

1 **Deciphering Distinct Genetic Risk Factors for FTLD-TDP**

2 **Pathological Subtypes via Whole-Genome Sequencing**

3 Cyril Pottier^{1,2,3,4,5,*}, Fahri Küçükali^{2,3}, Matt Baker¹, Anthony Batzler⁶, Gregory D. Jenkins⁶,
4 Marka van Blitterswijk¹, Cristina T. Vicente^{2,3}, Wouter De Coster^{2,3}, Sarah Wynants^{2,3}, Pieter
5 Van de Walle^{2,3}, Owen A. Ross¹, Melissa E. Murray¹, Júlia Faura^{2,3}, Stephen J. Haggarty⁷,
6 Jeroen GJ. van Rooij⁸, Merel O. Mol⁸, Ging-Yuek R. Hsiung⁹, Caroline Graff^{10,11}, Linn
7 Öijerstedt^{10,11}, Manuela Neumann^{12,13}, Yan Asmann¹⁴, Shannon K. McDonnell⁶, Saurabh
8 Baheti⁶, Keith A. Josephs¹⁵, Jennifer L. Whitwell¹⁶, Kevin F. Bieniek^{1,17}, Leah Forsberg¹⁵,
9 Hilary Heuer¹⁸, Argentina Lario Lago¹⁸, Ethan G. Geier¹⁸, Jennifer S. Yokoyama¹⁸, Alexis P.
10 Oddi¹⁸, Margaret Flanagan¹⁷, Qinwen Mao¹⁹, John R. Hodges²⁰, John B. Kwok^{21,22}, Kimiko
11 Domoto-Reilly²³, Matthis Synofzik^{12,24}, Carlo Wilke^{12,24}, Chiadi Onyike²⁵, Bradford C.
12 Dickerson²⁶, Bret M. Evers²⁷, Brittany N. Dugger²⁸, David G. Munoz^{29,30}, Julia Keith^{30,31},
13 Lorne Zinman³¹, Ekaterina Rogaeva³², EunRan Suh³³, Tamar Gefen³⁴, Changiz Geula³⁴,
14 Sandra Weintraub³⁴, Janine Diehl-Schmid^{35,36}, Martin R. Farlow³⁷, Dieter Edbauer³⁸, Bryan K.
15 Woodruff³⁹, Richard J. Caselli³⁹, Laura L. Donker Kaat⁴⁰, Edward D. Huey⁴¹, Eric M.
16 Reiman⁴², Simon Mead⁴³, Andrew King^{44,45}, Sigrun Roeber⁴⁶, Alissa L. Nana¹⁸, Nilufer
17 Ertekin-Taner^{1,47}, David S. Knopman¹⁵, Ronald C. Petersen¹⁵, Leonard Petrucelli¹, Ryan J.
18 Uitti⁴⁷, Zbigniew K. Wszolek⁴⁷, Eliana Marisa Ramos⁴⁸, Lea T. Grinberg¹⁸, Maria Luisa Gorno
19 Tempini¹⁸, Howard J. Rosen⁴⁹, Salvatore Spina¹⁸, Olivier Piguet⁵⁰, Murray Grossman⁵¹, John
20 Q. Trojanowski³³, Dirk C. Keene⁵², Jin Lee-Way⁵³, Johannes Prudlo^{54,55}, Daniel H.
21 Geschwind⁴⁸, Robert A. Rissman⁵⁶, Carlos Cruchaga⁵⁷, Bernardino Ghetti⁵⁸, Glenda M.
22 Halliday²¹, Thomas G. Beach⁵⁹, Geidy E. Serrano⁵⁹, Thomas Arzberger^{46,60}, Jochen Herms^{38,46},
23 Adam L. Boxer¹⁸, Lawrence S. Honig⁶¹, Jean P. Vonsattel⁶², Oscar L. Lopez⁶³, Julia Kofler⁶⁴,
24 Charles L. White III²⁷, Marla Gearing⁶⁵, Jonathan Glass⁶⁵, Jonathan D. Rohrer⁶⁶, David J.
25 Irwin⁵¹, Edward B. Lee³³, Vivianna Van Deerlin³³, Rudolph Castellani⁶⁷, Marsel M.
26 Mesulam³⁴, Maria C. Tartaglia³², Elizabeth C. Finger⁶⁸, Claire Troakes⁴⁴, Safa Al-Sarraj^{44,69},
27 Bruce L. Miller¹⁸, Harro Seelaar⁸, Neill R. Graff-Radford⁴⁷, Bradley F. Boeve¹⁵, Ian RA.
28 Mackenzie⁷⁰, John C. van Swieten⁸, William W. Seeley¹⁸, Kristel Slegers^{2,2}, Dennis W.
29 Dickson¹, Joanna M. Biernacka^{6,71}, Rosa Rademakers^{1,2,3,*}

30

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

31 † In memoriam

32 * **Corresponding Authors**

33 Rosa Rademakers, Ph.D. (ORCID: 0000-0002-4049-0863)

34 VIB Center for Molecular Neurology

35 Universiteitsplein 1, 2610 Wilrijk, Belgium

36 Phone: +32 3 265 95 95

37 E-mail: rosa.rademakers@uantwerpen.vib.be

38

39 Cyril Pottier, Ph.D.

40 Department of Neurology, Washington University School of Medicine

41 St Louis, MO, United States of America

42 Phone: +1 314 747 2601

43 E-mail: cpottier@wustl.edu

44

45 ¹Department of Neuroscience, Mayo Clinic, 4500 San Pablo Road, Jacksonville, FL, 32224,
46 United States of America

47 ²Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium

48 ³VIB Center for Molecular Neurology, VIB, Antwerp, Belgium

49 ⁴Department of Neurology, Washington University School of Medicine, St Louis, MO,
50 United States of America

51 ⁵NeuroGenomics and Informatics Center, Washington University School of Medicine, St
52 Louis, MO, United States of America

53 ⁶Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, 55905, United
54 States of America

55 ⁷Department of Neurology, Massachusetts General Hospital and Harvard Medical School,

56 Boston, MA, 02114, United States of America

57 ⁸Department of Neurology, Erasmus Medical Center, Wytemaweg 80, Rotterdam, 3015 CN,
58 The Netherlands

59 ⁹Department of Medicine, Division of Neurology, University of British Columbia,
60 Vancouver, BC, V6T 2B5, Canada

61 ¹⁰Division of Neurogeriatrics, Karolinska Institutet, Bioclinicum J10:20, Visionsgatan 4,
62 Solna, 171 64, Sweden

63 ¹¹Unit for Hereditary Dementias, Karolinska University Hospital, Solna, 171 64, Sweden

64 ¹²German Center for Neurodegenerative Diseases (DZNE), Tübingen, 72076, Germany

65 ¹³Department of Neuropathology, University of Tübingen, Tübingen, 72076, Germany

66 ¹⁴Department of Health Sciences Research, Mayo Clinic, 4500 San Pablo Road, Jacksonville,
67 FL, 32224, United States of America

68 ¹⁵Department of Neurology, Mayo Clinic, Rochester, MN, 55905, United States of America

69 ¹⁶Department of Radiology, Mayo Clinic, Rochester, MN, 55905, United States of America

70 ¹⁷University of Texas Health Science Center San Antonio, San Antonio, TX, 78229, United
71 States of America

72 ¹⁸Department of Neurology, UCSF Weill Institute for Neurosciences, University of
73 California, San Francisco, San Francisco, CA, 94158, United States of America

74 ¹⁹Department of Pathology, University of Utah, Salt Lake City, UT, 84112, United States of
75 America

76 ²⁰Central Clinical School and Brain and Mind Centre, University of Sydney, Sydney, 2050,
77 Australia

78 ²¹University of Sydney, Sydney, 2050, Australia

79 ²²NeuRA, University of New South Wales, Randwick, 2031, Australia

80 ²³Department of Neurology, University of Washington, 325 9th Ave, Seattle, WA, 98104,
81 United States of America

82 ²⁴Division Translational Genomics of Neurodegenerative Diseases, Center for Neurology and
83 Hertie-Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

84 ²⁵Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, Baltimore,
85 MD, 21218, United States of America

86 ²⁶Department of Neurology, Case Western Reserve University, Cleveland, OH, 44106,
87 United States of America

88 ²⁷Division of Neuropathology, University of Texas Southwestern Medical Center, 5323 Harry
89 Hines Blvd, Dallas, TX, 75390-9073, United States of America

90 ²⁸Department of Pathology and Laboratory Medicine, University of California, Davis
91 Medical Center, Sacramento, CA, 95817, United States of America
92 ²⁹St. Michael's Hospital, 30 Bond St, Toronto, ON, M5B 1W8, Canada
93 ³⁰Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto,
94 ON, M5S 1A1, Canada
95 ³¹Sunnybrook Health Sciences Centre, Toronto, ON, M4N 3M5, Canada
96 ³²Krembil Discovery Tower, Tanz Centre for Research in Neurodegenerative Disease,
97 University of Toronto, 60 Leonard Av, Toronto, ON, M5T 0S8, Canada
98 ³³Department of Pathology and Laboratory Medicine, Center for Neurodegenerative Disease
99 Research, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA,
100 19104, United States of America
101 ³⁴Mesulam Center for Cognitive Neurology and Alzheimer's Disease, Northwestern
102 University, Chicago, IL, 60611, United States of America
103 ³⁵Department of Psychiatry and Psychotherapy, Technical University of Munich, Munich,
104 80333, Germany
105 ³⁶kbo-Inn-Salzach-Klinikum, Clinical Center for Psychiatry, Psychotherapy, Psychosomatic
106 Medicine, Geriatrics and Neurology, Wasserburg/Inn, 83512, Germany
107 ³⁷Department of Neurology, Indiana University School of Medicine, 355 West 16th Street,
108 Indianapolis, IN, 46202, United States of America
109 ³⁸German Center for Neurodegenerative Diseases (DZNE), Feodor-Lynen-Str 17, Munich,
110 81377, Germany
111 ³⁹Department of Neurology, Mayo Clinic, Scottsdale, AZ, 85259, United States of America
112 ⁴⁰Department of Clinical Genetics, Erasmus Medical Center, Wytemaweg 80, Rotterdam,
113 3015 CN, The Netherlands
114 ⁴¹Department of Psychiatry and Human Behavior, Brown Alpert Medical School, Brown
115 University, Providence, RI, 02912, United States of America
116 ⁴²Banner Alzheimer's Institute, Phoenix, AZ, 85006, United States of America
117 ⁴³MRC Prion Unit at University College London, Institute of Prion Diseases, London, United
118 Kingdom
119 ⁴⁴Department of Basic and Clinical Neuroscience, London Neurodegenerative Diseases Brain
120 Bank, Institute of Psychiatry, Psychology and Neuroscience, King's College London,
121 London, SE5 8AF, United Kingdom
122 ⁴⁵Department of Clinical Neuropathology, King's College Hospital NHS Foundation Trust,
123 London, SE5 9RS, United Kingdom

- 124 ⁴⁶Centre for Neuropathology and Prion Research, Ludwig-Maximilians-University of
125 Munich, Feodor-Lynen-Straße 23, Munich, 81377, Germany
- 126 ⁴⁷Department of Neurology, Mayo Clinic, 4500 San Pablo Road, Jacksonville, FL, 32224,
127 United States of America
- 128 ⁴⁸Department of Neurology, David Geffen School of Medicine, University of California, Los
129 Angeles, Los Angeles, CA, 90095, United States of America
- 130 ⁴⁹Department of Pathology, UCSF Weill Institute for Neurosciences, University of
131 California, San Francisco, San Francisco, CA, 94158, United States of America
- 132 ⁵⁰School of Psychology and Brain and Mind Centre, University of Sydney, Sydney, 2050,
133 Australia
- 134 ⁵¹Department of Neurology, Penn Frontotemporal Degeneration Center, Perelman School of
135 Medicine at the University of Pennsylvania, Philadelphia, PA, 19104, United States of
136 America
- 137 ⁵²University of Washington BioRepository and Integrated Neuropathology (BRaIN) lab,
138 Harborview Medical Center, 325 9th Ave, Seattle, WA, 98104, United States of America
- 139 ⁵³M.I.N.D. Institute Laboratory, University of California, Davis Medical Center, 2805 50th
140 St, Sacramento, CA, 95817, United States of America
- 141 ⁵⁴German Center for Neurodegenerative Diseases (DZNE), Rostock, 18147, Germany
- 142 ⁵⁵Department of Neurology, Rostock University Medical Center, Rostock, 18147, Germany
- 143 ⁵⁶Alzheimer's Therapeutic Research Institute, Keck School of Medicine of the University of
144 Southern California, San Diego, CA, USA
- 145 ⁵⁷Department of Psychiatry, Knight Alzheimer Disease Research Center, Washington
146 University School of Medicine, Saint Louis, MO, 63108, United States of America
- 147 ⁵⁸Department of Pathology and Laboratory Medicine, Indiana University School of Medicine,
148 635 Barnhill Drive, Indianapolis, IN, 46202, United States of America
- 149 ⁵⁹Civin Laboratory for Neuropathology, Banner Sun Health Research Institute, Sun City, AZ,
150 85351, United States of America
- 151 ⁶⁰Department of Psychiatry and Psychotherapy, University Hospital, Ludwig-Maximilians-
152 University of Munich, Munich, 81377, Germany
- 153 ⁶¹Department of Neurology, Taub Institute for Research on Alzheimer's Disease and the
154 Aging Brain, Columbia University Irving Medical Center, 630 West 168th St, New York,
155 NY, 10032, United States of America
- 156 ⁶²Department of Pathology, Taub Institute for Research on Alzheimer's Disease and the
157 Aging Brain, Columbia University Irving Medical Center, 630 West 168th St, New York,

158 NY, 10032, United States of America

159 ⁶³Department of Neurology, University of Pittsburgh, Pittsburgh, PA, 15213, United States of
160 America

161 ⁶⁴Department of Pathology, University of Pittsburgh, Pittsburgh, PA, 15213, United States of
162 America

163 ⁶⁵Department of Pathology and Laboratory Medicine and Department of Neurology, Emory
164 University, Atlanta, GA, 30322, United States of America

165 ⁶⁶Department of Neurodegenerative Disease, Dementia Research Centre, University College
166 London Queen Square Institute of Neurology, London, WC1N 3BG, United Kingdom

167 ⁶⁷Department of Pathology, Feinberg School of Medicine, Northwestern University, Chicago,
168 IL, 60611, United States of America

169 ⁶⁸Department of Clinical Neurological Sciences, Schulich School of Medicine and Dentistry,
170 University of Western Ontario, London, ON, N6A 2E2, Canada

171 ⁶⁹King's College Hospital NHS Foundation Trust, London, SE5 9RS, United Kingdom

172 ⁷⁰Department of Pathology and Laboratory Medicine, University of British Columbia,
173 Vancouver, BC, V6T 1Z7, Canada

174 ⁷¹Department of Psychiatry & Psychology, Mayo Clinic, Rochester, MN, 55905, United
175 States of America

176

177

178 **Abstract**

179 Frontotemporal lobar degeneration with neuronal inclusions of the TAR DNA-binding protein

180 43 (FTLD-TDP) is a fatal neurodegenerative disorder with only a limited number of risk loci

181 identified. We report our comprehensive genome-wide association study as part of the

182 International FTLD-TDP Whole-Genome Sequencing Consortium, including 985 cases and

183 3,153 controls, and meta-analysis with the Dementia-seq cohort, compiled from 26

184 institutions/brain banks in the United States, Europe and Australia. We confirm *UNC13A* as

185 the strongest overall FTLD-TDP risk factor and identify *TNIP1* as a novel FTLD-TDP risk

186 factor. In subgroup analyses, we further identify for the first time genome-wide significant loci

187 specific to each of the three main FTLD-TDP pathological subtypes (A, B and C), as well as
188 enrichment of risk loci in distinct tissues, brain regions, and neuronal subtypes, suggesting
189 distinct disease aetiologies in each of the subtypes. Rare variant analysis confirmed *TBKI* and
190 identified *VIPRI*, *RBPJL*, and *L3MBTL1* as novel subtype specific FTLD-TDP risk genes,
191 further highlighting the role of innate and adaptive immunity and notch signalling pathway in
192 FTLD-TDP, with potential diagnostic and novel therapeutic implications.

