

# UCSF

## UC San Francisco Previously Published Works

### Title

Frontotemporal lobar degeneration targets brain regions linked to expression of recently evolved genes

### Permalink

<https://escholarship.org/uc/item/1hg3z83q>

### Journal

Brain, 147(9)

### ISSN

0006-8950

### Authors

Pasquini, Lorenzo  
Pereira, Felipe L  
Seddighi, Sahba  
[et al.](#)

### Publication Date

2024-09-03

### DOI

10.1093/brain/awae205





### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



# Frontotemporal lobar degeneration targets brain regions linked to expression of recently evolved genes

Lorenzo Pasquini,<sup>1,2</sup> Felipe L. Pereira,<sup>1</sup> Sahba Seddighi,<sup>3</sup> Yi Zeng,<sup>4</sup>  Yongbin Wei,<sup>5</sup>  Ignacio Illán-Gala,<sup>1,6,7</sup> Sarat C. Vatsavayai,<sup>1</sup> Adit Friedberg,<sup>1,6</sup> Alex J. Lee,<sup>1</sup> Jesse A. Brown,<sup>1</sup>  Salvatore Spina,<sup>1</sup>  Lea T. Grinberg,<sup>1,8</sup> Daniel W. Sirkis,<sup>1</sup> Luke W. Bonham,<sup>1,9</sup> Jennifer S. Yokoyama,<sup>1,9</sup> Adam L. Boxer,<sup>1</sup> Joel H. Kramer,<sup>1</sup> Howard J. Rosen,<sup>1</sup> Jack Humphrey,<sup>10</sup> Aaron D. Gitler,<sup>4</sup> Bruce L. Miller,<sup>1</sup> Katherine S. Pollard,<sup>11,12,13,14</sup> Michael E. Ward<sup>3</sup> and William W. Seeley<sup>1,8</sup>

See Shen et al. (<https://doi.org/10.1093/brain/awae267>) for a scientific commentary on this article.

In frontotemporal lobar degeneration (FTLD), pathological protein aggregation in specific brain regions is associated with declines in human-specialized social-emotional and language functions. In most patients, disease protein aggregates contain either TDP-43 (FTLD-TDP) or tau (FTLD-tau). Here, we explored whether FTLD-associated regional degeneration patterns relate to regional gene expression of human accelerated regions (HARs), conserved sequences that have undergone positive selection during recent human evolution.

To this end, we used structural neuroimaging from patients with FTLD and human brain regional transcriptomic data from controls to identify genes expressed in FTLD-targeted brain regions. We then integrated primate comparative genomic data to test our hypothesis that FTLD targets brain regions linked to expression levels of recently evolved genes. In addition, we asked whether genes whose expression correlates with FTLD atrophy are enriched for genes that undergo cryptic splicing when TDP-43 function is impaired.

We found that FTLD-TDP and FTLD-tau subtypes target brain regions with overlapping and distinct gene expression correlates, highlighting many genes linked to neuromodulatory functions. FTLD atrophy-correlated genes were strongly enriched for HARs. Atrophy-correlated genes in FTLD-TDP showed greater overlap with TDP-43 cryptic splicing genes and genes with more numerous TDP-43 binding sites compared with atrophy-correlated genes in FTLD-tau. Cryptic splicing genes were enriched for HAR genes, and vice versa, but this effect was due to the confounding influence of gene length. Analyses performed at the individual-patient level revealed that the expression of HAR genes and cryptically spliced genes within putative regions of disease onset differed across FTLD-TDP subtypes.

Overall, our findings suggest that FTLD targets brain regions that have undergone recent evolutionary specialization and provide intriguing potential leads regarding the transcriptomic basis for selective vulnerability in distinct FTLD molecular-anatomical subtypes.

1 Department of Neurology, Memory and Aging Center, University of California, San Francisco, CA 94158, USA

2 Department of Neurology, Neuroscape, University of California, San Francisco, CA 94158, USA

3 National Institute of Neurological Disorders and Stroke, Neurogenetics Branch, Bethesda, MD 20892, USA

4 Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA

5 School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876, China

6 Global Brain Health Institute, University of California, San Francisco, San Francisco, CA, 94158 USA and Trinity College Dublin, Dublin D02 X9W9, Ireland

Received October 27, 2023. Revised May 08, 2024. Accepted May 20, 2024. Advance access publication June 28, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of the Guarantors of Brain.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

- 7 Department of Neurology, Hospital de la Santa Creu i Sant Pau, Biomedical Research Institute, Universitat Autònoma de Barcelona, Barcelona, Catalunya, 08041, Spain
- 8 Department of Pathology, University of California, San Francisco, CA 94158, USA
- 9 Department of Radiology, University of California, San Francisco, CA 94158, USA
- 10 Nash Family Department of Neuroscience and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
- 11 Gladstone Institute of Data Science and Biotechnology, San Francisco, CA 94158, USA
- 12 Institute for Human Genetics, University of California San Francisco, San Francisco, CA 94158, USA
- 13 Department of Epidemiology and Biostatistics and Bakar Institute for Computational Health Sciences, University of California San Francisco, San Francisco, CA 94158, USA
- 14 Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

Correspondence to: Lorenzo Pasquini, PhD  
Department of Neurology, Memory and Aging Center  
University of California, San Francisco, 675 Nelson Rising Ln  
San Francisco, CA, 94158, USA  
E-mail: Lorenzo.pasquini@ucsf.edu

Correspondence may also be addressed to: William W. Seeley, MD  
E-mail: bill.seeley@ucsf.edu

**Keywords:** frontotemporal lobar degeneration; cryptic exon; human accelerated regions; TDP-43; tau; gene expression

## Introduction

Frontotemporal dementia (FTD) refers to a group of clinical syndromes linked to underlying frontotemporal lobar degeneration (FTLD) pathology. These FTD syndromes present with deficits in recently evolved social-emotional and language functions,<sup>1,2</sup> resulting from degeneration within networks anchored by frontal, insular and anterior temporal brain regions.<sup>3,4</sup> The most common FTD syndromes are the behavioural variant (bvFTD) characterized by social-emotional deficits<sup>2</sup>; semantic variant primary progressive aphasia (svPPA), which presents with loss of knowledge related to words, objects and emotions<sup>1,5</sup>; non-fluent variant PPA (nfvPPA), characterized by speech production difficulties<sup>1,5</sup>; corticobasal syndrome, an asymmetric akinetic-rigid syndrome associated with a progressive loss of limb controls<sup>4</sup>; and progressive supranuclear palsy-Richardson syndrome (PSP-RS), associated with oculomotor deficits and gait instability.<sup>4</sup> FTLD pathology, in turn, can be divided into three major molecular classes based on whether the neuronal and glial inclusions are composed of tau (FTLD-tau), TDP-43 (FTLD-TDP) or FUS/EWS/TAF15 (FET) family proteins.<sup>6</sup> Each major molecular class can be further divided into distinct subtypes based on the pathomorphology and distribution of the inclusions across regions, layers and cell types. Factors driving the selective predilection of each FTLD subtype for distinct, yet overlapping, anatomical structures remain unknown.

Comparative genomic studies have the potential to shed light on the relationship between brain evolution and the erosion of human-specialized functions in FTLD. Novel genome-wide studies have identified a unique set of loci that are conserved across mammalian evolution, indicating important biological roles, but significantly diverged in the human lineage, suggesting changes to those roles in humans compared to chimpanzees and other primates.<sup>7,8</sup> These short sequences, referred to as human accelerated regions (HARs), are commonly located in non-coding DNA, often near genes associated with transcription and DNA binding.<sup>9</sup> Several studies have linked HARs to the emergence of distinctive human traits,

such as the opposable thumb,<sup>8</sup> language and social behaviour,<sup>10</sup> and mutations in HAR genes are observed in neuropsychiatric<sup>11,12</sup> and neurodevelopmental disorders.<sup>7</sup>

Because FTLD is pathologically heterogeneous, it is possible that FTLD-TDP and FTLD-tau intersect with brain evolution through distinct, yet overlapping mechanisms. The role of protein aggregation has been a major focus of research, but in FTLD-TDP, TDP-43 aggregation is almost always accompanied, and may be preceded, by loss of nuclear TDP-43 expression.<sup>13–16</sup> TDP-43, encoded by *TARDBP*, is an RNA-binding protein predominantly expressed in the nucleus of healthy neurons,<sup>17–19</sup> where it plays diverse roles in transcription regulation by binding to RNAs rich in GU repeats. One role is to suppress incorporation of cryptic exons, short stretches of non-conserved intronic RNA that are normally removed during pre-mRNA splicing.<sup>20–23</sup> Upon TDP-43 loss-of-function, cryptic exons are spliced into messenger RNAs, often introducing premature stop codons or frame shifts that result in nonsense mediated decay and loss of normal protein expression.<sup>24–26</sup> To date, >1000 potential genes subject to cryptic splicing have been identified, yet, despite some compelling candidates,<sup>17,24</sup> it remains unknown which cryptic splicing events, if any, are most relevant to FTLD-TDP pathogenesis. Crucially, although TDP-43 is a highly conserved protein, cryptic splicing genes show almost no overlap when comparing mice to humans,<sup>20</sup> suggesting that TDP-43 may have adopted a unique regulatory role across the human evolutionary lineage.

Here, we combined neuroimaging, neuropathological, average regional brain transcriptomic, comparative genomic and cell biological data to identify genes whose expression pattern, in the healthy brain, correlates with cortical atrophy patterns in FTLD and to test the hypothesis that HAR genes are expressed in patterns that correlate with FTLD atrophy. We further explored associations between expression patterns for TDP-43 cryptic splicing genes and FTLD-associated atrophy and the overlap between HAR and cryptic splicing genes.

## Materials and methods

### Participants

We searched the University of California, San Francisco (UCSF) Neurodegenerative Disease Brain Bank database for participants autopsied between 2008 and 2020 who had at least one high quality structural MRI scan prior to death; this search yielded 322 participants. From this pool, we included patients with a primary neuropathological diagnosis of FTLD-TDP (types A–C) or FTLD-tau [corticobasal degeneration (CBD) or Pick's disease]. Although PSP is a common subtype of FTLD-tau, we omitted it here because it is characterized by a predominantly subcortical atrophy pattern not well suited to our transcriptomic data (see later). Patients were excluded if they had: (i) Braak neurofibrillary tangle stage<sup>27</sup> >3; (ii) Alzheimer's disease neuropathological change<sup>28</sup> >intermediate; (iii) Lewy body disease<sup>29</sup> >brainstem predominant; or (iv) major territorial ischaemic infarcts or intracranial haemorrhages.<sup>30</sup> Based on these stringent criteria, 28 patients with TDP-A, 35 with TDP-B, 29 with TDP-C, 45 with CBD and 27 with Pick's disease were included in the study. In accordance with the Declaration of Helsinki, patients or their surrogates provided written informed consent prior to participation, including consent for brain donation. The UCSF Committee on Human Research approved the study.

All patients received clinical diagnoses by a multidisciplinary team following thorough neurological, neuroimaging and neuropsychological assessments. Clinical severity was assessed using the Mini-Mental State Examination<sup>31</sup> and the Clinical Dementia Rating scale total and sum of boxes scores, using a version of the Clinical Dementia Rating scale adapted for FTD.<sup>32</sup> Patient demographics, as well as clinical and neuropathological data, are provided in Table 1. Race and ethnicity, handedness and years of education were self-reported by patients or surrogates.

