Transcriptomic analysis of frontotemporal lobar degeneration with TDP-43 pathology reveals cellular alterations across multiple brain regions

Rahat Hasan[1*], Jack Humphrey[2-5*], Conceição Bettencourt[6,7], Jia Newcombe[6], NYGC ALS Consortium[#], Tammaryn Lashley[6,7], Pietro Fratta[8], Towfique Raj[2-5]

[1.]Graduate School of Biomedical Sciences; [2.]Nash Family Department of Neuroscience & Friedman Brain Institute; [3.]Ronald M. Loeb Center for Alzheimer's disease; [4.]Department of Genetics and Genomic Sciences & Icahn Institute for Data Science and Genomic Technology; [5.]Estelle and Daniel Maggin Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[6.]Department of Neurodegenerative Disease; [7.]Queen Square Brain Bank for Neurological Disorders; [8.]Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology, London, UK

#. Full list of contributors in Supplementary Acknowledgements.

*. These authors contributed equally.

Correspondence:

Jack Humphrey (jack.humphrey@mssm.edu); Towfique Raj (towfique.raj@mssm.edu)

Abstract

Frontotemporal lobar degeneration (FTLD) is a group of heterogeneous neurodegenerative disorders affecting the frontal and temporal lobes of the brain. Nuclear loss and cytoplasmic aggregation of the RNA-binding protein TDP-43 represents the major FTLD pathology, known as FTLD-TDP. To date, there is no effective treatment for FTLD-TDP due to an incomplete understanding of the molecular mechanisms underlying disease development. Here we compared post-mortem tissue RNA-seq transcriptomes from the frontal cortex, temporal cortex and cerebellum between 28 controls and 30 FTLD-TDP patients to profile changes in cell-type composition, gene expression and transcript usage. We observed downregulation of neuronal markers in all three regions of the brain, accompanied by upregulation of microglia, astrocytes, and oligodendrocytes, as well as endothelial cells and pericytes, suggesting shifts in both immune activation and within the vasculature. We validate our estimates of neuronal loss using neuropathological atrophy scores and show that neuronal loss in the cortex can be mainly attributed to excitatory neurons, and that increases in microglial and endothelial cell expression are highly correlated with neuronal loss. All our analyses identified a strong involvement of the cerebellum in the neurodegenerative process of FTLD-TDP. Altogether, our data provides a detailed landscape of gene expression alterations to help unravel relevant disease mechanisms in FTLD.

Introduction

Frontotemporal dementia (FTD) is a spectrum of neurodegenerative disorders affecting the frontal and anterior temporal lobes of the brain, manifesting clinically as disturbances in behaviour or language. The neuropathological correlate of FTD, Frontotemporal Lobar Degeneration (FTLD), can be subclassified into three distinct subgroups based on specific protein inclusions. Around 50% of patients present with cytoplasmic inclusions of transactive response DNA-binding protein 43 kDa (TDP-43), leading to a histopathological diagnosis of FTLD-TDP [76]. TDP-43 pathology is present in a range of other neurodegenerative diseases, most notably amyotrophic lateral sclerosis (ALS) [76]. Approximately 40% of FTD patients have a positive family history of disease [37]. Mutations in two genes, *C9orf72* and *GRN*, explain the majority of familial FTLD-TDP [3, 19, 22, 83], but several other causal genes have been identified that explain a smaller fraction of cases. These genes include *TARDBP,* the gene encoding TDP-43, as well as *VCP, SQSTM1,* and *TBK1 [8, 11, 34, 58]*.

Much research has focused on understanding the formation and consequences of TDP-43 aggregation in FTLD. TDP-43 is an RNA binding protein that regulates many aspects of mRNA processing, including alternative splicing and polyadenylation [79, 85, 99]. Under conditions not currently understood, TDP-43 is mislocalized from the nucleus to the cytoplasm, where it undergoes post-translational modifications such as hyperphosphorylation, ubiquitination, and N-terminal truncation [76]. Both nuclear loss of function (LOF) and cytoplasmic gain of function (GOF) mechanisms have been proposed for the consequences of TDP-43 mislocalization [63]. Supporting a potential role for LOF in FTLD, recent studies have demonstrated that loss of nuclear TDP-43 alters the expression and splicing of its target mRNAs, many of which are involved in synapse organization and plasticity [45, 79]. Nuclear loss of TDP-43 leads to aberrant cryptic splicing of these targets, including key neuronal genes *STMN2* and *UNC13A* [13, 54, 71, 84]. Thus, TDP-43 nuclear loss may play a role in the neurodegenerative process of FTLD.

FTLD is characterised by neuronal loss and an inflammatory response driven by glial cells. Both excitatory glutamatergic pyramidal cells and inhibitory GABAergic neurons are lost or dysregulated in post-mortem FTLD-TDP brains [28, 46, 74, 88]. A unique subgroup of excitatory projection neurons known as von Economo neurons (VENs) has been shown to be selectively vulnerable to TDP-43 pathology, dying particularly early in disease progression [31, 75, 87]. Although the cerebellum is spared from TDP-43 pathology [12], it is an open question whether neurodegeneration also occurs there in FTLD. One of the suggested mechanisms of cerebellar neuronal loss is toxicity from dipeptide repeat proteins, which are generated via unconventional translation of the *C9orf72* gene in cases caused by *C9orf72* mutations. [32, 106]. Microglia and astrocytes are essential for clearing of debris and maintaining brain homeostasis [50]. However, in response to neurodegeneration, these cells are believed to switch to an activated state, producing pro-

inflammatory cytokines, chemokines, and reactive oxygen species that contribute to a state of neuroinflammation [36, 61, 81]. Although this may initially be beneficial for removing surrounding protein aggregates and dying neurons, sustained inflammatory responses can damage neurons and synapses, ultimately exacerbating neurodegeneration.

Although recent advancements have improved our understanding of the molecular basis of FTLD-TDP, it is still unclear how mutations in the different causal genes lead to disease, and what the underlying pathways are. Transcriptome profiling of human post-mortem brain tissue is necessary to gain a deeper insight into the functional pathological changes in all FTLD-TDP subtypes. To address this need, we have assembled a cohort of RNA-seq generated from post-mortem brain samples from FTLD-TDP and non-neurological disease control patients. We assessed the shared and distinct transcriptomic changes associated with FTLD-TDP across the frontal cortex, temporal cortex, and cerebellum. We performed differential gene expression, differential transcript usage, and cell-type deconvolution analysis, identifying changes in cellular composition and cellular pathways across all three brain regions.

Methods

Table 1. Clinical and pathological characteristics of the FTLD-TDP cohort

	Control	FTLD-TDP		P-value		
		non-*C9orf72*	*C9orf72*	non-*C9orf72* vs control	*C9orf72* vs control	non-*C9orf72* vs *C9orf72*
Donors	28	21	9	-	-	-
RNA-seq samples						
Frontal Cortex	22	16	7	-	-	-
RIN	6.3 (5.3-7.9)	5.9 (5.1-7.0)	6.2 (5.6-7.7)	0.11	0.82	0.48
Temporal Cortex	17	14	8	-	-	-
RIN	6.2 (5.1-7.8)	5.8 (5.1-7.6)	6.0 (5.1-7.5)	0.073	0.38	1
Cerebellum	16	20	9	-	-	-
RIN	6.3 (5.2-7.9)	6.1 (5.2-7.0)	6.6 (5.5-7.7)	0.46	0.24	**0.028**
Clinical demographics						
% Female	46.4%	52.4%	44.4%	0.44	1	1
Disease duration, months	-	125 (36-276)	90 (60-132)	-	-	0.25

Age at death	72 (38-90)	69 (53-87)	64 (45-74)	0.19	**0.032**	0.22
PMI	23 (7-61)	55 (19-115)	70 (26-107)	**2.3e-6**	**6.2e-6**	0.17
Pathology						
Braak Tau	44% > 0	76% > 0	89% > 0	**0.040**	**0.026**	0.64
Thal Phase	30% > 0	52 % > 0	22% > 0	0.21	1	0.23
CERAD	23% > 0	62% > 0	11% > 0	0.34	0.56	0.21
CAA	7.4% > 0	67% > 0	0% > 0	0.081	0.56	0.071
Alpha-synuclein	12% > 0	14% > 0	22% > 0	0	0.59	0.62

CAA: Cerebral amyloid angiopathy, CERAD: Consortium to Establish a Registry for Alzheimer's Disease, PMI: Post-mortem interval, RIN: RNA integrity number. Continuous variables, presented as mean and range, were compared with Wilcoxon non-parametric tests. Categorical variables, presented as the percentage of donors with a score above 0, were compared with Fisher's exact tests. P-values are uncorrected for multiple testing.

NYGC ALS Consortium – FTLD-TDP cohort

RNA-seq samples were downloaded from the January 2020 data freeze of the New York Genome Center (NYGC) ALS Consortium. The consortium includes a collection of 169 postmortem brain samples from the frontal cortex, temporal cortex, and cerebellum of 41 non-neurological controls and 37 FTLD-TDP donors. After strict quality control, this study assembled a cohort using 55 samples from 28 controls, and 74 samples from 30 FTLD-TDP donors. The 30 FTLD-TDP donors consist of 16 sporadic cases and 14 genetic cases, of which 9 carry mutations in *C9orf72,* 4 in *GRN*, and 1 in *TBK1* (**Table 1)**. De-identified clinical information for the cohort is presented in **Supplementary Table 1.** Distributions of RNA integrity number (RIN), post-mortem interval (PMI), and age at death are visualized in **Supplementary Fig. 1.**

The NYGC ALS Consortium samples presented in this work were acquired through various institutional review board (IRB) protocols from member sites and the Target ALS postmortem tissue core and transferred to the NYGC in accordance with all applicable foreign, domestic, federal, state, and local laws and regulations for processing, sequencing, and analysis. The Biomedical Research Alliance of New York (BRANY) IRB serves as the central ethics oversight body for NYGC ALS Consortium. Ethical approval was given and is effective through 08/22/2022.

Brain tissue dissection was performed by pathologists at the UCL Queen Square Brain Bank, the National Hospital for Neurology and Neurosurgery, and Oxford University. Cortical and cerebellar

regions were removed from each subject and divided into left and right hemispheres. One hemisphere was flash-frozen for transcriptome sequencing while the other was sectioned for histopathological evaluation. All donors were subject to a standard neuropathological workup.

The RNA sequencing procedures of the NYGC have been previously described [94]. RNA was isolated from the frozen brain tissue with TRIzol reagent and purified using RNeasy mini columns (Qiagen). RNA integrity numbers (RIN) for the brain samples were estimated on a Bioanalyzer (Agilent Technologies). RNA-Seq libraries were generated starting from 500 ng of total RNA using the KAPA Stranded RNA-Seq Kit with RiboErase (KAPA Biosystems) to remove rRNA and Illumina-compatible indexes (NEXTflex RNA-Seq Barcodes, NOVA-512915, PerkinElmer, and IDT for Illumina TruSeq UD Indexes, 20022370). Pooled libraries (average insert size: 375 bp) were then sequenced on either an Illumina HiSeq 2500 (125 bp paired end) or Illumina Novaseq (100 bp paired end). The sequenced samples were then subjected to extensive quality control protocols to confirm variables such as sex, tissue, and *C9orf72* repeat expansion status. *C9orf72* repeat expansions were identified on genotyped samples using the Asuragen AmplideX PCR/CE C9orf72 Kit and ExpansionHunter. Only samples with RNA integrity number (RIN) greater than 5 were chosen for study, due to the impact of low RIN on gene expression [90]. Of the 169 samples from the NYGC FTLD-TDP cohort, 25 FTLD-TDP and 14 control samples did not meet the RIN cutoff.

RNA-seq data processing and sample selection

RNA-seq samples were uniformly processed using RAPiD-nf, a processing pipeline implemented in the NextFlow framework. After trimming adapter sequences with Trimmomatic (version 0.36) [10], the samples were aligned to the hg38 build of the human reference genome (GRCh38.primary_assembly) using STAR (2.7.a) [25]. Gene counts were generated using RSEM (1.3.1) [59]. Quality control was performed using SAMtools [62] and Picard (https://broadinstitute.github.io/picard/), modeling the criteria of the Genotype Tissue Expression Consortium [38]. To identify outliers, we performed principal component analysis (PCA) on the normalised RNA-seq expression matrix, checking for points that did not cluster with their tissue type. One outlying sample was identified and removed from the study. Repeating PCA with this sample removed did not show any additional outliers or changes in clustering. For the present study, we employed samples from the frontal cortex, temporal cortex and cerebellum, spanning 74 samples from 30 FTLD-TDP donors and 55 samples from 28 non-neurological controls.

Covariate adjustment

Before performing DGE analysis, the RNA-seq expression matrix was normalised and adjusted for covariates. Normalisation was performed using trimmed mean of M values and transformed with the limma::voom() function [57]. Lowly expressed genes were removed using a threshold of >1 counts

per million in at least 90% of the samples. Covariate adjustment was performed separately for each brain region, and the steps described here were modeled from a large differential expression study [30]. First, clinical variables were combined with sequencing variables and technical metrics from Picard. Then, potential confounders were determined by ranking the variables based on their contributions to gene expression variance. The variance contributions of the covariates to each gene were scored using the limma::selectModel() function, which returned the number of genes with a lower Bayesian Information Criterion (BIC) as a result of adding the covariate to a base model containing only disease as the predictor (**Supplementary Fig. 2)**. For a given gene, a lower BIC indicated an improvement over the base model, or a larger portion of the variance explained by the covariate. The top ten covariates that improved the BIC for the largest number of genes were considered for model building**.** Correlating technical factors with each other allowed us to select a set of distinct factors for modelling (**Supplementary Fig. 3**).