193

194

195 **Main**

196 **Introduction**

197 Frontotemporal lobar degeneration (FTLD) is one of the leading causes of dementia in
198 individuals younger than 65 years but can also affect individuals later in life. The predominant
199 clinical presentations of FTLD are behavior and language dysfunction resulting in behavioral
200 variant frontotemporal dementia (bvFTD)¹, semantic variant primary progressive aphasia
201 (svPPA), or nonfluent variant primary progressive aphasia (nfvPPA)². The diagnosis of FTLD
202 can be established with certainty only with neuropathologic postmortem examination and is
203 characterized neuropathologically with significant atrophy of the frontal and temporal lobes
204 and accumulation of abnormal neuronal and/or glial inclusions upon immunohistochemical
205 analysis. FTLD-TDP, characterized by neuronal and cytoplasmic aggregates of the DNA and
206 RNA-binding protein TDP-43, is one of the two main pathological subtypes (the other being
207 FTLD-Tau) and can be further classified into five FTLD-TDP subtypes (FTLD-TDP A-E)
208 based on the distribution of the neuronal cytoplasmic TDP-43-positive inclusions and
209 dystrophic neurites in the cortical layers^{3,4}. In general, an accurate prediction of the underlying
210 neuropathological FTLD subtype of individual patients constitutes a diagnostic challenge;
211 however, a few clinicopathological correlations exist. Specifically, over 80% of patients
212 clinically presenting with svPPA are diagnosed as FTLD-TDP C at autopsy⁵, and patients with
213 bvFTD with concomitant amyotrophic lateral sclerosis (ALS) almost invariably present as
214 FTLD-TDP B at autopsy^{2,4,6}.

215 A small number of autosomal dominant genes and risk factors associated with FTLD-TDP have
216 been reported⁷⁻¹³. The first FTLD-TDP genome-wide association study (GWAS) identified the
217 *TMEM106B* locus (rs1990622), supporting lysosomal dysfunction in FTLD-TDP; however,
218 this signal was strongly driven by *GRN* mutation carriers included in that study. Three

219 additional FTLD-TDP loci, *UNC13A*, *DPP6* and *HLA-DQA2*, were identified in phase I of the
220 International FTLD-TDP whole-genome sequencing (WGS) consortium and require
221 replication in larger datasets¹¹. Importantly, most FTLD-TDP patients are not yet genetically
222 explained, and the relatively small sample size precluded rare variant analyses in phase I.

223 To replicate and identify new genetic risk factors, we doubled the original sample size of the
224 FTLD-TDP WGS consortium by not only sequencing more pathologically confirmed FTLD-
225 TDP cases but also including clinically defined FTLD subtypes enriched for specific FTLD-
226 TDP pathological subtypes at autopsy. GWAS analyses of both common and rare variants,
227 followed by comprehensive gene-prioritization, enrichment analyses, and co-localization
228 studies identified novel FTLD-TDP risk loci, including novel risk genes and loci specific to
229 FTLD-TDP pathological subtypes. Our study highlights similarities and differences between
230 FTLD-TDP and other neurodegenerative diseases while unique biological processes in specific
231 tissues, brain regions, and cell types were found to characterize individual FTLD-TDP
232 pathological subtypes.

233

234 **Results**

235 **GWAS analysis**

236 **Common variant genome-wide association study**

237 To identify novel common FTLD-TDP genetic risk factors, we performed single variant
238 GWAS using an additive disease risk model for 6,568,099 common variants in 985 patients
239 and 3,153 controls free of neurodegenerative disorder that passed quality control (QC).
240 Combining all patients (FTLD-TDP All) we identified one genome-wide significant signal at
241 the *UNC13A* locus (rs8111424, OR=1.37, $P=1.17 \times 10^{-8}$). We also performed separate GWAS
242 within the FTLD-TDP A, FTLD-TDP B, and FTLD-TDP C pathological subtypes (**Fig. 1**,

243 **Tables 1-2).** The most significant locus identified in FTLD-TDP A was *GRN* (rs5848;
244 OR=1.89, $P=5.57 \times 10^{-9}$). In phase I, this locus only reached genome-wide significance under
245 an exploratory recessive model¹¹, and also now, the recessive model provided an even stronger
246 association (OR=4.12, $P=8.28 \times 10^{-15}$). We further detected 3 additional new genome-wide
247 significant loci in FTLD-TDP A: *TINAG* (rs138698596), *MZT1* (rs138959102) and *FARP2*
248 (rs886815). In FTLD-TDP B, we detected a genome-wide significant association at the
249 *UNC13A* locus (rs12973192). The lead variant rs12973192 is in linkage disequilibrium (LD)
250 with rs8111424 ($D'=1$; $r^2=0.43$) identified in the FTLD-TDP All analysis. We further detected
251 3 new genome-wide significant risk loci in FTLD-TDP B: *TNIP1* (rs871269), *RCL1*
252 (rs7674221), *PDS5B* (rs5277499), and one in FTLD-TDP C *C19orf52* (also known as
253 *TIMM29*, rs576561313).

254 In order to prioritize risk genes and identify possible biological mechanisms, we applied a range
255 of variant annotation and molecular quantitative trait loci (QTL)-GWAS integration analyses
256 as previously described¹⁴ (**Supplementary Tables 2-13**). We integrated different levels of
257 evidence using a weighting scheme and obtained a weighted sum of the hits in different
258 subcategories for each gene. We grouped candidate risk genes in genome-wide significant loci
259 and in subthreshold loci and prioritized them at two levels of confidence for being a likely risk
260 gene as tier 1 (higher confidence) and tier 2 (lower confidence).

261 The gene prioritization analyses nominated a total of 59 tier 1 and 246 tier 2 genes in 258
262 different loci for 4 different GWAS analyses (**Fig. 2, Supplementary Table 3,**
263 **Supplementary Fig. 2-5**). Our results showed that the nearest protein-coding genes were
264 prioritized as tier 1 ($n=8$) and tier 2 ($n=1$) risk genes in the genome-wide significant loci for
265 the distinct FTLD-TDP subtypes. Of the 8 tier 1 prioritized genes, 3 were found in common
266 variant loci where molecular QTL-GWAS analyses aided their prioritization (**Fig. 2**). First, in
267 locus A4, *GRN* was prioritized through consistent expression QTL (eQTL) domain hits in bulk

268 brain regions (eQTL $P_{\text{ROSMAP DLPFC}}=6.32 \times 10^{-38}$ and $\text{beta}_{\text{ROSMAP DLPFC}}=-0.25$, eQTL
269 colocalization (coloc) PPH4s of 81.8%-99.7%, and fine-mapped [posterior inclusion
270 probability being 100%] expression transcriptome wide association study [eTWAS]
271 associations with P from 1.74×10^{-8} to 5.15×10^{-9} and Z-scores from -5.63 to -5.84) and in
272 oligodendrocytes (cell type specific eQTL (ct-eQTL) coloc PPH4=90%) where genetic
273 downregulation of *GRN* gene expression was associated with the FTLD-TDP A risk signal
274 (**Supplementary Tables 5, 6, and 11**), which was also observed in brain proteome-wide
275 association study (PWAS) with the same effect direction for the FTLD-TDP A risk (P_{ROSMAP}
276 $\text{DLPFC} = 3.32 \times 10^{-6}$, $Z_{\text{ROSMAP DLPFC}} = -4.65$, **Supplementary Table 13**). Second, in locus B1,
277 *TNIP1* was prioritized because the minor allele was associated with decreased *TNIP1*
278 expression ($P_{\text{ROSMAP DLPFC}} = 2.40 \times 10^{-4}$, $\text{beta}_{\text{ROSMAP DLPFC}} = -0.10$), the GWAS signal colocalized
279 with a microglia splicing QTL (sQTL) associated with *TNIP1* chr5:151032383-151035002
280 known splice junction (coloc PPH4=82.2%) and because the methylation QTL (mQTL)
281 variants for cg03340667, a CpG ~3.7 kb upstream of the transcription start site (TSS) of the
282 canonical transcript of *TNIP1*, colocalized with the GWAS variants (coloc PPH4=70%) in
283 dorsolateral prefrontal cortex (DLPFC) (**Supplementary Tables 5, 7, and 8**). Third, in locus
284 B4, *UNC13A* was prioritized through an eQTL-GWAS colocalization in temporal cortex (coloc
285 PPH4=81.82%, **Supplementary Table 6**). Furthermore, beyond genome-wide significant loci,
286 we identified additional candidate prioritized risk genes in subthreshold regions through
287 molecular QTL-GWAS coloc and TWAS analyses, one important example being *TMEM106B*
288 as the prioritized risk gene in locus A_S14. The FTLD-TDP A GWAS signal near *TMEM106B*
289 colocalized with eQTL variants regulating *TMEM106B* gene expression in bulk brain regions
290 (eQTL coloc PPH4s=81.40% in MayoRNASeq temporal cortex and 89.66% in GTEx brain
291 cortex, **Supplementary Table 6**). We also observed a significant eTWAS association in GTEx
292 cortex ($P=4.13 \times 10^{-7}$, $Z=-5.06$), together with a significant PWAS hit ($P=2.01 \times 10^{-8}$,

293 Zscore=5.61) (**Supplementary Tables 11, 13**). Finally, a significant hit in splicing TWAS in
294 cortex (sTWAS, $P=6.66 \times 10^{-7}$, $Z=-4.97$) predicted a decreased preference for the *TMEM106B*
295 splice junction chr7:12224385-12229679 with the increased FTLD-TDP A GWAS risk, while
296 we also observed methylation QTL (mQTL) coloc hits for two CpGs for *TMEM106B* (~500 bp
297 upstream cg23422036 coloc PPH4=94.25% and intronic cg09613507 coloc PPH4=94.09%)
298 (**Supplementary Tables 12, 8**). We summarized our gene prioritization results in **Fig. 2** for
299 the genome-wide significant loci and a selection of the suggestive loci (with genes having
300 GWAS evidence of $P \leq 5 \times 10^{-6}$ in 1 Mb extended regions), and full results are presented in
301 **Supplementary Tables 3-13 and Supplementary Fig. 2-5**.

302 Next, we performed gene ontology analyses on tier1 prioritized genes. The most significant
303 term in the nominated genes in FTLD-TDP All was positive regulation of defense response to
304 bacterium ($P=3.21 \times 10^{-5}$). Lysosomal function appeared to be strongly affected in FTLD-TDP
305 A with several genes such as *GRN* and *TMEM106B* (lysosomal organization GO term,
306 $P=4.12 \times 10^{-4}$) as well as cathepsin B (*CTSB*). We further detected enriched terms for retrograde
307 transport in FTLD-TDP B ($P=2.21 \times 10^{-3}$) driven by *DENND2A* and *VPS53* genes and for
308 excitatory postsynaptic potential in FTLD-TDP C ($p=1.48 \times 10^{-3}$) driven by *DMPK* and *P2RX5*
309 genes (**Fig. 3, Supplementary Table 14**). Importantly, except for lysosomal transport, no
310 terms overlapped between subtypes of FTLD-TDP, suggesting mostly distinct genetic
311 etiologies in the different FTLD-TDP groups.

312 To further characterize genetic factors associated with FTLD-TDP, we performed gene-based
313 analyses on common variants with $P < 10^{-5}$ using MAGMA. Analyses of FTLD-TDP All did not
314 yield exome-wide significant loci; however, FTLD-TDP A showed exome-wide significant
315 signals for the two genes located at the *GRN* locus (*FAM171A2*, *ITGA2B*) and for *TMEM106B*
316 ($P=4.74 \times 10^{-7}$). The *TMEM106B* signal was driven by the rs10281425 variant (OR=0.54,

317 $P=2.12 \times 10^{-7}$). No exome-wide significant signal was detected for the other two FTLD-TDP
318 pathological subtypes.

319 **Expression of the top suggestive signals in cell types and brain regions**

320 To find tissues and cell types for which gene expression profiles were enriched for genes within
321 FTLD-risk loci, we combined gene-based association statistics calculated using MAGMA with
322 gene expression patterns from the Genotype–Tissue Expression (GTEx) project in a gene set
323 enrichment analysis. We observed an enrichment in genes expressed in brain tissue
324 (cerebellum, frontal cortex, and cortex) in FTLD-TDP A and B. This was strikingly different
325 from the signature observed in FTLD-TDP C for which significant enrichment was only
326 detected in non-central nervous system tissue, in particular small intestine terminal ileum (**Fig.**
327 **4, Supplementary Table 15**). We also compared FTLD gene expression loci with similar data
328 obtained from Alzheimer’s disease and related disorders (ADRD)¹⁴ and ALS GWAS¹⁵. FTLD-
329 TDP subtypes presented with a distinct genetic signature as compared to these related disorders
330 highlighting the importance of regional specificity in FTLD-TDP.

