). Notably, we identified motif sequences for EGR1, KLF5, and ZNF263 in DARs found in all nonneuronal cell types (Fig. 1C). In contrast, NFIC motif sequences were predominantly present in terminally differentiated glial cells (MG, ASC, ODC), but not in OPCs. Interestingly, CTCF emerged as one of the prominent TF motifs in DARs present in microglia, OPCs, and ODCs. Furthermore, SOX10, crucial for oligodendrocyte specification, was enriched in DARs identified in oligodendrocyte lineage cells. These findings suggest that the commonly observed abnormalities in glial cells, such as astrogliosis, microglial dysfunction, and oligodendrocyte dysregulation, may arise from widespread changes in transcription factor occupancy in *C9orf72* ALS/FTD.

We then performed a comparison of transcriptomes to identify alterations between control and *C9orf72* samples. Initially, when analyzing the snRNA-seq and snATAC-seq datasets separately using UMAP, a distinctive batch effect surfaced in the snRNA-seq data, whereas the paired snATAC-seq data remained unaffected (SI Appendix, Fig. S3 A and B, Top panels). To mitigate this issue, we applied Harmony (11) batch correction, adjusting for sample,

groups categorized by pTDP-43 levels, and sample preparation batch covariates, effectively eliminating the batch effect in both individual datasets and the combined datasets (SI Appendix, Fig. S3 A–C, Bottom panels). For robustness and to reduce false positives in gene expression analysis, we employed the linear mixed-effect model implemented in MAST (16). This allowed us to meticulously consider technical covariates using the generalized mixed-effect models and model cells individually using the two-hurdle model implemented in MAST (Methods). Differentially expressed genes were identified across all *C9orf72* ALS/FTD donor groups and cortical cell types, with the greatest number observed in the TDPhigh donor group (Fig. 1D and Dataset S4B). These findings suggest that both the transcriptome and epigenome are most affected during the late disease stages characterized by high levels of pTDP-43. Although numerous differentially expressed genes are altered in both neurons and nonneuronal cells, a higher number of genes show differential expression in nonneuronal cells (Fig. 1D and E), consistent with the findings from the differentially accessible region analysis. These differentially expressed genes are involved in synaptic transmission, myelination, and encode for microtubule proteins. (Fig. 1F and Dataset S4C).

To leverage the paired multiomics datasets, we analyzed genome-wide peak-to-gene links utilizing the integrated single-nucleus RNA-seq and ATAC-seq data captured simultaneously in our study. The strength and specificity of each link are determined by the correlation between chromatin accessibility

levels at a given peak and gene expression levels for a specific gene in a single cell (Fig. 1*J*). Therefore, a peak and a gene are considered linked if both are altered in the same cell. This approach enabled us to assess how DARs, presumably present at transcriptional regulatory elements, might influence the differential expression of genes in a cell type–specific manner. We found a number of DARs linked to DEGs (Fig. 1*G*). However, not all DARs have a correlated DEG, possibly because DEGs were identified using a more stringent hurdle model-based analysis considering the generalized linear mixed effect. Specifically, we found enrichment of differentially expressed genes with linked DARs involved in peptide binding, lamin binding, and chaperone binding in non-neuronal cells (Fig. 1*H*). For instance, the downregulation of the *CCK* gene is associated with an up-regulated DAR in TDPhigh samples in astrocytes, the upregulation of *SUN2* is associated with an up-regulated DAR in TDPhigh samples in microglia, and the upregulation of *SOX10* is associated with a downregulated DAR in TDPhigh samples in oligodendrocytes (Fig. 1*I*). This analysis

suggests that DARs may act as transcriptional enhancers or silencers in different genomic contexts and cell types.

C9orf72 ALS/FTD Is Associated with Impaired Oligodendrocyte Maturation in Late Disease Stages. Seven distinct cell populations were identified in the oligodendrocyte lineage in the Emory cohort ($n = 15,341$ nuclei), including OPCs and differentiated oligodendrocytes (ODCs) (Fig. 2 *A* and *B* and *SI Appendix*, Fig. S4). Oligodendrocyte lineage cells are also the largest cell population in the Mayo cohort (*SI Appendix*, Fig. S1D), where six distinct cell populations were identified (Fig. 2*G*). Clusters in each cohort were named independently; for example, the Emory ODC-1 cluster is not the same as Mayo ODC-1. Oligodendrocytes function in the central nervous system by establishing the myelin layer and providing metabolic support to neurons. Importantly, gray matter demyelination has been observed in the motor cortex and the spinal cord of ALS patients (17), and pTDP-43 inclusions in oligodendrocytes are a characteristic neuropathological finding