To find the subset of covariates resulting in the best fitting linear regression model for DGE analysis, stepwise regression was performed by successively adding each variable to the base model. After evaluating each successive model with the limma::selectModel() function **(Supplementary Fig. 4a)**, we chose the model that improved the BIC for the largest number of genes. Using an orthogonal approach, variancePartition [43] was run on the covariates to quantify their contributions to gene expression variance **(Supplementary Fig. 4b)**. The following models were fitted for each brain region:

Table 2 - Selected models for FTLD-TDP vs control differentiation expression

Region	**Model**
Frontal Cortex	*expression ~ disease + sex + age + median 3' bias*
Temporal Cortex	*expression ~ disease + sex + age + median 3' bias + % intronic bases + % ribosomal bases*
Cerebellum	*expression ~ disease + sex + age + % R2 transcript strand reads*

We note that RIN is absent from the final models as it did not rank among the top ten contributors to gene expression variance **(Supplementary Fig. 2a)**. This is most likely due to our use of a RIN threshold during sample selection, thus reducing the variance explained by RIN in the final set of RNA-seq samples. Importantly, we show that RIN cannot be included in the final regression models as it correlates significantly with technical variables **(Supplementary Fig. 5).**

Differential gene expression analyses

Adjusting for the covariates described above, we performed differential gene expression (DGE) analysis to compare the FTLD-TDP cases with controls. The limma package [57] was used to compute the log_2-fold changes, t-statistics, and P-values for all genes tested in each brain region **(Supplementary Table 2)**. Genes having an adjusted $P < 0.05$ were considered differentially expressed. We then correlated the log_2-fold change effect sizes of each gene tested between each pair of brain regions. Next, we repeated DGE analysis, but split the FTLD-TDP cases by *C9orf72* repeat expansion status. After comparing the *C9orf72* and non-*C9orf72* cases with controls, we correlated the log_2-fold changes of all genes tested between each disease group in each region. DEGs (adjusted $P < 0.05$) were overlapped between the two groups. To check for confounding effects of post-mortem interval (PMI) on gene expression, we performed DGE analysis on the cases and controls respectively by using PMI as a continuous variable. We then associated the log_2-fold changes of the PMI-related genes and the FTD genes in each brain region and calculated Pearson correlation coefficients to quantify the effect of PMI.

Gene set enrichment analyses

Gene set enrichment analyses (GSEA) were conducted using the clusterProfiler package [107]. The inputs included the differentially expressed genes (DEGs), ranked by t-statistic, and pre-annotated sets of marker genes from multiple sources. Mouse marker genes for the five brain cell-types, microglia, astrocytes, oligodendrocytes, neurons, and endothelial cells were obtained from Neuroexpresso [65] and converted to their human homologues using homologene [66]. Marker genes for cellular pathways included the hallmark gene sets from the molecular signatures database (MSigDB) [60]. The six sets of glial activation genes, microglial activation, disease-associated microglia (DAM), disease-associated astrocytes (DAA), astrocyte reactivity (MCAO and LPS), and plaque-induced genes were obtained from their respective websites and supplementary materials [15, 40, 53, 65, 108]. For each combination of brain region DEGs and marker genes, GSEA ranks the DEGs by t-statistic, which has the sign of the direction of differential expression, and takes a running cumulative tally of the overlap with the genes in the set. The maximal score is the enrichment score (ES), which is a measure of how enriched a gene set is at the top or bottom of the ranked DEG list. The DEG list is randomly shuffled and the process repeated to generate an empirical null ES distribution to calculate a P-value. To compare between each gene set tested, each ES is divided by the mean null ES to create a normalised enrichment score (NES). A positive NES means a gene set is enriched at the top of the ranked list. All results from GSEA are provided in **Supplementary Table 4.**

Expression-weighted cell-type enrichment analysis

Expression-weighted cell-type enrichment analysis was performed using the EWCE package [68, 91]. We performed two separate analyses using the reference single-cell RNA-seq datasets by Mathys et al. [68] and Darmanis et a.l [20]. Using the reference single-cell data, cell-type specificity scores were calculated for the top 250 upregulated and downregulated genes for each brain region, ordered by t-statistic. The specificity scores of each set were then compared to the mean of the empirical null distribution from 10,000 randomly sampled gene lists. For each comparison, enrichment was expressed as the number of standard deviations from the mean. P-values were Bonferroni-corrected before applying a significance threshold of adjusted $P < 0.05$.

Cell-type deconvolution

Cell-type deconvolution was performed on the voom-normalised RNA-seq data using the reference datasets by Mathys et al. and Darmanis et al. [20, 68]. We first ran the dtangle package [47] to estimate cell-type proportions in the FTLD patients and controls. After regressing out the same clinical and technical variables as in the differential expression modelling, we applied the Wilcoxon rank sum test to compare the estimated proportions of each cell-type between patients and controls. For all comparisons, P-values were Bonferroni-corrected for multiple-testing, and significance was set at adjusted $P < 0.05$. Correlating the dtangle estimates from the two reference datasets, we identified strong associations across the four overlapping cell-types. Using the Darmanis reference, we also ran the MuSiC algorithm, a state-of-the-art deconvolution method that accounts for cross-subject variance in gene expression [102]. Deconvolution results from dtangle and Mathys et al. are presented in **Supplementary Table 7.**

Correlations with neuropathological atrophy scores

Microscopic atrophy was assessed on hematoxylin and eosin stained slides. Sections were examined for the cortical thickness and neuronal loss in the frontal and temporal cortices in a four point grading system and compared to a neurological normal control with no underlying neurodegenerative changes: 0 – the cortical thickness was within normal limits and no neuronal loss; 1 – reduction in cortical thickness but the number of neurons comparable to normal levels; 2 – reduction in cortical thickness and reduction in numbers of neurons; 3 – severe reduction in cortical thickness and no neurons observed. The microscopic atrophy in each region was scored semi-quantitatively at an objective magnification of x20. Scoring for all cases was performed by a single experienced observer blinded to clinical, histopathological and genetic status. We note that none of the FTLD-TDP samples in this study warranted a microscopic atrophy score of 0.

Macroscopic atrophy was determined from observations of gyri and sulci from the coronal slices observed during brain cutting procedures. Levels of atrophy were graded according to four stages (none, mild, moderate, and severe) as set out in [89]. Representative example photos are provided in the Supplement. **Supplementary Table 8** lists the atrophy scores for the subset of 45 FTLD-TDP samples with RNA-seq data, and includes whether atrophy is symmetrical between both hemispheres or is less or more severe on the hemisphere used for sequencing.

We then correlated the neuropathological scores of the frontal and temporal regions with the estimated proportions of each cell-type (excitatory neurons, inhibitory neurons, endothelial cells, pericytes, astrocytes, oligodendrocytes, and microglia). For both microscopic and macroscopic atrophy we used Wilcoxon rank sum tests to compare the proportions between each successive atrophy stage.

Correlations with hippocampal TDP-43 burden

Hippocampal TDP-43 burden scores were obtained for 21 of the 30 FTLD-TDP donors used in this study. Slides with 8 µm mounted tissue sections were incubated at 60 °C overnight. Sections were deparaffinized in Xylene and rehydrated in decreasing grades of alcohol. Slides were incubated in methanol/hydrogen peroxide (0.3%) solution for 10 min to block endogenous peroxidase activity. For heat-induced antigen retrieval, slides were then transferred to a boiling solution of 0.1 M citrate buffer (pH 6.0) and pressure cooked at maximum pressure for 10 min. Slides were then incubated in 10% non-fat milk for 30 min at room temperature to block non-specific binding. Sections were incubated with phospho-TDP43 antibody (Cosmo Bio (TIP-PTD-P02)) for 1 hour at room temperature, washed and then incubated for 30 mins in biotinylated goat anti-rabbit IgG secondary antibody (Vector Laboratories). Slides were washed and then incubated in Streptavidin–Biotin Complex (ABC; DAKO). The slides were then washed for a final time before being submerged in 3,3'-Diaminobenzidine (DAB) chromogen and then counterstained in Mayer's haematoxylin (BDH). Finally, slides were dehydrated in increasing grades of alcohol (70, 90 and 100% IMS), cleared in xylene and mounted. To construct the burden scores, stained sections were digitally scanned using an Olympus slide scanner. A percentage of phospho-TDP-43 positive inclusions was determined from the number of inclusions counted in the total number of neurons in the granule cell layer of the hippocampus. Hippocampal burden scores were correlated with the estimated cellular proportions from the frontal, temporal, and cerebellar samples of the 16 matching donors.

Correlations with microglial burden scores

Microglia burden scores for a subset of the FTLD-TDP samples (**Supplementary Table 10)** were extracted from the supplementary data provided by Woollacott et al. [104]. Briefly, frontal and temporal tissue was fixed and immunohistochemically stained with antibodies against IBA1, CD68

and CR3/43 and 3,3′-Di-aminobenzidine (DAB) was used as the chromogen. Slides were scanned and digitised for analysis with ImageJ. For each of the three microglial markers, the number of positive cells were divided by the total number of DAB-positive cells in each section to construct burden scores. These scores were then correlated with the estimated microglial proportions.

Correlations with DNA methylation estimates of neuronal proportion

Genomic DNA was extracted from flash-frozen frontal cortex grey matter tissue using standard protocols. A bisulfite conversion was performed with 500 ng of genomic DNA using the EZ DNA Methylation Kit (Zymo Research). Genome-wide methylation profiling was performed using the Infinium HumanMethylationEPIC BeadChip (Illumina), as per the manufacturer's instructions. Beta values were used to estimate the methylation levels of each CpG site using the ratio of intensities between methylated and unmethylated alleles. Beta values range from 0 to 1, representing approximately 0% to 100% methylation, respectively. Data analysis was performed using several R Bioconductor packages. Briefly, raw data (idat files) were imported into R for pre-processing and quality control using minfi [2], ChAMP [98], and watermelon [78]. Probes that met one or more of the following criteria were excluded from further analysis: 1) poor quality, 2) cross-reactive, 3) overlapped common genetic variants, and 4) mapped to X or Y chromosome. Samples were dropped during quality control if: 1) presenting with high failure rate, 2) the predicted sex was not matching the phenotypic sex, and 3) inappropriately clustering on multidimensional scaling analysis. Beta values were normalised with ChAMP using the Beta-Mixture Quantile (BMIQ) Normalisation method, and neuronal proportions were then estimated using the CETS package [39]. For the subset of 17 FTLD-TDP samples with DNA methylation and RNA-seq data, the neuronal proportion estimates from DNA methylation **(Supplementary Table 9)** were correlated with excitatory and inhibitory neuron estimates from gene expression.

Differential Transcript Usage Analysis

Transcript expression was estimated in each sample using RSEM [59] with the GENCODE v30 transcript reference. Lowly expressed transcripts were removed with the threshold transcript counts per million > 1 in at least 30% of all samples. Differential transcript usage (DTU) was compared between FTLD-TDP cases and controls in each brain region using satuRn [35], a fast method for computing differential transcript usage. The same clinical and technical covariates were used as in the differential expression modelling. Pairwise comparisons between cases and controls were extracted using the limma::makeContrasts() function. We then filtered the data using a FDR threshold of 0.05, which yielded totals of 4637, 5098, and 8249 transcripts from the frontal cortex, temporal cortex, and cerebellum respectively. DTU results are presented in **Supplementary Table 3,** which includes effect sizes, P-values, and adjusted P-values for all transcripts tested.

To understand the functional profiles of the DTU genes (gDTUs), we performed a series of hypergeometric enrichment tests. We split the genes into two sets: gDTUs only, and gDTUs shared with DEGs. Using the clusterProfiler::enricher() function, we compared each set against a panel of marker genes representing the same cell-types, glial activation states, and cellular pathways as in the gene-set enrichment analyses. P-values from these enrichment tests are presented in **Supplementary Tables 5 & 6.**

Comparisons with TDP-43 knockdown and overexpression genes

Differential expression results were obtained from a TDP-43 knockdown study in IPS-derived cortical neurons [13]. The results included P-values and expression $\log_2$ fold-changes of all genes tested, as well as Boolean values (TRUE/FALSE) indicating whether or not a gene is predicted to contain a cryptic exon. DEGs (adjusted $P < 0.05$) from TDP-43 knockdown were compared with DEGs from the FTLD-TDP vs Control comparison in each brain region. Furthermore, genes flagged as having cryptic exons were compared with DTU genes. P-values and odds ratios of all overlaps were computed using a one-sided Fisher's exact test. Log_2-fold changes between the FTLD-TDP and TDP-43 knockdown genes were compared using Spearman correlations. In addition, comparisons were made with genes with impaired RNA stability following TDP-43 overexpression in IPS-derived neurons [95]. This study used bromouridine labelling of nascent transcripts to measure RNA stability. Destabilized genes were overlapped with the FTLD-TDP DEGs and DTU genes. Similarly to the previous comparisons, one-sided Fisher's exact tests were used to calculate all P-values and odds ratios.