### Neuropathology

Post-mortem research-oriented autopsies were performed at the UCSF Neurodegenerative Disease Brain Bank. Neuropathological diagnoses were made following consensus diagnostic criteria<sup>6,29,33</sup> based on standard histological and immunohistochemical methods.<sup>34,35</sup>

### Structural neuroimaging

#### Acquisition and preprocessing

Because patients were evaluated over an 18-year interval, MRI scans were obtained on different scanners over time. For 71 patients, T<sub>1</sub>-weighted magnetization prepared rapid gradient echo (MPRAGE) MRI sequences were acquired at the UCSF Neuroscience Imaging Center, either on a 3 T Siemens Tim Trio (*n* = 67) or a 3 T Siemens Prisma Fit scanner (*n* = 4). Both scanners had similar acquisition parameters (sagittal slice orientation; slice thickness = 1.0 mm; slices per slab = 160; in-plane resolution = 1.0 × 1.0 mm; matrix = 240 × 256; repetition time = 2300 ms; inversion time = 900 ms; flip angle = 9°), although echo time differed slightly (Trio: 2.98 ms; Prisma: 2.9 ms). For the remaining 93 patients, MRI scans were obtained at the San Francisco Veterans Affairs Medical Center using MPRAGE sequences acquired either on a Siemens 1.5 T Magnetom scanner (*n* = 73; voxel resolution 1.0 × 1.0 × 1.5 mm; repetition time = 10 ms; echo time = 4 ms; inversion time = 300 ms; flip angle = 15°) or on a Siemens 4 T Magnetom scanner (*n* = 20; voxel resolution 1.0 × 1.0 × 1.0 mm; repetition time = 2.3 ms; echo time = 3 ms; inversion time = 950 ms; flip angle = 7°). The average interval between scanning date and autopsy was 4.8 ± 2.6 years.

All structural magnetic resonance images underwent a voxel-based morphometry analysis<sup>36</sup> after being visually inspected for motion and scanning artefacts. Structural images were segmented in grey matter, white matter and CSF, and normalized to MNI space using SPM12 (<http://www.fil.ion.ucl.ac.uk/spm/software/spm12/>). Grey matter images were modulated by dividing the tissue probability values by the Jacobian of the warp field. The resulting images of voxel resolution 2.0 × 2.0 × 2.0 mm were smoothed with an isotropic Gaussian kernel with a full-width at half-maximum of 8 mm.

### W-score maps

To generate participant-specific atrophy maps, the smoothed grey matter images were transformed to W-score maps. W-score maps are voxel-wise statistical maps that reflect levels of atrophy for each individual after adjustment for relevant covariates.<sup>37,38</sup> This approach is particularly useful when performing analyses with heterogeneous samples as grey matter atrophy estimates are adjusted for demographical and methodological nuisance covariates, including age, gender, scanner type and total intracranial volume.<sup>38</sup> The W-score model used for this study was published<sup>39</sup>; it involved 397 healthy older controls (Supplementary Table 1) and included age at MRI, sex, years of education, handedness, scanner type and total intracranial volume as covariates. Briefly, multiple regression analyses at the voxel-wise level were performed on the normative dataset to estimate grey matter intensity as adjusted for age, sex, years of education, handedness, MRI scanner and total intracranial volume. This model was then used to estimate the expected grey matter intensity for a patient by using their demographical data. The discrepancy between the raw and the expected maps is used to derive the W-score map:

$$W = \frac{(Raw - Expected)}{SD_{\epsilon}} \quad (1)$$

where *Raw* is the raw value of a voxel from the smoothed image of a patient; *Expected* is the expected value for the voxel of that specific patient based on the healthy control model; and *SD<sub>ε</sub>* is the standard deviation of the residuals from the healthy control model. Additional analyses were performed to assess whether the W-score model effectively mitigated the impact of scanner type in the generation of W-score maps (Supplementary Fig. 1, Supplementary Table 2 and Supplementary material, 'Findings' file).

### Gene expression

#### Allen human brain atlas

Average human brain gene expression data were derived from microarray data available through the Allen Human Brain Atlas (AHBA) (<http://human.brain-map.org/static/download>). As described in detail elsewhere,<sup>40</sup> tissue samples were extracted across both hemispheres from two human brain donors, as well as the left hemisphere of four additional donors, totalling 3702 tissue samples. Microarray analysis quantified gene expression across 58 692 probes, providing an estimate of the relative expression of 20 734 genes within the tissue samples.<sup>41</sup> The publicly available toolbox *abagen* (<https://github.com/rmarkello/abagen>)<sup>42</sup> was then used for: (i) updating probe-to-gene annotations using the latest available data; (ii) data filtering, where expression values that do not exceed background are removed; (iii) probe selection, which, for genes indexed by multiple probes, involves selecting a single representative measure to represent the expression of that gene across all

Table 1 Pathological, clinical and demographic characteristics of patients with FTL D

Pathological diagnoses	TDP-A	TDP-B	TDP-C	CBD	Pick's disease
n	28	35	29	45	27
Age at MRI scan in years	63.1 (6.7)	62.1 (9.1)	64.3 (7.4)	64.5 (6.7)	64.4 (7.1)
Scan-death interval in years	4.7 (3.2)	2.6 (2.2)	7.4 (2.8)	3.2 (1.6)	5.1 (3.4)
Sex (female:male)	16:12	15:20	11:18	25:20	11:16
Handedness (ambidextrous:left:right)	0:2:26	0:2:33	0:3:26	0:0:45	0:4:23
Education in years	16.1 (3.4)	16.1 (3.0)	16.6 (3.0)	16.2 (2.9)	16.0 (2.7)
Race (Asian:Black:White:n.a.)	1:0:24:3	0:0:34:1	0:0:28:1	4:1:36:4	2:0:22:3
Ethnicity (Hispanic/Latino:n.a.)	0:13	2:12	1:3	1:16	2:6
Scanner type (1.5 T:Trio:Prisma:4 T)	13:12:0:3	9:16:3:7	21:5:0:3	16:24:1:4	14:10:0:3
Total intracranial volume (l)	1.4 (0.1)	1.4 (0.1)	1.5 (0.2)	1.5 (0.1)	1.5 (0.2)
Clinical Dementia Rating (CDR)	1.1 (0.8)	1.4 (0.8)	0.7 (0.6)	0.8 (0.6)	0.8 (0.6)
CDR Sum of Boxes	6.1 (4.3)	7.4 (4.2)	3.8 (3.0)	4.1 (3.6)	5.3 (3.4)
MMSE	22.4 (7.2)	24.2 (5.4)	23.3 (7.2)	24.3 (5.8)	22.8 (8.0)
<b>Clinical diagnoses</b>					
Alzheimer's disease-type dementia	0	1	0	2	1
Behavioural variant FTD	16	30	2	11	16
Corticobasal syndrome	3	0	0	15	2
Mild cognitive impairment	0	1	1	0	0
Non-fluent variant PPA	4	1	0	11	6
Posterior supranuclear palsy syndrome	0	0	0	6	0
Semantic variant PPA	5	2	26	0	2

CBD = corticobasal degeneration; FTD = frontotemporal dementia; FTL D = frontotemporal lobar degeneration; MMSE = Mini-Mental State Examination; n.a. = not available; PPA = primary progressive aphasia.

donor brains; and (iv) sample assignment, where tissue samples from the AHBA were mapped to 273 parcels of the Brainnetome brain regional parcellation atlas (<https://atlas.brainnetome.org/>).<sup>43</sup> Expression values of regions in the left hemisphere were not mirrored to homologous regions in the right hemisphere, due to the predilection of distinct FTL D subtypes to target one hemisphere versus the other (e.g. FTL D-TDP-C for the left hemisphere). See the online [Supplementary material 'Appendix' section](#) for a more detailed report on the procedures using abagen. This atlas was selected because it was derived using functional and structural anatomy and connectivity, and it includes paired homologous bilateral regions. Steps (i–iv) were followed by (v) normalization of expression measures to account for inter-individual differences and outlying values; (vi) gene-set filtering, to remove genes that are inconsistently expressed across the six brains; and (vii) averaging gene expression across donors. This procedure resulted in 15 655 genes ([Supplementary material](#)) and their average expression in 273 brain parcels. Finally, we removed subcortical regions (cerebellum, thalamus, basal ganglia and hippocampus) due to substantial differences in their gene expression values when compared to cortical parcels,<sup>44</sup> yielding a 202 regions × 15 655 genes matrix ([Supplementary Fig. 2 and Supplementary material, 'Data' section](#)).

### Human accelerated region genes

HARs were taken from a comparative genomic analysis that identified, from a list of several recent publications, loci with accelerated divergence in humans when compared to chimpanzees.<sup>7</sup> A total of 2737 HARs were identified, representing 2164 unique HAR-associated genes.<sup>7</sup> Of these 2164 HAR genes, 1373 were identified as sufficiently expressed in the brain based on the AHBA brain-expressed gene dataset (probes used if they exceeded background signal for >50% of all samples) and used in our analyses, referred to simply as HAR genes<sup>45</sup> ([Supplementary material, 'Data' section](#)).

### Cryptic splicing genes

Three prior studies contributed to the list of TDP-43-dependent cryptic splicing genes<sup>24–26</sup>. These studies were chosen because they identified cryptic splicing targets either based on FTL D-TDP human post-mortem tissue data<sup>24</sup> or induced pluripotent stem cell (iPSC)-derived neuronal cell lines,<sup>25,26</sup> omitting studies based on non-neuronal cells. For the tissue-based study, splicing analyses were performed on RNA sequencing data from TDP-43-positive and TDP-43-negative neuronal nuclei isolated from frontal cortices of seven patients with FTL D-TDP. This approach identified 66 cryptic splicing genes,<sup>24</sup> of which 63 genes exceeded the brain expression threshold imposed on the AHBA data. In a separate study, RNA sequencing was performed on human iPSC-derived cortical-like neurons, in which TDP-43 expression was reduced using clustered regularly interspaced short palindromic repeats interference (CRISPRi).<sup>25</sup> Differential splicing and expression analyses based on this dataset followed by splice junction evaluation via visual inspection of sashimi plots revealed 107 genes with putative cryptic splicing events, of which 97 also exceeded the brain expression threshold in the AHBA dataset. Finally, a third study performed a comprehensive analysis of differential splicing events in TDP-43 depleted versus control iPSC-derived neurons to develop a high-quality neuronal cryptic exon atlas.<sup>26</sup> This analysis revealed 233 cryptic splicing genes, of which 201 exceeded the brain expression threshold in the AHBA brain-expressed gene dataset, as previously defined. By combining the two iPSC-derived neuronal cryptic splicing gene sets, 216 unique cryptic splicing genes were identified and combined with the post-mortem tissue-derived cryptic splicing genes to yield a final brain-expressed cryptic splicing gene set of 257 unique genes ([Supplementary material, 'Data' section](#)).

In addition, we identified brain-expressed genes containing higher amounts of GU repeats, which serve as putative regulatory binding sites for TDP-43. The *Bioconductor* *BioMART* package in R was used to retrieve the genetic sequences of brain-expressed genes, which were available for 15 244 of 15 655 brain-expressed

genes. The presence of GU repeats (GU tetramers, pentamers and hexamers) in these brain-expressed genes was estimated using the *universalmotiv* function in R, which queried for GT repeats in the DNA sequences. As GU repeats were found in a majority of brain-expressed genes (14 117/15 244 or 93% of genes), we increased the likelihood that genes are regulated by TDP-43 by identifying those GU repeat-containing genes that contain higher amounts of putative TDP-43 binding sites. This was achieved by identifying a set of 7792 genes that contain higher amounts of GU repeats, defined as above-median content of GU tetramers, pentamers or hexamers.

### Spatial correlation analyses

We derived group-averaged, voxel-wise *W*-score maps for each FTLT pathological subtype (TDP-A, TDP-B, TDP-C, CBD and Pick's disease). For each subtype-averaged map, we then averaged the *W*-score values of voxels contained in each cortical Brainnetome parcel to derive a vector reflecting the level of regional cortical atrophy in each FTLT subtype. These FTLT subtype-specific atrophy vectors were then separately correlated with the cortical gene expression levels derived from the AHBA brain-expressed gene dataset using Pearson's correlation coefficients. This procedure enabled us to assess the spatial similarity between each FTLT subtype atrophy pattern and the cortical expression levels of 15 655 brain-expressed genes from the AHBA.<sup>44,46,47</sup> Similar analyses were performed with a publicly available *W*-score map of cortical atrophy from 147 amyloid- and tau-confirmed patients with Alzheimer's disease-type dementia<sup>48</sup> (Supplementary material, 'Data' section and Supplementary Fig. 7).