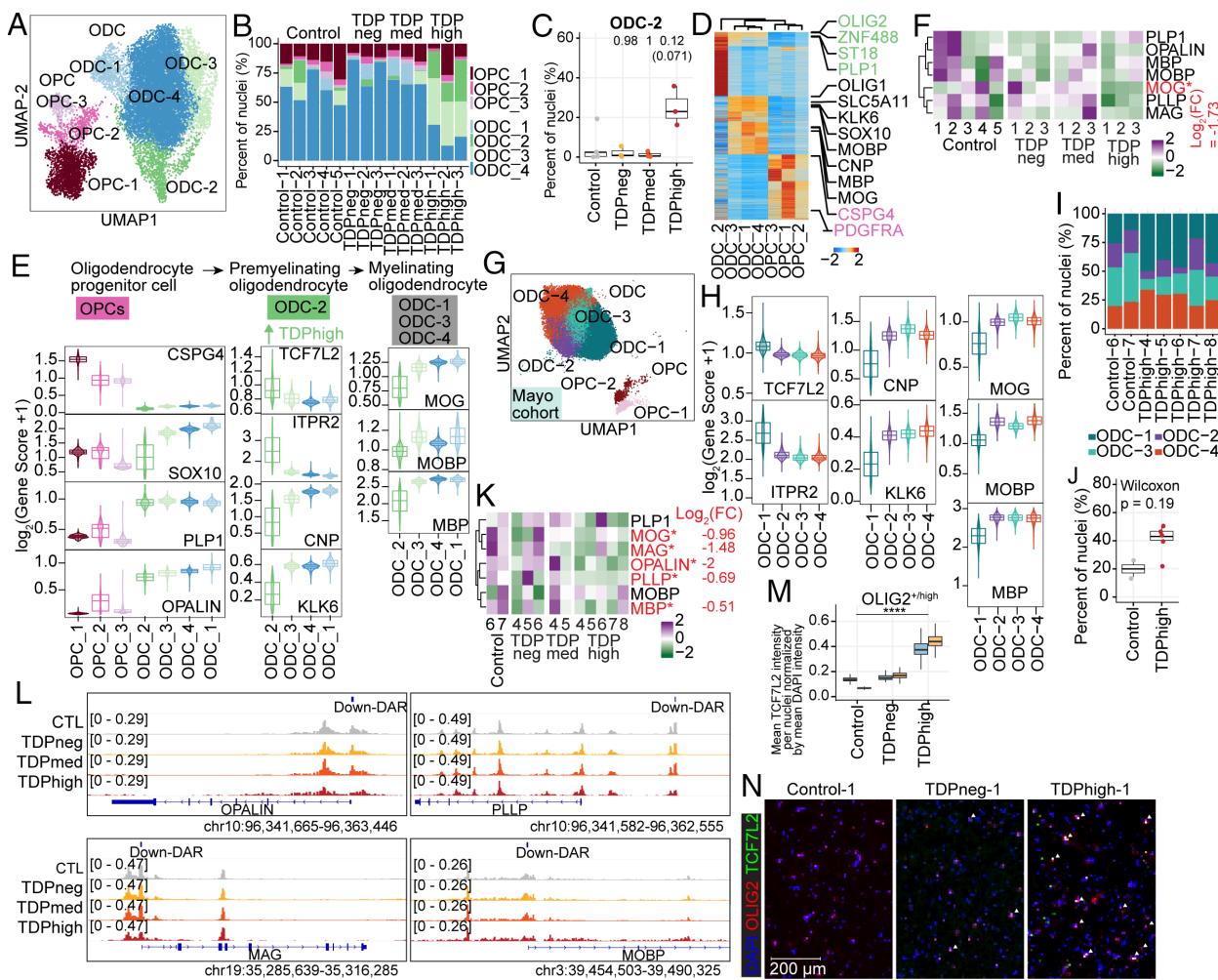


Fig. 2. Premature and premyelinating oligodendrocytes are unique to high pTDP-43 donors in late disease stages. (A) UMAP plot of oligodendrocyte lineage cells for Emory samples. (B) Proportion of OPCs and oligodendrocytes (ODC) clusters in each sample, including donors with different levels of pTDP-43 and cognitively normal controls of the Emory cohort. (C) Proportion of ODC-2 in different Emory cohort pTDP-43 donor groups (Kruskal-Wallis test with Benjamini-Hochberg correction (P -adj=0.0771) and without correction (P -value=0.0129). (D) Plot of snATAC-seq gene scores ordered by hierarchical clustering with marker genes distinguishing each ODC cell cluster for Emory cohort samples. (E) Illustration of developmental stages of oligodendrocyte lineage cells for Emory samples. Developmental stage-specific genes and their gene scores are shown for each cluster (Bottom), highlighting the unique characteristics of ODC-2 with high expression of premyelinating oligodendrocyte genes. (F) Average expression of myelin-associated genes in Emory samples. (G) UMAP plot of oligodendrocyte lineage cells for Mayo cohort samples. (H) ODC-1 in Mayo samples exhibit high expression of premyelinating oligodendrocyte genes, similar to ODC-2 in Emory samples. (I) Proportion of ODC-1 in Mayo control and TDPhigh samples (P -value=0.19, Wilcoxon rank sum). (J) Proportion of ODC clusters in Mayo cohort control and TDPhigh samples. (K) Average expression of myelin-associated genes in Mayo samples. (M) Quantification of mean TCF7L2 intensity per OLIG2-high nucleus normalized by mean DAPI intensity after nuclei segmentation. One-way ANOVA, **** P value < 0.0001. (N) Immunostaining of human postmortem cortical tissue for the oligodendrocyte lineage marker OLIG2 (red), premature oligodendrocyte marker TCF7L2 (green) and DAPI in blue. Overlapping OLIG2 and TCF7L2 staining were marked with white arrowhead.

in brains of *C9orf72* ALS/FTD patients (15). These observations suggest that oligodendrocyte dysfunction plays an important role in *C9orf72* ALS/FTD. Emory clusters OPC-1, OPC-2, and OPC-3 contain OPCs with high expression of *PDGFRA* and *CSPG4* (Fig. 2 D and E and *SI Appendix*, Fig. S6E). The remaining four clusters are differentiated oligodendrocytes with higher levels of *OPALIN* and *PLP1*. We noticed that there are proportionally more cells in Emory ODC-C2 and Emory ODC-C3 in TDPhigh compared to other donor groups and controls (Fig. 2B). Specifically, an average of 25% of oligodendrocyte lineage cells in TDPhigh donors are found in the Emory ODC-2 cluster (Fig. 2C), and less than 2% of oligodendrocyte lineage cells are present in this cluster in TDPmed and TDPneg donors. This suggests that Emory ODC-2 is unique to the late disease stages with high pTDP-43 burden. Based on gene score and hierarchical clustering of marker genes, Emory ODC-C2 cells are transcriptionally distinct from other Emory ODC clusters (Fig. 2D), with lower expression of *MOG*, *MOBP*, and *MBP*, which encode for the major protein components of myelin (Fig. 2E and *SI Appendix*, Fig. S6E). Emory ODC-2 cells also have higher expression levels of *TCF7L2* and *ITPR2*, and lower expression of *CNP* and *KLK6* (Fig. 2E and *SI Appendix*, Fig. S6E). Expression of these two genes in ODCs is an indication of newly differentiated premyelinating oligodendrocytes, which is typically a transient stage during adult oligodendrogenesis that survives 2 d in the adult mouse brain (18). Most of these cells undergo apoptosis while some survive and mature into myelinating oligodendrocytes (19). Therefore, these results suggest that Emory ODC-2 cells represent newly formed premyelinating oligodendrocytes that should not typically be present in high ratios in the adult brain, suggesting that either they failed to enter programmed cell death or to proceed into maturation. In contrast, the rest of ODC clusters are composed of mature myelinating ODCs with strong expression of genes involved in myelinating processes (Fig. 2E). Emory ODC-3 is also present in large proportion in pTDP-43 high samples (Fig. 2B); although cells in this cluster express high levels of myelination genes, their levels are slightly lower compared to Emory ODC-1 and ODC-4. These normal mature myelinating ODCs with high expression of myelinating genes are found mainly in TDPmed and TDPneg donor groups in earlier disease stages and lower in pTDP-43 high samples (Fig. 2B). Interestingly, compared with oligodendrocytes from control donors, oligodendrocytes from the TDPhigh donor group exhibit downregulation of *MOG*, the myelin oligodendrocyte glycoprotein (Fig. 2F and *Dataset S4*). When we analyzed oligodendrocyte lineage clusters from the Mayo cohort (Fig. 2G), the Mayo ODC-1 cluster exhibits the same premature premyelinating oligodendrocyte markers as the Emory ODC-2 cluster, with high expression of *TCF7L2* and *ITPR2* and lower expression of *CNP* and *KLK6* compared to other clusters (Fig. 2H). There are also proportionally more cells in the Mayo ODC-1 cluster in TDPhigh compared to controls samples (Fig. 2 I and J) and oligodendrocytes from the Mayo TDPhigh donor group exhibit downregulation of the same myelin-associated genes (Fig. 2K and *Dataset S4*). The Mayo ODC-3 cluster expresses high level of myelination genes, similar to what was found in the Emory ODC-1 cluster. These findings, observed in the two cohorts studied, further strengthen the conclusion that a large portion of oligodendrocytes in the dorsolateral prefrontal cortex in late FTD disease stages with high pTDP-43 burden remain in the typically transient premyelinating stage and defective in myelination. We also found down-regulated DARs located near the promoter region of genes involved in myelination (Fig. 2L and *Dataset S4*), including *OPALIN*, *MAG*, *PLLP*, and *MOBP*.

These findings suggest that downregulation of these genes could be a direct consequence of cytoplasmic accumulation of pTDP-43 accompanied by its loss from the nucleus in the late disease stage.

To validate the high abundance of premature oligodendrocytes found in the TDPhigh samples based on results from the single-cell analysis, we performed immunofluorescence microscopy for *TCF7L2*, *OLIG2*, and *NeuN* using control, TDPneg, and TDPhigh samples from the Emory cohort. Nuclei were segmented and classified based on *OLIG2* and *NeuN* staining intensities into “*OLIG2-high*”, “*NeuN-high*”, or “other” categories (*SI Appendix*, Fig. S6 A–C and Methods). *OLIG2-high* nuclei represent cells in oligodendrocyte lineages. Analysis of DAPI-normalized *TCF7L2* intensity in *OLIG2-high* nuclei revealed significantly higher *TCF7L2* signals in TDPhigh samples compared to control and pTDPneg samples (Fig. 2M). Examination of images from upper cortical regions confirmed the quantification results, showing a higher number of overlapping *TCF7L2*-positive and *OLIG2*-positive nuclei (Fig. 2N). To confirm these findings, we employed a second approach using TrueBlack Lipofuscin Autofluorescence Quencher on paraformaldehyde-fixed floating tissue sections, followed by confocal imaging. Consistently, we found a higher number of *OLIG2+TCF7L2+* nuclei in TDPhigh samples compared to control and TDPneg samples (*SI Appendix*, Fig. S6D). These independent validation methods confirm an overabundance of *OLIG2+TCF7L2+* nuclei in the gray matter region of pTDPhigh samples, reinforcing the single-nucleus analysis findings. The identification of highly abundant premature oligodendrocytes specifically in pTDPhigh samples presents a unique insight into *C9orf72* ALS/FTD disease progression and TDP-43 proteinopathy. We speculate that the presence of pTDP-43 aggregates, or the correlated absence of nuclear TDP-43, may play a direct role in dysregulating mRNAs encoding myelin components, thus affecting the maturation of oligodendrocytes and their ability to myelinate neurons.