Data availability

All raw RNA-seq data can be accessed via the NCBI GEO database (GEO GSE137810, GSE124439, GSE116622, and GSE153960). All RNA-seq data generated by the NYGC ALS Consortium are made immediately available to all members of the Consortium and with other consortia with whom we have a reciprocal sharing arrangement. To request immediate access to new and ongoing data generated by the NYGC ALS Consortium and for samples provided through the Target ALS Postmortem Core, complete a genetic data request form at ALSData@nygenome.org.

Code availability

All code written for this project was written in the R language (v4.0.3), making extensive use of the tidyverse and Bioconductor suites of packages [33, 103], along with UpsetR [18], [105], and ggpubR [51] for visualisation and descriptive statistics. All code is hosted as Rmarkdown files on Github: https://github.com/jackhump/FTLD-TDP_analysis

Results

Differential gene expression analyses reveal changes across multiple brain regions

We analysed RNA-seq data generated from three post-mortem brain regions (frontal cortex, temporal cortex, and cerebellum) from 30 FTLD-TDP patients and 28 non-neurological controls as part of the New York Genome Center ALS Consortium. After extensive quality control, a total of 74 FTLD-TDP and 55 control RNA-seq samples were used to perform differential gene expression (DGE) analysis **(Table 1)**. On average, the post-mortem interval (PMI) was longer in the FTLD-TDP samples ($P < 1e\text{-}5$), whereas the average age of death of the FTLD-TDP donors was slightly lower than that of healthy controls ($P < 0.05$); **Supplementary Fig. 1)**. FTLD-TDP donors had higher Braak Tau stage pathology than controls ($P < 0.05$), but otherwise did not differ on any other neuropathological measurement. Full demographic and clinical details of the cohort are listed in **Supplementary Table 1**.

We first identified general gene expression patterns in the FTLD-TDP patients by performing principal component analysis (PCA) **(Fig. 1a)**. The cerebellar samples formed a distinct cluster from the cortical samples, as expected due to cell-type differences between the cortex and cerebellum. The 2nd principal component (16.4% variance explained) indicated a clear separation between the FTD cases and controls in all three brain regions. Differentially expressed genes (DEGs) were then computed (adjusted $P < 0.05$) for each brain region by comparing the FTLD-TDP patients with controls, while adjusting for clinical and technical variation **(see Methods)**. We observed widespread changes in gene expression across the brain, with the largest number of unique genes observed in the frontal cortex (1711 DEGs), followed by temporal cortex (709 DEGs) and cerebellum (438 DEGs). For all the brain regions, we overlapped the FTLD-TDP DEGs **(Fig. 1b)**, and found that the frontal and temporal cortex shared the largest number of genes (FDR < 0.05 in both regions), while overlaps between cortical tissues and the cerebellum were lower. We correlated the $\log_2$-fold change effect sizes of each gene tested between each pair of brain regions **(Fig. 1d)**. The frontal and temporal cortex showed a strong positive correlation ($R = 0.840$), suggesting that these two areas have similar gene expression profiles in FTLD-TDP. In the correlations involving the cerebellum, the directionality of the expression changes were concordant, but the effect sizes were weaker ($R = 0.476$ with frontal cortex; $R = 0.468$ with temporal cortex), implicating both shared and distinct gene expression changes in the cerebellum.

We identified genes that were previously implicated with FTLD and/or ALS **(Fig. 1c)** including *TARDBP, CHCHD10*, *DCTN1*, *GFAP* and *ADAMTS2*. *TARDBP,* the gene encoding TDP-43, was upregulated in both cortical regions but not the cerebellum. *CHCHD10* encodes a mitochondrial protein, and missense mutations in this gene have been observed in familial FTD cases and

associated with impaired oxidative phosphorylation [5]. *DCTN1,* which encodes a motor protein involved in vesicular transport, has been shown to be downregulated in FTD and ALS [56, 72]. *GFAP* is a marker of astrocytes previously reported to be upregulated in FTD [41], and the *ADAMTS2* gene encodes an enzyme responsible for collagen production [48]. A recent study has identified *ADAMTS2* as a marker of Von-economo neurons, a neuronal subclass vulnerable to TDP-43 pathology [42]. In the cerebellum, we highlight two downregulated genes linked to cerebellar Purkinje neurons, *CALB1* and *HOMER3*, which have been implicated in motor disorders such as Spinocerebellar ataxia and Huntington's disease [1, 6, 73, 86].

To check for confounding effects due to post-mortem interval (PMI), we performed DGE analysis using PMI as a continuous variable. In neither control nor FTLD samples could we find any significant PMI-related genes (FDR < 0.05) in any brain region. In addition, we correlated the $\log_2$-fold changes between all FTLD-related and PMI-related genes **(Supplementary Fig. 6)**. Although we found positive correlations in each region, the effect sizes associated with PMI were negligible.

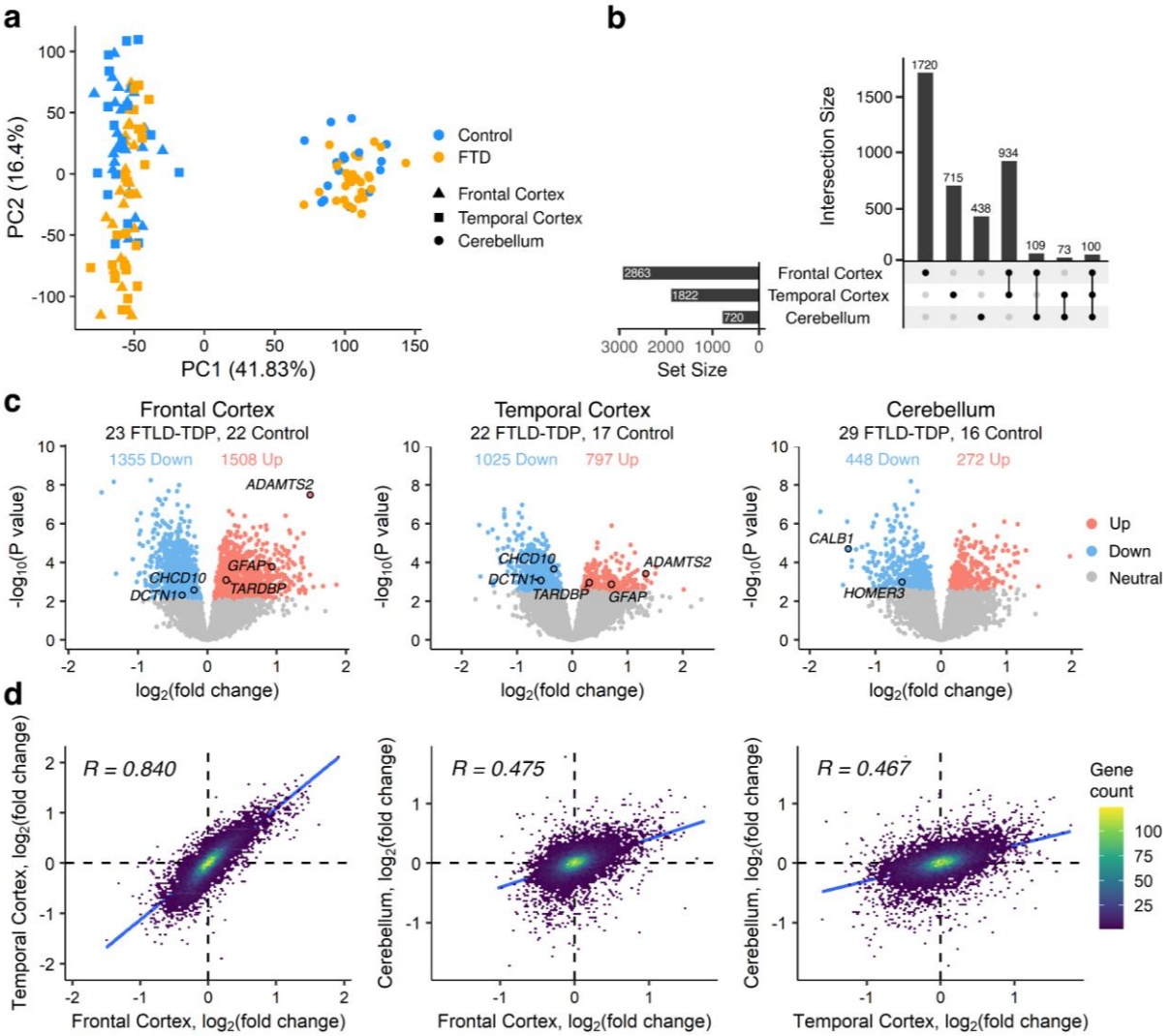

Fig. 1 | Overview of differential expression analyses. a. Principal component analysis of the RNA-seq expression matrix, following TMM-normalisation and covariate adjustment. Colour corresponds to disease status and shape corresponds to brain region. Distinct clusters can be seen between the cortical and cerebellar samples. **b.** Upset plot showing the number of distinct and overlapping differentially expressed genes in each brain region. **c.** Volcano plots comparing FTLD-TDP samples with controls. Red and blue dots represent genes that are upregulated and downregulated respectively (FDR adjusted $P < 0.05$), while gray dots are genes that are not differentially expressed. Key genes related to FTLD or ALS are labelled. **d.** Scatter plots comparing log_2-fold changes of all genes tested between each pair of brain regions. Each point is a gene, coloured by the density of overlapping points. Log_2-fold changes are highly concordant between the frontal and temporal cortex, but less so between the cerebellum.

C9orf72 and non-C9orf72 cases have similar gene expression profiles

We repeated DGE analysis but split the FTLD-TDP cases by *C9orf72* repeat expansion status. This allowed us to compute the gene expression changes, with respect to controls, in the 9 *C9orf72* and 21 non-*C9orf72* FTLD-TDP donors, the latter mostly consisting of sporadic disease. In the frontal cortex, most DEGs were shared between the two disease groups **(Supplementary Fig. 7a)**. Gene expression changes between the two groups were strongly concordant across the regions (frontal cortex: $R = 0.89$; $P < 2.2e\text{-}16$; temporal cortex: $R = 0.90$; $P < 2.2e\text{-}16$; cerebellum: $R = 0.76$; $P < 2.2e\text{-}16$) **(Supplementary Fig. 7b)**. Taken together, these findings are contrary to studies that have reported distinct sets of differentially expressed genes in *C9orf72* repeat expansion carriers and sporadic FTD or ALS patients [24, 80]. Furthermore, we observed that the *C9orf72* gene itself was significantly downregulated in the *C9orf72* cases, which could be explained by hypermethylation of the *C9orf72* promoter locus [49].

Pathway analysis finds upregulated inflammatory response and circulatory system

Next, we performed gene set enrichment analysis (GSEA) [93] to examine affected cellular pathways in the FTLD-TDP patients **(Fig. 2a; Supplementary Fig. 8)**. In all three brain regions, the upregulated genes were strongly enriched for epithelial mesenchymal transition, an extracellular matrix (ECM) remodelling process, and circulatory system pathways such as angiogenesis, heme metabolism and coagulation. We also examined the expression-fold changes of matrix metalloproteinases (MMPs), which have emerged as important regulators of the ECM and circulatory system [82]. Increased expression levels of MMPs have been previously reported in various neurodegenerative diseases [26, 64]. Supporting a strong involvement in FTLD, we have found that most MMP genes were differentially upregulated in the cortex but not the cerebellum, and that two genes, *MMP2* and *MMP14*, were among the DEGs with the largest log_2-fold changes **(Supplementary Fig. 9)**.

Furthermore, we observed that pathways related to immune signaling and inflammatory response were positively enriched in the cortex, but not in the cerebellum. TNF-alpha signaling via NFkB, which is believed to trigger microglia-induced neurodegeneration [69], displayed the strongest enrichment among the immune signaling pathways. Accordingly, we found strong involvement of

p53 pathways and apoptosis, which may act together with TDP-43 to mediate neuronal cell death [100]. Cell-proliferation pathways such as MTORC1 and PI3K/AKT/MTOR signaling were upregulated in the frontal cortex, while mitotic spindle and cholesterol homeostasis were upregulated in the cerebellum.