331 We subsequently queried PsychENCODE frontal-cortex single-cell RNA-seq datasets of
332 human-derived brain samples to specify further which brain-specific enriched cell types
333 express the genetic loci associated with FTLD-TDP risk (**Fig. 4, Supplementary Table 16**).
334 We observed a significant enrichment in genes expressed in excitatory neurons for FTLD-TDP
335 A loci (Ex4 $P=3.55 \times 10^{-2}$, Ex5b $P=2.72 \times 10^{-2}$), and FTLD-TDP B loci (Ex8 $P=1.27 \times 10^{-4}$), while
336 no other cell type reached significance. While FTLD-TDP C loci were also significantly
337 enriched in genes expressed in excitatory neurons (Ex3e $P=2.10 \times 10^{-2}$), they were additionally
338 enriched in genes expressed in astrocytes and oligodendrocyte progenitor cells ($P=4.69 \times 10^{-2}$,
339 $P=2.53 \times 10^{-2}$). Genes expressed in microglia were enriched only in ADRD gene loci
340 ($P=1.90 \times 10^{-2}$). Overall, loci comprising genes expressed in excitatory neurons were enriched

341 in the three FTLD-TDP subtypes with stronger specificity for specific neuron types in each
342 FTLD-TDP subtype as compared to what was observed for ALS gene loci.

343 ***UNC13A* and *TNIP1* loci**

344 To provide further support for the identified FTLD-TDP risk loci, we performed a meta-
345 analysis of our FTLD-TDP cohort with the Dementia-seq study (phs001963.v1.p1) that
346 includes 2,102 clinical FTLD patients and 1,748 controls. Given that this cohort lacks details
347 on the FTLD pathology underlying each patient, pathological subgroup analyses could not be
348 performed. Meta-analysis confirmed *UNC13A* and identified the new *TNIP1* locus as genome-
349 wide significantly associated with FTLD ($P_{rs12973192}=8.85\times 10^{-10}$; $P_{rs871269}=3.42\times 10^{-8}$,
350 respectively). Note that the most significant single nucleotide variant (SNV) at the *UNC13A*
351 locus was rs12608932 ($p=9.13\times 10^{-11}$), in strong LD with rs12973192 ($r^2=0.96$ $D'=0.99$).

352 Both *UNC13A* and *TNIP1* were previously associated with other neurodegenerative
353 diseases^{14,16}. Colocalization analyses showed that our *UNC13A* signal was shared with ALS
354 (coloc PPH4=95.71%), strongly confirming the genetic overlap between both diseases (**Fig.**
355 **5A**). Interestingly, for *TNIP1* we found strong colocalization with the ADRD association
356 signal^{14,16} (coloc PPH4=99.2%) while its colocalization with ALS was weaker (71.5%), which
357 was confirmed in a specificity analysis (coloc PPH4=20.1%, for $p12=10e-06$), possibly
358 reflecting multiple independent association signals in FTLD in this locus (**Fig. 5B-C**).

359 **Rare variant analysis**

360 To identify genes carrying rare variants contributing to FTLD-TDP, we performed a burden
361 test in genes with variants likely to affect protein function. In the overall FTLD-TDP cohort,
362 no exome wide significant gene was detected (**Supplementary Table 17**). We did detect five
363 exome-wide significant signals within FTLD-TDP pathological subtypes (**Table 3**,
364 **Supplementary Table 18**). *TBKI* was associated with disease status in FTLD-TDP A and B
365 ($P=1.27\times 10^{-11}$, $P=3.17\times 10^{-12}$, respectively). The signal was driven by 3 carriers in FTLD-TDP

366 A patients (3/193=1.5%) and 5 carriers in FTL D-TDP B patients (5/288=1.7%). We also
367 detected an enrichment in rare variants in *VIPRI* in FTL D-TDP B ($P=4.65e-07$, 3/288 FTL D-
368 TDP B and 1/3153 control; **Fig. 6**) and 2 exome wide significant signals in FTL D-TDP C
369 *L3MBTL1* ($P=2.87 \times 10^{-7}$, 8/467 FTL D-TDP C and 3/3153 controls) and *RBPJL* ($P=6.39 \times 10^{-7}$,
370 5/467 FTL D-TDP C and 3/3153 controls). Weighted gene coexpression network analysis using
371 the ROSMAP dataset and the BrainExp database¹⁷ revealed that *L3MBTL1* and *RBPJL*
372 belonged to the same module (yellow, $P_{L3MBTL1}=1.32 \times 10^{-45}$, $P_{RBPJL}=1.00 \times 10^{-79}$;
373 **Supplementary Figure 6**) that is enriched in neuroactive ligand-receptor interaction and the
374 cytokine-cytokine receptor interaction gene-ontology terms ($P_{FDR}=3.7 \times 10^{-12}$, $P_{FDR}=5.8 \times 10^{-12}$,
375 respectively). While expression of *L3MBTL1* was throughout the central nervous system cells,
376 *RBPJL* expression was restricted to inhibitory neurons and in particular to Parvalbumin
377 neurons (**Supplementary Figure 6**).

378 Discussion

379 In this work, we report 8 new genome-wide significant FTL D-TDP risk loci and 3 new genes
380 harboring rare variants contributing to FTL D-TDP risk, by performing the largest FTL D-TDP
381 WGS study to date, including 985 patients and 3,153 controls. A comprehensive analysis of
382 our data highlights the genetic overlap between FTL D-TDP, ADRD, and ALS while also
383 defining tissue and cell type enrichment unique to FTL D-TDP. Most importantly, we highlight
384 distinct genetic aetiologies for each of the three main FTL D-TDP pathological subtypes (A, B
385 and C) suggesting that multiple distinct pathomechanisms underlie the TDP-43 dysfunction
386 and deposition in FTL D-TDP (**Fig. 7**).

387 We confirm and replicate for the first time our previously reported GWAS signal at the
388 *UNC13A* locus in FTL D-TDP patients¹¹. The *UNC13A* risk haplotype tagged by rs12973192
389 and rs12608932 was previously shown to increase cryptic splicing of *UNC13A* in brain tissue
390 by modulating TDP-43 binding^{18,19}. The cryptic splicing leads to transcripts with premature

391 stop codons and the subsequent loss of UNC13A protein, significantly impacting the release
392 of vesicles in glutamatergic synapses²⁰. Since *UNC13A* thus represents a shared risk factor
393 between ALS and FTLD-TDP additional genetic or environmental mechanisms likely drive
394 disease presentation.

395 We further establish and replicate a novel genetic association between the *TNIP1* locus and
396 FTLD-TDP. Recently, Restuadi *et al.* deeply characterized the *GPX3/TNIP1* locus associated
397 with ALS and suggested that *GPX3* should be prioritized for deeper exploration into disease
398 mechanisms related to this region²¹. *GPX3*, encoding for glutathione peroxidase 3, is a secreted
399 enzyme involved in the regulation of oxidative damage, and its levels were found to be reduced
400 in ALS sera²². Interestingly, however, the risk variant associated with FTLD-TDP (rs871269)
401 is an expression quantitative trait locus for *TNIP1* in the dorsolateral prefrontal cortex, and
402 along with the fact that we only observed a weak colocalization signal with the ALS locus, we
403 highlight *TNIP1* as the most likely gene candidate for FTLD-TDP. In fact, we observed a
404 shared signal at this locus between our FTLD-TDP GWAS and the recent large ADRD
405 GWAS¹⁴, suggesting genetic and/or clinical overlap between AD and FTLD-TDP. *TNIP1* is a
406 ubiquitin-binding adaptor protein that regulates cell death and innate immune response through
407 NF- κ B activation²³⁻²⁵. *TNIP1* undergoes phosphorylation by TBK1 and interacts with OPTN²⁶,
408 two proteins associated with FTLD-TDP etiology^{12,27}. While this functional connection further
409 supports *TNIP1* as FTLD-TDP risk gene, more work is needed to understand the mechanisms
410 underlying disease onset. Overall, we substantiate the genetic overlap between ALS, ADRD,
411 and FTLD-TDP and emphasize the need for deeper exploration into pathways underlying
412 disease-specific risk.

413 One of the most striking conclusions from this phase II FTLD-TDP GWAS is the distinct
414 association signals among FTLD-TDP pathological subtypes. Even the *UNC13A* and *TNIP1*
415 risk loci, which reach genome-wide significance in the meta-analysis stage, show stronger

416 association in FTLD-TDP B alone, and for the first time, genome-wide significant common
417 risk loci are reported for each of the individual pathological FTLD-TDP subtypes.

418 In FTLD-TDP A, in addition to individual genome-wide significant common variants assigned
419 to *GRN*, *TINAG*, *MZT1*, and *FARP2* risk loci, we identified exome-wide significant association
420 with the burden of common variants in *GRN* and *TMEM106B*, in addition to multiple QTL-
421 based analyses prioritizing *TMEM106B* as a tier 1 risk gene, re-enforcing the specific
422 connection of these genes with FTLD-TDP A, even in patients without loss-of-function *GRN*
423 mutations^{11,28}. While *GRN* and *TMEM106B* are also reported as AD risk genes¹⁴, an even
424 stronger connection exists between these genes and limbic-predominant age-related TDP-43
425 encephalopathy (LATE)^{29,30}, which has a more restricted neuroanatomical distribution of TDP-
426 43 pathology as compared to FTLD-TDP but with some characteristics of FTLD-TDP type
427 A^{31,32}. The *TMEM106B* signal is primarily influenced by rs10281425, a variant located in the
428 3'UTR of *TMEM106B*, which tags the previously reported *TMEM106B* risk haplotype³³
429 associated with an increase in *TMEM106B* mRNA expression³³ and a higher burden of
430 insoluble disease-associated *TMEM106B* C-terminal fragments³⁴. More broadly, also
431 including prioritized genes from the subthreshold regions, gene ontology analysis in FTLD-
432 TDP A revealed enrichment in genes implicated in lysosomal function driven by *GRN*,
433 *TMEM106B* but also *CSTB*, three genes which also had the highest individual gene scores in
434 the prioritization analysis in FTLD-TDP A. *CSTB* encodes one of the most abundant lysosomal
435 proteases in the brain³⁵, and has been reported as a progranulin protease^{36,37}. Genes involved
436 in lysosomal dysfunction were also overrepresented in FTLD-TDP B, including *GRN* and
437 *PPT1*. *PPT1*, is a lysosomal enzyme which facilitates the degradation of fatty-acylated proteins
438 by lysosomal hydrolases. Mutations in *PPT1* cause neuronal ceroid lipofuscinosis 1^{38,39} and
439 *Ppt1* knock-out mice displayed fewer lipid droplets (LD) than wild type, indicating impairment
440 of lipophagy, previously associated with FTLD/ALS⁴⁰⁻⁴³. Overall, our genetic data provides

441 compelling evidence that lysosomal dysfunction contributes to the pathobiology of FTLD-TDP
442 A, and, to a lesser extent FTLD-TDP B.

443 For FTLD-TDP B, additionally, we identified individual genome-wide significant associations
444 with variants in the *RCL1* and *PDS5* loci, and we observed enrichment for gene ontology terms
445 related to retrograde transport resulting from the *VPS53* and *DENND2A* loci. *VPS53* is part of
446 the Golgi-associated retrograde protein (GARP) complex^{44,45} involved in intracellular
447 cholesterol transport by targeting NPC2 to lysosomes⁴⁶. Recently, laser capture
448 microdissection and single-cell mass spectrometry-based proteomics in motor neurons of ALS
449 patients revealed a strong reduction in endolysosomal trafficking complexes such as the GARP
450 complexes⁴⁷. Limited information about DENND2A function is currently available, but
451 structural and functional analysis indicates it may be involved in intracellular vesicle
452 trafficking to the lysosome and to the Golgi through its guanine nucleotide exchange factor
453 activity and regulation of RAB family GTPases⁴⁸. However, retrograde transport has been
454 previously implicated in ALS with, for instance, mutations in *DCTN1*^{49,50} and *KIF5A*^{51,52},
455 highlighting functional connections of prioritized genes from the subthreshold loci with TDP-
456 43 dysfunction and ALS. Future GWAS with larger sample sizes, potentially combining
457 FTLD-TDP B and ALS, are required to firmly establish a genetic contribution of this pathway
458 to disease.

459 Focusing on rare variants, exome-wide significant association with *TBK1* was observed in both
460 FTLD-TDP A and B (but not FTLD-TDP C), confirming *TBK1* mutations as the most common
461 cause of FTLD-TDP after *GRN* and *C9orf72*¹². We further unveiled rare predicted pathogenic
462 variants associated with FTLD-TDP B within *VIPRI*, which encodes for the vasoactive
463 intestinal peptide receptor 1. The variants are predicted to lead to an alteration of VIPR1
464 function, impairing the vasoactive intestinal peptide (VIP) biological pathway. Indeed, VIPR1
465 is activated upon binding by VIP, which exerts a neuroprotective effect mainly through glial

466 cells^{53,54} even though neurons also express VIPRs^{55,56}. Notably, VIP is also a key regulator of
467 innate and adaptive immunity⁵⁷, making it an important therapeutic target for multiple
468 neurodegenerative diseases. Altogether, our studies suggest that lysosome dysfunction and/or
469 alterations in the innate and adaptive immune system are important contributors to both FTLD-
470 TDP A and B risk, yet to varying degrees in each pathological subtype and with likely
471 important variability in the contribution from each pathway among individual patients.