Because spatial autocorrelation inherent to neuroimaging data can inflate *P*-values in brain map analyses,<sup>49</sup> we corrected for this autocorrelation by applying a bootstrap-based approach (Supplementary Fig. 3), in which 5000 surrogate maps that preserve the autocorrelation properties of the FTLT atrophy maps were generated using the toolbox BrainSMASH (<https://brainsmash.readthedocs.io/en/latest/approach.html>). We then derived autocorrelation-corrected *P*-values by counting how often the correlation between surrogate atrophy maps and each gene map was equal to or higher than the true correlation and divided this number by the total number of bootstraps. Only false discovery rate (FDR) adjusted *P*-values < 0.05 were considered significant. For control analyses, we also generated lists with a more stringent threshold (FDR adjusted *P* < 0.01; Supplementary material, 'Data' section and Supplementary Fig. 6). This procedure enabled us to identify genes with spatial expression patterns resembling the FTLT atrophy maps.

To remove genes displaying spurious correlations to FTLT-TDP atrophy maps from subsequent analyses, we assessed how applying a threshold based on the absolute correlation value would affect the list of selected genes significantly correlating with FTLT-TDP atrophy maps. We applied absolute correlation thresholds between 0 and 0.5 in steps of 0.05 and derived an index assessing how each threshold impacted the ratio between the number of genes uniquely associated with each FTLT-TDP atrophy map and the number of genes that significantly correlated with at least two FTLT-TDP subtype atrophy maps:

$$Uniqueness_{index} = \frac{(Unique_{TDP-A} + Unique_{TDP-B} + Unique_{TDP-C})}{(Shared_{TDP-AB} + Shared_{TDP-BC} + Shared_{TDP-AC} + Shared_{TDP-ABC})} \quad (2)$$

This uniqueness index was used to identify 0.2 as the optimal absolute correlation threshold to better identify subtype-specific

correlating genes while at the same time preserving longer gene lists. Genes displaying correlation values between 0.2 and -0.2 were hence considered as not being sufficiently correlated with FTLT-TDP atrophy maps and discarded from further analyses regardless of the associated *P*-value. We then assessed the overlap between gene lists representing FTLT atrophy-correlated, HAR and cryptic splicing genes through one-tailed Fisher's exact tests using the *fisher.test* function in R. The 15 565 brain-expressed genes from the AHBA dataset were used as background genes, unless specified otherwise.

### Gene set enrichment analysis

Gene set enrichment analyses are computational methods that query large genomic databases to determine whether an *a priori* defined set of genes shows statistically significant associations with known biological pathways. We ran five separate gene set enrichment analyses<sup>50</sup> leveraging the correlation values (FDR adjusted *P* < 0.05) between brain genes and atrophy maps of FTLT-TDP and FTLT-tau subtypes (one analysis for each subtype). We performed gene set enrichment analyses using the *clusterProfiler* package in R software v4.2.1 for Gene Ontology<sup>51</sup> and for Kyoto Encyclopedia of Genes and Genomes (KEGG).<sup>52</sup> We used the 15 655 brain-expressed genes from AHBA as the background list. We then focused on terms and pathways with a Holm adjusted *P*-value of < 0.05 shared by a majority of FTLT subtypes. Correlation matrices were used to reflect the strength of correlation between genes associated with various KEGG terms and FTLT-TDP and FTLT-tau subtype atrophy patterns (FDR adjusted *P* < 0.05).

### Graph network

Over the course of the study, we identified a set of genes that correlated with FTLT-TDP atrophy and overlapped with the previously introduced cryptic splicing (CS) and HAR gene lists. Graph theoretical approaches were applied to this list of 23 CS-HAR genes correlating with FTLT-TDP atrophy to generate topographical representations of gene expression networks.<sup>53</sup> The regional expression values of selected genes were correlated with each other to generate a gene-to-gene regional co-expression matrix. This matrix was binarized at a Pearson's correlation coefficient threshold of 0.3 and the number of surviving edges between genes was counted for each gene of interest to derive a measure of nodal degree, reflecting the level of connectedness between a specific gene and each other gene selected for the analysis. Resulting data were rendered as a graph.

### Gene length

Standard *Biomart* code in R (<https://useast.ensembl.org/info/data/biomart/index.html>) leveraging the widely used Ensembl genome browser (<https://ensemblgenomes.org/>) was used to derive the full coordinates of exonic and intronic sequences of brain-expressed genes used in our analyses (BSgenome.Hsapiens.UCSC.hg38). This information was used to estimate the length of HAR genes used in our analyses and to identify a set of 1353 length-matched non-HAR genes through nearest neighbour matching using the *matchit* library in R.

### Disease epicentres

We used an approach previously applied by our group and others<sup>54-57</sup> to identify patient-tailored epicentres. This method defines an epicentre as the brain region whose normative functional

(task-free) connectivity map most closely resembles the patient's atrophy *W*-map. Normative connectivity maps were derived using task-free functional MRI data from a cohort of 75 healthy older subjects to generate a library of 194 intrinsic functional connectivity maps from seed regions of the Brainnetome atlas spanning the entire cerebral cortex. Characterization of the normative sample and methodological details are presented elsewhere.<sup>54</sup> We then compared each patient's atrophy *W*-map with this library to select the seed that generated a functional connectivity map showing the highest spatial (Pearson) correlation to the subject-specific atrophy *W*-map. Subtype-specific frequency maps were used to quantify the spatial distribution of epicentres in FTL D-TDP, as group-level atrophy in FTL D-TDP was associated with both cryptic splicing and HAR genes. Finally, we used Wilcoxon signed-rank tests to assess whether the expression of cryptic splicing and HAR genes within disease epicentres differed across FTL D-TDP subtypes. For each gene, epicentres' gene expression levels were then averaged for each subtype to derive a cross-fold change score reflecting higher or lower gene expression in one group versus the other.

## Results

### FTL D subtypes target brain regions expressing shared and distinct genes

We first sought to explore the relationship between FTL D atrophy topography and regional gene expression in the healthy brain. We included the three major FTL D-TDP subtypes (A–C) and the two major sporadic FTL D-tau subtypes, CBD and Pick's disease, with predominant cortical involvement. First, we replicated the atrophy patterns associated with these disorders using *in vivo* brain MRI data from patients with an autopsy-based diagnosis.<sup>58,59</sup> Structural MRI volumes were segmented into grey matter tissue probability maps and further transformed to *W*-score maps,<sup>37,38</sup> where higher *W*-scores reflect more severe voxel-wise atrophy for each patient, adjusted for demographic variables. FTL D subtype-specific cortical atrophy maps were derived by averaging *W*-score maps within each group (Fig. 1A–E, left), revealing a more dorsal frontal and insular pattern in TDP-A; a more ventral pattern in TDP-B, with less severe atrophy overall; an anterior temporal-predominant pattern in TDP-C; a milder, predominantly frontal atrophy pattern in CBD; and a severe frontal and insular pattern of atrophy in Pick's disease.

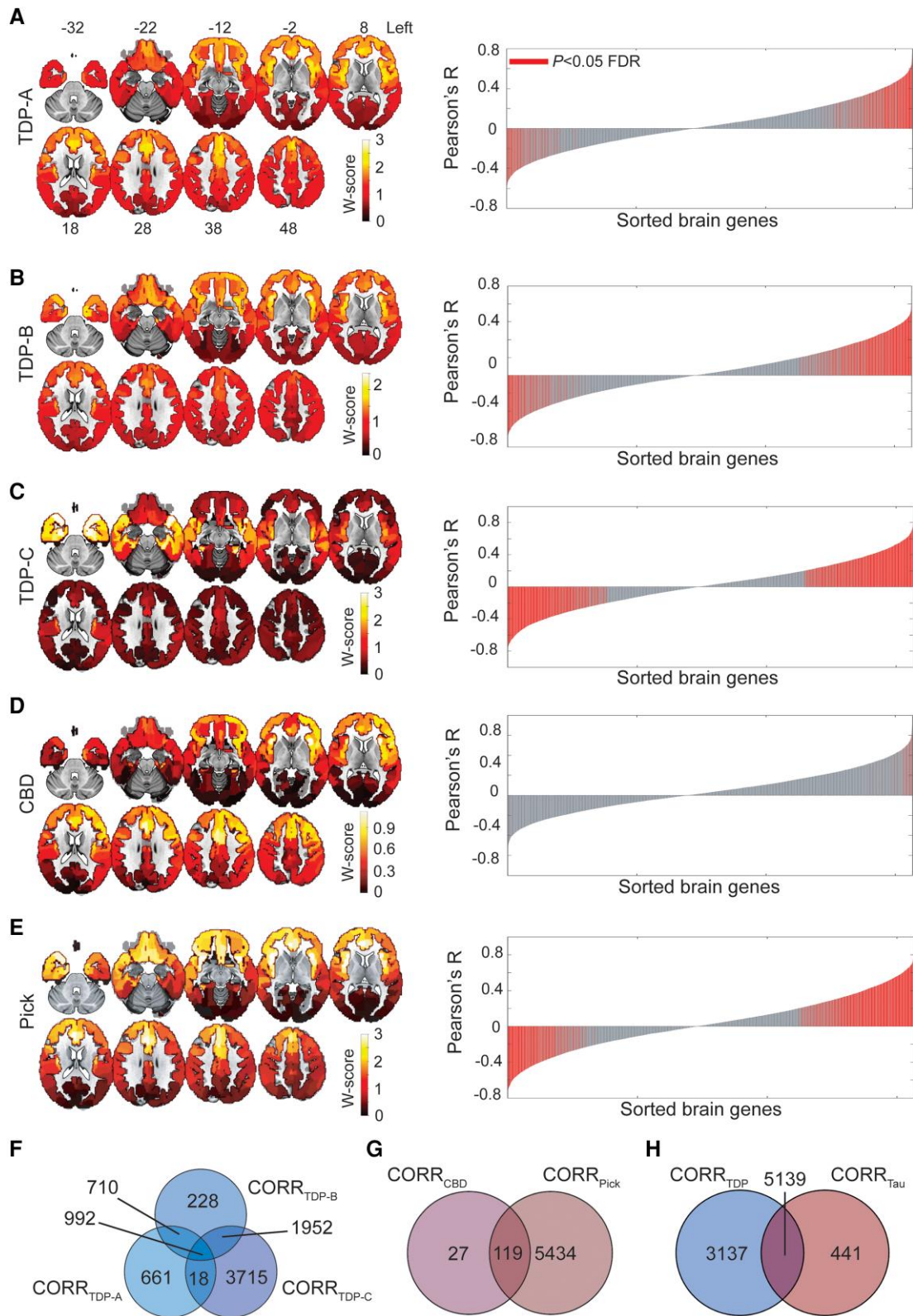
We next leveraged gene expression microarray data from the AHBA to identify genes whose cortical expression pattern resembles cortical grey matter atrophy characteristic of each FTL D subtype. Starting with a matrix of 15 655 genes and their regional expression in 202 cortical brain parcels (Supplementary Fig. 2), we performed spatial correlation analyses to associate group-level grey matter atrophy patterns with average regional expression levels of brain-expressed genes. We used a bootstrap-based test, correcting for the spatial autocorrelations inherent to neuroimaging data,<sup>49</sup> to delineate genes significantly correlated with atrophy in each FTL D subtype (FDR adjusted  $P < 0.05$ , Fig. 1A–E, red bars in the right panels). To remove spurious correlations from further analyses, we only considered absolute correlation values  $\geq 0.2$ , based on a uniqueness index maximizing the number of genes uniquely associated with each FTL D subtype (Supplementary Fig. 3). This procedure identified unique and shared genes correlating with atrophy across FTL D subtypes. TDP-C showed the highest number of uniquely correlated genes (3715) among the FTL D-TDP subtypes, followed by TDP-A (661 genes) and TDP-B (228 genes). A total of 992

genes were shared between all FTL D-TDP subtypes (Fig. 1F). Among FTL D-tau, CBD showed only 27 uniquely correlated genes, due to the mild frontal atrophy found in this sample highly resembling the inherent frontal-posterior spatial autocorrelation gradient inherent to neuroimaging data.<sup>49</sup> Pick's disease showed the most uniquely correlated genes (5434) among all FTL D subtypes. CBD shared 119 genes with Pick's disease, which yielded a total of 5434 uniquely correlated genes for FTL D-tau (Fig. 1G). When pooling both FTL D-TDP and FTL D-tau together, 3137 genes were uniquely correlated with at least one TDP subtype, 441 with at least one tau subtype, and 5139 were shared among TDP-43 and tau FTL D subtypes (Fig. 1H). While positively correlated genes showed high expression levels in atrophied brain regions and low expression levels in spared areas, negatively correlated genes showed the opposite pattern, showing high expression levels in spared brain regions and low expression levels in atrophied areas (Supplementary Fig. 4).