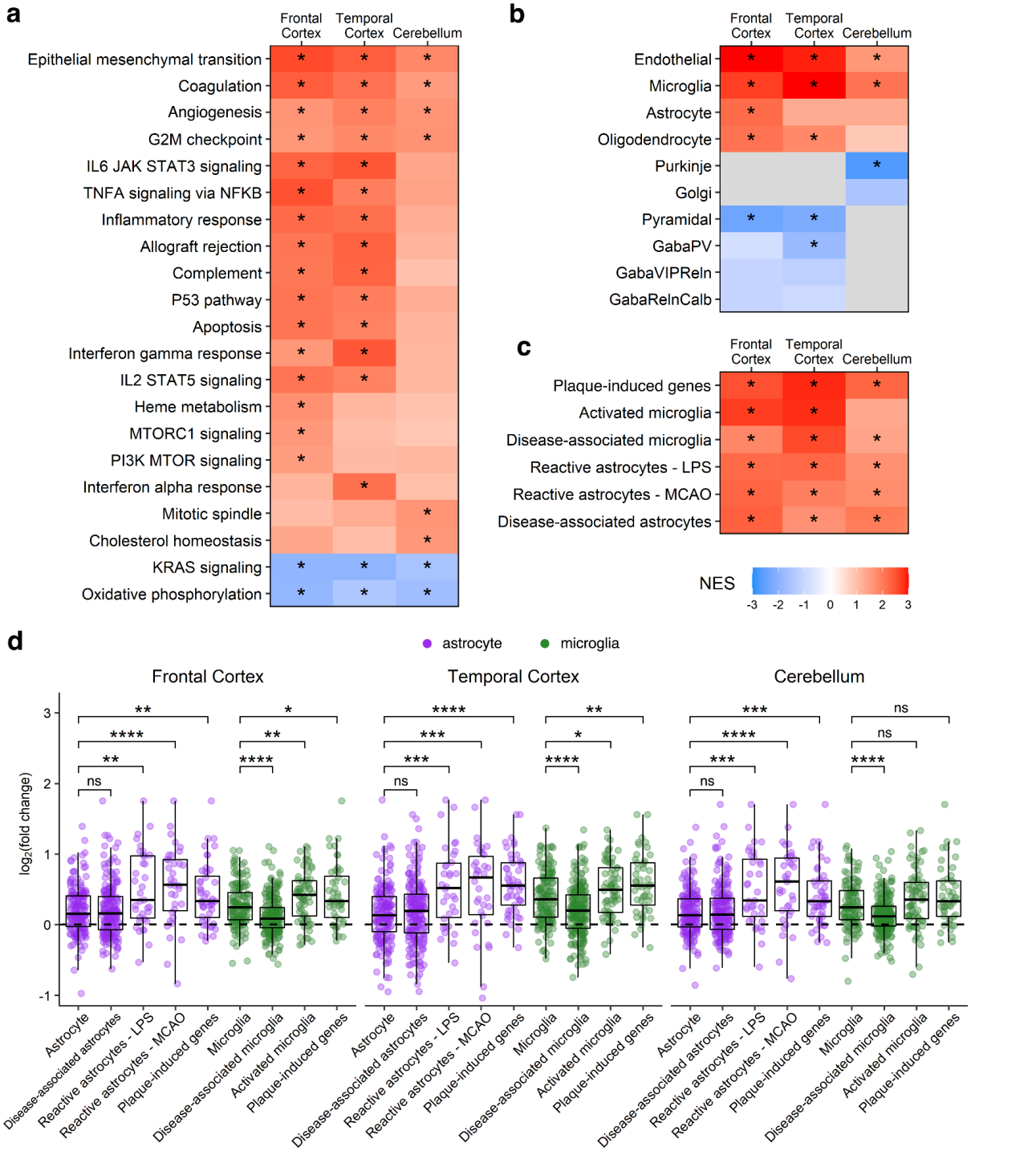

Fig. 2 | Cellular and pathway-level enrichment analyses. a-c. Heatmaps showing the GSEA results for cellular pathways, cell-types, and glial activation states. Coloured tiles represent the normalised-enrichment score (NES) for each term. Grey tiles were not tested. **d**. Boxplots comparing the $\log_2$-fold changes between the activation genes in **c** and standard marker genes in **b** for astrocytes and microglia. Comparisons were made using the Wilcoxon rank sum test. All P-values from **a-d** are Bonferroni-corrected. *** p < 1e-4; ** p < 1e-3; * p < 0.05; ns p > 0.05.

Significant alterations in endothelial, glial, and neuronal gene expression

To understand how gene expression in individual brain cell-types is altered in patients with FTLD-TDP compared to controls, we performed GSEA using marker genes [65] for five cell types (microglia, astrocytes, oligodendrocytes, neurons, and endothelial cells) **(Fig. 2b)**. Endothelial genes were upregulated in all three brain regions and showed the strongest enrichment of the five cell-types. Enrichment for glial-cell types varied across the brain, with microglial genes upregulated in all three regions, astrocyte genes upregulated in only the frontal cortex, and oligodendrocyte genes showing increases in both the frontal and temporal cortex.

We also observed an overall reduction in neuronal gene expression. Examining the individual neuronal subtypes, Purkinje neuron expression decreased in the cerebellum, and pyramidal neuron expression decreased in the cortex. Our expression-weighted cell-type enrichment (EWCE) [91] analysis yielded similar results **(Supplementary Fig. 10)**, showing strong loss of excitatory neuronal markers in the cortices and loss of inhibitory neuronal markers in the cortices and cerebellum. These changes could be explained by general nervous system defects caused by the disease and/or loss of neurons in the FTLD samples. A recent single-cell RNA-seq study has defined a set of marker genes for Von-economo neurons (VENs), a population of excitatory cortical neurons that are selectively reduced in TDP-43 pathology [42]. Contrary to the known vulnerability of VENs, we discovered that most of the marker genes were differentially upregulated, with the upregulated genes exhibiting larger expression fold changes than downregulated genes **(Supplementary Fig. 11)**.

To further characterise glial cell-type expression, we calculated enrichment for activated glia using a panel of immune marker genes from previous transcriptomic studies (**Fig. 2c**). The gene sets represent signatures of microglia and astrocyte responses in various mouse models of disease and/or inflammation. Disease-associated microglia (DAM) ([15, 40, 53, 65, 108]) and disease-associated astrocytes (DAA) ([15, 40, 53, 65, 108]) were identified using single-nucleus RNA-seq in the 5XFAD mouse model of Alzheimer's disease. Activated microglia genes were defined as genes upregulated in sorted mouse microglia after treatment with lipopolysaccharide (LPS) [44], as curated in Neuroexpresso [65], while middle cerebral artery occlusion- (MCAO) and lipopolysaccharide (LPS)-reactive astrocytes were identified in sorted astrocytes from mice subjected to brain ischemia and LPS stimulation respectively ([15, 40, 53, 65, 108]). Finally, plaque-induced genes were defined as genes enriched near amyloid plaques using spatial transcriptomics in the *App*$^{NL-G-F}$ mouse model of Alzheimer's disease and potentially represent a broad immune response that encompasses microglia, astrocytes, and oligodendrocytes ([15, 40, 53, 65, 108]). We found that the six gene sets only partially overlapped (**Supplementary Fig. 12**) and all were upregulated in the three brain regions. To understand which activation signatures disrupt glial homeostasis and are likely to play a role in inflammation, we compared the $\log_2$-fold changes of the activation genes with those of the

standard marker genes (**Fig. 2d**). Comparing the microglial activation sets with the microglia markers in each region, we observed higher log_2-fold changes for plaque-induced genes and activated microglia, but lower log_2-fold changes for DAM genes. In the comparisons involving astrocytes, plaque-induced genes and the two astrocyte reactivity sets (MCAO and LPS) exhibited higher fold changes, but DAA genes did not show any significant differences. These results point to the involvement of distinct microglia and astrocyte signatures which may play roles in the inflammatory response seen in FTLD-TDP.

Altered compositions of endothelial cells, microglia, and neurons

Next, to estimate cellular composition in the FTLD-TDP and control samples, we performed deconvolution analysis using a human cortical snRNA-seq dataset [20, 68]. Across all samples, we observed that neurons and astrocytes were among the most abundant cell-types, while microglia were the least abundant. Correlating the estimated cell-type proportions with each other revealed negative associations between neurons and glia which remained even after splitting cases and controls (**Supplementary Fig. 13)**. The correlations in the FTLD-TDP cases, moreover, displayed stronger effect sizes than controls. To assess changes in cellular composition, we compared the estimated proportions between all control and FTLD-TDP samples **(Fig. 3a; Supplementary Figs 14-16)**. Endothelial cells and pericyte proportions increased in all three regions, while excitatory and inhibitory neurons were decreased in the cortex. Inhibitory neurons were also reduced in the cerebellum, consistent with the marker gene expression changes for inhibitory Purkinje neurons. Notably, proportions of microglia, astrocytes, and oligodendrocytes were increased in the cortices, but not in the cerebellum, contrary to our GSEA results with marker genes. This could be due to the fact we used cellular markers from the human cortex, which may not be fully representative of the cerebellum, or due to the use of covariate adjustment for the GSEA but not the deconvolution. Overall, the composition changes were broadly consistent with the GSEA **(Fig. 2c)** and EWCE **(Supplementary Fig. 10)** results. Using a different deconvolution tool and another reference dataset [20, 68], we obtained similar results for the overlapping cell-types **(Supplementary Fig. 11-13)**. We also compared deconvolution estimates between all *C9orf72* and non-*C9orf72* samples, but failed to detect differences in cellular composition in any of the brain regions **(Supplementary Fig. 17)**.

To validate our deconvolution results, we correlated the estimated neuronal proportions for each sample with manually determined atrophy scores from a neuropathologist. Microscopic atrophy was assessed by grading both the amount of neuron loss and changes in cortical thickness at 20X magnification **(Fig. 3b; Supplementary Fig. 18)** whereas macroscopic atrophy assessed the gross brain atrophy of the frontal and temporal cortices **(Supplementary Fig. 19)**. Samples were scored 0-3 based on no, mild, moderate or severe atrophy. In the correlations involving macroscopic atrophy, we failed to identify associations in either the frontal or temporal cortex (**Supplementary**

Fig. 19). However, we observed significant correlations with the microscopic atrophy scores of the frontal cortex samples **(Fig. 3c)**. Both excitatory and inhibitory neuronal proportions displayed strong negative associations (**Fig. 3c)** with microscopic atrophy, thereby providing experimental confirmation of our computational predictions. Excitatory neurons were more strongly associated with the scores than inhibitory neurons, suggesting that in the frontal cortex, neuronal loss can be mainly attributed to a loss of excitatory neurons. Conversely, endothelial and microglial cells showed strong positive associations with atrophy scores. These findings may indicate a potential relationship between neuronal loss and microglial and endothelial cell changes in the cortex.

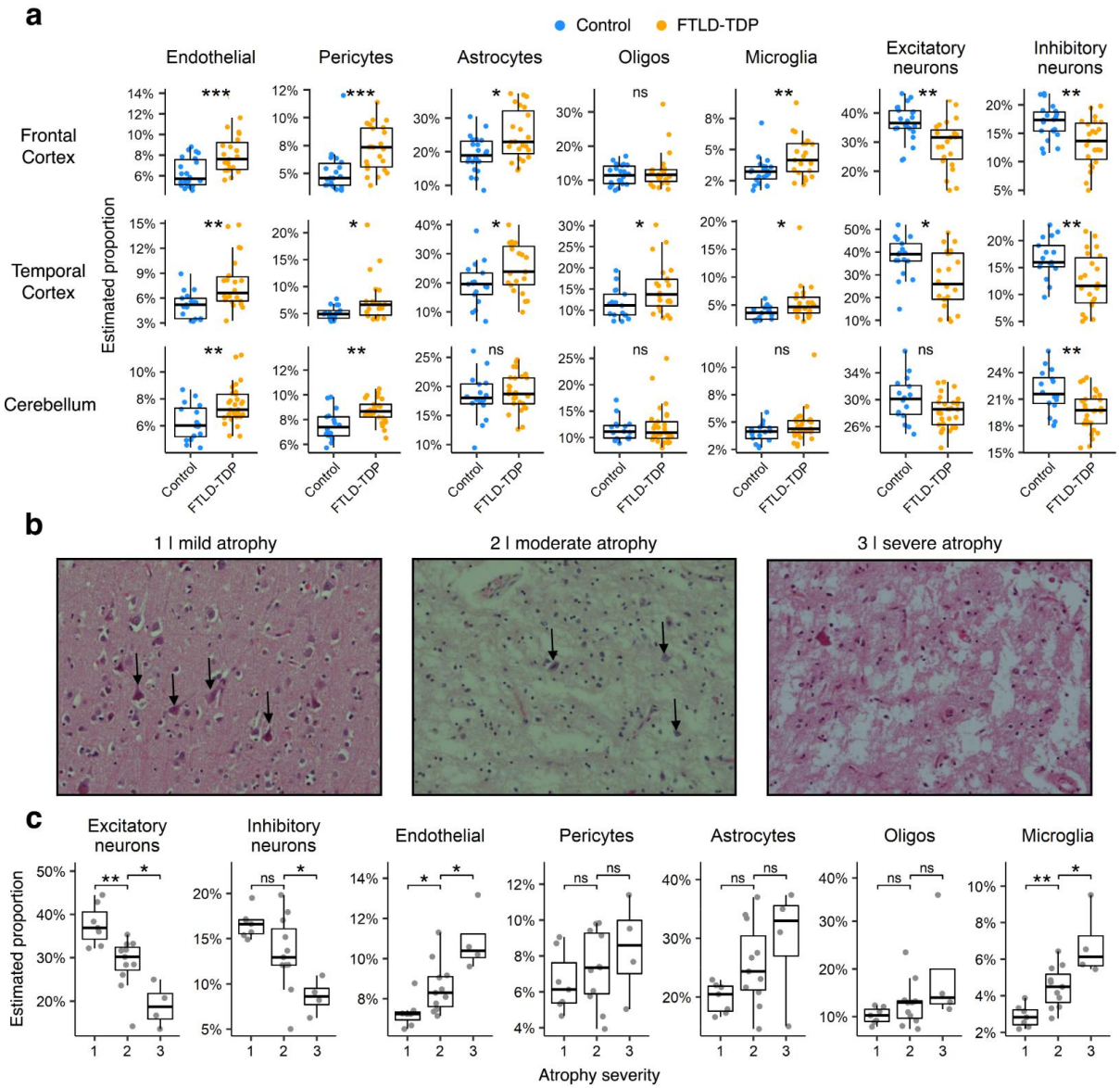

Fig. 3 | Cellular proportion changes and atrophy correlations. a. Comparisons of the cellular proportion estimates between control and FTLD-TDP samples. Estimates were generated with dtangle using single nuclear RNA-seq data from Mathys et al. **b**. Hematoxylin and eosin stained sections of select frontal cortex FTLD-TDP samples. From left to right, the slides depict cases with mild, moderate, and severe microscopic atrophy. Neurons are indicated by black arrows. **c.** Comparisons of the cellular proportion estimates of the frontal cortex FTLD-TDP samples between each microscopic atrophy stage. Excitatory neurons are associated with neuronal loss as the estimates show significant decreases between all atrophy stages. Inhibitory neurons show decreases between stages 2 and 3, but not 1 and 2, indicating a partial negative association. Positive

associations with neuronal loss are observed for endothelial cells and microglia. Asterisks in **a** and **c** represent Bonferroni-adjusted P-values from Wilcoxon rank sum tests. *** p < 1e-4; ** p < 1e-3; * p < 0.05; ns p > 0.05.