472 FTLD-TDP C was previously recognized as a clinicopathological entity distinct from FTLD-
473 TDP A and B⁵⁸, and our genetic studies support this notion showing no overlap in common or
474 rare risk genes with the other FTLD-TDP types. Importantly, however, while often considered
475 a sporadic FTLD subtype^{13,59,60}, we implicate several genes and risk loci in FTLD-TDP C and
476 uncover a potential role for mitochondrial membrane dysfunction and the notch signaling
477 pathway. *C19orf52* (*TIMM29*), which mediates the import and insertion of multi-pass
478 transmembrane proteins into the mitochondrial inner membrane, was identified as the first
479 genome-wide significant risk locus for FTLD-TDP C. Rare-variant burden analyses further
480 associated rare predicted pathogenic variants in the *RBPJL* gene with FTLD-TDP C. *RBPJL*
481 encodes for the recombination signal binding protein for immunoglobulin kappa J region like
482 transcription factor. *RBPJL* is able to repress Notch target gene expression (*Hey1*, *Hey2*, *HeyL*
483 and *Notch3*)⁶¹. As such, our findings align with a previous analysis of sub-genome-wide
484 significant genes in clinical svPPA patients, enriched in FTLD-TDP C, which highlighted an
485 overrepresentation of the Notch pathway⁶². Interestingly *RBPJL* and *L3MBTL1*, the second
486 gene carrying rare predicted pathogenic variants in FTLD-TDP C, are part of the same co-
487 expression module, suggesting that they are functionally related. Moreover, *L3MBTL1*, a
488 histone methyl-lysine binding protein, is a key regulator of proteotoxicity associated with
489 *C9orf72* dipeptide repeats and mutant *SOD1*⁶³ and was found to be increased in spinal cord of
490 ALS patients. Furthermore, reduction of *L3MBTL1* expression in drosophila models with the

491 C9orf72-associated dipeptides poly(PR) or poly(GR) ameliorated the rough-eye phenotype⁶³,
492 suggesting that loss of L3MBTL1 expression is beneficial. While no RNA samples were
493 accessible from rare variant carriers, nonsense-mediated decay escape has been reported in
494 other genes linked to ALS⁶⁴. It is thus possible that the *L3MBTL1* variants lead to the generation
495 of truncated proteins with toxic gain-of-function, but additional work is necessary to
496 understand the disease etiology fully.

497 When analyzed in sum, common variants associated with the different FTLD-TDP pathological
498 subtypes appeared to be located in genes expressed in excitatory neurons in contrast to AD risk
499 variants, which are enriched in microglia. Interestingly, glutamatergic transmission impairment
500 has been reported in FTLD⁶⁵⁻⁶⁹ and voxel-based brain changes have been significantly
501 associated with spatial distribution of mGluR5 in symptomatic *C9orf72* and *GRN* carriers⁷⁰.
502 Therefore, and in line with previously reported studies, our data suggest that neurons are the
503 major players in disease etiology, as compared to what has been observed in ADRD.
504 Interestingly, the distribution of risk loci was specific to the cerebellar hemisphere and the
505 frontal-cortex for both FTLD-TDP A and B, as opposed to FTLD-TDP C where genes
506 expressed in small intestine were enriched in risk loci. While the link between gut microbiome
507 and FTLD remains limited⁷¹, our data suggest that the gut-brain axis might be of interest for
508 future studies. In fact, emerging evidence also supports a role for the gut-brain axis in
509 autoimmune diseases⁷², a group of disorders that were found to be enriched in svPPA patients⁷³.

510 In prior studies, besides *UNC13A*, common variants in the *HLA* and *DPP6* loci were reported
511 to be associated with FTLD^{11,13}. The *HLA-DR5* locus was identified as associated with clinical
512 FTLD but was not replicated in phase I of our International FTLD-TDP WGS Consortium.
513 *HLA-DQA2* and *DPP6* loci were reported as overall FTLD-TDP risk loci in phase I¹¹ but were
514 not replicated in the current study. The relative composition of patients with FTLD-TDP
515 pathological subtypes in phase I and II (e.g., less FTLD-TDP A in phase II) and inclusion of

516 clinically diagnosed individuals in phase II may have contributed to this; however, it is also
517 possible that the increase in sample size reduced type I errors from phase I. Importantly, we
518 did identify and replicate in two independent cohorts the *UNC13A* and *TNIP1* loci associated
519 with FTLD-TDP. Replication of the newly identified risk loci, each specific to distinct
520 neuropathological FTLD-TDP subtypes, will require additional GWAS studies in the future.
521 Obtaining sufficient samples will, however, be challenging, especially for FTLD-TDP A,
522 which lacks a clear clinical correlate of the pathological phenotype. Functional characterization
523 of the newly identified genes and loci may also provide mechanistic insight.

524 In conclusion, we confirmed *UNC13A* and identified 8 new genetic loci, i.e. *TNIP1*, *GRN*,
525 *TINAG*, *MZT1*, *FARP2*, *RCL1*, *PDS5B*, and *C19orf52*, and 3 new genes with rare variants
526 associated with FTLD-TDP risk, i.e. *VIPR1*, *RBPJL*, and *L3MBTL1*. By enriching in
527 neuropathologically confirmed patients and substantially increasing our cohort size, we gained
528 important knowledge in our understanding of FTLD-TDP pathophysiology. The recognition
529 that individual FTLD-TDP subtypes could in fact be considered separate diseases with distinct
530 pathomechanisms has significant implications for the design of clinical trials and therapeutic
531 interventions.

532

533 **Methods**

534 **Samples**

535 Our current dataset includes previously generated data through the International FTLD-TDP
536 WGS consortium phase I¹¹ with 554 persons with clinicopathologically defined FTLD-TDP
537 and newly generated phase II sequencing data from 32 FTLD-TDP A, 43 FTLD-TDP B, 66
538 FTLD-TDP C, 4 FTLD-TDP E, and 9 with unclassifiable FTLD-TDP pathology (abbreviated
539 as FTLD-TDP U, **Supplementary Table 19**). To increase statistical power, we also sequenced
540 70 persons with clinical diagnosis of bvFTD/ALS, a clinical subtype associated with FTLD-

541 TDP B, and 283 persons with svPPA, a clinical subtype associated with FTLN-TDP C. Overall,
542 the total cohort pre-quality control was a combined FTLN-TDP cohort of 202 FTLN-TDP A,
543 237 FTLN-TDP B, 225 FTLN-TDP C, 4 FTLN-TDP D, 11 FTLN-TDP E, 29 FTLN-TDP U
544 persons, 70 persons with bvFTD/ALS and 283 persons with svPPA. All persons clinically or
545 pathologically diagnosed with FTLN are referred to as patients throughout the manuscript. We
546 further used WGS data from 982 participants from the Mayo Clinic Biobank (from phase I)^{11,74},
547 322 new controls free of neurodegenerative disorder from Mayo Clinic with WGS available,
548 and 2,037 controls derived from the Alzheimer's disease sequencing project (ADSP). *C9orf72*
549 repeat expansions were assessed in all patients using our previously reported two-step protocol
550 and Sanger sequencing was used to perform mutation analyses of *GRN*^{7,9}. Patients with
551 *C9orf72* repeat expansion and LOF mutations in *GRN* were removed prior to WGS. Study
552 protocols were reviewed and approved by the appropriate institutional review boards.

553 **Whole genome sequencing**

554 In phase I of the International FTD-TDP WGS consortium, WGS was generated on 554 patients
555 with FTLN-TDP (512 passed QC in that study)¹¹. Briefly, whole blood- or brain-derived DNA
556 from 499 unrelated FTLN-TDP patients and 982 participants from the Mayo Clinic Biobank
557 Study were sequenced at HudsonAlpha using the standard library preparation protocol using
558 NEBNext® DNA Library Prep Master Mix Set for Illumina® (New England BioLabs Inc.,
559 Ipswich, MA, USA), Concentration of the libraries was assessed by Qubit® 2.0 Fluorometer,
560 and the quality of the libraries was estimated by a DNA 5 K chip on a Caliper GX. Accurate
561 quantification was determined using the qPCR-based KAPA Biosystems Library
562 Quantification kit (Kapa Biosystems, Inc., Woburn, MA, USA). Each sample was sequenced
563 on one lane of Illumina's HiSeq X instrument using v2 flow cells and reagents to target 30×
564 genomic coverage. Fastq files previously generated on an Illumina HiSeq X for 55 FTLN-TDP
565 patients were obtained from 3 sites: UCSF (n = 36), DZNE (n = 14) and NSW (n = 5). In phase

566 II, additional WGS of 507 patients with FTLD-TDP, svPPA, bvFTD/ALS and 322 controls
567 free of neurodegenerative disorders was performed at USUHS sequencing center or Mayo
568 Clinic Rochester using the TruSeq DNA PCR-Free Library preparation Kit (Illumina) followed
569 by Whole Genome Sequencing by synthesis (SBS) chemistry on HiSeq X Illumina platform
570 using the HiSeq X Ten Reag. kit v2.5. For all patients and controls, fastq files were transferred
571 to Mayo Clinic and processed through the Mayo Genome GPS v4.0 pipeline in batches of up
572 to 75 samples. Briefly, reads were mapped to the human reference sequence (GRCh38 build)
573 using the Burrows–Wheeler Aligner, and local realignment around indels was performed using
574 the Genome Analysis Toolkit (GATK). Variant calling was performed using GATK
575 HaplotypeCaller followed by variant recalibration (VQSR) according to the GATK best
576 practice recommendations. At the time of analysis, participants from the Mayo Clinic Biobank
577 with possible clinical diagnosis or family history of a neurodegenerative disorder were
578 removed. We also included genotypic data (vcf) from 2,037 controls from the Alzheimer’s
579 Disease Sequencing Project, leading to a total of 1,061 patients and 3,341 controls.

580 **Sample level quality control and definition of subgroups**

581 Samples with less than 30× coverage in more than 50% of the genome, call rate below 85%,
582 sex error, contamination defined by a FREEMIX score above 4 or non-Caucasian ethnicity
583 were removed. At this step, joint genotyping on all samples was performed, a final relatedness
584 measurement was calculated using PREST⁷⁵, and duplicates were removed, while only one
585 individual per family was kept. In total, 985 pathologically confirmed FTLD-TDP or presumed
586 FTLD-TDP patients clinically presenting with svPPA or bvFTD/ALS, as well as 3,153
587 neurologically normal controls passed all quality control measures (**Supplementary Table 1**).
588 Age at onset of svPPA and bvFTD/ALS did not differ from the age at onset of FTLD-TDP C
589 ($P=1$) and FTLD-TDP B patients ($P=1$), respectively. Based on these findings and the
590 previously established associations between the svPPA and bvFTD/ALS clinical diagnoses

591 with specific FTLD-TDP pathological subtypes, we combined svPPA with FTLD-TDP C and
592 bvFTD/ALS with FTLD-TDP B patients in all analyses. Within our overall cohort of 193
593 FTLD-TDP A, 288 FTLD-TDP B (defined as FTLD-TDP B and bvFTD/ALS) and 467 FTLD-
594 TDP C (defined as FTLD-TDP C and svPPA), the ages at onset and death differed significantly
595 between the pathological FTLD-TDP subtypes (**Table 2; Supplementary Figure 1**). FTLD-
596 TDP A patients had a later age at onset than FTLD-TDP B and FTLD-TDP C groups
597 ($P=4.73 \times 10^{-8}$, $P=1.37 \times 10^{-13}$, respectively), and a later age at death ($P=4.00 \times 10^{-15}$, $P=3.00 \times 10^{-6}$,
598 respectively). FTLD-TDP B had an earlier age at death as compared to FTLD-TDP C
599 ($P=5.60 \times 10^{-8}$). Differences in age distribution in between patient groups were assessed using
600 the Kruskal-Wallis test followed by Wilcoxon test correcting for multiple testing. Corrected P
601 are provided.

602 **Variant level quality control**

603 Genotype calls with genotype quality (GQ) < 20 and/or depth (DP) < 10 were set to missing,
604 and variants with edit-distance > 4 and call rate < 80% were removed from all subsequent
605 analyses leading to a total of 85,345,466 variants. For all analyses, only variants that pass
606 VQSR (127,658 variants removed) and with a call rate > 95% in cases or controls were
607 considered (591,431 variants removed). Functional annotation of variants was performed using
608 ANNOVAR (version2016Feb01). Rare loss of function variants (frameshift
609 insertion/deletion/block substitution, stopgain, stoploss and splicing single nucleotide
610 variants—SNVs) identified in exome-wide significant associated genes in the rare variant
611 analyses (**Supplementary Table 17**) were confirmed in patients by Sanger sequencing
612 (primers available upon request). For the known neurodegenerative disease genes (*GRN*,
613 *MAPT*, *TBK1*, *OPTN*, *VCP*, *TARDBP*, *CHCHD10*, *SQSTM1*, *UBQLN2*, *hnRNPA1*,
614 *hnRNPA2B1*, *CSF1R*, *FUS*, *CHMP2B*, and *LRRK2*), potentially pathogenic rare variants were
615 also identified and confirmed by Sanger sequencing (n=25; **Supplementary Table 20**).

616 **Generation of principal components**

617 Prior to running genetic association analyses, principal component (PC) analysis was
618 performed using a subset of variants meeting the following criteria: minor allele frequency
619 (MAF)>5% and full sample Hardy–Weinberg Equilibrium (HWE) $P>1\times 10^{-5}$. Influential
620 regions such as the *HLA* region were removed, and variants were pruned by linkage
621 disequilibrium with r^2 threshold of 0.1 prior to PC analysis. This analysis identified 13 PCs that
622 were significantly associated with patient/control status, which were subsequently used as
623 covariates in all genetic association analyses.

624 **Variant-level analysis of common variants**

625 For the common variant GWAS, single nucleotide variants (SNV) with MAF>0.01 in patients
626 or controls (n=7,178,250 variants) and HWE $P>1.00\times 10^{-6}$ in controls were analyzed (17,450
627 variants removed). In addition, since whole genome sequencing of FTLN-TDP patients and
628 controls was performed at multiple sites, a test was performed to identify variants with
629 significant differences in genotype distributions between sequencing batches, and 592,701
630 SNVs showing evidence of batch effects ($p<0.05$) were removed leading to a total of
631 6,568,099 variants analyzed.

632 For all remaining variants, association of genotypes with the patient/control status was assessed
633 using logistic regression with allele dosage as the predictor assuming log-additive allele effects.
634 Sex and the first thirteen PCs were included as covariates in the models. The SNV-level
635 analyses were performed using PLINKv.00a23LM2 combining all FTLN-TDP cases (FTLN-
636 TDP All) and in FTLN-TDP pathological subtypes. Meta-analyses of FTLN-TDP phase II with
637 publicly available dataset from the Dementia-seq project (phs001963.v2.p1) was performed
638 under a fixed-effects model comparing our data with 2,102 FTLN cases and 1,748 controls
639 from the Dementia-seq project using Metal⁷⁶. Dementia-seq vcf were processed the exact same

640 way as our data except that 10 PCs were included in the model to perform common variant
641 association analysis.

642 **Colocalization analyses**

643 We performed colocalization analysis for *UNC13A* and *TNIP1* loci (top SNVs ± 100 kb) with
644 ALS (GCST90027164) and ADRD (GCST90027158) using the ‘coloc’ package version 4.0.4
645 in R using our meta-analyses data. When the summary statistics of the other trait was expressed
646 on another build than GRCh38, the variant alleles and positions were converted. We set the
647 prior probabilities to $\pi_1 = 1 \times 10^{-4}$, $\pi_2 = 1 \times 10^{-4}$ and $\pi_{12} = 1 \times 10^{-5}$ for a causal variant in trait 1
648 or trait 2 and a shared causal variant between traits 1 and 2, respectively (default parameters).
649 Sensitivity analysis was performed at $\pi_{12} = 1 \times 10^{-6}$. $P < 0.05$ was considered statistically
650 significant.