To better understand the biological function of FTL D atrophy-correlated genes, we performed gene set enrichment analyses<sup>50</sup> against Gene Ontology and KEGG.<sup>51</sup> While no gene ontology was considered enriched, the gene set enrichment analysis revealed only one KEGG pathway, associated with neuroactive ligand-receptor interactions (has04080),<sup>52</sup> shared across FTL D-TDP and FTL D-Pick subtypes. This KEGG pathway consisted of common and unique associations between FTL D atrophy-correlated genes and major classes of neuromodulatory and homeostatic neurotransmitter systems (Fig. 2 and Supplementary Fig. 5). Genes associated with acetylcholine receptors did not, for the most part, correlate with atrophy in any FTL D subtype (Fig. 2A), as expected based on the lack of a cholinergic deficit in FTL D.<sup>60</sup> Conversely, genes associated with receptors within the epinephrine/norepinephrine, opioid, somatostatin and neuropeptide Y systems correlated across most FTL D subtypes (Fig. 2B and F–H). Some receptor systems showed more specific associations with FTL D subtypes, as in the case of serotonin and dopamine to TDP-C (Fig. 2C and D), or oxytocin to TDP-B, TDP-C and Pick's disease (Fig. 2E). The lack of associations for CBD is likely attributable to a low number of genes associated with atrophy in this subtype and the high spatial autocorrelation of its atrophy map.

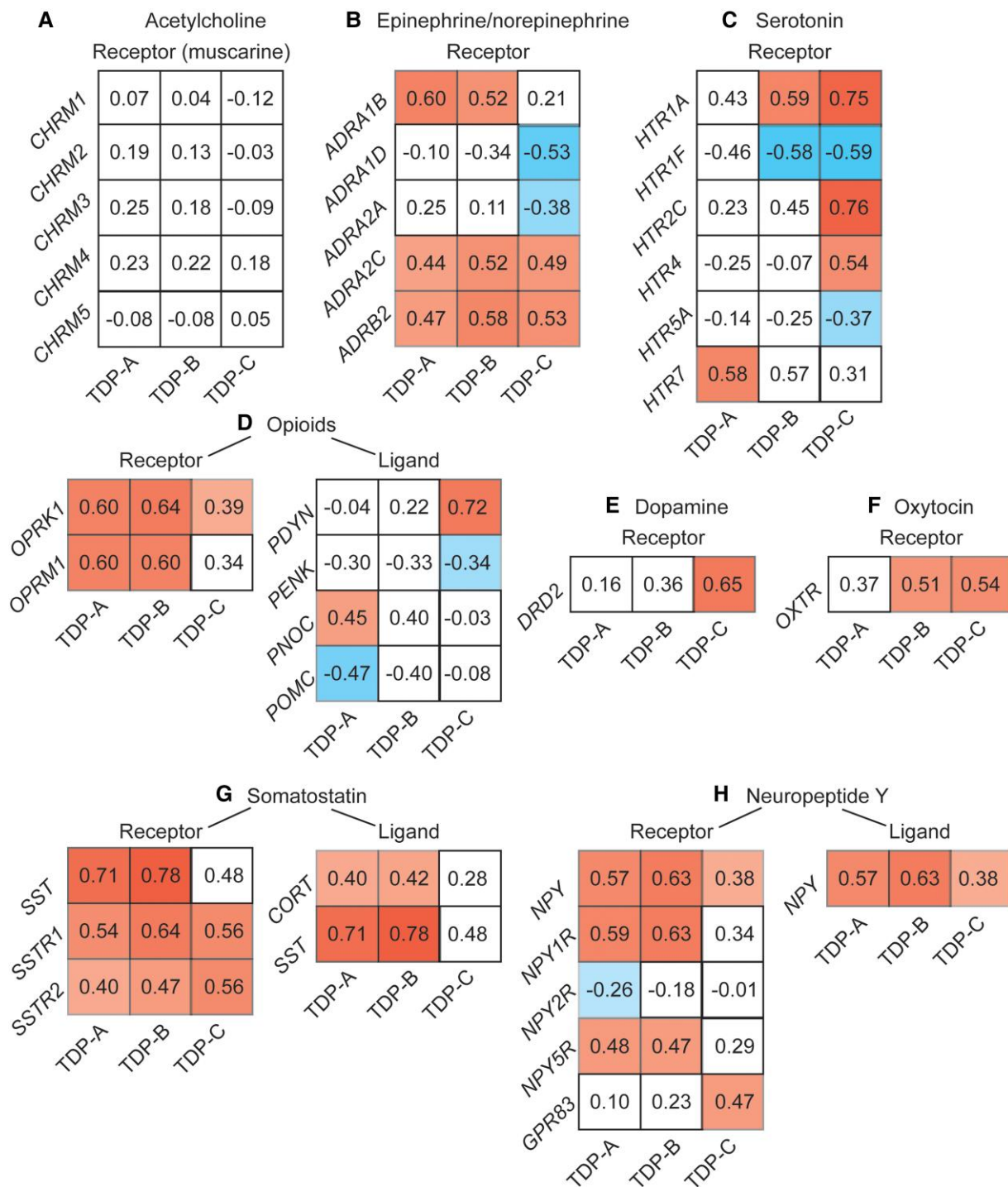
### Genes correlated with FTL D atrophy are enriched for HAR genes

First, we identified 1373 brain-expressed HAR genes derived from a comparative genomic study of conserved loci with elevated divergence in humans versus chimpanzees and other mammals (Fig. 3A and Supplementary material).<sup>7</sup> We next assessed the overlap between brain-expressed HAR genes and genes whose expression patterns correlated with atrophy in FTL D using Fisher's exact test. Atrophy-correlated genes were enriched for HAR genes in both FTL D-TDP (808 genes,  $P < 0.0005$ , Fig. 3B) and FTL D-tau (560 genes,  $P < 0.0005$ , Fig. 3C). Of the 808 FTL D-TDP atrophy-correlated HAR genes, 339 correlated negatively and 470 correlated positively with atrophy (ADCYAP1 correlated both positively and negatively with atrophy in at least one FTL D-TDP subtype). Of the 560 FTL D-tau atrophy-correlated HAR genes, 202 correlated negatively and 358 correlated positively with atrophy. Of the HAR genes correlating with at least one FTL D subtype, 283 were unique to FTL D-TDP, 525 were shared between FTL D-TDP and FTL D-tau (Supplementary Fig. 6A), with only 35 HAR genes correlating uniquely with at least one FTL D-tau subtype. The association between HAR genes and atrophy-correlated genes in FTL D-TDP and FTL D-tau remained significant also after applying more stringent statistical thresholds to derive atrophy-correlated genes (FDR adjusted  $P < 0.01$ ; Supplementary Fig. 7).



**Figure 1 Regional gene expression correlating with atrophy in FTL D subtypes.** (A–E) Maps on the left show group-averaged W-score maps for each frontotemporal lobar degeneration (FTL D) subtype; warmer colours reflect greater cortical grey matter atrophy. Bar plots on the right show the Pearson correlation coefficients between the regional gene expression and regional atrophy in FTL D subtypes; red bars indicate significantly correlated genes, False Discovery Rate (FDR) adjusted  $P < 0.05$ . The left hemisphere is shown on the right. (F) Several genes were shared among FTL D-TDP subtypes, although a conspicuous number of genes was uniquely correlated to each subtype, particularly to FTL D-TDP-C. (G) While FTL D-Pick showed many uniquely correlated genes, most genes correlating with atrophy in FTL D-CBD were shared with FTL D-Pick. (H) FTL D-TDP had many uniquely correlated genes, while a large proportion of genes correlating with atrophy in FTL D-tau were shared with FTL D-TDP.  $CORR_{FTL D-subtype}$  indicates lists of genes correlating with that specific FTL D subtype. CBD = corticobasal degeneration.

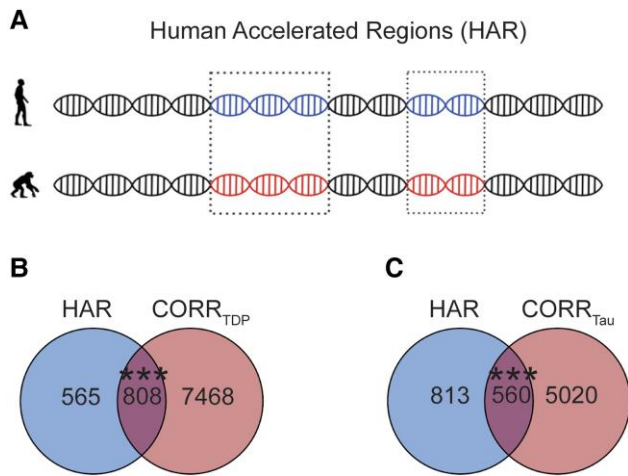




**Figure 2 Genes correlating with atrophy in FTL D are enriched for neuromodulatory terms.** Gene set enrichment analyses using the spatial correlation values between brain gene expression levels and frontotemporal lobar degeneration (FTLD) atrophy severity revealed a Kyoto Encyclopedia of Genes and Genomes (KEGG) neuroligand receptor interaction pathway in common among most FTL D subtypes. Whereas genes associated with acetylcholine neurotransmission (A), a system often preserved in FTL D, did, for the most part, not correlate with FTL D atrophy. Shared and unique associations were observed between FTL D subtypes and genes associated with epinephrine (B), serotonin (C), dopamine (D), oxytocin (E), opioids (F), somatostatin (G) and neuropeptide Y receptor and ligand systems (H). Matrices reflect the strength and sign of spatial correlation between genes associated with ligand and receptor systems and atrophy maps in FTL D subtypes. White squares reflect genes whose correlation with FTL D atrophy maps did not reach significance based on the bootstrap test correcting for spatial autocorrelation (FDR adjusted  $P < 0.05$ ).