Similarly to using gene expression, DNA methylation of purified cell-types can be used to deconvolve the cell-type composition of mixed samples. We used methylation calls for 17 of the FTLD-TDP frontal cortex donors to estimate neuronal proportions in those samples. However, these estimates showed no correlation with estimates derived from gene expression (R = 0.15; P = 0.56; **(Supplementary Fig. 20)**).

Additionally, we compared the estimated microglial proportions with histologically determined microglial proliferation scores on a subset of the same samples from a previous study [104]. 14 frontal cortex and 13 temporal cortex samples were shared between the two studies. Each sample was stained for IBA1, a constitutive marker for microglia, CD68, a marker of activated phagocytic microglia, and CR3/43, which detects major compatibility complex class II molecules including HLA-DR, HLA-DP and HLA-DQ, present in activated antigen-presenting microglia. When correlating counts and combined counts of cells positive for any of the three stains, no measure significantly correlated with microglia proportion **(Supplementary Fig. 21)**. Similarly, we compared estimated cell-type proportions with histologically determined TDP-43 burden scores. The burden scores measured the number of neurons with TDP-43 inclusions in hippocampal tissue from 21 of the 30 FTLD-TDP donors. No correlation was observed between these scores and the estimated cell-type proportions **(Supplementary Fig. 22)**.

Differential transcript usage analysis highlights additional genes and pathways

To further characterise differences in gene expression between the FTLD-TDP patients and controls, we applied differential transcript usage (DTU) analysis. This allowed us to identify expression changes, at the level of transcripts, caused by alterations in polyadenylation, promoter usage, and alternative splicing. Adjusting for clinical and technical variation **(see Methods)**, the analysis revealed 2630, 2891, and 4045 genes with differential transcript usage (gDTU) in the frontal cortex, temporal cortex, and cerebellum respectively **(Fig. 4a)**. Most gDTUs did not overlap with the DEGs in any brain region **(Fig. 4a)**. Interestingly, the gene encoding TDP-43, *TARDBP,* was found within the gDTU-DEG overlap for the frontal cortex. Examining the DTU signals of *TARDBP*, we found downward shifts in transcript usage for ENST00000613864, a predicted nonsense-mediated decay transcript, and ENST00000315091, a transcript employing an alternative transcription start site. **(Supplementary Fig. 23a)**. Although splicing of intron 7 within the TARDBP 3'UTR has been suggested to be the mechanism of TDP-43 regulating its own expression [55], none of the 3'UTR splicing transcripts were sufficiently expressed to test for DTU. We also detected significant DTU for *HNRNPK*, which encodes hnRNPK, an RNA-binding protein shown to colocalize with cytoplasmic TDP-43 aggregates in the frontal cortex [52]. hnRNPK nuclear mislocalization has recently been

observed to be a common neuropathology in FTLD and aged brains [4]. For *HNRNPK,* we observed that usage of ENST00000376263 rose in the cerebellum, while that of ENST00000481820, ENST00000376281, ENST00000483135, and ENST00000376256 decreased in one or more of the brain regions, suggesting a complex shift in promoter usage, polyadenylation site selection and intron retention **(Supplementary Fig. 23b)**.

To understand the functional consequences of TDP-43 pathology, we performed a series of enrichment tests on the gDTUs. We split the genes into two sets: gDTUs only, and gDTUs shared with DEGs. We then compared each set against a panel of marker genes representing the same cell-types, glial activation states, and cellular pathways as before **(Fig. 4b-c)**. gDTUs, either alone or in combination with DEGs, showed significant enrichment (adjusted $P < 0.05$) for inhibitory neurons and oligodendrocyte progenitor cells in at least one brain region. Inputting only gDTUs, we found a strong presence of DAM genes in the frontal cortex, and DAA genes in the temporal cortex and cerebellum. We also identified several pathways in certain regions that were overlooked by our previous enrichment analysis using DEGs, such as mitotic spindle and apical junction, two pathways used by non-neuronal cell types. We show the transcript-level expression changes of *UNC13B*, a synaptic gene which may have important functional relevance in FTLD-TDP **(Fig. 4d)**. Brown and colleagues have previously reported a loss of *UNC13B* expression due to splicing changes caused by TDP-43 knockdown in neuronal cell lines [13]. Elaborating on this finding, we observed that *UNC13B* was downregulated in the cortices, and that four of the six *UNC13B* transcripts were altered in at least one brain region. The four transcripts differ by cassette exon inclusion, 3'UTR length and alternative promoter usage. Usage of ENST00000619578 decreased across all three brain regions, and the effect sizes associated with this isoform were among the largest in the DTU results. Usage for ENST00000378495 also decreased within the temporal cortex and cerebellum, while that of ENST000061708 and ENST0000036694 increased.

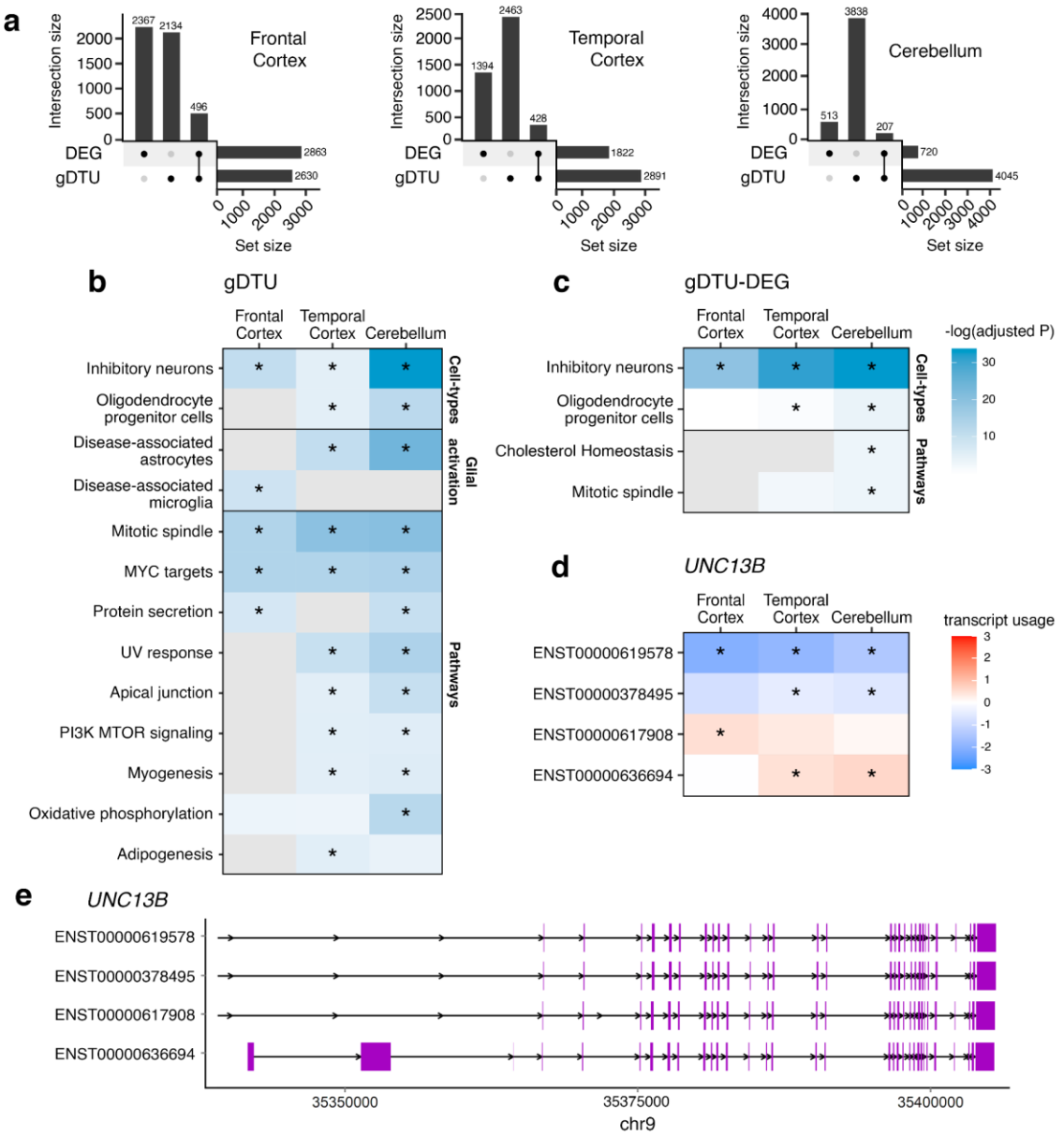

Fig. 4 | Differential transcript usage analysis. **a.** Upset plots showing overlaps between the gDTUs (adjusted P < 0.05) and DEGs (adjusted P < 0.05) for each brain region. **b-c.** Cell-types, glial activation states, and pathways enriched in gDTUs **(b)** and gDTUs + DEGs **(c)**. Coloured tiles represent the negative logarithm of the adjusted P-value for each term. Grey tiles were not tested. **d.** Transcript usage of the four *UNC13B* isoforms with significant effect sizes in at least one brain region. **e.** Alignment chart of the *UNC13B* isoforms in **d**, showing the region with greatest variation in transcript structure. Purple rectangles correspond to protein-coding exons. Asterisks in **b-d** refer to P-values that meet the FDR threshold of 0.05. *** p < 1e-4; ** p < 1e-3; * p < 0.05; ns p > 0.05.

The finding of DTU of *UNC13B* led us to consider whether changes in gene expression and splicing directly related to TDP-43 pathology might be observed in our bulk brain samples. TDP-43 nuclear loss pathology should be present in both frontal and temporal cortex but is not generally observed in the cerebellum [76]. Recently, RNA-seq on human iPSC-derived neurons identified 3,910 differentially expressed genes resulting from TDP-43 knockdown, as well as 161 genes with possible cryptic exon (CE) inclusions [13]. We compared the DEGs from our FTLD-TDP vs control comparison with those found in the study by Brown et al **(Supplementary Fig. 24a)**. These TDP-43 knockdown

genes overlapped with at least 25% of the FTLD-TDP DEGs found in each tissue (frontal cortex: OR = 1.23; P = 1.65e-5 ; temporal cortex: OR = 1.42; P = 1.12e-9 ; cerebellum: OR = 1.64; P = 1.02e-8). Furthermore, we overlapped our gDTUs with genes with possible CE inclusions and identified 52, 58, and 72 genes in the frontal cortex, temporal cortex, and cerebellum respectively **(Supplementary Fig. 24b)**. By correlating the expression-fold changes between the two gene sets **(Supplementary Fig. 24c)**, we also found weak, but nonetheless significant associations (P < 0.05) in all three regions (frontal cortex: R = 0.14; P < 2.2e-16 ; temporal cortex: R = 0.13; P < 2.2e-16 ; cerebellum: R = 0.058, P = 3.2e-11). In contrast, using a set of 1,243 genes found to be destabilized in IPS-derived neurons by TDP-43 overexpression [95], we found that only DEGs from the temporal cortex showed a significant overlap (OR = 1.52; P = 1.4e-4) (**Supplementary Fig. 25a)**. DTU genes failed to show significant overlaps in any of the brain regions (**Supplementary Fig. 25b**.

Discussion

In this study, we assembled a large cohort of RNA-seq from post-mortem FTLD-TDP brains to understand the cellular mechanisms underlying FTLD and TDP-43 pathology. We provide a detailed landscape of gene expression alterations from multiple brain regions, highlighting the roles of specific glial cell-types, the vulnerability of excitatory neurons, and a strong involvement of the cerebellum in the neurodegenerative process of FTLD. Our data represents a transcriptomic resource to help accelerate future research towards a better understanding of the pathogenic mechanisms in FTD.

By profiling transcriptomic changes in FTLD brains using multiple techniques (GSEA, EWCE, deconvolution), we identified robust shifts in cell-type abundance and expression across both the cortex and cerebellum. Notably, we show that microglia, astrocytes, and oligodendrocytes, cell types believed to play key roles in the neurodegenerative process of FTLD, were prominently upregulated in the frontal and temporal cortex. We found that genes associated with some but not all glial activation sets exhibited stronger expression fold-changes than standard marker genes, suggesting a possible shift towards activated glia in FTLD brains. Furthermore, in line with the known consequences of glial activation [70], we have identified enrichment in neuroinflammation and apoptosis pathways in the cortical regions. However, we are aware that these interpretations are limited by our use of bulk-tissue sections. A more thorough analysis of glial cell-type changes would involve quantifying the relative cellular proportions between intrinsic cell-states at the level of single cells. Moreover, we were unable to assess whether the observed changes in microglia and astrocyte gene expression are at least partly driven by peripheral monocytes, which are known to infiltrate through the blood brain barrier [29].