Loss of Frontal Cortical Neurosurveillance Microglia Is a Hallmark of Both Early and Late Stages of *C9orf72* ALS/FTD and Is Also Prevalent in Early and Late AD. Microglia typically account for 5% of all brain cells and have the highest expression of *C9orf72* compared to other cortical cell types (20). We also find that the gene score activity and gene expression of *C9orf72* are highest in microglia compared to other cortical cell types, in samples from both the Emory and Mayo cohorts (*SI Appendix*, Fig. S4F). We first analyzed microglia from the Emory cohort; 4 cortex cell clusters with a total of 3,438 nuclei have microglia identity and express known microglia markers (21) (Fig. 3 A and B and *SI Appendix*, Figs. S4 D and S7 F). Each of these four microglia clusters exhibits a distinct set of expressed genes, snATAC-seq peaks, and transcription factor binding motifs (Fig. 3 B and C), suggesting they correspond to distinct identities of microglia cells. Additionally, based on GO enrichment analysis, the four microglia clusters identified in the Emory cohort each have a distinct set of marker genes related to their molecular functions (*SI Appendix*, Fig. S7H and *Dataset S3C*), which we will discuss in detail below. MG-1 is the largest microglia cluster ($n = 1,751$ nuclei) and appears to be in a combination of homeostatic and activating states based on the expression of marker genes in the multiome data. This cluster exhibits the highest expression of microglia homeostatic marker genes, including *CX3CR1*, *TMEM119*, and *CSF1R* (Fig. 3B). Cells in this cluster also express genes characteristic of the activating state, including inflammatory genes involved in antigen presentation (*CD86*, *CD80*; MHC II – *C1QA*, *C1QB*, *C1QC*), reactive chemokines (*CCL2*, *CCL3*), and interleukin (*IL-1a*, *IL18*) (Fig. 3B). TF binding motifs for SPI1 (also known as PU.1), a TF essential for microglia activation (22), are specifically enriched

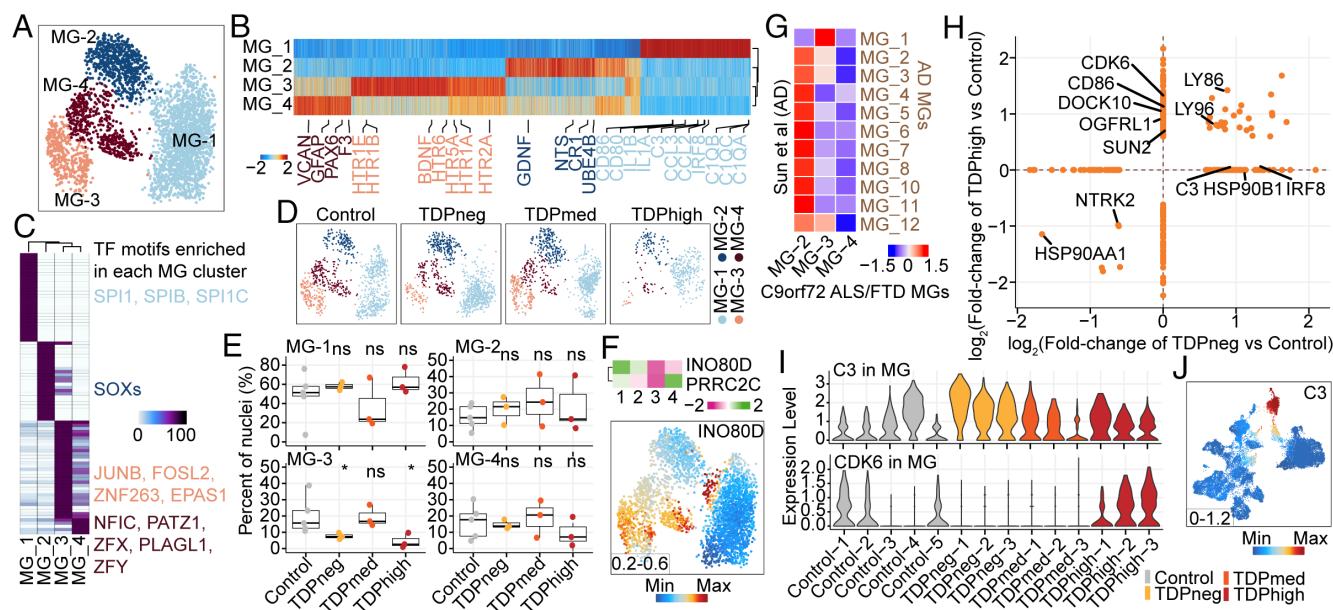


Fig. 3. Loss of neuronal surveillance microglia in *C9orf72* ALS/FTD donors with low and high pTDP-43. (A) UMAP plots of the four microglia clusters. (B) Heatmap showing the row-normalized pseudo-bulk gene score in each snATAC-seq cluster split by nuclei from each of the four MG clusters; rows are organized based on hierarchical clustering and the key genes that define the microglia lineages are marked (Bottom). (C) Heatmap of motif enrichment at differential marker peaks of each microglia cluster. Color indicates the motif enrichment [$-\log_{10}(P \text{ value})$] based on the hypergeometric test. TFs specifically enriched for each MG cluster are highlighted using the same cell cluster-specific colors as in (A). (D) UMAP plots of the distribution of each pTDP-43 sample group for the four microglia clusters. (E) Fraction of each MG cluster in control and pTDP-43 donor groups (Kruskal–Wallis test with Benjamini–Hochberg correction; $P > 0.05$, n.s.) relative to the total number of microglia. (F) Gene activity of MG-1 marker genes from the AD dataset. (G) Heatmap showing the similarities of marker genes between Sun et al. and MG-2, MG-3, and MG-4 microglia clusters. The Jaccard score indicates the percentage of pairwise overlapping genes. (H) DEGs found in pTDPhigh compared to control and pTDPneg compared to control samples. (I) Violin plots showing gene expression levels of the *C3* and *CDK6* genes in microglia of all Emory samples. (J) Gene activity score for the *C3* gene.

in the MG-1 cluster at chromatin-accessible regions (Fig. 3C). IRF8, another critical TF that transforms microglia into a reactive phenotype (22), is uniquely highly expressed in the MG-1 cluster. Distinct from MG-1, the other three MG clusters exhibit lower expression of genes involved in microglia homeostasis, which may be responsible for the shift away from the homeostatic state (Fig. 3A–C). Clusters MG-2 and MG-3 exhibit moderate expression of markers for an alternative M2-like microglia state (Fig. 3B), which has been proposed to be an anti-inflammatory state that plays a protective role in the brain, in contrast to reactive microglia (23). Specifically, MG-2 and MG-3 are defined by different sets of neurotrophic factors (MG-2: BDNF/GDNF/ NTS; MG-3: BDNF/GDNF/ NGF) that have established roles in supporting neuron survival (Fig. 3B). In addition, MG-2 is marked by genes involved in cell adhesion, proliferation (*UBE4B*), interferon type I interferon receptor binding, and the Complement receptor gene *CR1* (Fig. 3B). In contrast, MG-3 cells are marked by genes encoding serotonin receptors and genes involved in G-protein-coupled receptor signaling. MG-4 cells express marker genes of microglia and astrocytes, such as *CSF1R*, *GFAP*, *VCAN*, and *AQP4* (Fig. 3B and SI Appendix, Figs. S4 C and S5 E), suggesting this cluster might correspond to a specific subset of microglia cells that is phenotypically transitioning into astrocyte-like cells. The expression of *PAX6* in MG-4 further confirms the similarity of cells in this cluster with astrocytes (Fig. 3B). This type of cell is present in an inherited model of ALS (24). Although MG-4 shows no observable differences in doublet scores derived from snATAC-seq data, cells in this cluster exhibit a higher number of UMIs and genes compared to nuclei in other MG clusters (SI Appendix, Fig. S7G). Samples from the Mayo cohort captured 2723 nuclei with microglia identity forming 2 unsupervised clusters (SI Appendix, Fig. S7A). However, the distinct groups of microglia cells found in the Emory cohort can also be identified in the Mayo cohort, sharing the same marker genes based on

module gene scores (SI Appendix, Fig. S7B). For example, the Mayo MG-1 cluster is the largest microglia cluster ($n = 2,027$ nuclei) and expresses the same marker genes as the Emory MG-1 cluster (SI Appendix, Fig. S7B). Pseudo-bulk PCA analysis confirms the similarity between these two clusters (SI Appendix, Fig. S5G). Although markers of the other three Emory microglia clusters can be found in different groups of nuclei within the Mayo MG-2 cluster ($n = 553$ nuclei), the separation between these microglia cell types is not as distinctive as in the Emory cohort (SI Appendix, Fig. S7B).