651 **Tissue and cell type enrichment analysis**

652 Tissue and cell type enrichment analyses were performed using the summary statistics and
653 FUMA. Briefly, FUMA aggregates summary statistics per gene to calculate gene-wise
654 association signals using MAGMA version 1.6 and subsequently tests whether tissues and cell
655 types are enriched for expression of these genes. For tissue enrichment analysis, we used the
656 GTEx version 8 reference set. $P < 0.05$ across all tissues ($n = 54$) were considered statistically
657 significant. For cell type enrichment analyses, we used human-derived single-cell RNA-seq
658 data from major brain cell types (PsychENCODE). Excitatory and inhibitory neurons from the
659 PsychENCODE dataset were labeled based on their transcriptional profile from 1 to 8⁷⁷.
660 $P < 0.05$ were considered statistically significant.

661 **Gene prioritization and functional interpretation of GWAS**

662 We performed the gene prioritization and functional interpretation analyses for FTLT-TDP All
663 and each FTLT-TDP pathological subtype separately by using the subtype-specific GWAS
664 summary statistics. We adapted a systematic gene prioritization and functional interpretation

665 strategy (as previously described in Bellenguez et al.¹⁴) to prioritize GWAS-implicated
666 candidate risk genes and nominate possible downstream biological mechanisms. Briefly, six
667 distinct domains, that are related to lead variant annotation and molecular QTL-GWAS
668 integration analyses (e.g., colocalization and TWAS) in FTLD-relevant tissues and cell types
669 were systematically assessed: (1) variant annotation, (2) eQTL-GWAS integration, (3) sQTL-
670 GWAS integration, (4) protein expression QTL (pQTL)-GWAS integration, (5) mQTL-GWAS
671 integration, and (6) histone acetylation QTL (haQTL)-GWAS integration; for which detailed
672 information on categories and subcategories is provided in **Supplementary Table 2**.

673 In the variant annotation domain, for each lead variant at each locus, we queried which
674 candidate risk genes were the nearest protein-coding genes with respect to the genomic position
675 of the lead variants, and/or whether the lead variant was a rare (MAF < 1% in gnomAD v4 non-
676 Finnish European samples) and/or protein-altering (missense or predicted LOF) variant for the
677 same nearest protein-coding genes. In the molecular QTL-GWAS integration domains, we
678 leveraged molecular cis-QTL catalogues for different molecular phenotypes (i.e., gene
679 expression, splicing, protein expression, methylation, and histone acetylation) in FTLD-
680 relevant tissues and cell types, we performed genetic colocalization analyses between
681 molecular cis-QTL and GWAS signals, TWAS, and proteome-wide association studies
682 (PWAS). For these analyses, we processed and used publicly available molecular QTL
683 catalogues; namely, FTLD-relevant bulk brain regions from AMP-AD⁷⁸⁻⁸¹ (as reanalyzed in
684 Bellenguez et al.¹⁴) and GTEx v8⁸² cohorts for the bulk brain eQTLs and sQTLs, eight major
685 brain cell types (excitatory neurons, inhibitory neurons, astrocytes, oligodendrocytes,
686 microglia, oligodendrocyte precursor cells/committed oligodendrocyte precursors
687 [OPCs/COPs], pericytes, and endothelial cells) from Bryois et al.⁸³ and primary microglia from
688 Young et al.⁸⁴ and the MiGA study⁸⁵ for the brain cell-type-specific eQTLs (ct-eQTL) and for
689 microglia sQTLs (from the MiGA study), dorsolateral prefrontal cortex (DLPFC) pQTLs from

690 Wingo et al.⁸⁶ (v2), and DLPFC mQTLs and haQTLs from Brain xQTL serve (June 2021
691 release)^{86,87}. Finally, we also included naïve state monocyte and macrophage eQTL
692 catalogues⁸⁸⁻⁹³ reanalyzed by eQTL Catalogue (Release 6)⁹⁴ and lymphoblastoid cell line
693 (LCL) eQTLs from GTEx v8⁸² and the European Alzheimer & Dementia Biobank (EADB)
694 Belgian LCL cohorts¹⁴. Using each of these molecular QTL catalogues, we first investigated
695 whether the reported lead variants in this study were significant molecular QTLs for the
696 quantified levels of molecular phenotypes in tissues and cell types of interest. Moreover, for
697 each quantified molecular phenotype in these catalogues we performed molecular QTL-GWAS
698 coloc (v5.2.2) analyses to determine if specific molecular QTL signals are colocalized (at coloc
699 $PP4 \geq 70\%$) with FTLD subtype GWAS signals. Finally, we conducted TWAS (using FUSION
700 and S-PrediXcan [implemented in MetaXcan] tools) for each heritable feature modelled in gene
701 expression (eTWAS; followed by eTWAS fine mapping with FOCUS⁹⁵ [v0.803] within 1 Mb
702 extended genome-wide significant lead variant genetic regions in each FTLD-TDP subtype
703 GWAS), splicing (sTWAS), and PWAS reference panels derived from AMP-AD bulk brain<sup>78-
704 81</sup>, GTEx bulk brain and LCL⁸², EADB Belgian LCL¹⁴, and Wingo et al. DLPFC data⁸⁶, to
705 identify the significant associations (after Bonferroni correction) between predicted levels of
706 gene expression, splicing, and protein expression with each FTLD subtype-specific genetic
707 risk. Detailed description and details (e.g., number of samples, significance criteria, references
708 and sources) of these molecular QTL catalogues used in this study for the systematic gene
709 prioritization strategy and functional interpretation of FTLD-TDP GWAS results can be found
710 in **Supplementary Table 5**.

711 Using a predetermined weighting scheme for each type of evidence (see **Supplementary Table**
712 **3**), we computed a gene prioritization score (between 0 and 87) for each gene which was
713 constructed by the weighted sum of the hits in different subcategories within six distinct
714 domains described above. As described in Bellenguez et al.¹⁴ in detail, we gave higher weights

715 for the hits obtained through the brain QTLs rather than other tissue QTLs, for the replicated
716 hits across multiple catalogues or reference panels, and for the fine-mapped eTWAS hits. After
717 obtaining weighted gene prioritization scores in each FTLD-TDP subtype-specific gene
718 prioritization analysis, we first assigned each candidate risk gene (with gene prioritization score
719 >0) to the genome-wide significant loci if their gene coordinates (based on GENCODE v24)
720 are positioned within a ± 1 Mb window of the identified lead variants (**Table 2**). The rest of the
721 candidate risk genes in subthreshold regions (nominated by coloc and TWAS analyses only)
722 were grouped together if they were positioned together (<1 Mb), and these subthreshold regions
723 were indexed and named as subthreshold loci. The candidate risk genes in genome-wide
724 significant and subthreshold loci were also annotated by the evidence of minimum P observed
725 within 1 Mb of the gene coordinates in related FTLD-TDP subtype-specific GWAS summary
726 statistics. We then ranked the protein-coding genes per locus in each FTLD subtype-specific
727 analysis based on their total weighted scores, and investigated the relative score differences
728 between the highest ranked protein-coding gene and the other candidate risk genes in each
729 locus, together with the overall total weighted score of the top-ranked gene. We then classified
730 candidate risk genes in each locus as tier 1 and tier 2 prioritized risk genes, respectively having
731 a higher and lower level of confidence for being a true risk gene in a given locus (see
732 Bellenguez et al.¹⁴ for detailed description). As also described in Bellenguez et al.¹⁴, the gene
733 prioritization pipeline determines a single tier 1 prioritized risk gene in each locus if there is
734 adequate evidence, meanwhile additional tier 2 prioritized risk genes in the same loci or
735 multiple tier 2 prioritized risk genes in a locus can also be assigned based on the score
736 distribution of candidate genes in the investigated loci.

737 **Gene Ontology analyses**

738 Gene ontology on tier 1 genes identified in FTLD-TDP All or in individuals FTLD-TDP
739 subtype analyses were performed using an R package which aggregates summary

740 statistics and assesses gene ontology term enrichment. Gene ontology terms were collapsed
741 using the rrvigo R package. Only terms with 2 or more genes were considered in the analyses.
742 $P < 0.05$ was considered statistically significant.

743 **Gene-level analysis of rare variants**

744 Association of rare variants with the patient/control status was assessed using an unweighted
745 burden test implemented using the SKAT_1.2.1 R package. Only VQSR pass variants with call
746 rate $> 90\%$, $ED \leq 4$, and $MAF < 0.01$ in either patients or controls were included. We included
747 only frameshift (insertion/deletion/block substitution), stopgain, stoploss and splicing SNVs
748 (jointly defined as loss-of-function (LOF) variants), and non-synonymous SNVs with REVEL
749 score above 0.75^{96} . Sex and the first thirteen PCs were used as covariates. Genome-wide
750 significance was defined as $P < 5 \times 10^{-8}$ and exome-wide significance as a p value $< 2.5 \times 10^{-6}$
751 (Bonferroni correction for 20,000 genes).

752 ***RBPJL* and *L3MBTL1* RNA expression**

753 Assessment of module membership of *RBPJL* and *L3MBTL1* was performed using the gene
754 co-expression analysis from the BrainEXP-NPD¹⁷ website using default parameters. Single
755 nuclei RNA expression was assessed using the transcriptomic comparative viewer if the Seattle
756 Alzheimer's Disease Brain cell Atlas from middle temporal gyrus of 84 aged donors (42
757 cognitively normal and 42 with dementia).

758 **Data Availability**

759 Summary statistics will be available on dbGAP platform post publication.

760 Datasets and molecular QTLs used in the gene prioritization are publicly available (see also

761 **Supplementary Table 4**):

762 eQTLs and eTWAS reference panels in AD-relevant bulk brain regions from AMP-AD
763 cohorts and in LCLs from the EADB Belgian cohort, as analyzed by Bellenguez et
764 al.¹⁴: <https://doi.org/10.5281/zenodo.5745927>;

765 sQTLs and sTWAS reference panels in AD-relevant bulk brain regions from AMP-AD
766 cohorts and in LCLs from the EADB Belgian cohort, as analyzed by Bellenguez et
767 al.¹⁴.: <https://doi.org/10.5281/zenodo.5745929>;

768 Bryois et al.⁸³ ct-eQTL catalogues (<https://doi.org/10.5281/zenodo.5543734>);

769 eQTL Catalogue database (<https://www.ebi.ac.uk/eqtl/>);

770 Brain xQTL serve mQTL and haQTL catalogues
771 (https://mostafavilab.stat.ubc.ca/xqtl/xQTL_updated_data/);

772 GTEx v8 eQTL and sQTL catalogues (<https://www.gtexportal.org/>);

773 GTEx v8 expression and splicing prediction models for eTWAS/sTWAS
774 ([https://predictdb.org/post/2021/07/21/gtex-v8-models-on-eqtl-and-sqtl/#mashr-based-](https://predictdb.org/post/2021/07/21/gtex-v8-models-on-eqtl-and-sqtl/#mashr-based-models)
775 [models](https://predictdb.org/post/2021/07/21/gtex-v8-models-on-eqtl-and-sqtl/#mashr-based-models));

776 MiGA eQTLs (<https://doi.org/10.5281/zenodo.4118605>);

777 MiGA sQTLs (<https://doi.org/10.5281/zenodo.4118403>);

778 MiGA meta-analysis (<https://doi.org/10.5281/zenodo.4118676>); and

779 Wingo et al.⁸⁶ pQTLs v2 (<https://www.synapse.org/#!Synapse:syn23627957>).

780

781 **References**

- 782 1. Rascovsky, K. *et al.* Sensitivity of revised diagnostic criteria for the behavioural
783 variant of frontotemporal dementia. *Brain* **134**, 2456-77 (2011).
- 784 2. Gorno-Tempini, M.L. *et al.* Classification of primary progressive aphasia and its
785 variants. *Neurology* **76**, 1006-14 (2011).
- 786 3. Lee, E.B. *et al.* Expansion of the classification of FTLT-DTP: distinct pathology
787 associated with rapidly progressive frontotemporal degeneration. *Acta Neuropathol*
788 **134**, 65-78 (2017).
- 789 4. Mackenzie, I.R. *et al.* A harmonized classification system for FTLT-DTP pathology.
790 *Acta Neuropathol* **122**, 111-3 (2011).
- 791 5. Mackenzie, I.R. *et al.* Heterogeneity of ubiquitin pathology in frontotemporal lobar
792 degeneration: classification and relation to clinical phenotype. *Acta Neuropathol* **112**,
793 539-49 (2006).
- 794 6. Mackenzie, I.R. & Neumann, M. Reappraisal of TDP-43 pathology in FTLT-U
795 subtypes. *Acta Neuropathol* **134**, 79-96 (2017).
- 796 7. Baker, M. *et al.* Mutations in progranulin cause tau-negative frontotemporal dementia
797 linked to chromosome 17. *Nature* **442**, 916-9 (2006).
- 798 8. Cruts, M. *et al.* Null mutations in progranulin cause ubiquitin-positive frontotemporal
799 dementia linked to chromosome 17q21. *Nature* **442**, 920-4 (2006).
- 800 9. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC hexanucleotide repeat in
801 noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron*
802 **72**, 245-56 (2011).
- 803 10. Renton, A.E. *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of
804 chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257-68 (2011).
- 805 11. Pottier, C. *et al.* Genome-wide analyses as part of the international FTLT-DTP
806 whole-genome sequencing consortium reveals novel disease risk factors and increases
807 support for immune dysfunction in FTLT. *Acta Neuropathol* **137**, 879-899 (2019).
- 808 12. Pottier, C. *et al.* Whole-genome sequencing reveals important role for TBK1 and
809 OPTN mutations in frontotemporal lobar degeneration without motor neuron disease.
810 *Acta Neuropathol* **130**, 77-92 (2015).
- 811 13. Ferrari, R. *et al.* Frontotemporal dementia and its subtypes: a genome-wide
812 association study. *Lancet Neurol* **13**, 686-99 (2014).
- 813 14. Bellenguez, C. *et al.* New insights into the genetic etiology of Alzheimer's disease and
814 related dementias. *Nat Genet* **54**, 412-436 (2022).
- 815 15. Project Min, E.A.L.S.S.C. Project MinE: study design and pilot analyses of a large-
816 scale whole-genome sequencing study in amyotrophic lateral sclerosis. *Eur J Hum*
817 *Genet* **26**, 1537-1546 (2018).
- 818 16. Benyamin, B. *et al.* Cross-ethnic meta-analysis identifies association of the GPX3-
819 TNIP1 locus with amyotrophic lateral sclerosis. *Nat Commun* **8**, 611 (2017).
- 820 17. Jiao, C. *et al.* BrainEXP: a database featuring with spatiotemporal expression
821 variations and co-expression organizations in human brains. *Bioinformatics* **35**, 172-
822 174 (2019).
- 823 18. Brown, A.L. *et al.* TDP-43 loss and ALS-risk SNPs drive mis-splicing and depletion
824 of UNC13A. *Nature* **603**, 131-137 (2022).
- 825 19. Ma, X.R. *et al.* TDP-43 represses cryptic exon inclusion in the FTD-ALS gene
826 UNC13A. *Nature* **603**, 124-130 (2022).
- 827 20. Augustin, I., Rosenmund, C., Sudhof, T.C. & Brose, N. Munc13-1 is essential for
828 fusion competence of glutamatergic synaptic vesicles. *Nature* **400**, 457-61 (1999).