Our enrichment analyses have detected widespread upregulation of endothelial cells and angiogenesis pathways, suggesting increased blood vessel abundance and growth in FTLD brains.

It is generally not known how or if the circulatory system is involved in FTLD pathogenesis, although vascular abnormalities have been found in postmortem brains with TDP-43 pathology [27]. It is possible that these findings are a general consequence of disease, as vascular defects have been previously linked to various pathological processes such as Alzheimer's disease, cerebral amyloid angiopathy, infection, and infarction [14, 92, 97, 101]. Across the FTLD-TDP samples, we also detected strong enrichment for extracellular matrix (ECM) remodelling, a cellular process that is poorly understood in FTLD. One of the few lines of evidence supporting ECM dysregulation in FTLD comes from a transcriptomic study done on individual genetic subgroups - *C9orf72, GRN,* and *MAPT* [72]. Within these subgroups, Menden et al. detected increased expression of regulatory ECM genes, such as matrix metalloproteinases (MMPs). In other neurodegenerative diseases, MMPs have been shown to increase production of growth factors that promote blood vessel development, providing a potential causal link between the observed enrichment of endothelial cells [82]. Given this relationship, we think the ECM and circulatory system represent important subjects to study for future FTD research.

From our deconvolution analysis, we have identified robust decreases in excitatory and inhibitory neuron proportions in the cortical regions. These findings fall in line with initial studies on postmortem FTLD brains that have found deficits to glutamatergic and GABAergic systems [28]. Recently, evidence from multiple studies has accumulated in support of the selective vulnerability of excitatory neurons to TDP-43 pathology [7, 75, 87, 109]. This led us to examine the expression changes linked to Von Economo neurons (VENs), a group of selectively targeted excitatory neurons, using marker genes from the recent study by Hodge et. al [42]. However, we found that the marker genes were mainly upregulated. Given that we used bulk-tissue sections, our dataset is likely affected by the relatively low proportion of VEN neurons within each sample. In support of the vulnerability of excitatory neurons in FTLD-TDP, our correlations with neuropathological atrophy show stronger associations for excitatory neurons than inhibitory neurons. However, we note that significant associations were only observed for the frontal cortex.

We identified a decrease in inhibitory neurons in the cerebellum that agrees with recent reports of Purkinje neuron loss in FTD-ALS mouse models [17], despite lacking a cerebellum-specific single-cell RNA-seq panel. We also detected a reduction in expression of known Purkinje neuron marker genes in our enrichment analyses. These findings add to the growing body of evidence supporting the involvement of the cerebellum in FTLD. Imaging and neuropathology studies have shown that the cerebellum may be selectively targeted in both sporadic FTLD-TDP and in cases caused by *C9orf72* mutations [9, 16, 21, 96]. In *C9orf72* cases, accumulation of *C9orf72* dipeptide repeat protein, rather than TDP-43, may be responsible for cerebellar neuronal loss [32, 106], although a direct causal link has not yet been established. Interestingly, by comparing FTLD-TDP cases with and without *C9orf72* mutations, we found concordant changes in both gene expression and

estimated cellular composition in the cerebellum. Potentially, this could suggest that mechanisms driving cerebellar neuronal loss are shared between *C9orf72* and sporadic cases.

We did not observe significant correlations with atrophy scores from the temporal cortex. Nor did we observe concordance between neuronal proportions estimated from DNA methylation, nor between estimated microglia proportions and histological counts of microglia, nor an estimate of hippocampal TDP-43 burden. These failed associations are presumably partly due to sample size as well as the difficulties of accurate deconvolution of brain bulk RNA-seq samples [77]. There is the additional confounder of asymmetric changes between the hemispheres, as the deconvolution estimates and neuropathological observations were determined from opposite hemispheres of the same donor. To resolve this in future, it would be useful to fix and freeze adjacent sections of the same hemisphere.

Applying DTU analysis has allowed us to detect changes in transcript usage in FTLD-TDP affecting a largely distinct set of genes from the DGE analysis. These genes were less enriched in specific cell-types, which suggests they may reflect cell-intrinsic changes, rather than simply cellular composition changes exhibited by the DEGs. For example, the largest number of genes with DTU was observed for the cerebellum, in which we previously found the lowest number of DEGs and the weakest cell-type enrichments. Multiple studies have demonstrated that reductions in nuclear TDP-43 alter the expression and splicing of its target mRNAs, many of which regulate neuronal function [79, 85, 99]. In support of this, we observed significant overlaps between DTU genes and genes known to be altered by TDP-43 loss [13], but not TDP-43 over-expression [95]. However, as widespread DTU has been observed in multiple neurodegenerative diseases, including Alzheimer's [67] and Parkinson's [23], as well as similar enrichment with DTU genes found in the cerebellum, we are wary of linking the majority of transcript changes to TDP-43 pathology. For example, the DTU study of Alzheimer's brains also observed DTU in genes and pathways related to immune activation and synaptic transmission [67]. The DTU genes may therefore reflect a general consequence of neurodegenerative disease rather than TDP-43 pathology. Another example of this is the observation of DTU in both *UNC13A* and *UNC13B*, two synaptic genes whose splicing is altered in TDP-43 loss-of-function models [13, 84]. For *UNC13B,* the prominent DTU signals observed in the frontal and temporal cortex would suggest that this gene is affected by loss of TDP-43. However, given that DTU signals were also observed in the cerebellum, a region that lacks TDP-43 pathology, we think some other factor, in addition to TDP-43 pathology, may converge on these critical neuronal genes.

A major limitation of this study is the sample size of the FTLD-TDP group. We lacked a sufficient number of donors to analyze differences between individual genetic or TDP-43 pathology subtypes. A larger cohort would allow us to identify cell-types and pathways that are affected between subgroups of FTLD, and grant us statistical power to perform additional analyses such as gene co-

expression networks. Hopefully, in future studies we will be able to meta-analyse across FTLD cohorts.

To conclude, we highlight potential consequences of FTLD-TDP, including vascular dysfunction, RNA mis-splicing, and glial activation. We emphasize the vulnerability of excitatory neurons, and a strong involvement of the cerebellum in the FTLD neurodegenerative process. We hope our data stimulates further research that will lead toward a better understanding of the relevant disease mechanisms in FTLD.

Author contributions

JH and TR conceived and designed the project. RH led the main data analysis, under the supervision of JH. CB performed methylation analysis and neuronal estimation. JN provided pathological and clinical information on the samples. TL provided pathological, genetic, clinical information, microglial analysis scores on the samples and performed the atrophy analyses. JH and TR oversaw all aspects of the study, with contributions from TL and PF. The NYGC ALS Consortium and the Target ALS Human Postmortem Tissue Core provided human tissue samples. RH and JH wrote the manuscript with input from all co-authors.

Acknowledgements

We thank Professor Francesco Scaravilli, Professor Margaret Esiri, Professor Tamas Revesz and Professor Janice Holton for providing neuropathological reports. We thank all members of the Raj lab for their feedback on the manuscript. JH and TR are funded by grants from the US National Institutes of Health (NIH NIA R56-AG055824 and NIA U01-AG068880). PF is funded by the UK MRCl (MR/M008606/1 and MR/S006508/1), the UK Motor Neurone Disease Association, Rosetrees Trust and the UCLH NIHR Biomedical Research Centre. CB is funded by the Alzheimer's Research UK (ARUK-RF2019B-005) and the Multiple System Atrophy Trust. TL is supported by an Alzheimer's Research UK Senior fellowship. This work was supported in part through the computational resources and staff expertise provided by Scientific Computing at the Icahn School of Medicine at Mount Sinai. Research reported in this paper was supported by the Office of Research Infrastructure of the National Institutes of Health under award number S10OD018522 and S10OD026880. All NYGC ALS Consortium activities are supported by the ALS Association (ALSA, 19-SI-459) and the Tow Foundation. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

1. Afshar P, Ashtari N, Jiao X, Rahimi-Balaei M, Zhang X, Yaganeh B, Del Bigio MR, Kong J, Marzban H (2017) Overexpression of Human SOD1 Leads to Discrete Defects in the Cerebellar Architecture in the Mouse. Front Neuroanat 11:22

2. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA (2014) Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 30:1363–1369

3. Baker M, Mackenzie IR, Pickering-Brown SM, Gass J, Rademakers R, Lindholm C, Snowden J, Adamson J, Sadovnick AD, Rollinson S, Cannon A, Dwosh E, Neary D, Melquist S, Richardson A, Dickson D, Berger Z, Eriksen J, Robinson T, Zehr C, Dickey CA, Crook R, McGowan E, Mann D, Boeve B, Feldman H, Hutton M (2006) Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. Nature 442:916–919

4. Bampton A, Gatt A, Humphrey J, Cappelli S, Bhattacharya D, Foti S, Brown A-L, Asi Y, Low YH, Foiani M, Raj T, Buratti E, Fratta P, Lashley T (2021) HnRNP K mislocalisation is a novel protein pathology of frontotemporal lobar degeneration and ageing and leads to cryptic splicing. Acta Neuropathol 142:609–627

5. Bannwarth S, Ait-El-Mkadem S, Chaussenot A, Genin EC, Lacas-Gervais S, Fragaki K, Berg-Alonso L, Kageyama Y, Serre V, Moore DG, Verschueren A, Rouzier C, Le Ber I, Augé G, Cochaud C, Lespinasse F, N'Guyen K, de Septenville A, Brice A, Yu-Wai-Man P, Sesaki H, Pouget J, Paquis-Flucklinger V (2014) A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. Brain 137:2329–2345

6. Barski JJ, Hartmann J, Rose CR, Hoebeek F, Mörl K, Noll-Hussong M, De Zeeuw CI, Konnerth A, Meyer M (2003) Calbindin in cerebellar Purkinje cells is a critical determinant of the precision of motor coordination. J Neurosci 23:3469–3477

7. Benussi A, Alberici A, Buratti E, Ghidoni R, Gardoni F, Di Luca M, Padovani A, Borroni B (2019) Toward a Glutamate Hypothesis of Frontotemporal Dementia. Front Neurosci 13:304

8. Bersano A, Del Bo R, Lamperti C, Ghezzi S, Fagiolari G, Fortunato F, Ballabio E, Moggio M, Candelise L, Galimberti D, Virgilio R, Lanfranconi S, Torrente Y, Carpo M, Bresolin N, Comi GP, Corti S (2009) Inclusion body myopathy and frontotemporal dementia caused by a novel VCP mutation. Neurobiol Aging 30:752–758

9. Bocchetta M, Cardoso MJ, Cash DM, Ourselin S, Warren JD, Rohrer JD (2016) Patterns of regional cerebellar atrophy in genetic frontotemporal dementia. Neuroimage Clin 11:287–290

10. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120

11. Borroni B, Bonvicini C, Alberici A, Buratti E, Agosti C, Archetti S, Papetti A, Stuani C, Di Luca M, Gennarelli M, Padovani A (2009) Mutation within TARDBP leads to frontotemporal dementia without motor neuron disease. Hum Mutat 30:E974–83

12. Brettschneider J, Del Tredici K, Irwin DJ, Grossman M, Robinson JL, Toledo JB, Fang L, Van Deerlin VM, Ludolph AC, Lee VM-Y, Braak H, Trojanowski JQ (2014) Sequential distribution of pTDP-43 pathology in behavioral variant frontotemporal dementia (bvFTD). Acta Neuropathol 127:423–439

13. Brown A-L, Wilkins OG, Keuss MJ, Hill SE, Zanovello M, Lee WC, Lee FCY, Masino L, Qi YA, Bryce-Smith S, Bampton A, Gatt A, Phatnani H, NYGC ALS Consortium, Schiavo G, Fisher EMC, Raj T, Secrier M, Lashley T, Ule J, Buratti E, Humphrey J, Ward ME, Fratta P (2021)

Common ALS/FTD risk variants in UNC13A exacerbate its cryptic splicing and loss upon TDP-43 mislocalization. bioRxiv 2021.04.02.438170

14. Cadavid D, Mena H, Koeller K, Frommelt RA (2000) Cerebral beta amyloid angiopathy is a risk factor for cerebral ischemic infarction. A case control study in human brain biopsies. J Neuropathol Exp Neurol 59:768–773

15. Chen W-T, Lu A, Craessaerts K, Pavie B, Sala Frigerio C, Corthout N, Qian X, Laláková J, Kühnemund M, Voytyuk I, Wolfs L, Mancuso R, Salta E, Balusu S, Snellinx A, Munck S, Jurek A, Fernandez Navarro J, Saido TC, Huitinga I, Lundeberg J, Fiers M, De Strooper B (2020) Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. Cell 182:976–991.e19

16. Chen Y, Kumfor F, Landin-Romero R, Irish M, Piguet O (2019) The Cerebellum in Frontotemporal Dementia: a Meta-Analysis of Neuroimaging Studies. Neuropsychol Rev 29:450–464