We observed a reduction in microglia in TDPhigh samples from *C9orf72* ALS/FTD donors compared to control samples from the Emory cohort (Fig. 3D). We analyzed the proportion of each microglia cluster relative to total microglia identified per sample within each sample group and found that the proportion of cells in the MG-3 cluster is significantly lower in both the TDPneg and TDPhigh samples compared to control samples (Fig. 3E). This analysis revealed that the MG-3 cluster shares marker genes, such as *INO80D* and *PRRC2C* (Fig. 3F), with the MG-1 cluster found in patients with AD, which has been implicated in neuronal surveillance function (25), and shows decreased proportion of cells in both early and late AD stages (26). Furthermore, we found that the MG-3 cluster exhibits high expression levels of various neurotransmitter receptors (SI Appendix, Fig. S7C). While MG-1 shares signatures with all AD microglia clusters due to its mixed state (SI Appendix, Fig. S7D), the Emory MG-3 cluster demonstrates the closest transcriptome resemblance to the MG-1 cluster in samples from AD donors (Fig. 3G). The Mayo MG-2 cluster most closely resembles Emory MG-3 (SI Appendix, Fig. S7B). However, there are too few nuclei in the Mayo MG-2 cluster (Dataset S2B), which limits our ability to confidently conclude whether the same loss of neuroprotective microglia detected in the Emory cohort is also observed in the pTDPhigh samples of the Mayo cohort. When examining differential gene expression,

we find that the neurotransmitter receptor genes *OGFR1* and *DOCK10* are upregulated in the pTDPhigh group (Fig. 3*H*). These two genes are more highly expressed in the MG-1 homeostatic cluster, suggesting that the reduction in neuronal surveillance microglia occurs in the early disease stage, preceding the accumulation of pTDP-43, perhaps as a consequence of the high expression of the *C9orf72* gene in microglia. Furthermore, the decrease in the proportion of cells in this cluster becomes more pronounced with high levels of pTDP-43 accumulation. This trend is also observed in AD donors in late stages with extensive AD pathology and severe dementia.

Although the decrease in neuronal surveillance microglia is common to the pTDPneg and pTDPhigh groups, the upregulated DEGs in these groups are distinct whereas the downregulated DEGs share some properties (*SI Appendix*, Fig. S7*I* and **Dataset S4**). Specifically, genes involved in phagocytosis, response to interferon, and immune response activation are upregulated in the pTDPneg samples compared to control samples. In contrast, genes related to immune response-regulating signaling pathways, positive regulation of kinase activity, and immunological synapse formation are upregulated in the pTDPhigh samples compared to controls, including upregulation *CDK6*, *CD86*, and *SUN2* (Fig. 3 *H* and *I*). A similar trend, although not statistically significant, was observed in the Mayo pTDPhigh samples (*SI Appendix*, Fig. S7*E*). We observed upregulation of the neuroinflammation genes *C3* and *IRF8* in the pTDPneg group in the pseudobulk microglia analysis as well as in the MG-1 cluster in the Emory samples (Fig. 3 *H* and *I* and **Dataset S4**). We found similar findings in the Mayo cohort, where *IRF8* is also significantly upregulated in pTDPneg samples (**Dataset S4**). *C3* expression shows a similar trend, although it is not statistically significant (*SI Appendix*, Fig. S7*E*). The *C3* gene encodes a complement protein involved in phagocytosis and synapse pruning (27). Elevated *C3* levels in early Alzheimer's suggest a shared feature between early *C9orf72* ALS/FTD and AD (28). Early synaptic loss in AD is thought to be mediated by complement pathways and microglia before amyloid plaques accumulate. However, in late-stage AD, amyloid plaques trigger *C3* release from astrocytes, which interacts with microglia and neurons, leading to further synaptic loss and worsening cognitive decline (29). Our results indicate that *C3* is highly expressed in microglia compared to other cortical cell types (Fig. 3*J*) and that upregulation of *C3* is only observed in pTDPneg samples (Fig. 3*J*). This suggests that in the early stages of

C9orf72 ALS/FTD, prior to pTDP-43 accumulation, increased *C3* expression and release from microglia could lead to early-stage synapse pruning, similar to early AD. This activation could result from *C9orf72* repeat expansions, which are highly expressed in microglia. Surprisingly, we observed no changes in *C3* or *C1QA* in late-stage FTD with pTDP-43 accumulation, suggesting the complement cascade may not play a significant role in late-stage *C9orf72* ALS/FTD.

Astrocyte Dysregulation Becomes More Pronounced in Advanced Disease Stages.

Astrocytes represent another cortical cell type known to become reactive and to respond to disease state in neurodegenerative diseases, particularly via dysregulation of metabolic pathways (30). Cell clusters ASC-1 to ASC-4 with a total of 3,703 nuclei in the Emory cohort and ASC-1 to ASC-3 with a total of 4,162 nuclei in the Mayo cohort can be identified as having astrocyte identity based on high expression of *GFAP*, *AQP4*, and *SLC1A2* (Fig. 4*A* and *SI Appendix*, Fig. S4 *C* and *D*). Each astrocyte subpopulation exhibits a distinct set of expressed genes (**Dataset S3**) and the ASC-3 cluster in the Emory cohort and the ASC-2 cluster in the Mayo cohort have higher levels of *GFAP*, a marker for reactive astrocytes (31). The gene activity of *MT2A*, which encodes a metallothionein protein associated with neuronal injury, is higher in these two reactive astrocyte clusters found in the Emory and Mayo cohorts, and lower in ASC-2 in the Emory cohort (Fig. 4*B*).

We observed various DEGs in astrocytes, distinguishing astrocyte reactivity in early and late disease stages, particularly for samples in the Emory cohort. However, due to the low number of astrocytes in our dataset, we could not perform astrocyte cluster-specific differential gene expression analysis. Therefore, we analyzed changes of gene expression found in all astrocytes using a linear mixed effect model. The astrocyte-reactive genes *CRYAB* and *NTRK2* are downregulated in the pTDPneg group, while another astrocyte-reactive gene, *KCNJ10*, is upregulated in the pTDPhigh group (Fig. 4*C*). *NTRK2* encodes the TrkB neurotrophic tyrosine receptor kinase 2, which interacts with brain-derived neurotrophic factor (BDNF). BDNF regulates neuronal survival and synaptic plasticity. When TrkB is activated, it triggers a positive feedback loop and upregulates the transcription of BDNF through MAPK pathways (32). One of the mechanisms by which astrocytes provide neurotrophic support is releasing BDNF. Deprivation of BDNF and the TrkB signaling pathway increases inflammatory cytokines

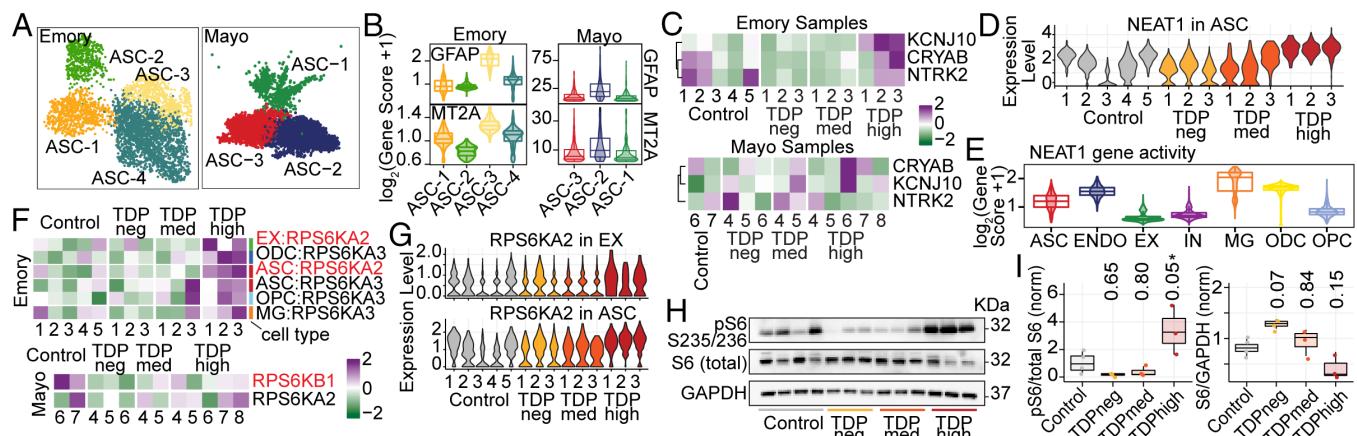


Fig. 4. Changes of gene expression in astrocytes are more pronounced in pTDPhigh samples. (A) UMAP plots of astrocyte (ASC) clusters for Emory (Left) and Mayo (Right) cohorts. (B) Astrocyte clusters exhibit differential levels of *GFAP* and *MT2A* in both cohorts. (C) Changes in gene expression of astrocyte reactivity marker genes. (D) *NEAT1* expression in astrocytes for all samples in the Emory cohort. (E) Cell type specificity of *NEAT1* gene activity. (F) Changes of *RPS6KA2* and *RPS6KA3* gene expression are more significant in pTDPhigh samples. (G) *RPS6KA2* is upregulated both in astrocytes and excitatory neurons. (H) Western immunoblot images of phosphorylated ribosomal protein S6, total ribosomal protein S6, and GAPDH (I) Quantification of the ratio of phosphorylated ribosomal protein S6 to total protein (Left) and total ribosomal protein S6 to GAPDH (norm) (Right). 1-way ANOVA with Tukey's post hoc test, adjusted P-values are shown, and *P < 0.05.

and promotes neuronal cell death (33). A decrease in TrkB expression is also found in the postmortem brains from AD patients (34), suggesting a common dysregulation shared between *C9orf72* ALS/FTD and AD. This prior evidence together with our findings suggests a potential impact of astrocytes on neuronal survival not just in late disease stages but also early before the accumulation of pTDP-43.