To evaluate the specificity of the overlap between HAR genes and FTL D atrophy-correlated genes, we leveraged a publicly available W-score map of cortical atrophy from 147 amyloid- and tau-confirmed patients with Alzheimer’s disease-type dementia

(henceforth, AD).<sup>48</sup> This analysis revealed an overlap between HAR genes and AD atrophy-correlated genes, but this overlap did not reach significance and was statistically smaller than the overlap between HAR genes and random FTL D-TDP and FTL D-tau

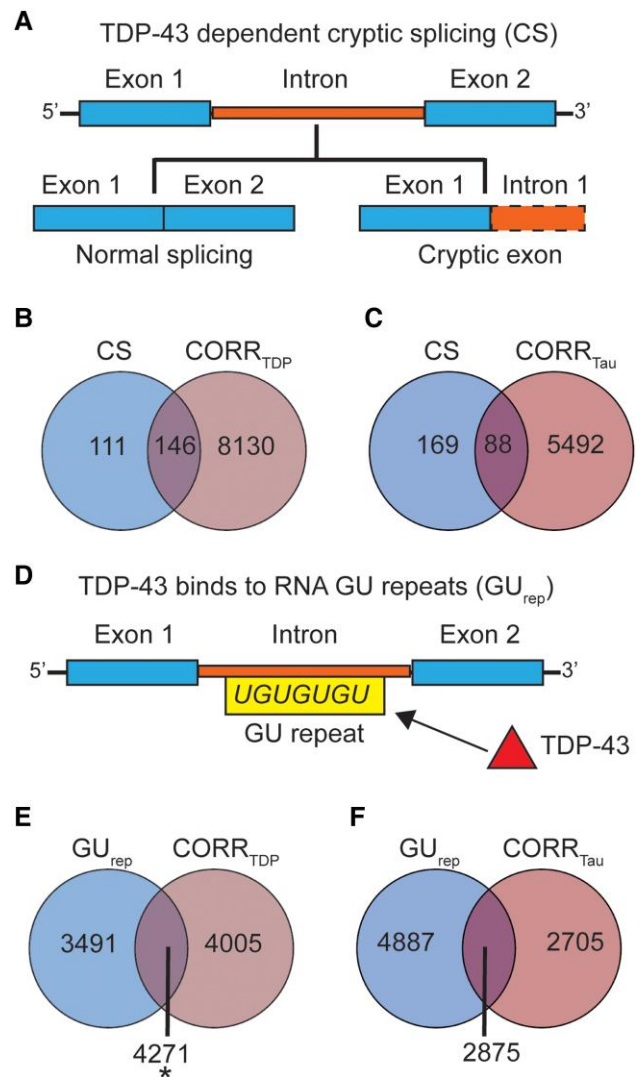


**Figure 3** Overlap between FTLD atrophy correlated genes and HAR genes. (A) HARs are conserved genomic loci which have undergone accelerated divergence in the human evolutionary lineage. (B) Genes correlating with atrophy in FTLD-TDP had 808 genes in common with HAR genes, (C) while genes correlating with atrophy in FTLD-tau had 560 genes in common with HAR genes. Both overlaps were unlikely to occur by chance when compared to a background set of brain-expressed genes. \*\*\* $P < 0.0005$ . (A) adapted with permission from Doan et al.<sup>7</sup> and Wei et al.<sup>45</sup> FTLD = frontotemporal lobar degeneration.

atrophy-correlated gene lists of same size as the AD atrophy-correlated genes (Supplementary Fig. 8).

### FTLD-TDP atrophy-correlated genes are linked to TDP-43 regulated genes

Recent studies have identified a corpus of genes that undergo cryptic splicing upon TDP-43 loss-of-function<sup>24–26</sup> (Fig. 4A). Given the strong enrichment of FTLD-TDP atrophy-correlated genes for HAR genes, and that humans show non-conserved patterns of TDP-43 related cryptic splicing, we hypothesized that TDP-43 cryptic splicing genes might be enriched among genes correlated with FTLD-TDP (but not FTLD-tau) atrophy. To explore this possibility, we consolidated published and unpublished TDP-43 cryptic splicing gene lists, filtering for human brain-expressed cryptic splicing genes identified using human cell models (216 genes)<sup>25,26</sup> or patient brain tissues (63 genes<sup>24</sup>; Supplementary Table 3), resulting in 257 unique cryptic splicing genes (Supplementary material). One hundred and forty-six genes were shared between FTLD-TDP atrophy-correlated and cryptic splicing genes, with 82 cryptic splicing genes correlating negatively and 64 correlating positively with atrophy, yet this overlap was not significant ( $P = 0.11$ , Fig. 4B). To evaluate the biological relevance of the observed trend, we compared the overlap between cryptic splicing genes and FTLD-TDP atrophy-correlated genes to the overlap between cryptic splicing genes and FTLD-tau atrophy-correlated genes. We identified 88 FTLD-tau correlated cryptic splicing genes, of which 51 correlated negatively and 37 correlated positively with atrophy. In contrast to FTLD-TDP, however, this overlap did not approach significance ( $P = 0.71$ , Fig. 4C). The proportion of cryptic splicing genes overlapping with FTLD-TDP atrophy-correlated genes was significantly higher than the proportion of cryptic splicing genes overlapping with FTLD-tau atrophy-correlated genes [ $\chi^2(2,514) = 26.39$ ,  $P < 0.0001$ ]. To control these analyses for the smaller number of FTLD-tau atrophy-correlated genes, we randomly re-sampled (5000 times) lists of 5580 FTLD-TDP atrophy-correlated genes



**Figure 4** Overlap between FTLD atrophy correlated genes and genes regulated by TDP-43. (A) Cryptic splicing (CS) genes are genes that incorporate novel intronic RNA into mature mRNA when TDP-43 is knocked down experimentally or depleted from diseased neuronal nuclei. (B) One hundred and forty-six genes were shared between cryptic splicing genes and genes correlating with atrophy in FTLD-TDP. This overlap did not reach significance ( $P = 0.11$ ). (C) Eighty-eight genes were shared between cryptic splicing genes and genes correlating with atrophy in FTLD-tau, with this overlap being far from significant ( $P = 0.75$ ). (D) TDP-43 binds to GU repeats located on the intronic portions of the pre-messenger RNA. GU repeats are ubiquitous among brain-expressed genes, we hence identified a set of 7792 genes with above-median GU repeat content, increasing the likelihood that these genes are regulated by TDP-43. (E) Genes ( $n = 4271$ ) were shared between genes with above-median GU repeat content and genes correlating with atrophy in FTLD-TDP. This overlap was significant ( $P = 0.03$ ). (F) Genes ( $n = 2875$ ) were shared between genes with above-median GU repeat content and genes correlating with atrophy in FTLD-tau, with this overlap being not significant ( $P = 0.13$ ). \* $P < 0.05$ . FTLD = frontotemporal lobar degeneration.

(matching the number of FTLD-tau atrophy-correlated genes). This approach revealed a significantly higher overlap with cryptic splicing genes for FTLD-TDP when compared to FTLD-tau atrophy-correlated genes (Supplementary Fig. 9,  $P < 0.05$ ), suggesting that the differential overlap between FTLD-TDP and FTLD-tau atrophy-correlated genes with cryptic splicing genes is not merely a

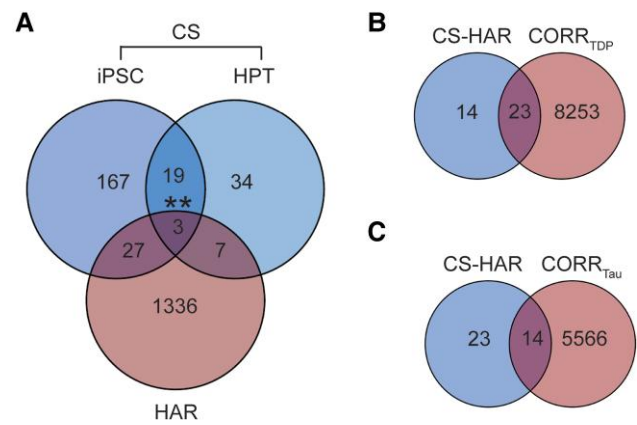
reflection of gene list length. Cryptic splicing genes correlated with at least one FTL D subtype were either unique to FTL D-TDP (63 genes) or shared between FTL D-TDP and FTL D-tau (83 genes) (Supplementary Fig. 6B). Overall, these findings suggest that although both FTL D-TDP and FTL D-tau target brain regions expressing recently evolved genes, FTL D-TDP exhibits a closer relationship to regions that express cryptic splicing genes.

To strengthen the association between TDP-43 transcriptional regulation functions and genes linked to atrophy in FTL D-TDP, we next determined brain-expressed genes with a higher likelihood of being regulated by TDP-43, operationally defined here as having an above-median number of GU repeats, the putative binding site of TDP-43 (Fig. 4D).<sup>24</sup> This approach yielded 7792 genes containing higher amounts of GU repeats compared to other brain-expressed genes [GU tetramers, median (range) = 324 (46–11 313); GU pentamers, median (range) = 107 (23–4087); GU hexamers, median (range) = 37 (5–1663)]. We found a significant overlap between genes with above-median GU repeat content and genes correlating with atrophy in FTL D-TDP ( $P = 0.03$ ), but not for genes correlating with atrophy in FTL D-tau ( $P = 0.13$ ) (Fig. 4E and F).

### HAR genes overlap with TDP-43 cryptic splicing genes

Among the TDP-43 CS genes, 37 were also HAR genes (Fig. 5A and Supplementary material), an overlap that was higher than expected by chance (Fisher's exact test). Identified CS-HAR genes included genes linked to social behaviour and language (e.g. *ERC2*, *KIF26B*, *CBLN2*, *SEMA6D*),<sup>61–65</sup> as well as neurodevelopment and neuropsychiatric diseases (e.g. *CBLN2SPRY3*, *CAMTA1*, *LRP8*).<sup>66–70</sup> Some CS-HAR genes identified (e.g. *CBLN2*, *ICA1*, *PTPRT* and *PTPRD*) have been shown to accurately distinguish FTL D-TDP post-mortem cortex from non-neurological disease control tissue based on the frontal/temporal cortex expression of cryptic exons.<sup>26</sup> CS-HAR genes tended to be longer than other brain-expressed genes [ $t(39) = 4.74$ ;  $P < 0.0001$ ], but their length fell within the range found for 85% of brain-expressed genes (Supplementary Fig. 10A). CS-HAR genes contained mostly intronic HAR sequences (76%, Supplementary Fig. 10B), followed by intergenic (21%) and exonic (3%) sequences, and cryptic splicing events affecting CS-HAR genes mainly included cryptic exon incorporation<sup>20</sup> (71%, Supplementary Fig. 10C), consistent with other (non-HAR) cryptic splicing genes. We performed control analyses assessing the overlap between cryptic splicing genes and a set of non-HAR genes matched for length to HAR genes (Supplementary Fig. 11A). This approach yielded an overlap of 42 genes (Supplementary Fig. 11B) [ $\chi^2(2,2726) = 0.41$ ,  $P = 0.52$ ], suggesting that gene size is a confounding factor driving the overlap between cryptic splicing and HAR genes. Therefore, although HAR and TDP-43 cryptic splicing genes overlap, our analyses do not provide evidence for a special, gene length-independent relationship between HAR and TDP-43 cryptic splicing genes.

We next sought to better understand relationships between FTL D atrophy-correlated genes and the 37 CS-HAR genes (Fig. 5A). Of these 37 genes, 23 overlapped with FTL D-TDP atrophy-correlated genes ( $P = 0.17$ , Fig. 5B) and 14 with FTL D-tau atrophy-correlated genes ( $P = 0.98$ , Fig. 5C), but these overlaps were not greater than expected by chance. Of the CS-HAR genes correlating with at least one FTL D subtype, 10 were unique to FTL D-TDP and 13 were shared between FTL D-TDP and FTL D-tau (Supplementary Fig. 6C). Accordingly, to generate hypotheses, we examined the relationships among the 23 FTL D-TDP atrophy-associated CS-HAR genes in terms of their co-expression relationships in the healthy brain

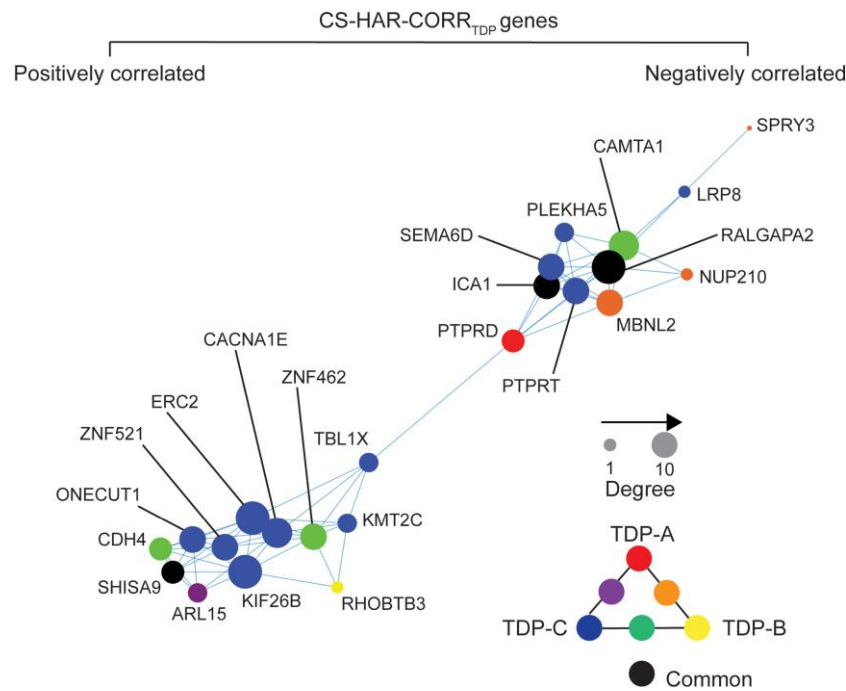


**Figure 5 HAR and cryptic splicing genes overlap.** (A) Comparing the human accelerated region (HAR) and cryptic splicing (CS) gene lists, including those derived from induced pluripotent stem cell (iPSC)-derived neurons and from human patient tissues (HPT), 37 total genes were shared. Fisher's exact test was used to determine whether gene sets overlapped at a rate greater than expected by chance. We then explored the overlap between CS-HAR genes and genes correlating with atrophy in FTL D-TDP or FTL D-tau. (B) Twenty-three genes were shared between CS-HAR genes and genes correlating with atrophy in FTL D-TDP. This overlap did not reach significance ( $P = 0.17$ ). (C) Fourteen genes were shared between CS-HAR genes and genes correlating with atrophy in FTL D-tau, with this overlap not being significant ( $P = 0.98$ ). \*\* $P < 0.005$ . FTL D = frontotemporal lobar degeneration.