- 829 21. Restuadi, R. *et al.* Functional characterisation of the amyotrophic lateral sclerosis risk
830 locus GPX3/TNIP1. *Genome Med* **14**, 7 (2022).
- 831 22. Tanaka, H. *et al.* ITIH4 and Gpx3 are potential biomarkers for amyotrophic lateral
832 sclerosis. *J Neurol* **260**, 1782-97 (2013).
- 833 23. Oshima, S. *et al.* ABIN-1 is a ubiquitin sensor that restricts cell death and sustains
834 embryonic development. *Nature* **457**, 906-9 (2009).
- 835 24. Su, Z. *et al.* ABIN-1 heterozygosity sensitizes to innate immune response in both
836 RIPK1-dependent and RIPK1-independent manner. *Cell Death Differ* **26**, 1077-1088
837 (2019).
- 838 25. Zhou, J. *et al.* A20-binding inhibitor of NF-kappaB (ABIN1) controls Toll-like
839 receptor-mediated CCAAT/enhancer-binding protein beta activation and protects
840 from inflammatory disease. *Proc Natl Acad Sci U S A* **108**, E998-1006 (2011).
- 841 26. Lee, Y. *et al.* Coordinate regulation of the senescent state by selective autophagy. *Dev*
842 *Cell* **56**, 1512-1525 e7 (2021).
- 843 27. Freischmidt, A. *et al.* Haploinsufficiency of TBK1 causes familial ALS and fronto-
844 temporal dementia. *Nat Neurosci* **18**, 631-6 (2015).
- 845 28. Mackenzie, I.R. *et al.* The neuropathology of frontotemporal lobar degeneration
846 caused by mutations in the progranulin gene. *Brain* **129**, 3081-90 (2006).
- 847 29. Katsumata, Y. *et al.* Multiple gene variants linked to Alzheimer's-type clinical
848 dementia via GWAS are also associated with non-Alzheimer's neuropathologic
849 entities. *Neurobiol Dis* **174**, 105880 (2022).
- 850 30. Katsumata, Y., Nelson, P.T., Ellingson, S.R. & Fardo, D.W. Gene-based association
851 study of genes linked to hippocampal sclerosis of aging neuropathology: GRN,
852 TMEM106B, ABCC9, and KCNMB2. *Neurobiol Aging* **53**, 193 e17-193 e25 (2017).
- 853 31. Nelson, P.T. *et al.* Limbic-predominant age-related TDP-43 encephalopathy (LATE):
854 consensus working group report. *Brain* **142**, 1503-1527 (2019).
- 855 32. Josephs, K.A. *et al.* Pathological, imaging and genetic characteristics support the
856 existence of distinct TDP-43 types in non-FTLD brains. *Acta Neuropathol* **137**, 227-
857 238 (2019).
- 858 33. Van Deerlin, V.M. *et al.* Common variants at 7p21 are associated with frontotemporal
859 lobar degeneration with TDP-43 inclusions. *Nat Genet* **42**, 234-9 (2010).
- 860 34. C, T.V. *et al.* C-terminal TMEM106B fragments in human brain correlate with
861 disease-associated TMEM106B haplotypes. *Brain* **146**, 4055-4064 (2023).
- 862 35. Hook, V. *et al.* Cathepsin B in neurodegeneration of Alzheimer's disease, traumatic
863 brain injury, and related brain disorders. *Biochim Biophys Acta Proteins Proteom*
864 **1868**, 140428 (2020).
- 865 36. Lee, C.W. *et al.* The lysosomal protein cathepsin L is a progranulin protease. *Mol*
866 *Neurodegener* **12**, 55 (2017).
- 867 37. Zhou, X. *et al.* Lysosomal processing of progranulin. *Mol Neurodegener* **12**, 62
868 (2017).
- 869 38. Vesa, J. *et al.* Mutations in the palmitoyl protein thioesterase gene causing infantile
870 neuronal ceroid lipofuscinosis. *Nature* **376**, 584-7 (1995).
- 871 39. Camp, L.A., Verkruyse, L.A., Afendis, S.J., Slaughter, C.A. & Hofmann, S.L.
872 Molecular cloning and expression of palmitoyl-protein thioesterase. *J Biol Chem* **269**,
873 23212-9 (1994).
- 874 40. Branchu, J. *et al.* Loss of spatascin function alters lysosomal lipid clearance leading to
875 upper and lower motor neuron degeneration. *Neurobiol Dis* **102**, 21-37 (2017).
- 876 41. Han, S.M. *et al.* VAPB/ALS8 MSP ligands regulate striated muscle energy
877 metabolism critical for adult survival in caenorhabditis elegans. *PLoS Genet* **9**,
878 e1003738 (2013).

- 879 42. Cingolani, F. & Czaja, M.J. Regulation and Functions of Autophagic Lipolysis.
880 *Trends Endocrinol Metab* **27**, 696-705 (2016).
- 881 43. Liu, Y. *et al.* A C9orf72-CARM1 axis regulates lipid metabolism under glucose
882 starvation-induced nutrient stress. *Genes Dev* **32**, 1380-1397 (2018).
- 883 44. Conibear, E. & Stevens, T.H. Vps52p, Vps53p, and Vps54p form a novel
884 multisubunit complex required for protein sorting at the yeast late Golgi. *Mol Biol*
885 *Cell* **11**, 305-23 (2000).
- 886 45. Liewen, H. *et al.* Characterization of the human GARP (Golgi associated retrograde
887 protein) complex. *Exp Cell Res* **306**, 24-34 (2005).
- 888 46. Wei, J. *et al.* The GARP Complex Is Involved in Intracellular Cholesterol Transport
889 via Targeting NPC2 to Lysosomes. *Cell Rep* **19**, 2823-2835 (2017).
- 890 47. Guise, A.J. *et al.* TDP-43-stratified single-cell proteomic profiling of postmortem
891 human spinal motor neurons reveals protein dynamics in amyotrophic lateral
892 sclerosis. *bioRxiv* (2023).
- 893 48. Yoshimura, S., Gerondopoulos, A., Linford, A., Rigden, D.J. & Barr, F.A. Family-
894 wide characterization of the DENN domain Rab GDP-GTP exchange factors. *J Cell*
895 *Biol* **191**, 367-81 (2010).
- 896 49. Deshimaru, M. *et al.* DCTN1 Binds to TDP-43 and Regulates TDP-43 Aggregation.
897 *Int J Mol Sci* **22**(2021).
- 898 50. He, J., Yu, W., Liu, X. & Fan, D. An identical DCTN1 mutation in two Chinese
899 siblings manifest as dHMN and ALS respectively: a case report. *Amyotroph Lateral*
900 *Scler Frontotemporal Degener* **23**, 149-153 (2022).
- 901 51. Nicolas, A. *et al.* Genome-wide Analyses Identify KIF5A as a Novel ALS Gene.
902 *Neuron* **97**, 1267-1288 (2018).
- 903 52. Xia, C.H. *et al.* Abnormal neurofilament transport caused by targeted disruption of
904 neuronal kinesin heavy chain KIF5A. *J Cell Biol* **161**, 55-66 (2003).
- 905 53. Brenneman, D.E., Hauser, J., Spong, C.Y. & Phillips, T.M. Chemokines released
906 from astroglia by vasoactive intestinal peptide. Mechanism of neuroprotection from
907 HIV envelope protein toxicity. *Ann N Y Acad Sci* **921**, 109-14 (2000).
- 908 54. Brenneman, D.E. *et al.* Vasoactive intestinal peptide. Link between electrical activity
909 and glia-mediated neurotrophism. *Ann N Y Acad Sci* **897**, 17-26 (1999).
- 910 55. Chneiweiss, H., Glowinski, J. & Premont, J. Vasoactive intestinal polypeptide
911 receptors linked to an adenylate cyclase, and their relationship with biogenic amine-
912 and somatostatin-sensitive adenylate cyclases on central neuronal and glial cells in
913 primary cultures. *J Neurochem* **44**, 779-86 (1985).
- 914 56. Hosli, E. & Hosli, L. Autoradiographic localization of binding sites for vasoactive
915 intestinal peptide and angiotensin II on neurons and astrocytes of cultured rat central
916 nervous system. *Neuroscience* **31**, 463-70 (1989).
- 917 57. Delgado, M. & Ganea, D. Vasoactive intestinal peptide: a neuropeptide with
918 pleiotropic immune functions. *Amino Acids* **45**, 25-39 (2013).
- 919 58. Baer, G.M. *et al.* Sporadic FTL D type C is a distinct clinicopathological entity
920 from FTL D types A/B (S7.007). *Neurology* **88**, S7.007 (2017).
- 921 59. Rohrer, J.D. *et al.* The heritability and genetics of frontotemporal lobar degeneration.
922 *Neurology* **73**, 1451-6 (2009).
- 923 60. Snowden, J., Neary, D. & Mann, D. Frontotemporal lobar degeneration: clinical and
924 pathological relationships. *Acta Neuropathol* **114**, 31-8 (2007).
- 925 61. Pan, L. *et al.* Transcription Factor RBPJL Is Able to Repress Notch Target Gene
926 Expression but Is Non-Responsive to Notch Activation. *Cancers (Basel)* **13**(2021).

- 927 62. Bonham, L.W. *et al.* Genetic variation across RNA metabolism and cell death gene
928 networks is implicated in the semantic variant of primary progressive aphasia. *Sci Rep*
929 **9**, 10854 (2019).
- 930 63. Lu, J. *et al.* L3MBTL1 regulates ALS/FTD-associated proteotoxicity and quality
931 control. *Nat Neurosci* **22**, 875-886 (2019).
- 932 64. Guissart, C. *et al.* Premature termination codons in SOD1 causing Amyotrophic
933 Lateral Sclerosis are predicted to escape the nonsense-mediated mRNA decay. *Sci*
934 *Rep* **10**, 20738 (2020).
- 935 65. Murley, A.G. & Rowe, J.B. Neurotransmitter deficits from frontotemporal lobar
936 degeneration. *Brain* **141**, 1263-1285 (2018).
- 937 66. Ferrer, I. Neurons and their dendrites in frontotemporal dementia. *Dement Geriatr*
938 *Cogn Disord* **10 Suppl 1**, 55-60 (1999).
- 939 67. Premi, E. *et al.* Unravelling neurotransmitters impairment in primary progressive
940 aphasias. *Hum Brain Mapp* **44**, 2245-2253 (2023).
- 941 68. Bowen, D.M. *et al.* Imbalance of a serotonergic system in frontotemporal dementia:
942 implication for pharmacotherapy. *Psychopharmacology (Berl)* **196**, 603-10 (2008).
- 943 69. Procter, A.W., Qurne, M. & Francis, P.T. Neurochemical features of frontotemporal
944 dementia. *Dement Geriatr Cogn Disord* **10 Suppl 1**, 80-4 (1999).
- 945 70. Premi, E. *et al.* Early neurotransmitters changes in prodromal frontotemporal
946 dementia: A GENFI study. *Neurobiol Dis* **179**, 106068 (2023).
- 947 71. Ghezzi, L., Cantoni, C., Rotondo, E. & Galimberti, D. The Gut Microbiome-Brain
948 Crosstalk in Neurodegenerative Diseases. *Biomedicines* **10**(2022).
- 949 72. Khawar, M.M., Sr. *et al.* The Gut-Brain Axis in Autoimmune Diseases: Emerging
950 Insights and Therapeutic Implications. *Cureus* **15**, e48655 (2023).
- 951 73. Miller, Z.A. *et al.* TDP-43 frontotemporal lobar degeneration and autoimmune
952 disease. *J Neurol Neurosurg Psychiatry* **84**, 956-62 (2013).
- 953 74. Olson, J.E. *et al.* The Mayo Clinic Biobank: a building block for individualized
954 medicine. *Mayo Clin Proc* **88**, 952-62 (2013).
- 955 75. McPeck, M.S. & Sun, L. Statistical tests for detection of misspecified relationships by
956 use of genome-screen data. *Am J Hum Genet* **66**, 1076-94 (2000).
- 957 76. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of
958 genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
- 959 77. Lake, B.B. *et al.* Integrative single-cell analysis of transcriptional and epigenetic
960 states in the human adult brain. *Nat Biotechnol* **36**, 70-80 (2018).
- 961 78. Allen, M. *et al.* Human whole genome genotype and transcriptome data for
962 Alzheimer's and other neurodegenerative diseases. *Sci Data* **3**, 160089 (2016).
- 963 79. Mostafavi, S. *et al.* A molecular network of the aging human brain provides insights
964 into the pathology and cognitive decline of Alzheimer's disease. *Nat Neurosci* **21**,
965 811-819 (2018).
- 966 80. Bennett, D.A. *et al.* Religious Orders Study and Rush Memory and Aging Project. *J*
967 *Alzheimers Dis* **64**, S161-S189 (2018).
- 968 81. Wang, M. *et al.* The Mount Sinai cohort of large-scale genomic, transcriptomic and
969 proteomic data in Alzheimer's disease. *Sci Data* **5**, 180185 (2018).
- 970 82. Consortium, G.T. The GTEx Consortium atlas of genetic regulatory effects across
971 human tissues. *Science* **369**, 1318-1330 (2020).
- 972 83. Bryois, J. *et al.* Cell-type-specific cis-eQTLs in eight human brain cell types identify
973 novel risk genes for psychiatric and neurological disorders. *Nat Neurosci* **25**, 1104-
974 1112 (2022).
- 975 84. Young, A.M.H. *et al.* A map of transcriptional heterogeneity and regulatory variation
976 in human microglia. *Nat Genet* **53**, 861-868 (2021).