Several genes exhibit distinct expression changes specifically in the pTDPhigh group in astrocytes. For example, NEAT1 shows elevated expression levels in astrocytes in pTDPhigh samples from the Emory cohort and in other nonneuronal cells compared to neurons (Fig. 4 D and E). NEAT1 is a long noncoding RNA that has been proposed to function as a structural scaffold for assembling paraspeckles (35). Many proteins present at paraspeckles are involved in RNA splicing or posttranscriptional regulation and nuclear retention of RNAs, including TDP-43. NEAT1 is bound strongly by TDP-43 (36) and depletion of TDP-43 in human embryonic stem cells results in increased number of paraspeckles and changes in cell differentiation (37). It has been shown that the G₄C₂ foci found in *C9orf72* ALS/FTD patient-derived fibroblast colocalize with paraspeckle proteins but not NEAT1, suggesting that *C9orf72* G₄C₂ RNA could form a distinctive paraspeckle class that is independent of NEAT1 (38). Dysregulation of NEAT1 and paraspeckle function has been observed in several neurodegenerative diseases (39). Increased expression of NEAT1 in astrocytes of pTDPhigh samples suggests that astrocytes in late disease stage might be reactive due to dysregulation of RNA retention and/or number of paraspeckles. We found NEAT1 to be downregulated in the pTDPneg and pTDPmed samples from the Mayo cohort, but not statistically altered in the pTDPhigh samples (*SI Appendix*, Fig. S8C and Dataset S4B), perhaps due to the relatively low number of astrocytes in the pTDPhigh samples (Dataset S2B).

We observed upregulation of *RPS6KA2*, which encodes RSK3, a ribosomal protein S6 kinase, in astrocytes of Emory pTDPhigh samples. We also detected upregulation of *RPS6KB1*, which encodes S6K1, another ribosomal protein S6 kinase, in the Mayo pTDPhigh samples (Fig. 4 F and G). Although not statistically significant, we also noted elevated expression of *RPS6KA2* in excitatory neurons of pTDPhigh samples (Fig. 4 F and G and Dataset S4B). Ribosomal protein S6 kinases, particularly S6K1, are well-known downstream effectors in the mTORC1 signaling pathway (40), which promotes astrocyte development (41). RSKs, including RSK3, target various substrates, including Raptor, suggesting its involvement in mTORC1 signaling and its potential to regulate translation and cell survival by phosphorylating eukaryotic translation initiation factor-4B (eIF4B). To validate these results, we performed immunoblotting on protein lysates extracted from frontal cortical tissues of Emory cohort samples for phosphorylated S6 and total S6 protein. We observed significant increases in phosphorylated S6 levels specifically in pTDPhigh samples compared to controls (Fig. 4 H and I and *SI Appendix*, Fig. S8), and reduced phosphorylated S6 in pTDPneg and pTDPmed samples, although this decrease is not statistically significant (Fig. 4 H and I). Total S6 protein levels showed variation (Fig. 4 J), suggesting a disease staging trend with higher levels in pTDPneg and lower levels in pTDPmed and pTDPhigh samples, indicating a decrease as pTDP-43 accumulation increases. Reduced RSK3 expression has been suggested to alleviate the neurodegenerative phenotype observed in Spinocerebellar Ataxia Type 1 (SCA1) (42). These findings collectively suggest differential activation or inactivation of signaling involving phosphorylated S6 protein in early and late disease stages of *C9orf72* ALS/FTD.

Heterogeneity of pTDP-43 Accumulation in Inhibitory and Excitatory Neurons. A total of 11,835 nuclei can be annotated as excitatory or inhibitory neurons using gene activity scores for key lineage genes in the Emory cohort, and each category consists of 9 and 6 clusters, respectively (Fig. 5 A and B). Thus, neurons are the most diverse cell type in the single nucleus multiome dataset in the Emory cohort. In contrast, all samples from the Mayo cohort have much fewer neuronal nuclei (*SI Appendix*, Fig. S4). Thus, we have excluded the Mayo cohort from the following neuronal specific analysis. A similar issue has been observed by others when using *C9orf72* FTD frontal cortex samples from the Mayo Clinic Brain bank (43). Multiple neuronal subtypes can be annotated based on known marker genes (44). Excitatory neurons can be categorized by their cortical layer position (layer 2–6) and their axonal projections (Fig. 5 C), whereas inhibitory interneurons can be grouped by their developmental origin from the medial, lateral, or caudal ganglionic eminences and classified based on their subtypes (Fig. 5 D). It is not known which neuronal cell types are more vulnerable to pTDP-43 accumulation and/or nuclear loss of TDP-43. NeuN+ cortical neurons from the neocortex of *C9orf72* ALS/FTD patients have been fractionated previously based on levels of nuclear TDP-43, allowing the characterization of nuclear TDP-43 positive and negative specific transcriptomes using bulk RNA sequencing (45). However, this study was not able to identify the neuronal cell types contributing to the ensemble of TDP-43 positive and negative RNA-seq profiles. To address this issue, we employed the cell composition deconvolution algorithm CIBERSORTx (46) and compared our single nucleus datasets with the published TDP-43 sorted bulk transcriptomes on NeuN-positive nuclei. We were thus able to quantify the contribution of each individual neuronal subtype identified in our multiome dataset to the published TDP43-negative and TDP43-positive transcriptomes. NeuN-positive neurons are typically composed of 70% excitatory neurons and 30% inhibitory neurons (44), and we found our deconvolution analysis performs as expected in that more than 70% of NeuN-positive transcriptomes correspond to excitatory neuronal clusters (Fig. 5 E). Among all neuronal clusters, EX-1, a cluster consisting of cortical projection neurons with high expression of CUX2 and LAMP5, has the most significant contribution to the nuclear TDP43-negative cells (Fig. 5 E). This result suggests that a significant proportion of excitatory neurons with high expression of CUX2 and LAMP5 have nuclear TDP-43 loss, distinct from other neuronal populations.

CUX2+ Cortical Projection Excitatory Neurons Are Significantly Reduced in the Frontal Cortex of *C9orf72* ALS/FTD Donors. We grouped neurons based on their excitatory projection classification and developmental origin for interneurons to avoid cell clusters with few nuclei. We found that the proportion of cortical projection neurons is more than three-fold lower in TDPhigh and TDPneg patient groups compared to control (Fig. 5 F), suggesting that these neurons are especially susceptible to *C9orf72* ALS/FTD frontal cortex degeneration regardless of the frontal cortical pTDP-43 levels. We systematically assessed the differential abundance between *C9orf72* ALS/FTD donor and control groups for all neuronal clusters. Cortical projection neurons showed significant proportional changes in both TDPhigh and TDPneg donor groups, while subcortical projection neurons and inhibitory neurons originating from the medial caudal ganglionic eminence showed significant proportional changes in pTDPneg and pTDPhigh groups, respectively (Fig. 5 G). To confirm the loss of CUX2 neurons in the upper cortical layers, we used immunofluorescence microscopy using antibodies to