and their links to specific FTL D-TDP subtypes (Fig. 6). While some genes correlated negatively with atrophy (e.g. *PTPRD*), suggesting lower average expression among FTL D-TDP atrophied regions, other genes correlated positively (e.g. *ERC2*), reflecting higher expression among targeted regions. Some FTL D-TDP subtype-specific associations emerged, such as *PTPRD* with TDP-A, *RHOBTB3* with TDP-B, or *ERC2* with TDP-C, while other genes, such as *RALGAP2*, *ICA1* and *SHISA9*, were shared among all FTL D-TDP subtypes.

### Expression of cryptic splicing and HAR genes varies across FTL D-TDP subtypes and patient epicentres

The preceding analyses used group-level methods to identify brain-expressed HAR and cryptic splicing genes associated with cortical grey matter atrophy in FTL D subtypes, but there is also prominent anatomical heterogeneity within each subtype,<sup>71</sup> leading us to explore a more individualized, patient-tailored approach. Neurodegenerative diseases begin within specific vulnerable and network-anchoring regions, which we have termed 'epicentres',<sup>54,56</sup> before spreading to other areas along large-scale brain network connections.<sup>4,56,72</sup> In previous work, we introduced a method for identifying patient-specific epicentres, which enhanced network-based predictions of longitudinal atrophy progression in FTL D.<sup>54,56</sup> This approach uses task-free functional MRI acquired in a normative sample to estimate each brain region's connectivity to all other brain regions (Supplementary Fig. 12A) and correlates these connectivity maps with an individual patient's atrophy pattern to identify the region (i.e. epicentre) whose normative connectivity most resembles the atrophy pattern.<sup>54,56,57</sup> Based on this approach, we identified a single best-fit epicentre for each patient, focusing on FTL D-TDP given its specific relationship to cryptic splicing (Fig. 7A). Frequency maps revealed common epicentres in TDP-A, including the pregenual and subgenual anterior



**Figure 6** Co-expression network of CS-HAR genes associated with FTLD-TDP atrophy. The regional expression covariance of CS-HAR genes associated with FTLD-TDP atrophy was used to generate a topographical network of gene expression. The size of each circle reflects the centrality of each gene in the network, while the colour of the circles reflects the FTLD-TDP subtype whose atrophy pattern correlated with the gene's expression pattern. CS = cryptic splicing; FTLD = frontotemporal lobar degeneration; HAR = human accelerated region.

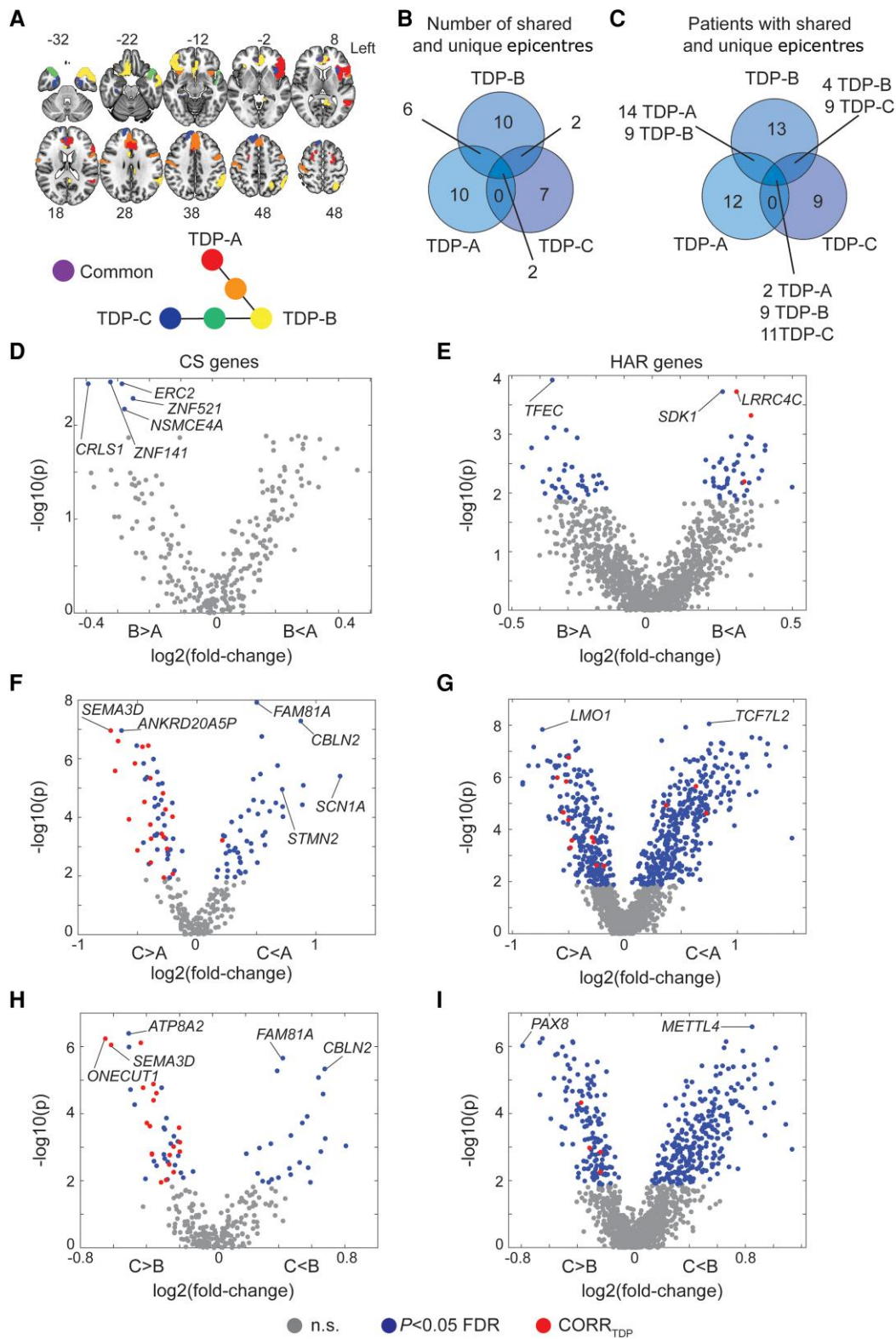
cingulate cortices and the anterior insula, with the anterior cingulate being the most common epicentre. The left anterior insula was the most common epicentre in TDP-B, which showed additional epicentres in the anterior cingulate and temporal pole. In TDP-C, the left entorhinal cortex was the most common epicentre, with additional epicentres in adjacent temporal polar regions (Supplementary Fig. 12B–D). While many epicentres were shared across subtypes (e.g. the left anterior cingulate cortex), several were associated with just one FTLD-TDP subtype, such as the left inferior frontal gyrus in TDP-A, the left angular gyrus in TDP-B, or the left fusiform gyrus with TDP-C (Fig. 7A). Overall, despite each FTLD-TDP subtype providing 7–10 uniquely associated epicentres (Fig. 7B), a substantial proportion of patients shared epicentres across subtypes (Fig. 7C), suggesting that any influence of gene expression on subtype-associated regional vulnerability is not likely to be deterministic.

To formulate hypotheses regarding potential HAR- and CS-related genes and pathways whose expression might influence the onset of FTLD-TDP in different regions across different FTLD-TDP subtypes, we then compared the average expression of HAR genes or TDP-43 cryptic splicing genes across FTLD-TDP subtype epicentres by using Wilcoxon signed-rank tests (FDR adjusted  $P < 0.05$ ). Resulting  $P$ -values were plotted against the cross-fold change, reflecting higher or lower gene expression in epicentres of one FTLD-TDP subtype versus another (Fig. 7D–I). These analyses revealed genes more highly expressed in regional epicentres found in TDP-B compared to TDP-A, such as cryptic splicing gene *ERC2* and HAR gene *TFCF* (Fig. 7D and E), and genes more highly expressed in epicentres for TDP-A versus TDP-B, such as HAR genes *LRR4C* and *SDK1* (Fig. 7E). When comparing TDP-A to TDP-C, the cryptic splicing gene *STMN2* was significantly more expressed in regional epicentres found in TDP-A, while the cryptic splicing gene

*SEMA3D* was more expressed in regional epicentres found in TDP-C (Fig. 7F). The HAR gene *LMO1* was significantly more expressed in epicentres of TDP-C, while the HAR gene *TCF7L2* was more expressed in epicentres of TDP-A (Fig. 7G). When comparing TDP-B to TDP-C, the cryptic splicing gene *CBLN2* and the HAR gene *METTL4* showed significantly higher expression among disease epicentres derived from TDP-B, while the cryptic splicing gene *ONECUT1* and the HAR gene *PAX8* were more highly expressed in epicentres of TDP-C (Fig. 7H and I).

## Discussion

The integrative anatomical biology data presented here identified overlapping and distinct genes associated with cortical atrophy in FTLD-TDP and FTLD-tau. Both gene sets were enriched for HAR genes, providing novel comparative transcriptomic evidence that FTLD selective vulnerability, regardless of the underlying molecular pathology, relates to genes that have undergone recent evolution as humans diverged from chimpanzees. Although HAR genes have been associated with neuropsychiatric and neurodevelopmental disorders,<sup>7</sup> FTLD is the first neurodegenerative disease that has been directly linked to regional patterns of HAR gene expression. Human studies leveraging chromatin interaction profiles in the fetal and adult cortex have revealed that specific cell types and laminae involved in cerebral cortical expansion are enriched for HAR elements.<sup>73</sup> These findings echo neuroimaging studies showing that the expression of HAR genes in the healthy human cerebral cortex is highest among brain networks that have undergone rapid evolutionary expansion in humans.<sup>45</sup> These recently expanded higher-order networks overlap with systems supporting language and social-emotional functions, the same systems that



**Figure 7** Average expression of cryptic splicing and HAR genes in FTL D-TDP disease epicentres. (A) Neuroanatomical distribution of FTL D-TDP subtype unique and shared epicentres. (B) Number of epicentres unique or shared across FTL D-TDP subtypes. (C) Number of patients with FTL D-TDP having unique or shared epicentres. Epicentres' average expression of cryptic splicing genes compared between (D) TDP-A and TDP-B, (E) TDP-A and TDP-C, and (F) TDP-B and TDP-C. Epicentres' average expression of human accelerated region (HAR) genes compared between (G) TDP-A and TDP-B, (H) TDP-A and TDP-C, (I) TDP-B and TDP-C. Grey circles depict genes with average expression within epicentres that do not significantly (n.s.) differ across FTL D-TDP subtypes, blue circles depict genes with average expression within epicentres that significantly differ (FDR adjusted  $P < 0.05$ ) across FTL D-TDP subtypes, red circles depict genes with average expression within epicentres that significantly differ (FDR adjusted  $P < 0.05$ ) across FTL D-TDP subtypes and also positively correlate with atrophy in the FTL D-TDP subtype where they are more expressed. FTL D = frontotemporal lobar degeneration.

undergo early and progressive neurodegeneration in FTLT.<sup>71</sup> Here, the expression of atrophy-associated HAR and non-HAR genes was either positively or negatively associated with FTLT atrophy. Genes showing a positively associated gradient are likely to be highly expressed, in the healthy brain, in regions subject to FTLT atrophy. Dependence on these genes could contribute to FTLT regional vulnerability, especially if protein expression is diminished, for example by TDP-43 loss-of-function and associated mis-splicing. Conversely, genes displaying a negative gradient are probably expressed at lower levels in the healthy brain among FTLT-atrophied regions, with neurodegeneration further compromising already low levels of expression. FTLT atrophy-associated genes were characterized by common and unique associations with neuromodulatory and homeostatic transmitter systems affected in FTLT.<sup>74</sup> Dopaminergic, norepinephrergic and serotonergic imbalances are common in FTD, while the acetylcholine system remains relatively intact.<sup>60</sup> Common antidepressants, such as selective serotonin reuptake inhibitors, have been shown to improve the behavioural symptoms of FTD,<sup>60</sup> reinforcing the notion of neuromodulatory imbalances in FTLT. Our analyses further revealed that FTLT-TDP atrophy-correlated genes are slightly enriched for genes that are targets of TDP-43 cryptic splicing repression, significantly more so than FTLT-tau atrophy-correlated genes. This observation, though tentative, suggests a potential FTLT-TDP-specific mechanism whereby incipient TDP-43 loss-of-function focuses vulnerability on specific brain networks by depleting regional levels of TDP-43 CS genes.