- 977 85. Lopes, K.P. *et al.* Genetic analysis of the human microglial transcriptome across brain
978 regions, aging and disease pathologies. *Nat Genet* **54**, 4-17 (2022).
- 979 86. Wingo, A.P. *et al.* Integrating human brain proteomes with genome-wide association
980 data implicates new proteins in Alzheimer's disease pathogenesis. *Nat Genet* **53**, 143-
981 146 (2021).
- 982 87. Ng, B. *et al.* An xQTL map integrates the genetic architecture of the human brain's
983 transcriptome and epigenome. *Nat Neurosci* **20**, 1418-1426 (2017).
- 984 88. Alasoo, K. *et al.* Shared genetic effects on chromatin and gene expression indicate a
985 role for enhancer priming in immune response. *Nat Genet* **50**, 424-431 (2018).
- 986 89. Nedelec, Y. *et al.* Genetic Ancestry and Natural Selection Drive Population
987 Differences in Immune Responses to Pathogens. *Cell* **167**, 657-669 e21 (2016).
- 988 90. Chen, L. *et al.* Genetic Drivers of Epigenetic and Transcriptional Variation in Human
989 Immune Cells. *Cell* **167**, 1398-1414 e24 (2016).
- 990 91. Momozawa, Y. *et al.* IBD risk loci are enriched in multigenic regulatory modules
991 encompassing putative causative genes. *Nat Commun* **9**, 2427 (2018).
- 992 92. Quach, H. *et al.* Genetic Adaptation and Neandertal Admixture Shaped the Immune
993 System of Human Populations. *Cell* **167**, 643-656 e17 (2016).
- 994 93. Fairfax, B.P. *et al.* Innate immune activity conditions the effect of regulatory variants
995 upon monocyte gene expression. *Science* **343**, 1246949 (2014).
- 996 94. Kerimov, N. *et al.* A compendium of uniformly processed human gene expression and
997 splicing quantitative trait loci. *Nat Genet* **53**, 1290-1299 (2021).
- 998 95. Mancuso, N. *et al.* Probabilistic fine-mapping of transcriptome-wide association
999 studies. *Nat Genet* **51**, 675-682 (2019).
- 1000 96. Ioannidis, N.M. *et al.* REVEL: An Ensemble Method for Predicting the Pathogenicity
1001 of Rare Missense Variants. *Am J Hum Genet* **99**, 877-885 (2016).
- 1002

1003

1004

1005

1006 **Declarations**

1007 **Ethics approval and consent to participate**

1008 This study was approved by the appropriate Mayo Clinic Institutional Review Board.

1009 **Consent for publication**

1010 Not applicable.

1011 **Competing interests**

1012 RR and IRM receive royalties from progranulin-related patent.

1013 **Funding**

1014 This work was made possible by the support of funding from the UG3/UH3 NS103870, P30
1015 AG062677, P01 AG003949, R01 AG037491, P01 NS084974, R01 AG37491, R01 DC12519,
1016 R21 NS94684, R01 DC010367, P30 AG062422, P01 AG019724, U01 AG057195, U19
1017 AG063911, P30 AG077444, P30 AG013854, P30 AG07297, P30 AG066468, P30
1018 AG012300, P30 AG072972, P30 AG010129, P50 AG005136, P30 AG066509, P30
1019 AG066511, P01 AG066597, P30 AG072979, R01 AG077444, and R01 DC008552.

1020 Supported by the ALLFTD Consortium (U19: AG063911, funded by the NIA and NINDS)
1021 and the former ARTFL & LEFFTDS Consortia (ARTFL: U54 NS092089, funded by the
1022 NINDS and NCATS; LEFFTDS: U01 AG045390, funded by the NIA and NINDS).
1023 Supported by the Mayo ADRC (P30 AG62677, funded by the NIA). A subset of samples
1024 were obtained through the National Centralized Repository for Alzheimer Disease and
1025 Related Dementias (NCRAD), which receives government support under a cooperative
1026 agreement grant (U24 AG021886) awarded by the National Institute on Aging (NIA), were
1027 used in this study. The work was supported by Mayo Clinic Center for Regenerative
1028 Medicine, the gifts from the Donald G. and Jodi P. Heeringa Family, the Haworth Family
1029 Professorship in Neurodegenerative Diseases fund, The Albertson Parkinson's Research
1030 Foundation, and PPND Family Foundation.

1031 This work was supported by TARGET ALS, the Rainwater Charitable Foundation and the
1032 Bluefield Project to Cure FTD, the DEC Brain and BioBank, CIHR grant 74580, the
1033 Canadian Consortium on Neurodegeneration in Aging, of the Canadian Institutes of Health

1034 Research, the University of Pittsburgh Brain Institute. This work was supported by the NIHR
1035 UCL/H Biomedical Research Centre, the UK MRC (MR/M008525/1; MR/M023664/1), the
1036 Bluefield Project and the JPND (2019-02248). This work was made possible, in part, by the
1037 Winspear Family Center for Research on the Neuropathology of Alzheimer disease, the
1038 McCune Foundation, and the Nancy and Buster Alvord endowment.

1039 Research was supported by Grants provided by Swedish FTD Initiative- Schörling foundation,
1040 the Swedish research Council (Dnr 521-2010-3134, 529-2014-7504, 2015-02926), Alzheimer
1041 Foundation Sweden, Brain Foundation Sweden, Swedish Brain Power, Gamla Tjänarinnor,
1042 Stohnes Foundation, Dementia Foundation Sweden and the Stockholm County Council (ALF-
1043 project). The brain pathology was provided through in part by the Brain Bank at Karolinska
1044 Institutet which was financially supported by Karolinska Institutet StratNeuro, Swedish Brain
1045 Power, Stockholm County Council core facility funding (CG) and Schörling Foundation.

1046 This work was supported in part by SAO-FRA 2021/0032. FK receives a postdoctoral
1047 fellowship (BOF 49758) from the University of Antwerp Research Fund.

1048 **Acknowledgements**

1049 We thank the Mayo Clinic Center of Individualized Medicine for collection and sequencing of
1050 the Mayo Clinic Biobank samples. CG would also like to acknowledge Dr Inger Nennesmo for
1051 the neuropathological characterization. Acknowledgment for the ADSP dataset can be found
1052 in **Supplementary Notes**.

1053

1054

1055

1056 **Legends**

1057 **Figure 1: Genome-wide association study on common variants.** A Manhattan plot of the
1058 FTLD-TDP All patients versus controls association study. **B** Manhattan plot of the FTLD-TDP
1059 A patients versus controls association study. **C** Manhattan plot of the FTLD-TDP B patients
1060 versus controls association study. **D** Manhattan plot of the FTLD-TDP C patients versus
1061 controls association study. The red-dotted line represents the genome-wide significance level
1062 ($p = 5 \times 10^{-8}$).

1063 **Figure 2: Gene prioritization results for FTLD-TDP subgroups.** A visual summary of
1064 weighted evidence category scores for the prioritized genes within genome-wide significant
1065 and subthreshold loci with candidate genes whose 1 Mb extended gene coordinates contain a
1066 minimal GWAS P evidence of $\leq 5 \times 10^{-6}$ in related FTLD subtype-specific GWAS summary
1067 statistics. Using the gene prioritization strategy in these selected loci, we prioritized a total of
1068 25 genes in 23 loci at two different confidence levels (10 tier 1 and 15 tier 2 prioritized genes).
1069 The leftmost squares which are colored in red for FTLD-TDP A, in blue for FTLD-TDP B, and
1070 in green for FTLD-TDP C specific analyses indicate the locus index numbers which contain
1071 additional “_S” patterns for the subthreshold loci, whereas others indicate the genome-wide
1072 significant loci. The types of evidence for each category are colored according to the six
1073 different domains to which they belong. Weighted scores for each evidence category are
1074 rescaled to a 0–100 scale based on the maximum score a candidate gene can obtain from a
1075 category (see **Supplementary Table 2**). The darker colors represent higher scores in
1076 categories, while tier 1 prioritized genes are displayed in dark green and tier 2 prioritized genes
1077 are displayed in light green. Only tier 1 and tier 2 genes are shown for each locus, and all
1078 candidate genes considered and scored can be found in **Supplementary Table 3**. MAFs (based
1079 on gnomAD v4 non-Finnish European samples) and CADD (v1.7) PHRED scores for rare
1080 and/or protein-altering rare variants are labeled in white within the respective squares. eQTL,

1081 expression QTL; sQTL, splicing QTL; mQTL, methylation QTL; pQTL, protein-expression
1082 QTL; haQTL, histone acetylation QTL; coloc, colocalization; eTWAS, expression
1083 transcriptome-wide association study; sTWAS, splicing transcriptome-wide association study;
1084 PWAS, proteome-wide association study; Mon. Mac., monocytes and macrophages; LCL,
1085 lymphoblastoid cell line; QTLCat, The eQTL Catalogue.

1086 **Figure 3: Top 5 Gene Ontology terms enriched in FTLD-TDP subgroups.** Hierarchical GO
1087 analysis of biological process terms considering genes in genetic loci prioritized for FTLD-
1088 TDP All, FTLD-TDP A, FTLD-TDP B and FTD-TDP C.

1089 **Figure 4: Enrichment of brain regions and cell types in FTLD subgroups.** **A** Enrichment
1090 of genes in multiple tissues, including 13 brain regions, and based on GTEX data in FTLD
1091 subgroups, ADRD and ALS. Color represents the enrichment coefficient, and size indicates
1092 two-sided $-\log_{10}$ (FDR adjusted P s) of enrichment obtained by the linear regression model in
1093 the MAGMA gene property analysis. **B** Central nervous system cell type enrichment analyses
1094 in FTLD subgroups, ADRD and ALS. Color represents the enrichment coefficient, and size
1095 indicates two-sided $-\log_{10}$ (FDR adjusted P s) of enrichment obtained by the linear regression
1096 model in the MAGMA gene property analysis. Excitatory neurons and glial cells are
1097 highlighted in blue. Excitatory and inhibitory neurons from the PsychENCODE dataset were
1098 labeled based on their transcriptional profile from 1 to 8. Asterisks denote brain regions or cell
1099 types enriched with FDR $P < 0.05$. Cx, cortex; Ex*, Excitatory neuron, In*, inhibitory neurons;
1100 Oligo, oligodendrocytes; OPCs, oligodendrocyte progenitor cells; Astro, astrocytes; Endo,
1101 endothelial cells; Per, pericytes.

1102 **Figure 5: Locus zoom plots for *UNC13A* and *TNIP1* loci.** **A** Genetic colocalization between
1103 the *UNC13A* locus in FTLD-TDP (meta-analysis) and ALS signal. **B** Genetic colocalization
1104 between the *TNIP1* locus in FTLD-TDP (meta-analysis) and ALS. **C** Genetic colocalization

1105 between the *TNIP1* locus in FTLD-TDP (meta-analysis) and ADRD. For **A**, **B** and **C**,
1106 chromosome position is located on the x axis and $-\log_{10}(P)$ is represented on the y axis. Each
1107 dot represents a SNV tested in the dataset for its association with disease status. Purple
1108 diamonds are the index SNVs reported. Linkage disequilibrium with index SNV is indicated
1109 by r^2 .

1110 **Figure 6: Rare loss of function and predicted pathogenic variants in proteins associated**
1111 **with FTLD.** Schematic representation of VIPR1, L3MBTL1 and RBPL protein structure
1112 (source Uniprot) showing a map of nonsense, splicing, frameshift and missense with a REVEL
1113 score > 0.75 rare variants in patients and controls. Variants identified in patients are colored in
1114 orange, variants identified in controls are colored in blue. n=number of carriers.

1115

1116 **Figure 7: Schematic representation of findings from the International FTLD-TDP WGS**
1117 **phase II.** Genome-wide significant single variant loci, exome-wide significant genes, enriched
1118 gene ontology pathways and tissues- and cell-types enriched for genome-wide significant risk
1119 loci are shown for each FTLD-TDP pathological subtype in rings moving from the center
1120 (genome-wide significant single variant loci in FTLD-TDP All) to the outer rings. Orange
1121 background shades correspond to FTLD-TDP A findings, green background shades to FTLD-
1122 TDP B findings and blue background shades to FTLD-TDP C findings. Gene names in green
1123 font were exome-wide significant using a gene-based approach with common variants while
1124 gene names in red font were exome-wide significant using a gene-based approach with rare
1125 variants. In addition to unique associations, some overlap between FTLD-TDP A and B exist
1126 (TBK1, lysosomal function and inflammatory response), whereas FTLD-TDP C showed
1127 unique and non-overlapping genetic profile.

Table 1: Demographics after quality control.

	FTLD-TDP A	FTLD-TDP B	FTLD-TDP C	FTLD-TDP U^a	Control
Number	193	288	467	37	3,153
(% Female)	(42.49)	(38.89)	(48.61)	(40.54)	(56.71)
Age at onset or age at collection (Standard Deviation, years)	68.00 (10.52)	62.00 (10.76)	60.50 (8.05)	60.50 (8.74)	64.08 (14.31)
Age at death (Standard Deviation, years)	78.00 (11.72)	67.00 (10.90)	72.00 (7.32)	67.84 (9.41)	82.00 (8.41)
Disease duration (Standard Deviation, years)	7.30 (4.65)	4.00 (3.61)	11.00 (7.46)	6.50 (5.20)	NA

^aFTLD-TDP U = FTLD-TDP unclassifiable

Table 2: Top variants associated with disease status

GROUP	rsID ^a	REF/ALT	LOCUS NAME ^b	GENOMIC POSITION ^c	ODDS RATIO (95% CI) ^d	P	MAF PATIENTS/CONTROLS	NEW LOCUS
FTLD-TDP All	rs8111424	A/G	<i>UNC13A</i>	19:17640336	1.37 (1.24-1.54)	1.17X10 ⁻⁸	0.376/0.141	
FTLD-TDP A	rs5848	C/T	<i>GRN</i>	17:44352876	1.89 (1.52-2.34)	5.57X10 ⁻⁹	0.442/0.292	New
FTLD-TDP A	rs138698596	T/A	<i>TINAG</i>	6:54591659	5.22 (2.91-9.36)	3.01X10 ⁻⁸	0.045/0.009	New
FTLD-TDP A	rs138959102	C/T	<i>MZT1</i>	13:72499532	8.41 (3.95-17.88)	3.22X10 ⁻⁸	0.029/0.004	New
FTLD-TDP A	rs886815	G/A	<i>FARP2</i>	2:241457011	9.55 (4.26-21.41)	4.26X10 ⁻⁸	0.026/0.003	New
FTLD-TDP B	rs76742217	G/A	<i>RCL1</i>	9:4821273	9.31 (4.30-20.18)	1.55X10 ⁻⁸	0.023/0.003	New
FTLD-TDP B	rs12973192	G/C	<i>UNC13A</i>	19:17642430	1.74 (1.46-2.08)	8.52X10 ⁻¹⁰	0.484/0.346	
FTLD-TDP B	rs527749954	C/T	<i>PDS5B</i>	13:32620689	7.66 (3.71-15.84)	3.89X10 ⁻⁸	0.025/0.003	New
FTLD-TDP B	rs871269	C/T	<i>TNIP1</i>	5:151052827	0.55 (0.44-0.68)	4.72X10 ⁻⁸	0.206/0.322	New
FTLD-TDP C	rs576561313	C/T	<i>C19orf52</i>	19:10945440	13.11 (5.33-32.23)	2.03X10 ⁻⁸	0.014/0.001	New

^ars number, according to dbSNP build 153.