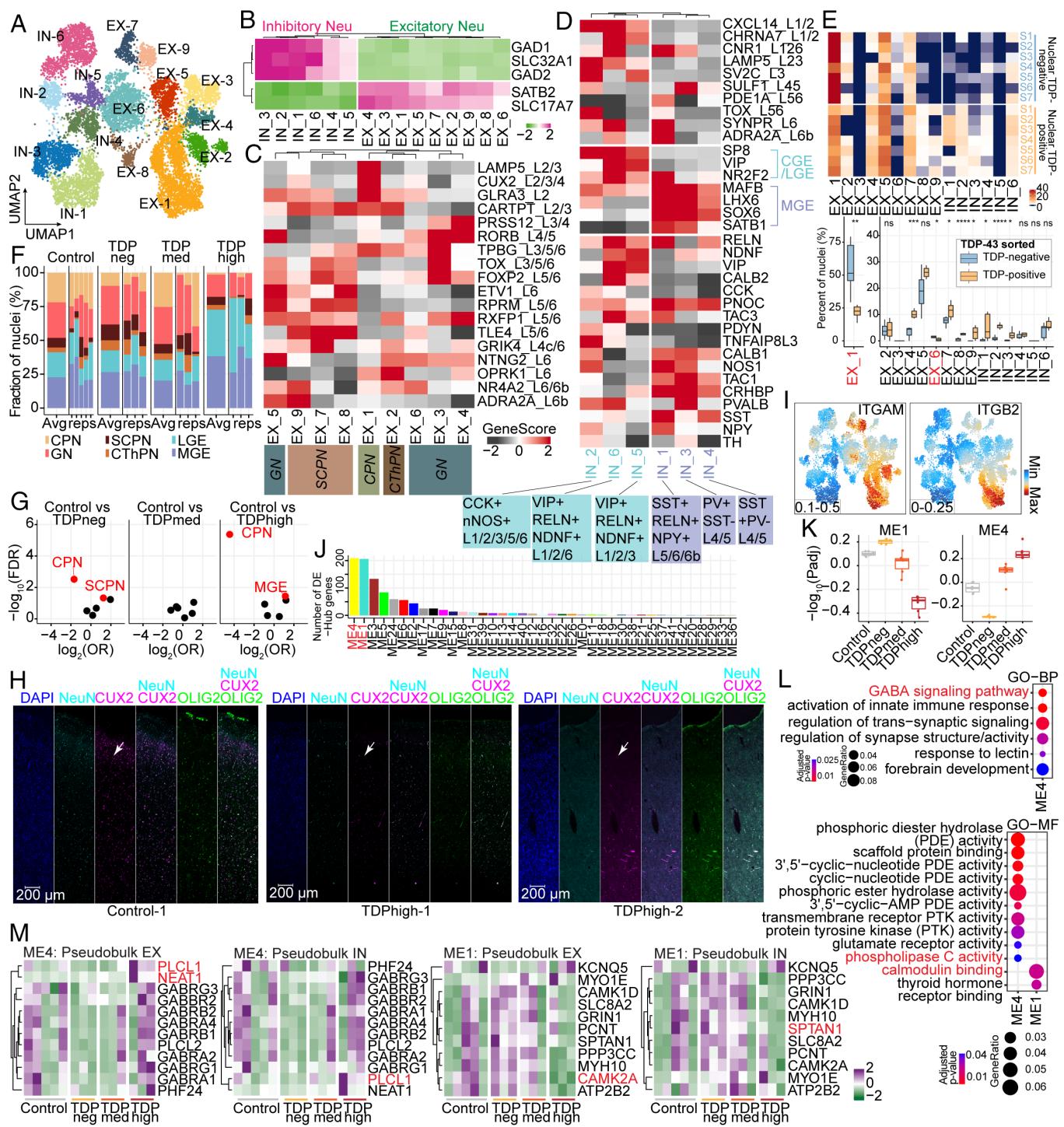


Fig. 5. Neuronal cell types in the prefrontal cortex of control and *C9orf72* ALS/FTD donors. (A) UMAP plots of neuronal clusters. (B) Gene activity scores for marker genes of excitatory and inhibitory neurons. (C) Heatmap of gene activity scores of cortical layer-specific marker genes for excitatory neurons. Axonal projection subclassification is indicated below. CPN, cortical projection neurons; GN, granule neurons; SCPN, subcortical projection neurons; CThPN, corticothalamic projection neurons. (D) Heatmap of gene activity scores of marker genes associated with inhibitory neurons of subpallial origin (Top), cortical layers (Middle), and subclassification (Bottom). CGE, caudal ganglionic eminence; MGE, medial caudal ganglionic eminence; LGE, lateral ganglionic eminence; SST, somatostatin; RELN, reelin; NPY, neuropeptide Y; PV, parvalbumin; VIP, vasoactive intestinal peptide; NDNF, neuron-derived neurotrophic factor; CCK, cholecystokinin; nNOS, neuronal nitric oxide synthase. (E) Top: summary of cell proportion deconvolution with pTDP-43 positive and negative nuclei (n.s. not statistically significant; $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$); Bottom: heatmap representation of cell proportion deconvolution data in each individual nuclear pTDP-43 positive and negative transcriptome. (F) Proportion of neuronal subtypes defined by cortical projection or developmental origins in all sample groups. (G) Volcano plots showing odds ratio (OR) and FDR computed by MASC (47) for all the neuronal subtypes. Red labeled neuronal subtypes that are significantly increased or depleted in association with specific *C9orf72* ALS/FTD donor groups (FDR-adjusted $P < 0.05$; absolute OR > 0). (H) Immunostaining of human postmortem cortical tissue for the pan-neuronal marker NeuN (cyan), oligodendrocyte lineage marker OLIG2 (green), CUX2 (magenta), and DAPI in blue. The location of CUX2-positive Upper layer cortical neurons is indicated by a white arrow in the panel labeled with the CUX2 antibody for the Control-1 sample. White arrows also mark the proximal location of the upper cortical layers in the CUX2 panels for the TDPhigh-1 and TDPhigh-2 samples. (I) Gene activity score for the *ITGAM* and *ITGB2* genes that encode the heterodimer C3 receptor. (J) Number of differential hub genes found in each module. (K) Significance of WGCNA modules with different levels of pTDP-43. (L) Gene ontology analysis of the different hub genes in the ME1 and ME4 modules. (M) Heatmap demonstrating the average gene expression of identified hub genes in modules ME1 and ME4 across all samples from the Emory cohort.

CUX2 and cell-type specific marker proteins in pTDPhigh samples (Fig. 5*H*). Loss of cortical projection neurons in TDPhigh donors is more extensive than in TDPNeg donors (Fig. 5 *F* and *G*). We found that the genes *ITGAM* and *ITGB2*, which encode for subunits of the CR3 complement factor C3 receptor, are specifically highly expressed in cortical projection neurons compared to other neuron types (Fig. 5*I*). This finding supports our earlier speculation that cortical projection neurons might be specifically tagged by C3 released from microglia in the early stages of disease, leading to microglia mediated synapse loss and phagocytosis of neurons. This finding could also explain the specific vulnerability of cortical projection neurons in early disease stages of *C9orf72* ALS/FTD. In addition to cortical projection neurons, inhibitory neuronal clusters originating from the medial caudal ganglionic eminence (MGE) are significantly increased in proportion in the TDPhigh donor group compared to control (Fig. 5 *F* and *G*). Based on the cell composition deconvolution analysis against pTDP-43 positive and negative specific transcriptomes, IN-1 and IN-3 MGE originated neurons have a higher contribution to the pTDP-43 positive compared to the pTDP43-negative transcriptome (Fig. 5*E*). The results suggest that inhibitory neurons originated from the MGE might be resistant to neurodegeneration, possibly because they are less vulnerable to nuclear TDP-43 loss.

To uncover gene regulatory networks linked to pTDP-43 accumulation in neurons, we used weighted gene coexpression network analysis (WGCNA) to cluster coexpressed genes in neurons into modules and identify highly correlated genes. We identified 43 modules (*SI Appendix*, Fig. S9*A*), two of which, ME1 and ME4, significantly correlate with pTDP-43 levels (Fig. 5 *J* and *K* and *SI Appendix*, Fig. S9*B*). ME4 positively correlates while ME1 negatively correlates with the amount of pTDP-43 (Fig. 5*K*). We found these two modules also have the most differentially expressed hub genes (Fig. 5*L* and Dataset S5*A*). Gene ontology analysis of the differential hub genes in the ME1 and ME4 modules reveals an enriched set of distinct genes. For example, genes involved in the gamma-aminobutyric acid (GABA) signaling pathway are enriched in the ME4 module whereas genes involved in calmodulin binding are enriched in the ME1 module (Fig. 5*L* and Dataset S5*B*). Specifically, *PLCL1* is found in multiple GO enriched terms, including modulation of the GABA signaling pathway, inositol lipid-mediated signaling, and phospholipase C activity (Dataset S7*B*). *PLCL1* is significantly upregulated in pTDPhigh samples in both excitatory and inhibitory neurons (Fig. 5*M*). *NEAT1* emerges as another prominent hub gene in the ME4 module (Fig. 5*M*), showing specific upregulation during the late disease stage in pTDPhigh samples. Therefore, dysregulation of paraspeckles in excitatory neurons could be another characteristic feature of the late disease stage with high levels of cortical pTDP-43. In contrast to ME4 hub genes, *SPTAN1* represents a strong ME1 module hub gene that is upregulated in early disease stages. *SPTAN1* encodes a spectrin family protein, a crucial component of the cytoskeleton (48). Furthermore, *CAMK2A*, another differential hub gene in the ME1 module, exhibits specific downregulation in the pTDPhigh group. Exon skipping of *CAMK2A* transcripts has been observed in TDP-43 knockdown mouse primary neurons (49), suggesting that the downregulation of *CAMK2A* could be a direct consequence of the loss of nuclear TDP-43 in pTDPhigh samples. Additionally, *CAMK2A* has been reported to have reduced protein abundance in the cerebrospinal fluid of ALS patients compared to controls (50). This comprehensive WGCNA in neurons further bolsters our earlier findings, indicating that distinct sets of genes are involved in early and late disease stages in a cell-type-specific manner.