Several lines of evidence derived from animal<sup>75,76</sup> and cellular models,<sup>77</sup> as well as human neuroimaging<sup>54,56</sup> and tissue studies,<sup>30</sup> support the notion that neuropathological changes start within vulnerable brain regions, often referred to as epicentres, before spreading via network pathways.<sup>78,79</sup> In this context, our study not only shows that the expression of HAR and cryptic splicing genes resembles atrophy patterns in FTLT-TDP but also that the expression of these genes within putative epicentres differs among FTLT-TDP subtypes. Our findings raise the possibility that the regional expression of specific HAR and TDP-43 cryptic splicing genes may predispose local circuits to cultivate specific FTLT-TDP subtypes, each characterized by a distinct TDP-43 aggregation pattern.<sup>80</sup> After accumulating within vulnerable epicentres, TDP-43 aggregates could then spread to other brain regions via functional and structural connectivity pathways. Histopathological studies suggest that neurodegenerative diseases may start within susceptible cell-types harboured within vulnerable brain regions.<sup>4,14,81,82</sup> For example, von Economo neurons and fork cells have been shown to be selectively targeted by TDP-43 depletion and inclusions in bvFTD due to FTLT-TDP.<sup>83</sup> These large, elongated neuron types are almost exclusively found within the anterior cingulate and frontoinsula of large-brained, highly social mammals<sup>84</sup>—including humans, greater apes, elephants and cetaceans—raising the possibility that these neurons contribute to circuits linked to the emergence of complex social behaviours.<sup>84</sup> Further, a recent neuroimaging genomics study revealed that atrophied regions in FTD genetic mutation carriers express astrocyte and endothelial cell-related genes.<sup>47</sup> Future histopathological studies and cell-type enrichment analyses could shed light on how expression levels of HAR and cryptic splicing genes within vulnerable neuronal, such as von Economo neurons, interact with glial cells and contribute to the emergence of distinct FTLT-TDP subtypes.

Our findings should be interpreted in light of several methodological considerations. First, our study used average brain transcriptomic data derived from the AHBA, which is based on bulk tissue

microarray data from the brains of six donors. Both hemispheres were collected in only two of six donors (with the remainder providing only the left). Future studies leveraging transcriptomic data derived from larger samples are needed to confirm the transcriptomic correlates of regional grey matter atrophy in FTLT. Moreover, we used a W-model approach to correct for the influence of several confounding factors on grey matter atrophy, including scanner type. It is possible that residual influence from scanner type survived our approach, which could be mitigated in future studies by applying more sophisticated harmonization methods.<sup>85</sup> Further, our study focused on cortical gene expression, omitting subcortical areas despite the relevance of the striatum, thalamus and other subcortical areas in FTLT.<sup>86</sup> Gene expression levels from bulk tissue microarray dramatically differ depending on whether these are extracted from cortical or subcortical areas, with most studies, including ours, analysing either one set of structures or the other.<sup>42,45</sup> The overlap we observed between HAR and cryptic splicing genes, though interesting, appears to be driven primarily by gene length. This makes sense given that longer genes may be implicated in multiple processes including brain development and neurodegenerative diseases.<sup>87,88</sup> Further studies are needed to assess whether altered sequences within HAR genes mechanistically interact with genomic regions regulated or bound by TDP-43. Finally, our study focused on cryptic splicing genes identified in FTLT-TDP human post-mortem tissue data<sup>24</sup> or iPSC-derived neuronal cell lines,<sup>25,26</sup> omitting those identified in HeLa cell lines.<sup>89</sup> Although this choice was intended to enhance biological confidence regarding the neural relevance of these genes, future studies may expand the TDP-43 ‘cryptic spliceosome’ to explore a more exhaustive gene set.

## Data availability

In house Python 2.7.16 (<https://www.python.org/>), R 4.1.2 (<https://www.r-project.org/>) and MATLAB R2021a (<https://www.mathworks.com/products/matlab.html>) scripts were used for the analyses. W-score maps averaged for each FTLT subtype are publicly available on NeuroVault<sup>90</sup> (<https://identifiers.org/neurovault.collection:16591>). The gene lists used to generate the findings are available as [Supplementary material](#). Further data that support the findings of this study are available on request from the corresponding author. Correspondence should be directed to Lorenzo Pasquini or William W. Seeley.

## Acknowledgements

We thank the participants and their families for their invaluable contributions to neurodegeneration research. We thank Sean Whalen for useful feedback on data analysis.

## Funding

This work was supported by National Institutes of Health National Institute on Aging grants K99-AG065457 and R00-AG065457 and by a Feodor Lynen Research Fellowship from the Humboldt Foundation to L.P.; P30AG062422, P01AG019724, U01AG057195 and U19AG063911 (W.W.S.), as well as the Rainwater Charitable Foundation and the Bluefield Project to Cure FTD (W.W.S.); GBHI ALZ UK-21-720973 and AACSF-21-850193 as well as JR20/0018 and PI21/00791 to I.I.G.; 2021-A-023-FEL to A.F.; Gladstone Institutes and R01MH123178 to K.S.P.

## Competing interests

L.P. is a scientific advisor for AWEAR LLC. His role as an advisor is unrelated to the content of the current paper. I.I.-G. reported personal fees from Esteve S.L, KRKA farmacéutica S.L, Nutricia, and Sociedad Española de Neurología and grants from Instituto de Salud Carlos. H.J.R. reported personal fees from Genentech Pharmaceuticals, Eisai Pharmaceuticals, Takeda Pharmaceuticals, Biogen Pharmaceuticals, and Wave Pharmaceuticals outside the submitted work. B.L.M. serves as medical advisor for the John Douglas French Foundation; serves as scientific advisor for the Larry L. Hillblom Foundation, Association for Frontotemporal Degeneration, National Institute for Health Research Cambridge Biomedical Research Centre and its subunit, the Biomedical Research Unit in Dementia, and the Buck Institute for Research on Aging; serves as external advisor for University of Washington ADRC, Stanford University ADRC, Arizona Alzheimer's Disease Center, and Massachusetts Alzheimer Disease Research Center; and serves on the external advisory committee for University of Southern California P01 Urban Air Pollution and Alzheimer's Disease: Risk, Heterogeneity and Mechanisms. No other disclosures were reported.

## Supplementary material

[Supplementary material](#) is available at *Brain* online.

## References

- Gorno-Tempini ML, Hillis AE, Weintraub S, et al. Classification of primary progressive aphasia and its variants. *Neurology*. 2011;02:1-10.
- Rascovsky K. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*. 2011;6: 234-237.
- Schroeter ML, Raczka K, Neumann J, von Cramon DY. Neural networks in frontotemporal dementia—A meta-analysis. *Neurobiol Aging*. 2008;29:418-426.
- Seeley William W. Mapping neurodegenerative disease onset and progression. *Cold Spring Harb Perspect Biol* 2017;312: 1389-1392.
- Gorno-Tempini ML. Cognition and anatomy in three variants of primary progressive aphasia. *Ann Neurol*. 2004;119:2658-2666.
- Mackenzie IRA, Neumann M, Baborie A, et al. A harmonized classification system for FTL-DTP pathology. *Acta Neuropathol*. 2011; 122:111-113.
- Doan RN, Bae BI, Cubelos B, et al. Mutations in human accelerated regions disrupt cognition and social behavior. *Cell*. 2016; 167:341-354.e12.
- Prabhakar S, Visel A, Akiyama JA, et al. Human-Specific gain of function in a developmental enhancer. *Science*. 2008;321: 1346-1350.
- Pollard KS, Salama SR, King B, et al. Forces shaping the fastest evolving regions in the human genome. *PLoS Genet*. 2006;2:e168.
- Levchenko A, Kanapin A, Samsonova A, Gainetdinov RR. Human accelerated regions and other human-specific sequence variations in the context of evolution and their relevance for brain development. *Genome Biol Evol*. 2018;10:166-188.
- Cheung JP, Tubbs JD, Sham PC. Extended gene set analysis of human neuro-psychiatric traits shows enrichment in brain-expressed human accelerated regions across development. *Schizophr Res*. 2022;246:148-155.
- Guardiola-Ripoll M, Fatjó-Vilas M. A systematic review of the human accelerated regions in schizophrenia and related disorders: Where the evolutionary and neurodevelopmental hypotheses converge. *Int J Mol Sci*. 2023;24:3597.
- Vatsavayai SC, Yoon SJ, Gardner RC, et al. Timing and significance of pathological features in C9orf72 expansion-associated frontotemporal dementia. *Brain*. 2016;139(Pt 12):3202-3216.
- Nana AL, Sidhu M, Gaus SE, et al. Neurons selectively targeted in frontotemporal dementia reveal early stage TDP-43 pathobiology. *Acta Neuropathol*. 2019;137:27-46.
- Iguchi Y, Katsuno M, Niwa JI, et al. Loss of TDP-43 causes age-dependent progressive motor neuron degeneration. *Brain*. 2013;136:1371-1382.
- Yang C, Wang H, Qiao T, et al. Partial loss of TDP-43 function causes phenotypes of amyotrophic lateral sclerosis. *Proc Natl Acad Sci U S A*. 2014;111:E1121-E1129.
- Melamed Z, López-Erauskin J, Baughn MW, et al. Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat Neurosci*. 2019;22: 180-190.
- Klim JR, Williams LA, Limone F, et al. ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat Neurosci*. 2019;22:167-179.
- Polymenidou M, Lagier-Tourenne C, Hutt KR, et al. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci*. 2011;14: 459-468.
- Ling JP, Wong PC. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science*. 2015;349:650-655.
- Donde A, Sun M, Ling JP, et al. Splicing repression is a major function of TDP-43 in motor neurons. *Acta Neuropathol*. 2019; 138:813-826.
- Sun M, Bell W, LaClair KD, et al. Cryptic exon incorporation occurs in Alzheimer's brain lacking TDP-43 inclusion but exhibiting nuclear clearance of TDP-43. *Acta Neuropathol*. 2017;133: 923-931.
- Jeong YH, Ling JP, Lin SZ, et al. TDP-43 cryptic exons are highly variable between cell types. *Mol Neurodegener*. 2017;12:13.
- Ma XR, Prudencio M, Koike Y, et al. TDP-43 represses cryptic exon inclusion in the FTD-ALS gene UNC13A. *Nature*. 2022; 603:124-130.
- Brown AL, Wilkins OG, Keuss MJ, et al. TDP-43 loss and ALS-risk SNPs drive mis-splicing and depletion of UNC13A. *Nature*. 2022; 603:131-137.
- Seddighi S, Qi YA, Brown AL, et al. Mis-spliced transcripts generate de novo proteins in TDP-43-related ALS/FTD. *bioRxiv*. [Preprint] <https://doi.org/10.1101/2023.01.23.525149>.
- Braak H, Braak E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol*. 1991;82:239-259.
- Montine TJ, Phelps CH, Beach TG, et al. National institute on aging-Alzheimer's association guidelines for the neuropathologic assessment of Alzheimer's disease: A practical approach. *Acta Neuropathol*. 2012;123:1-11.
- McKeith IG, Dickson DW, Lowe J, et al. Diagnosis and management of dementia with Lewy bodies: Third report of the DLB consortium. *Neurology*. 2005;65:1863-1872.
- Kim EJ, Hwang JHL, Gaus SE, et al. Evidence of corticofugal tau spreading in patients with frontotemporal dementia. *Acta Neuropathol*. 2019;139:27-43.
- Folstein MF, Folstein SEMP. Mini-mental state. A grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12:189-198.
- Knopman DS, Kramer JH, Boeve BF, et al. Development of methodology for conducting clinical trials in frontotemporal lobar degeneration. *Brain*. 2008;131(Pt 11):2957-2968.