^bNearest protein-coding gene according to GENCODE release 33.

^cGRCh38 assembly.

^dApproximate OR calculated with respect to the alternative allele.

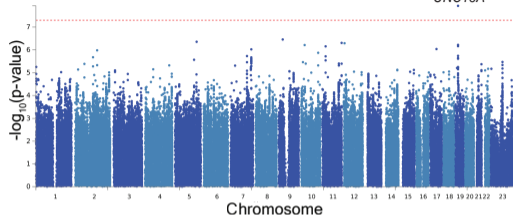
Table 3: Genes harboring rare variants associated with FTLN-TDP

GROUP	GENE NAME	NUMBER OF PATIENTS	MAF^a IN PATIENTS	NUMBER OF CONTROLS	MAF^a IN CONTROLS	P
FTLD-TDP A	<i>TBK1</i>	3	7.77X10 ⁻⁰³	0	0	1.27X10 ⁻¹¹
FTLD-TDP B	<i>TBK1</i>	5	8.68X10 ⁻⁰³	0	0	3.17X10 ⁻¹²
FTLD-TDP B	<i>VIPR1</i>	3	5.21X10 ⁻⁰³	1	1.59X10 ⁻⁰⁴	4.65X10 ⁻⁷
FTLD-TDP C	<i>RBPJL</i>	5	5.35X10 ⁻⁰³	3	4.76X10 ⁻⁰⁴	6.39X10 ⁻⁷
FTLD-TDP C	<i>L3MBTL1</i>	8	8.57X10 ⁻⁰³	3	4.76X10 ⁻⁰⁴	2.38X10 ⁻⁷

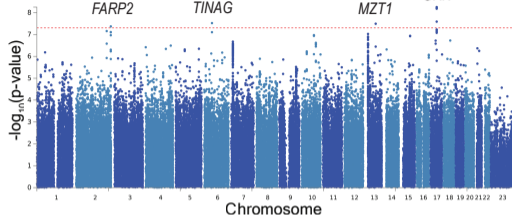
^aMAF = minor allele frequency

A

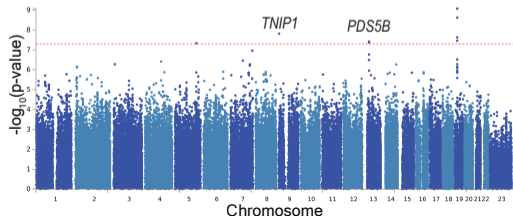
FTLD-TDP All

UNC13A**B**

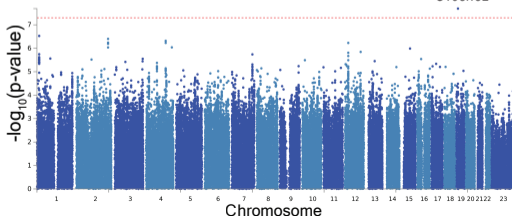
FTLD-TDP A

*FARP2**TINAG**MZT1**GRN***C**

FTLD-TDP B

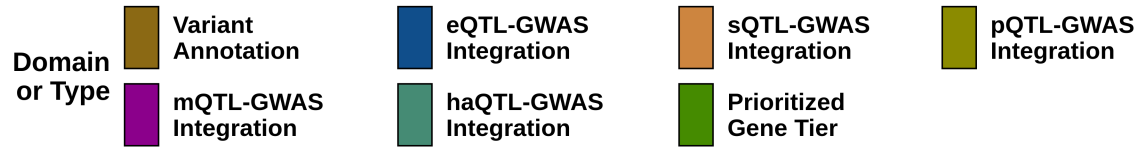
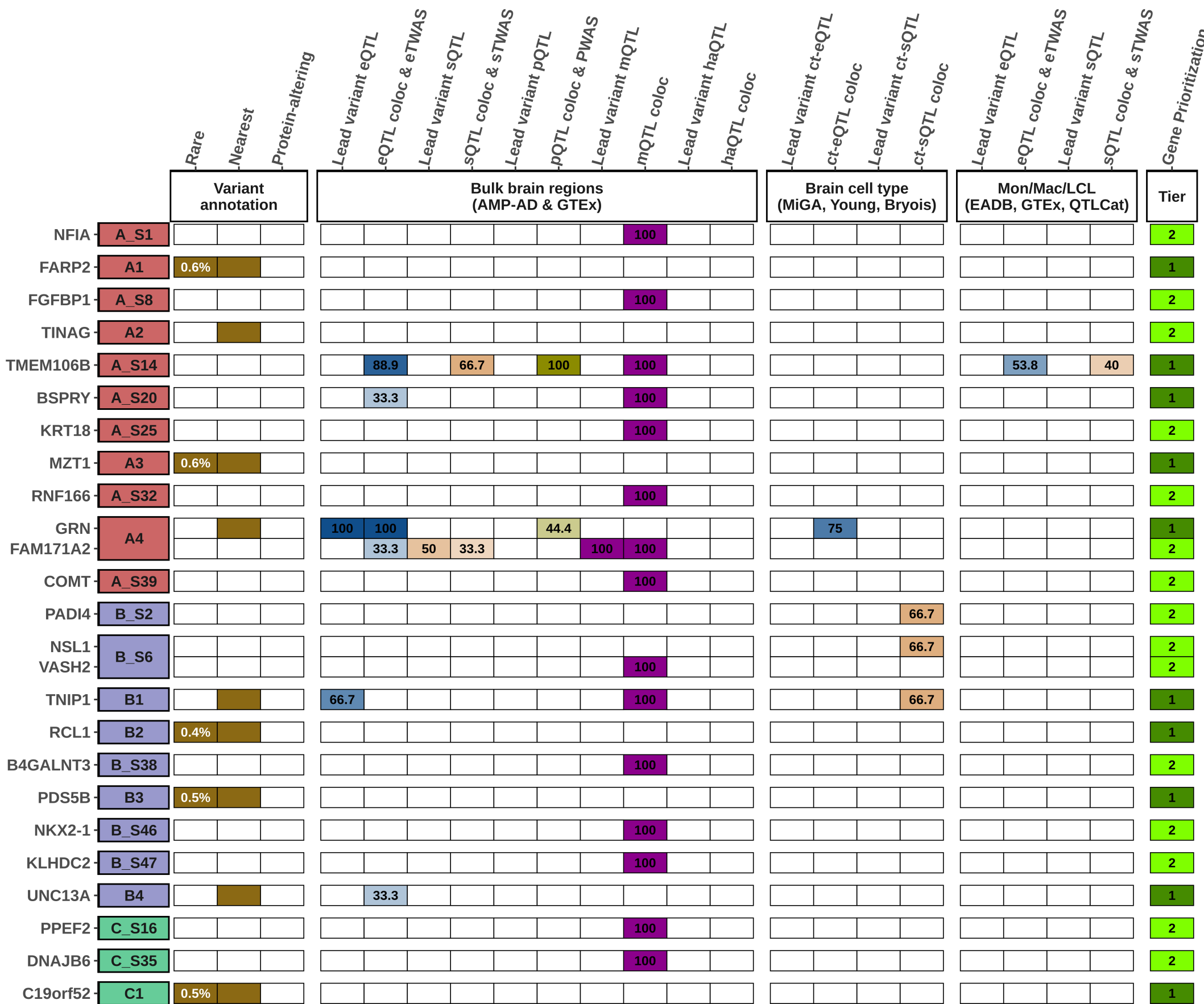
*UNC13A**TNIP1**PDS5B***D**

FTLD-TDP C

C19orf52

Category

Prioritized Gene

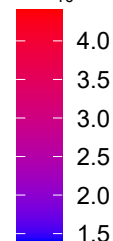


GO Terms across Groups

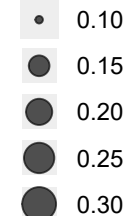
GO Term



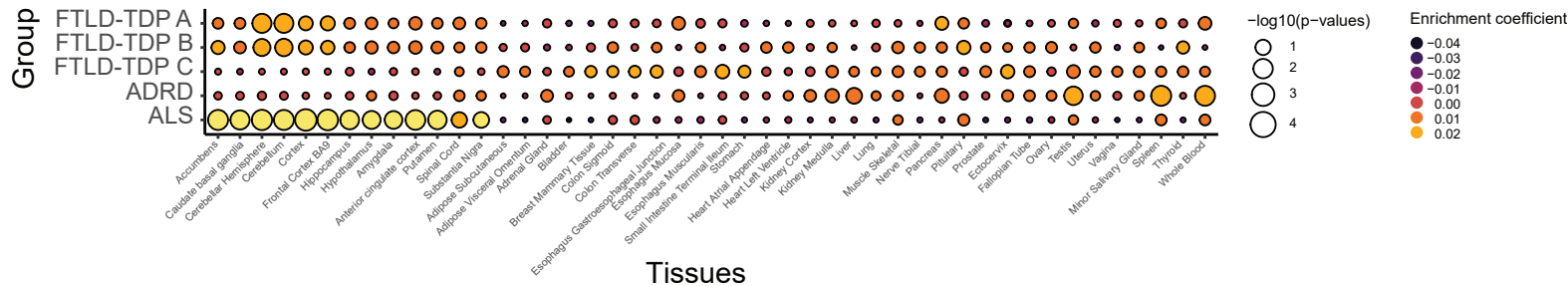
\log_{10} p-value



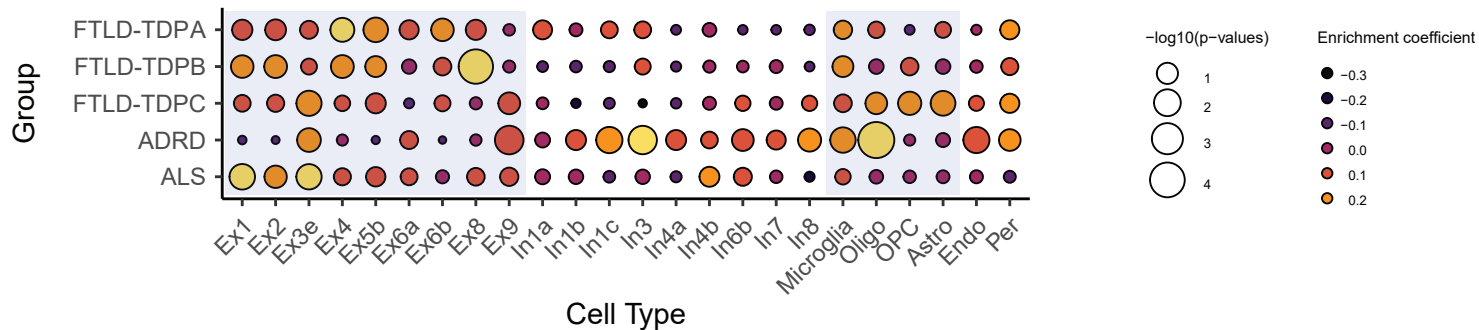
Gene ratio



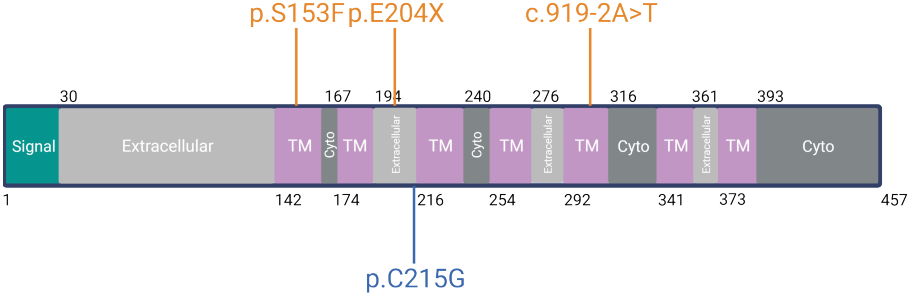
A



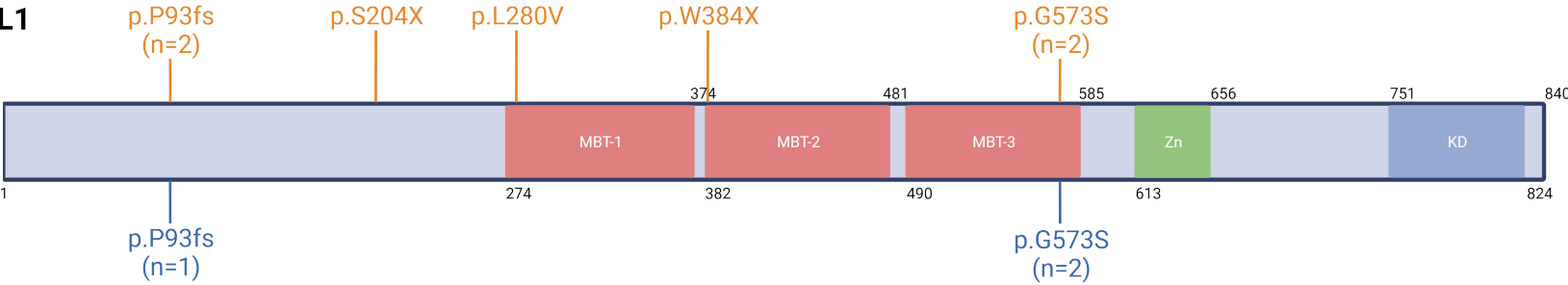
B



VIPR1



L3MBTL1



RBPJL