Discussion

Here, we utilized a unique *C9orf72* ALS/FTD staging paradigm by selecting cases based on the abundance of pTDP-43. Cortical cytoplasmic accumulation of pTDP-43 has been found to correlate with neuropathological burden and severity of FTD clinical symptoms, and the progression of pTDP-43 distribution in the CNS has been proposed to stage patients in different phases of the disease (51). The results of our study suggest several systematic changes in the early and late stages of disease not previously reported. These include the loss of neurosurveillance microglia, significant increases in phosphorylated ribosomal S6 protein, and global dysregulation of chromatin accessibility uniquely found in nonneuronal cells associated with high pTDP-43 levels. We also observed abnormalities in oligodendrocytes, microglia, and astrocytes specifically associated with late disease stages. Interestingly, cortical projection neurons appear to be selectively vulnerable to *C9orf72* ALS/FTD progression.

Changes observed in *C9orf72* ALS/FTD donors can be a consequence of the transcriptional misregulation of *C9orf72*, presence of repeats in the RNA, or the presence of peptides translated from these repeats. Our findings suggest that microglia are the first to respond in the early stages of FTD, before pTDP-43 aggregation, by activating the complement cascade, increasing phagocytosis, and altering neuronal surveillance. This pattern mirrors observations in Alzheimer's disease, albeit with different genetic contributions, suggesting a common activation of the microglia immune response in early neurodegeneration in both diseases. However, in late disease stages, the interferon response is activated in microglia instead. Astrocytes exhibit alterations both early and late in disease progression, including the downregulation of astrocyte-reactive genes before the formation of pTDP-43 inclusions. During late disease stages, astrocytes significantly upregulate *NEAT1* and *RPS6KA2*. This raises the question of the role of paraspeckles in late-stage pathology, since TDP-43 protein is known to bind to *NEAT1* and colocalize in paraspeckles with other splicing regulatory proteins (37). The increased expression of *RPS6KA2* and phosphorylation of ribosomal protein S6 specifically in astrocytes during late disease stages are particularly intriguing. Since *RPS6KA2* encodes RSK3, the only RSK with a potential nuclear localization signal, it is possible that upregulation of *RPS6KA2* may lead to the phosphorylation of additional nuclear proteins with roles in the regulation of gene expression. Furthermore, the upregulation of both *RPS6KA2* and *NEAT1* in astrocytes and neurons suggests a shared molecular mechanism between these cell types, potentially expanding our current understanding of their interactions.

The most significant finding unique to donors in late disease stages with high levels of pTDP-43 is the high proportion of newly differentiated/premyelinating oligodendrocytes, which is not observed in control or *C9orf72* ALS/FTD samples with low pTDP-43 accumulation. Dysregulation of oligodendrocyte maturation and function may be a direct consequence of the formation of pTDP-43 cytoplasmic inclusions or the nuclear loss of TDP-43. The relatively low expression of genes encoding myelin protein components in this cell cluster may be due to high cytoplasmic pTDP-43 accumulation, since nuclear TDP-43 binds to transcripts encoding for myelin proteins (52) and soluble cytoplasmic TDP-43 is involved in the posttranscriptional regulation of myelin proteins (53). We do not observe a significant change in oligodendrocyte progenitor cells, suggesting that the cluster of premature oligodendrocytes is likely a result of its inability to become mature due to the downregulation of myelin components and failure to undergo the typical programmed cell death observed for the majority of premyelinating

oligodendrocytes during adult oligodendrogenesis. It is possible that the cause of cortical projection neuron loss in TDPhigh donors is the lack of sufficient mature oligodendrocytes, which are essential for neuron myelination and metabolic support. The impairment of myelination is not limited to *C9orf72* FTD TDP-43 pathology. In AD donors carrying two copies of the APOE4 variant, cholesterol homeostasis is responsible for the downregulation of myelin-associated genes in oligodendrocytes (54). Further dissection of oligodendrocyte–neuron interactions may give additional insights into the mechanisms underlying the progression of FTD.

While our study comprises a relatively small number of samples, the approach of grouping samples based on cortical pTDP-43 levels allowed us to identify cellular and molecular changes associated with different stages during disease progression. Ultimately, the systematic identification of cell-type-specific defects in pathways common to all *C9orf72* ALS/FTD donors, as well as disease stage-specific alterations, will inform the identification of targets and timing of therapeutic interventions.

Methods

Characterization of Human Tissue Samples. Postmortem brain samples from the dorsolateral prefrontal cortex (DLPFC, Brodmann area 9; BA9) of *C9orf72* ALS/FTD patients and controls were obtained from the Goizueta Emory Alzheimer's Disease Center Brain Bank and Mayo Clinic Brain Bank with approval from the respective Institutional Review Board. TDP-43 levels were determined by sequential biochemical fractionation was performed first and followed by Meso-Scale Discovery immunoassay (4). Immunohistochemistry for pTDP-43 was done as previously described (55).

Single Nucleus Experiments. Briefly, libraries were generated from frozen tissues using the 10x Genomics Chromium Single Cell Multiome ATAC + Gene Expression kit following the manufacturer's instructions and sequenced to 20,000 read-pairs per nucleus on Illumina's NovaSeq 6000 and NovaSeq X Plus instruments. Sequences were processed using the 10x Genomics Cell Ranger ARC

pipeline and aligned to the hg38 human genome assembly. ArchR v1.0.2 (9) and Seurat v4.1.0 (10) were used for processing the paired snATAC-seq fragment data and snRNA-seq gene expression data for each sample as described in *SI Appendix*.

Immunostaining, Imaging, and Quantifications Using MERSCOPE. Frozen human brain tissue (prefrontal cortex, BA9 region) was sectioned at 10 μm thickness and the MERSCOPE protein stain verification protocol (Vizgen, 10400112) was employed following the manufacturer's instructions. Output TIFF files were utilized for nucleus segmentation and staining intensity quantification. For confocal microscopy, frozen human brain tissue was sectioned at 40 μm thickness using a frozen vibratome and processed as described in *SI Appendix*.

Data, Materials, and Software Availability. All data generated in this work are available through GEO Accession No: GSE212630 (56). All scripts used for analyzing the data in this manuscript can be found in the GitHub repository, https://github.com/wanghlv/c9alsftd_multime (57).

ACKNOWLEDGMENTS. We would like to thank Dr. Cynthia Vied at the Translational Science Laboratory of Florida State University for help with Illumina sequencing; Dr. Steven Sloan from the Department of Human Genetics at Emory University for assistance with the initial stage of astrocyte lineage analysis; Dr. Ryan Corces at the Gladstone Institute of Neurological Disease for assistance in the initial analysis of snATAC-seq data. This work was supported by U.S. Public Health Service Awards R35 GM139408 (V.G.C.); R35 NS111602, R01 HG008935, U01 MH116441 (P.J.) from the NIH. H.-LVW. was supported by NIH F32 ES031827. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Author affiliations: ^aDepartment of Human Genetics, Emory University School of Medicine, Atlanta, GA 30322; ^bEmory Center for Neurodegenerative Diseases, Emory University School of Medicine, Atlanta, GA 30322; ^cDepartment of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA 30322; ^dDepartment of Neuroscience, Mayo Clinic, Jacksonville, FL 32224; ^eCenter for Translational Research in Neurodegenerative Disease, University of Florida, Gainesville, FL 32607; ^fNorman Fixel Institute for Neurological Diseases, University of Florida, Gainesville, FL 32607; ^gDepartment of Neurology, Emory University School of Medicine, Atlanta, GA 30322; and ^hDepartment of Pathology and Laboratory Medicine, Emory University School of Medicine, Atlanta, GA 30322