33. MacKenzie IRA, Neumann M, Bigio EH, et al. Nomenclature and nosology for neuropathologic subtypes of frontotemporal lobar degeneration: An update. *Acta Neuropathol.* 2010;119:1-4.
34. Kim EJ, Sidhu M, Gaus SE, et al. Selective frontoinsular von Economo neuron and fork cell loss in early behavioral variant frontotemporal dementia. *Cerebral Cortex.* 2012;22:251-259.
35. Tartaglia MC, Sidhu M, Laluz V, et al. Sporadic corticobasal syndrome due to FTL-D-TDP. *Acta Neuropathol.* 2010;119:365-374.
36. Ashburner J, Friston KJ. Voxel-based morphometry—The methods. *Neuroimage.* 2000;11(6):805-821.
37. La Joie R, Visani A V, Baker SL, et al. Prospective longitudinal atrophy in Alzheimer's disease correlates with the intensity and topography of baseline tau-PET. *Sci Transl Med.* 2020;12:1-13.
38. Pasquini L, Nana AL, Toller G, et al. Salience network atrophy links neuron type-specific pathobiology to loss of empathy in frontotemporal dementia. *Cerebral Cortex.* 2020;30:5387-5399.
39. Friedberg A, Pasquini L, Diggs R, et al. Prevalence, timing, and network localization of emergent visual creativity in frontotemporal dementia. *JAMA Neurol.* 2023;80:377-387.
40. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature.* 2012;489:391-399.
41. Arnatkevičiūtė A, Fulcher BD, Fornito A. A practical guide to linking brain-wide gene expression and neuroimaging data. *Neuroimage.* 2019;189:353-367.
42. Markello RD, Arnatkevičiūtė A, Poline JB, Fulcher BD, Fornito A, Misić B. Standardizing workflows in imaging transcriptomics with the abagen toolbox. *Elife.* 2021;10:e72129.
43. Fan L, Li H, Zhuo J, et al. The human brainnetome atlas: A new brain atlas based on connective architecture. *Cerebral Cortex.* 2016;26:3508-3526.
44. Brown JA, Lee AJ, Pasquini L, Seeley WW. A dynamic gradient architecture generates brain activity states. *Neuroimage.* 2022;261:119526.
45. Wei Y, de Lange SC, Scholtens LH, et al. Genetic mapping and evolutionary analysis of human-expanded cognitive networks. *Nat Commun.* 2019;10:1-11.
46. Vogel JW, La Joie R, Grothe MJ, et al. A molecular gradient along the longitudinal axis of the human hippocampus informs large-scale behavioral systems. *Nat Commun.* 2020;11:960.
47. Altmann A, Cash DM, Bocchetta M, et al. Analysis of brain atrophy and local gene expression in genetic frontotemporal dementia. *Brain Commun.* 2020;2:1-13.
48. Strom A, Iaccarino L, Edwards L, et al. Cortical hypometabolism reflects local atrophy and tau pathology in symptomatic Alzheimer's disease. *Brain.* 2022;145:713-728.
49. Burt JB, Helmer M, Shinn M, Anticevic A, Murray JD. Generative modeling of brain maps with spatial autocorrelation. *Neuroimage.* 2020;220:117038.
50. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102:15545-15550.
51. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: Tool for the unification of biology. *Nat Genet.* 2000;25:25-29.
52. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28:27-30.
53. Bassett DS, Sporns O. Network neuroscience. *Nat Neurosci.* 2017;20:353-364.
54. Brown JA, Neuhaus J, Miller BL, et al. Patient-Tailored, connectivity-based forecasts of spreading brain atrophy. *Neuron.* 2019;104:856-868.e5.
55. Brown JA, Hua AY, Trujillo A, et al. Advancing functional dysconnectivity and atrophy in progressive supranuclear palsy. *Neuroimage Clin.* 2017;16:564-574.
56. Zhou J, Gennatas ED, Kramer JH, Miller BL, Seeley WW. Predicting regional neurodegeneration from the healthy brain functional connectome. *Neuron.* 2012;73:1216-1227.
57. Franzmeier N, Dewenter A, Frontzkowski L, et al. Patient-centered connectivity-based prediction of tau pathology spread in Alzheimer's disease. *Sci Adv.* 2023;6:eabd1327.
58. Perry DC, Brown JA, Possin KL, et al. Clinicopathological correlations in behavioural variant frontotemporal dementia. *Brain.* 2017;140:3329-3345.
59. Rohrer JD, Geser F, Zhou J, et al. TDP-43 subtypes are associated with distinct atrophy patterns in frontotemporal dementia. *Neurology.* 2010;75:2204.
60. Huey ED, Putnam KT, Grafman J. A systematic review of neurotransmitter deficits and treatments in frontotemporal dementia. *Neurology.* 2006;66:17-22.
61. Satoh Y, Endo S, Nakata T, et al. ERK2 contributes to the control of social behaviors in mice. *J Neurosci.* 2011;31:11953-11967.
62. Scekcic-Zahirovic J, Sanjuan-Ruiz I, Kan V, et al. Cytoplasmic FUS triggers early behavioral alterations linked to cortical neuronal hyperactivity and inhibitory synaptic defects. *Nat Commun.* 2021;12:3028.
63. Seigneur E, Wang J, Dai J, Polepalli J, Südhof TC. Cerebellin-2 regulates a serotonergic dorsal raphe circuit that controls compulsive behaviors. *Mol Psychiatry.* 2021;26:7509-7521.
64. Ercan-Sencicek AG, Davis Wright NR, Sanders SJ, et al. A balanced t(10;15) translocation in a male patient with developmental language disorder. *Eur J Med Genet.* 2012;55:128-131.
65. Thomas T, Perdue M V, Khalaf S, et al. Neuroimaging genetic associations between SEMA6D, brain structure, and reading skills. *J Clin Exp Neuropsychol.* 2021;43:276-289.
66. Quesnel-Vallières M, Dargaei Z, Irimia M, et al. Misregulation of an activity-dependent splicing network as a common mechanism underlying autism Spectrum disorders. *Mol Cell.* 2016;64:1023-1034.
67. Ning Z, McLellan AS, Ball M, et al. Regulation of SPRY3 by X chromosome and PAR2-linked promoters in an autism susceptibility region. *Hum Mol Genet.* 2015;24:7450.
68. Shinawi M, Coorg R, Shimony JS, Grange DK, Al-Kateb H. Intragenic CAMTA1 deletions are associated with a spectrum of neurobehavioral phenotypes. *Clin Genet.* 2015;87:478-482.
69. Xiao X, Yu H, Li J, et al. Further evidence for the association between LRP8 and schizophrenia. *Schizophr Res.* 2020;215:499-505.
70. Shibata M, Pattabiraman K, Lorente-Galdos B, et al. Regulation of prefrontal patterning and connectivity by retinoic acid. *Nature.* 2021;598:483-488.
71. Kim EJ, Vatsavayi S, Seeley WW. Neuropathology of dementia. In: Miller BL, Boeve BF, eds. *The behavioral neurology of dementia.* Cambridge University Press; 2016:94-122.
72. Brettschneider J, Del Tredici K, Lee VMY, Trojanowski JQ. Spreading of pathology in neurodegenerative diseases: A focus on human studies. *Nat Rev Neurosci.* 2015;16:109-120.
73. Won H, Huang J, Opland CK, Hartl CL, Geschwind DH. Human evolved regulatory elements modulate genes involved in cortical expansion and neurodevelopmental disease susceptibility. *Nat Commun.* 2019;10:2396.
74. Premi E, Dukart J, Mattioli I, et al. Unravelling neurotransmitters impairment in primary progressive aphasia. *Hum Brain Mapp.* 2023;44:2245-2253.
75. Porta S, Xu Y, Restrepo CR, et al. Patient-derived frontotemporal lobar degeneration brain extracts induce formation and spreading of TDP-43 pathology in vivo. *Nat Commun.* 2018;9:4220.



76. de Calignon A, Polydoro M, Suárez-Calvet M, et al. Propagation of tau pathology in a model of early Alzheimer's disease. *Neuron*. 2012;73:685-697.
77. Westergaard T, Jensen BK, Wen X, et al. Cell-to-Cell transmission of dipeptide repeat proteins linked to C9orf72-ALS/FTD. *Cell Rep*. 2016;17:645-652.
78. Raj A, Kuceyeski A, Weiner M. Article A network diffusion model of disease progression in dementia. *Neuron*. 2011;73:1204-1215.
79. Vogel JW, Iturria-Medina Y, Strandberg OT, et al. Spread of pathological tau proteins through communicating neurons in human Alzheimer's disease. *Nat Commun*. 2020;11:2612.
80. Arseni D, Chen R, Murzin AG, et al. TDP-43 forms amyloid filaments with a distinct fold in type A FTL D-TDP. *Nature*. 2023; 620:898-903.
81. Fu H, Hardy J, Duff KE. Selective vulnerability in neurodegenerative diseases. *Nat Neurosci*. 2018;21:1350-1358.
82. Bonham LW, Steele NZR, Karch CM, et al. Protein network analysis reveals selectively vulnerable regions and biological processes in FTD. *Neurol Genet*. 2018;4:e266.
83. Gami-Patel P, van Dijken I, van Swieten JC, et al. Von economo neurons are part of a larger neuronal population that are selectively vulnerable in C9orf72 frontotemporal dementia. *Neuropathol Appl Neurobiol*. 2019;45:671-680.
84. Allman JM, Tetreault NA, Hakeem AY, et al. The von economo neurons in the fronto-insular and anterior cingulate cortex. *Ann N Y Acad Sci*. 2011;1225:59-71.
85. Fortin JP, Cullen N, Sheline YI, et al. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage*. 2018;167:104-120.
86. Lee SE, Sias AC, Kosik EL, et al. Thalamo-cortical network hyperconnectivity in preclinical progranulin mutation carriers. *Neuroimage Clin*. 2019;22:101751.
87. Lopes I, Altab G, Raina P, de Magalhães JP. Gene size matters: An analysis of gene length in the human genome. *Front Genet*. 2021; 12:559998.
88. Soheili-Nezhad S, van der Linden RJ, Olde Rikkert M, Sprooten E, Poelmans G. Long genes are more frequently affected by somatic mutations and show reduced expression in Alzheimer's disease: Implications for disease etiology. *Alzheimer's Dement*. 2021;17: 489-499.
89. Humphrey J, Emmett W, Fratta P, Isaacs AM, Plagnol V. Quantitative analysis of cryptic splicing associated with TDP-43 depletion. *BMC Med Genom*. 2017;10:1-17.
90. Gorgolewski KJ, Varoquaux G, Rivera G, et al. NeuroVault.org: A web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Front Neuroinform*. 2015;9:1-9.