17. Chew J, Gendron TF, Prudencio M, Sasaguri H, Zhang Y-J, Castanedes-Casey M, Lee CW, Jansen-West K, Kurti A, Murray ME, Bieniek KF, Bauer PO, Whitelaw EC, Rousseau L, Stankowski JN, Stetler C, Daughrity LM, Perkerson EA, Desaro P, Johnston A, Overstreet K, Edbauer D, Rademakers R, Boylan KB, Dickson DW, Fryer JD, Petrucelli L (2015) Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. Science 348:1151–1154

18. Conway JR, Lex A, Gehlenborg N (2017) UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics 33:2938–2940

19. Cruts M, Gijselinck I, van der Zee J, Engelborghs S, Wils H, Pirici D, Rademakers R, Vandenberghe R, Dermaut B, Martin J-J, van Duijn C, Peeters K, Sciot R, Santens P, De Pooter T, Mattheijssens M, Van den Broeck M, Cuijt I, Vennekens K 'l, De Deyn PP, Kumar-Singh S, Van Broeckhoven C (2006) Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21. Nature 442:920–924

20. Darmanis S, Sloan SA, Zhang Y, Enge M, Caneda C, Shuer LM, Hayden Gephart MG, Barres BA, Quake SR (2015) A survey of human brain transcriptome diversity at the single cell level. Proc Natl Acad Sci U S A 112:7285–7290

21. Davidson Y, Robinson AC, Liu X, Wu D, Troakes C, Rollinson S, Masuda-Suzukake M, Suzuki G, Nonaka T, Shi J, Tian J, Hamdalla H, Ealing J, Richardson A, Jones M, Pickering-Brown S, Snowden JS, Hasegawa M, Mann DMA (2016) Neurodegeneration in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9orf72 is linked to TDP-43 pathology and not associated with aggregated forms of dipeptide repeat proteins. Neuropathol Appl Neurobiol 42:242–254

22. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung G-YR, Karydas A, Seeley WW, Josephs KA, Coppola G, Geschwind DH, Wszolek ZK, Feldman H, Knopman DS, Petersen RC, Miller BL, Dickson DW, Boylan KB, Graff-Radford NR, Rademakers R (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 72:245–256

23. Dick F, Nido GS, Alves GW, Tysnes O-B, Nilsen GH, Dölle C, Tzoulis C (2020) Differential transcript usage in the Parkinson's disease brain. PLoS Genet 16:e1009182

24. Dickson DW, Baker MC, Jackson JL, DeJesus-Hernandez M, Finch NA, Tian S, Heckman MG, Pottier C, Gendron TF, Murray ME, Ren Y, Reddy JS, Graff-Radford NR, Boeve BF, Petersen RC, Knopman DS, Josephs KA, Petrucelli L, Oskarsson B, Sheppard JW, Asmann

YW, Rademakers R, van Blitterswijk M (2019) Extensive transcriptomic study emphasizes importance of vesicular transport in C9orf72 expansion carriers. Acta Neuropathol Commun 7:150

25. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21

26. Duits FH, Hernandez-Guillamon M, Montaner J, Goos JDC, Montañola A, Wattjes MP, Barkhof F, Scheltens P, Teunissen CE, van der Flier WM (2015) Matrix metalloproteinases in Alzheimer's disease and concurrent cerebral microbleeds. J Alzheimers Dis 48:711–720

27. Ek Olofsson H, Englund E (2019) A cortical microvascular structure in vascular dementia, Alzheimer's disease, frontotemporal lobar degeneration and nondemented controls: a sign of angiogenesis due to brain ischaemia? Neuropathol Appl Neurobiol 45:557–569

28. Ferrer I (1999) Neurons and their dendrites in frontotemporal dementia. Dement Geriatr Cogn Disord 10 Suppl 1:55–60

29. da Fonseca ACC, Matias D, Garcia C, Amaral R, Geraldo LH, Freitas C, Lima FRS (2014) The impact of microglial activation on blood-brain barrier in brain diseases. Front Cell Neurosci 8:362

30. Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, Ruderfer DM, Oh EC, Topol A, Shah HR, Klei LL, Kramer R, Pinto D, Gümüş ZH, Cicek AE, Dang KK, Browne A, Lu C, Xie L, Readhead B, Stahl EA, Xiao J, Parvizi M, Hamamsy T, Fullard JF, Wang Y-C, Mahajan MC, Derry JMJ, Dudley JT, Hemby SE, Logsdon BA, Talbot K, Raj T, Bennett DA, De Jager PL, Zhu J, Zhang B, Sullivan PF, Chess A, Purcell SM, Shinobu LA, Mangravite LM, Toyoshiba H, Gur RE, Hahn C-G, Lewis DA, Haroutunian V, Peters MA, Lipska BK, Buxbaum JD, Schadt EE, Hirai K, Roeder K, Brennand KJ, Katsanis N, Domenici E, Devlin B, Sklar P (2016) Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat Neurosci 19:1442–1453

31. Gami-Patel P, van Dijken I, van Swieten JC, Pijnenburg YAL, Netherlands Brain Bank, Rozemuller AJM, Hoozemans JJM, Dijkstra AA (2019) Von Economo neurons are part of a larger neuronal population that are selectively vulnerable in C9orf72 frontotemporal dementia. Neuropathol Appl Neurobiol 45:671–680

32. Gendron TF, van Blitterswijk M, Bieniek KF, Daughrity LM, Jiang J, Rush BK, Pedraza O, Lucas JA, Murray ME, Desaro P, Robertson A, Overstreet K, Thomas CS, Crook JE, Castanedes-Casey M, Rousseau L, Josephs KA, Parisi JE, Knopman DS, Petersen RC, Boeve BF, Graff-Radford NR, Rademakers R, Lagier-Tourenne C, Edbauer D, Cleveland DW, Dickson DW, Petrucelli L, Boylan KB (2015) Cerebellar c9RAN proteins associate with clinical and neuropathological characteristics of C9ORF72 repeat expansion carriers. Acta Neuropathol 130:559–573

33. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JYH, Zhang J (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5:R80

34. Gijselinck I, Van Mossevelde S, van der Zee J, Sieben A, Philtjens S, Heeman B, Engelborghs S, Vandenbulcke M, De Baets G, Bäumer V, Cuijt I, Van den Broeck M, Peeters K, Mattheijssens M, Rousseau F, Vandenberghe R, De Jonghe P, Cras P, De Deyn PP, Martin J-J, Cruts M, Van Broeckhoven C, BELNEU Consortium (2015) Loss of TBK1 is a frequent cause of frontotemporal dementia in a Belgian cohort. Neurology 85:2116–2125

35. Gilis J, Vitting-Seerup K, Van den Berge K, Clement L (2021) satuRn: Scalable Analysis of differential Transcript Usage for bulk and single-cell RNA-sequencing applications. bioRxiv 2021.01.14.426636

36. Glass CK, Saijo K, Winner B, Marchetto MC, Gage FH (2010) Mechanisms underlying inflammation in neurodegeneration. Cell 140:918–934

37. Goldman JS, Farmer JM, Wood EM, Johnson JK, Boxer A, Neuhaus J, Lomen-Hoerth C, Wilhelmsen KC, Lee VM-Y, Grossman M, Miller BL (2005) Comparison of family histories in FTLD subtypes and related tauopathies. Neurology 65:1817–1819

38. GTEx Consortium (2020) The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science 369:1318–1330

39. Guintivano J, Aryee MJ, Kaminsky ZA (2013) A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. Epigenetics 8:290–302

40. Habib N, McCabe C, Medina S, Varshavsky M, Kitsberg D, Dvir-Szternfeld R, Green G, Dionne D, Nguyen L, Marshall JL, Chen F, Zhang F, Kaplan T, Regev A, Schwartz M (2020) Disease-associated astrocytes in Alzheimer's disease and aging. Nat Neurosci 23:701–706

41. Hallmann A-L, Araúzo-Bravo MJ, Mavrommatis L, Ehrlich M, Röpke A, Brockhaus J, Missler M, Sterneckert J, Schöler HR, Kuhlmann T, Zaehres H, Hargus G (2017) Astrocyte pathology in a human neural stem cell model of frontotemporal dementia caused by mutant TAU protein. Sci Rep 7:42991

42. Hodge RD, Miller JA, Novotny M, Kalmbach BE, Ting JT, Bakken TE, Aevermann BD, Barkan ER, Berkowitz-Cerasano ML, Cobbs C, Diez-Fuertes F, Ding S-L, McCorrison J, Schork NJ, Shehata SI, Smith KA, Sunkin SM, Tran DN, Venepally P, Yanny AM, Steemers FJ, Phillips JW, Bernard A, Koch C, Lasken RS, Scheuermann RH, Lein ES (2020) Transcriptomic evidence that von Economo neurons are regionally specialized extratelencephalic-projecting excitatory neurons. Nat Commun 11:1172

43. Hoffman GE, Schadt EE (2016) variancePartition: interpreting drivers of variation in complex gene expression studies. BMC Bioinformatics 17:483

44. Holtman IR, Noback M, Bijlsma M, Duong KN, van der Geest MA, Ketelaars PT, Brouwer N, Vainchtein ID, Eggen BJL, Boddeke HWG (2015) Glia Open Access Database (GOAD): A comprehensive gene expression encyclopedia of glia cells in health and disease. Glia 63:1495–1506

45. Honda D, Ishigaki S, Iguchi Y, Fujioka Y, Udagawa T, Masuda A, Ohno K, Katsuno M, Sobue G (2013) 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

46. Hughes LE, Rittman T, Robbins TW, Rowe JB (2018) Reorganization of cortical oscillatory dynamics underlying disinhibition in frontotemporal dementia. Brain 141:2486–2499

47. Hunt GJ, Freytag S, Bahlo M, Gagnon-Bartsch JA (2019) dtangle: accurate and robust cell type deconvolution. Bioinformatics 35:2093–2099

48. Hurskainen TL, Hirohata S, Seldin MF, Apte SS (1999) ADAM-TS5, ADAM-TS6, and ADAM-TS7, Novel Members of a New Family of Zinc Metalloproteases: GENERAL FEATURES AND GENOMIC DISTRIBUTION OF THE ADAM-TS FAMILY * 210. J Biol Chem 274:25555–25563

49. Jackson JL, Finch NA, Baker MC, Kachergus JM, DeJesus-Hernandez M, Pereira K, Christopher E, Prudencio M, Heckman MG, Thompson EA, Dickson DW, Shah J, Oskarsson

B, Petrucelli L, Rademakers R, van Blitterswijk M (2020) Elevated methylation levels, reduced expression levels, and frequent contractions in a clinical cohort of C9orf72 expansion carriers. Mol Neurodegener 15:7

50. Jung Y-J, Chung W-S (2018) Phagocytic roles of glial cells in healthy and diseased brains. Biomol Ther 26:350–357

51. Kassambara A (2018) ggpubr:“ggplot2” based publication ready plots. R package version 0 1 7

52. Kattuah W, Rogelj B, King A, Shaw CE, Hortobágyi T, Troakes C (2019) Heterogeneous Nuclear Ribonucleoprotein E2 (hnRNP E2) Is a Component of TDP-43 Aggregates Specifically in the A and C Pathological Subtypes of Frontotemporal Lobar Degeneration. Front Neurosci 13:551

53. Keren-Shaul H, Spinrad A, Weiner A, Matcovitch-Natan O, Dvir-Szternfeld R, Ulland TK, David E, Baruch K, Lara-Astaiso D, Toth B, Itzkovitz S, Colonna M, Schwartz M, Amit I (2017) A Unique Microglia Type Associated with Restricting Development of Alzheimer’s Disease. Cell 169:1276–1290.e17

54. Klim JR, Williams LA, Limone F, Guerra San Juan I, Davis-Dusenbery BN, Mordes DA, Burberry A, Steinbaugh MJ, Gamage KK, Kirchner R, Moccia R, Cassel SH, Chen K, Wainger BJ, Woolf CJ, Eggan K (2019) ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. Nat Neurosci 22:167–179

55. Koyama A, Sugai A, Kato T, Ishihara T, Shiga A, Toyoshima Y, Koyama M, Konno T, Hirokawa S, Yokoseki A, Nishizawa M, Kakita A, Takahashi H, Onodera O (2016) Increased cytoplasmic *TARDBP* mRNA in affected spinal motor neurons in ALS caused by abnormal autoregulation of TDP-43. Nucleic Acids Res gkw499

56. Kuźma-Kozakiewicz M, Chudy A, Kaźmierczak B, Dziewulska D, Usarek E, Barańczyk-Kuźma A (2013) Dynactin Deficiency in the CNS of Humans with Sporadic ALS and Mice with Genetically Determined Motor Neuron Degeneration. Neurochem Res. doi: 10.1007/s11064-013-1160-7

57. Law CW, Chen Y, Shi W, Smyth GK (2014) voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 15:R29

58. Le Ber I, Camuzat A, Guerreiro R, Bouya-Ahmed K, Bras J, Nicolas G, Gabelle A, Didic M, De Septenville A, Millecamps S, Lenglet T, Latouche M, Kabashi E, Campion D, Hannequin D, Hardy J, Brice A, French Clinical and Genetic Research Network on FTD/FTD-ALS (2013) SQSTM1 mutations in French patients with frontotemporal dementia or frontotemporal dementia with amyotrophic lateral sclerosis. JAMA Neurol 70:1403–1410

59. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323

60. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 1:417–425

61. Liddelow SA, Guttenplan KA, Clarke LE, Bennett FC, Bohlen CJ, Schirmer L, Bennett ML, Münch AE, Chung W-S, Peterson TC, Wilton DK, Frouin A, Napier BA, Panicker N, Kumar M, Buckwalter MS, Rowitch DH, Dawson VL, Dawson TM, Stevens B, Barres BA (2017) Neurotoxic reactive astrocytes are induced by activated microglia. Nature 541:481–487

62. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map

format and SAMtools. Bioinformatics 25:2078–2079

63. Ling S-C, Polymenidou M, Cleveland DW (2013) Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. Neuron 79:416–438

64. Lu P, Takai K, Weaver VM, Werb Z (2011) Extracellular matrix degradation and remodeling in development and disease. Cold Spring Harb Perspect Biol 3. doi: 10.1101/cshperspect.a005058

65. Mancarci BO, Toker L, Tripathy SJ, Li B, Rocco B, Sibille E, Pavlidis P (2017) Cross-Laboratory Analysis of Brain Cell Type Transcriptomes with Applications to Interpretation of Bulk Tissue Data. eNeuro 4. doi: 10.1523/ENEURO.0212-17.2017

66. Mancarci O (2019) Quick Access to Homologene and Gene Annotation Updates [R package homologene version 1.4.68.19.3.27]

67. Marques-Coelho D, da Cruz Carvalho LI, de Farias ARM, Lambert J-C, Costa MR, Bank NB Differential transcript usage unravels gene expression alterations in Alzheimer's disease human brains

68. Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, Menon M, He L, Abdurrob F, Jiang X, Martorell AJ, Ransohoff RM, Hafler BP, Bennett DA, Kellis M, Tsai L-H (2019) Single-cell transcriptomic analysis of Alzheimer's disease. Nature 570:332–337

69. Mattson MP, Meffert MK (2006) Roles for NF-κB in nerve cell survival, plasticity, and disease. Cell Death Differ 13:852–860

70. McCauley ME, Baloh RH (2019) Inflammation in ALS/FTD pathogenesis. Acta Neuropathol 137:715–730

71. Melamed Z, López-Erauskin J, Baughn MW, Zhang O, Drenner K, Sun Y, Freyermuth F, McMahon MA, Beccari MS, Artates JW, Ohkubo T, Rodriguez M, Lin N, Wu D, Frank Bennett C, Rigo F, Da Cruz S, Ravits J, Lagier-Tourenne C, Cleveland DW (2019) Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. Nat Neurosci 22:180–190

72. Menden K, Francescatto M, Niyma T, Blauwendraat C (2021) Integrated multi-omics analysis reveals common and distinct dysregulated pathways for genetic subtypes of Frontotemporal Dementia. bioRxiv

73. Mizutani A, Kuroda Y, Futatsugi A, Furuichi T, Mikoshiba K (2008) Phosphorylation of Homer3 by calcium/calmodulin-dependent kinase II regulates a coupling state of its target molecules in Purkinje cells. J Neurosci 28:5369–5382

74. Murley AG, Rouse MA, Jones PS, Ye R, Hezemans FH, O'Callaghan C, Frangou P, Kourtzi Z, Rua C, Carpenter TA, Rodgers CT, Rowe JB (2020) GABA and glutamate deficits from frontotemporal lobar degeneration are associated with disinhibition. Brain 143:3449–3462

75. Nana AL, Sidhu M, Gaus SE, Hwang J-HL, Li L, Park Y, Kim E-J, Pasquini L, Allen IE, Rankin KP, Toller G, Kramer JH, Geschwind DH, Coppola G, Huang EJ, Grinberg LT, Miller BL, Seeley WW (2019) Neurons selectively targeted in frontotemporal dementia reveal early stage TDP-43 pathobiology. Acta Neuropathol 137:27–46

76. Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, McCluskey LF, Miller BL, Masliah E, Mackenzie IR, Feldman H, Feiden W, Kretzschmar HA, Trojanowski JQ, Lee VM-Y (2006) Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. Science 314:130–133

77. Patrick E, Taga M, Ergun A, Ng B, Casazza W, Cimpean M, Yung C, Schneider JA, Bennett DA, Gaiteri C, De Jager PL, Bradshaw EM, Mostafavi S (2020) Deconvolving the contributions of cell-type heterogeneity on cortical gene expression. PLoS Comput Biol 16:e1008120

78. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC (2013) A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics 14:293

79. Polymenidou M, Lagier-Tourenne C, Hutt KR, Huelga SC, Moran J, Liang TY, Ling S-C, Sun E, Wancewicz E, Mazur C, Kordasiewicz H, Sedaghat Y, Donohue JP, Shiue L, Frank Bennett C, Yeo GW, Cleveland DW (2011) Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. Nature Neuroscience 14:459–468

80. Prudencio M, Belzil VV, Batra R, Ross CA, Gendron TF, Pregent LJ, Murray ME, Overstreet KK, Piazza-Johnston AE, Desaro P, Bieniek KF, DeTure M, Lee WC, Biendarra SM, Davis MD, Baker MC, Perkerson RB, van Blitterswijk M, Stetler CT, Rademakers R, Link CD, Dickson DW, Boylan KB, Li H, Petrucelli L (2015) Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. Nat Neurosci 18:1175–1182

81. Ramesh G, MacLean AG, Philipp MT (2013) Cytokines and chemokines at the crossroads of neuroinflammation, neurodegeneration, and neuropathic pain. Mediators Inflamm 2013:480739

82. Rempe RG, Hartz AMS, Bauer B (2016) Matrix metalloproteinases in the brain and blood-brain barrier: Versatile breakers and makers. J Cereb Blood Flow Metab 36:1481–1507

83. Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes AM, Kaganovich A, Scholz SW, Duckworth J, Ding J, Harmer DW, Hernandez DG, Johnson JO, Mok K, Ryten M, Trabzuni D, Guerreiro RJ, Orrell RW, Neal J, Murray A, Pearson J, Jansen IE, Sondervan D, Seelaar H, Blake D, Young K, Halliwell N, Callister JB, Toulson G, Richardson A, Gerhard A, Snowden J, Mann D, Neary D, Nalls MA, Peuralinna T, Jansson L, Isoviita V-M, Kaivorinne A-L, Hölttä-Vuori M, Ikonen E, Sulkava R, Benatar M, Wuu J, Chiò A, Restagno G, Borghero G, Sabatelli M, ITALSGEN Consortium, Heckerman D, Rogaeva E, Zinman L, Rothstein JD, Sendtner M, Drepper C, Eichler EE, Alkan C, Abdullaev Z, Pack SD, Dutra A, Pak E, Hardy J, Singleton A, Williams NM, Heutink P, Pickering-Brown S, Morris HR, Tienari PJ, Traynor BJ (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 72:257–268

84. Rosa Ma X, Prudencio M, Koike Y, Vatsavayai SC, Kim G, Harbinski F, Rodriguez CM, Broder Schmidt H, Cummings BB, Wyatt DW, Kurylo K, Miller G, Mekhoubad S, Sallee N, Jansen-West K, Cook CN, Pickles S, Oskarsson B, Graff-Radford NR, Boeve BF, Knopman DS, Petersen RC, Dickson DW, Green EM, Seeley WW, Petrucelli L, Gitler AD (2021) TDP-43 represses cryptic exon inclusion in FTD/ALS gene UNC13A. bioRxiv 2021.04.02.438213

85. Rot G, Wang Z, Huppertz I, Modic M, Lenče T, Hallegger M, Haberman N, Curk T, von Mering C, Ule J (2017) High-Resolution RNA Maps Suggest Common Principles of Splicing and Polyadenylation Regulation by TDP-43. Cell Rep 19:1056–1067

86. Ruegsegger C, Stucki DM, Steiner S, Angliker N, Radecke J, Keller E, Zuber B, Rüegg MA, Saxena S (2016) Impaired mTORC1-Dependent Expression of Homer-3 Influences SCA1 Pathophysiology. Neuron 89:129–146

87. Santillo AF, Nilsson C, Englund E (2013) von Economo neurones are selectively targeted in frontotemporal dementia. Neuropathol Appl Neurobiol 39:572–579

88. Šarac H, Žagar M, Vranješ D, Henigsberg N, Bilić E, Pavliša G (2008) Magnetic resonance imaging and magnetic resonance spectroscopy in a patient with amyotrophic lateral sclerosis

and frontotemporal dementia. Coll Antropol 32:205–210

89. Schofield E, Kersaitis C, Shepherd CE, Kril JJ, Halliday GM (2003) Severity of gliosis in Pick's disease and frontotemporal lobar degeneration: tau-positive glia differentiate these disorders. Brain 126:827–840

90. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol 7:3

91. Skene NG, Grant SGN (2016) Identification of Vulnerable Cell Types in Major Brain Disorders Using Single Cell Transcriptomes and Expression Weighted Cell Type Enrichment. Front Neurosci 10:16

92. de Sousa JR, Azevedo RSS, Martins Filho AJ, Araujo MTF, Moutinho ERC, Baldez Vasconcelos BC, Cruz ACR, Oliveira CS, Martins LC, Baldez Vasconcelos BH, Casseb LMN, Chiang JO, Quaresma JAS, Vasconcelos PFC (2018) Correlation between Apoptosis and in Situ Immune Response in Fatal Cases of Microcephaly Caused by Zika Virus. Am J Pathol 188:2644–2652

93. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102:15545–15550

94. Tam OH, Rozhkov NV, Shaw R, Kim D, Hubbard I, Fennessey S, Propp N, NYGC ALS Consortium, Fagegaltier D, Harris BT, Ostrow LW, Phatnani H, Ravits J, Dubnau J, Gale Hammell M (2019) Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia. Cell Rep 29:1164–1177.e5

95. Tank EM, Figueroa-Romero C, Hinder LM, Bedi K, Archbold HC, Li X, Weskamp K, Safren N, Paez-Colasante X, Pacut C, Thumma S, Paulsen MT, Guo K, Hur J, Ljungman M, Feldman EL, Barmada SJ (2018) Abnormal RNA stability in amyotrophic lateral sclerosis. Nat Commun 9:2845

96. Tan RH, Devenney E, Dobson-Stone C, Kwok JB, Hodges JR, Kiernan MC, Halliday GM, Hornberger M (2014) Cerebellar integrity in the amyotrophic lateral sclerosis-frontotemporal dementia continuum. PLoS One 9:e105632

97. Terada T (2010) Cryptococcosis in the central nervous system in a 36-year-old Japanese man: an autopsy study. Tohoku J Exp Med 222:33–37

98. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, Teschendorff AE (2017) ChAMP: updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics 33:3982–3984

99. Tollervey JR, Curk T, Rogelj B, Briese M, Cereda M, Kayikci M, Konig J, Hortobagyi T, Nishimura AL, Zupunski V, Patani R, Chandran S, Rot G, Zupan B, Shaw CE, Ule J (2011) Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nat Neurosci 14:452–458

100. Vogt MA, Ehsaei Z, Knuckles P, Higginbottom A, Helmbrecht MS, Kunath T, Eggan K, Williams LA, Shaw PJ, Wurst W, Floss T, Huber AB, Taylor V (2018) TDP-43 induces p53-mediated cell death of cortical progenitors and immature neurons. Sci Rep 8:8097

101. Wang J, Ohno-Matsui K, Nakahama K-I, Okamoto A, Yoshida T, Shimada N, Mochizuki M, Morita I (2011) Amyloid beta enhances migration of endothelial progenitor cells by upregulating CX3CR1 in response to fractalkine, which may be associated with development

of choroidal neovascularization. Arterioscler Thromb Vasc Biol 31:e11–8

102. Wang X, Park J, Susztak K, Zhang NR, Li M (2019) Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. Nat Commun 10:380

103. Wickham H (2017) Tidyverse: Easily install and load'tidyverse'packages. R package version 1

104. Woollacott IOC, Toomey CE, Strand C, Courtney R, Benson BC, Rohrer JD, Lashley T (2020) Microglial burden, activation and dystrophy patterns in frontotemporal lobar degeneration. J Neuroinflammation 17:234

105. Yin T, Cook D, Lawrence M (2012) ggbio: an R package for extending the grammar of graphics for genomic data. Genome Biol 13:R77

106. Yousef A, Robinson JL, Irwin DJ, Byrne MD, Kwong LK, Lee EB, Xu Y, Xie SX, Rennert L, Suh E, Van Deerlin VM, Grossman M, Lee VM-Y, Trojanowski JQ (2017) Neuron loss and degeneration in the progression of TDP-43 in frontotemporal lobar degeneration. Acta Neuropathol Commun 5:68

107. Yu G, Wang L-G, Han Y, He Q-Y (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16:284–287

108. Zamanian JL, Xu L, Foo LC, Nouri N, Zhou L, Giffard RG, Barres BA (2012) Genomic analysis of reactive astrogliosis. J Neurosci 32:6391–6410

109. Zhang J, Velmeshev D, Hashimoto K, Huang Y-H, Hofmann JW, Shi X, Chen J, Leidal AM, Dishart JG, Cahill MK, Kelley KW, Liddelow SA, Seeley WW, Miller BL, Walther TC, Farese RV Jr, Taylor JP, Ullian EM, Huang B, Debnath J, Wittmann T, Kriegstein AR, Huang EJ (2020) Neurotoxic microglia promote TDP-43 proteinopathy in progranulin deficiency. Nature 588:459–465