1. O. Hardiman *et al.*, Amyotrophic lateral sclerosis. *Nat. Rev. Dis. Primers* **3**, 17071 (2017).
2. J. D. Warren, J. D. Rohrer, M. N. Rossor, Clinical review Frontotemporal dementia. *BMJ* **347**, f4827 (2013).
3. M. DeJesus-Hernandez *et al.*, Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245–256 (2011).
4. M. Neumann *et al.*, Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).
5. J. Brettschneider *et al.*, TDP-43 pathology and neuronal loss in amyotrophic lateral sclerosis spinal cord. *Acta Neuropathol.* **128**, 423–437 (2014).
6. I. Kawakami, T. Arai, M. Hasegawa, The basis of clinicopathological heterogeneity in TDP-43 proteinopathy. *Acta Neuropathol.* **138**, 751–770 (2019).
7. I. R. Mackenzie *et al.*, Dipeptide repeat protein pathology in C9ORF72 mutation cases: Clinicopathological correlations. *Acta Neuropathol.* **126**, 859–879 (2013).
8. R. S. Wilson *et al.*, TDP-43 pathology, cognitive decline, and dementia in old age. *JAMA Neurol.* **70**, 1418–1424 (2013).
9. J. M. Granja *et al.*, ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat. Genet.* **53**, 403–411 (2021).
10. Y. Hao *et al.*, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e3529 (2021).
11. I. Korsunsky *et al.*, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
12. T. Stuart *et al.*, Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e1821 (2019).
13. M. D. Young, S. Behjati, SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* **9**, giaa151 (2020).
14. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
15. M. E. Murray *et al.*, Clinical and neuropathologic heterogeneity of c9FTD/ALS associated with hexanucleotide repeat expansion in C9ORF72. *Acta Neuropathol.* **122**, 673–690 (2011).
16. G. Finak *et al.*, MAST: A flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278 (2015).
17. S. H. Kang *et al.*, Degeneration and impaired regeneration of gray matter oligodendrocytes in amyotrophic lateral sclerosis. *Nat. Neurosci.* **16**, 571–579 (2013).
18. F. Guo, Y. Wang, TCF7L2, a nuclear marker that labels premyelinating oligodendrocytes and promotes oligodendroglial lineage progression. *Glia* **71**, 143–154 (2023).
19. E. G. Hughes, M. E. Stockton, Premyelinating oligodendrocytes: Mechanisms underlying cell survival and integration. *Front Cell Dev. Biol.* **9**, 714169 (2021).
20. J. G. O'Rourke *et al.*, *C9orf72* is required for proper macrophage and microglial function in mice. *Science* **351**, 1324–1329 (2016).
21. R. Sankowski *et al.*, Mapping microglia states in the human brain through the integration of high-dimensional techniques. *Nat. Neurosci.* **22**, 2098–2110 (2019).
22. T. Masuda *et al.*, IRF8 is a critical transcription factor for transforming microglia into a reactive phenotype. *Cell Rep.* **1**, 334–340 (2012).
23. A. M. Jurga, M. Paleczna, K. Z. Kuter, Overview of general and discriminating markers of differential microglia phenotypes. *Front Cell Neurosci.* **14**, 198 (2020).
24. E. Trias *et al.*, Phenotypic transition of microglia into astrocyte-like cells associated with disease onset in a model of inherited ALS. *Front Cell Neurosci.* **7**, 274 (2013).
25. H. Liu, R. K. Leaf, X. Hu, Neurotransmitter receptors on microglia. *Stroke Vasc. Neurol.* **1**, 52–58 (2016).
26. N. Sun *et al.*, Human microglial state dynamics in Alzheimer's disease progression. *Cell* **186**, 4386–4403.e4329 (2023).
27. C. Gao, J. Jiang, Y. Tan, S. Chen, Microglia in neurodegenerative diseases: Mechanism and potential therapeutic targets. *Signal Trans. Target Ther.* **8**, 359 (2023).
28. T. Wu *et al.*, Complement C3 is activated in human AD brain and is required for neurodegeneration in mouse models of amyloidosis and tauopathy. *Cell Rep.* **28**, 2111–2123.e2116 (2019).
29. A. K. Vandenberg, H. Offner, S. Matejk, A. Matejk, Microglia and astrocyte involvement in neurodegeneration and brain cancer. *J. Neuroinflamm.* **18**, 298 (2021).
30. C. Escartin *et al.*, Reactive astrocyte nomenclature, definitions, and future directions. *Nat. Neurosci.* **24**, 312–325 (2021).
31. M. Orre *et al.*, Isolation of glia from Alzheimer's mice reveals inflammation and dysfunction. *Neurobiol. Aging* **35**, 2746–2760 (2014).
32. J. Tuvinen, P. Pruuksild, E. Orav, E. E. Esvald, T. Timmus, AP-1 transcription factors mediate BDNF-positive feedback loop in cortical neurons. *J. Neurosci.* **36**, 1290–1305 (2016).
33. Z. H. Wang *et al.*, Deficiency in BDNF/TrkB neurotrophic activity stimulates delta-secretase by upregulating C/EBPbeta in Alzheimer's disease. *Cell Rep.* **28**, 655–669.e655 (2019).
34. K. L. Bharani, A. Ledreux, A. Gilmore, S. L. Carroll, A. C. Granholm, Serum pro-BDNF levels correlate with phospho-tau staining in Alzheimer's disease. *Neurobiol. Aging* **87**, 49–59 (2020).
35. C. M. Clemson *et al.*, An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* **33**, 717–726 (2009).
36. J. R. Tollervey *et al.*, Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* **14**, 452–458 (2011).
37. M. Modic *et al.*, Cross-regulation between TDP-43 and paraspeckles promotes pluripotency-differentiation transition. *Mol. Cell* **74**, 951–965.e913 (2019).
38. A. Bajc Cesnik *et al.*, Nuclear RNA foci from C9ORF72 expansion mutation form paraspeckle-like bodies. *J. Cell Sci.* **132**, jcs224303 (2019).

39. K. Li, Z. Wang, lncRNA NEAT1: Key player in neurodegenerative diseases. *Ageing Res. Rev.* **86**, 101878 (2023).
40. O. Meyuhas, Ribosomal protein S6 phosphorylation: Four decades of research. *Int. Rev. Cell Mol. Biol.* **320**, 41–73 (2015).
41. A. J. Voss *et al.*, Identification of ligand-receptor pairs that drive human astrocyte development. *Nat. Neurosci.* **26**, 1339–1351 (2023).
42. W. S. Lee *et al.*, Dual targeting of brain region-specific kinases potentiates neurological rescue in Spinocerebellar ataxia type 1. *EMBO J.* **40**, e106106 (2021).
43. J. Li *et al.*, Divergent single cell transcriptome and epigenome alterations in ALS and FTD patients with C9orf72 mutation. *Nat. Commun.* **14**, 5714 (2023).
44. B. B. Lake *et al.*, Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586–1590 (2016).
45. E. Y. Liu *et al.*, Loss of nuclear TDP-43 is associated with decondensation of LINE retrotransposons. *Cell Rep.* **27**, 1409–1421.e1406 (2019).
46. A. M. Newman *et al.*, Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782 (2019).
47. C. Y. Fonseka *et al.*, Mixed-effects association of single cells identifies an expanded effector CD4(+) T cell subset in rheumatoid arthritis. *Sci. Transl. Med.* **10**, eaq0305 (2018).
48. D. N. Lorenzo, R. J. Edwards, A. L. Slavutsky, Spectrins: Molecular organizers and targets of neurological disorders. *Nat. Rev. Neurosci.* **24**, 195–212 (2023).
49. D. Honda *et al.*, The ALS/FTLD-related RNA-binding proteins TDP-43 and FUS have common downstream RNA targets in cortical neurons. *FEBS Open Bio.* **4**, 1–10 (2013).
50. S. Oh, Y. Jang, C. H. Na, Discovery of biomarkers for amyotrophic lateral sclerosis from human cerebrospinal fluid using mass-spectrometry-based proteomics. *Biomedicines* **11**, 1250 (2023).
51. J. Brettschneider *et al.*, Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. *Ann. Neurol.* **74**, 20–38 (2013).
52. M. Polymenidou *et al.*, Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* **14**, 459–468 (2011).
53. D. Heo *et al.*, Stage-specific control of oligodendrocyte survival and morphogenesis by TDP-43. *eLife* **11**, e75230 (2022).
54. J. W. Blanchard *et al.*, APOE4 impairs myelination via cholesterol dysregulation in oligodendrocytes. *Nature* **611**, 769–779 (2022). 10.1038/s41586-022-05439-w.
55. Z. T. McEachin *et al.*, Chimeric peptide species contribute to divergent dipeptide repeat pathology in c9ALS/FTD and SCA36. *Neuron* **107**, 292–305.e296 (2020).
56. H.-L. V. Wang *et al.*, Data from "pTDP-43 levels correlate with cell type-specific molecular alterations in the prefrontal cortex of C9orf72 ALS/FTD patients." GEO. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE212630>. Deposited 2 September 2022.
57. H.-L. V. Wang *et al.*, Data from "pTDP-43 levels correlate with cell type-specific molecular alterations in the prefrontal cortex of C9orf72 ALS/FTD patients." GitHub. https://github.com/wanghlv/c9alsftd_multioime. Deposited 1 December 2022.