Title: Modules of genotypic variance reflect heterogeneity across TDP-43 proteinopathies

Authors: Barbara E. Spencer[1], David J. Irwin[1], Vivianna M. Van Deerlin[2], EunRan Suh[2], Edward B. Lee[2], Lauren B. Elman[1], Colin C. Quinn[1], Defne Amado[1], Michael Baer[1], Murray Grossman[1], David A. Wolk[1], and Corey T. McMillan[1]*

Affiliations:

[1] Department of Neurology, University of Pennsylvania; Philadelphia, PA, USA.

[2] Department of Pathology and Laboratory Medicine, University of Pennsylvania; Philadelphia, PA, USA.

*Corresponding author.

Email: cmcmilla@pennmedicine.upenn.edu

One Sentence Summary: We identified genotypic drivers of individual-level clinical and pathological features across heterogeneous TDP-43 proteinopathies.

Abstract: The contribution of genotypes to the shared and disparate clinical and pathologic features observed across TDP-43 proteinopathies is unclear. TDP-43 pathology can contribute to very different neurodegenerative conditions, including frontotemporal lobar degeneration with TDP-43 (FTLD-TDP), amyotrophic lateral sclerosis (ALS), and hippocampal sclerosis of aging (HS-Aging). Despite mounting evidence for shared genetic risk across these TDP-43 proteinopathies, the drivers of individual-level presentations of these syndromes are unknown. We identified data-driven modules of correlated single nucleotide polymorphisms using GWAS summary statistics. These genomic features were either phenotype specific or shared across TDP-43 proteinopathy phenotypes and may contribute to phenotypic heterogeneity in these cases. Pathway analysis revealed biologically meaningful associations within the identified modules, including common nervous system processes and distinct, phenotype-specific associations. We derived module-specific polygenic scores and hypothesized that they would contribute to individual-level differences in the clinical presentation and corresponding anatomic distribution of pathology. Within characteristic FTLD-TDP brain regions, higher FTLD-TDP polygenic risk associated with higher TDP-43 burden. Further, module-specific ALS and FTLD-TDP polygenic risk associated with clinical phenotypic heterogeneity, even within the context of autosomal dominant mutation carriers. We suggest that genotypic variation across TDP-43 proteinopathies may contribute to individual-level presentations of these syndromes.

Main Text:

INTRODUCTION

TAR DNA-binding protein ~43kDa (TDP-43) inclusions are the pathological hallmark of frontotemporal lobar degeneration with TDP-43 (FTLD-TDP), >98% of amyotrophic lateral sclerosis (ALS), and ~90% of hippocampal sclerosis of aging (HS-Aging).*(1–4)* Despite shared pathological features, individuals with TDP-43 proteinopathies present with heterogeneous clinical features including cognitive/behavioral impairments (i.e., FTLD-TDP) and motor neuron dysfunction (i.e., ALS). TDP-43 pathology and neurodegeneration are observed in characteristic neuroanatomical regions that correlate with the clinical presentations across these phenotypes. Some individuals present with both cognitive/behavioral and neuromuscular impairment (i.e., ALS-FTD) which, together with the shared pathological features support the notion that FTLD-TDP and ALS are part of a clinicopathologic spectrum.*(5)* Alternatively, individuals with HS-Aging typically present with an amnestic syndrome with an older age of onset.*(6–8)* Many efforts have focused on the identification of genetic variants that predispose individuals to FTLD-TDP, ALS, or HS-Aging. However, the contribution of genotypes to the shared and disparate clinical and pathologic features observed across TDP-43 proteinopathies is unclear.

Large scale case-control genome wide association studies (GWAS) have focused on the identification of single nucleotide polymorphisms (SNPs) associated with risk for ALS,*(9, 10)* FTLD-TDP,*(11)* or HS-Aging.*(3, 12)* Several large-scale GWAS established ALS as a complex, polygenic disease.*(9, 10, 13)* The first GWAS in FTLD-TDP cases identified three SNPs spanning *TMEM106B* that reached genome-wide significance and several studies have validated the importance of these loci in FTLD pathogenesis.*(14–19)* GWAS of HS-Aging have also implicated several SNPs, including those in *TMEM106B.(3, 12)* Additional efforts have aimed to uncover a shared genetic basis to explain the biological mechanisms that drive susceptibility in FTLD-TDP and ALS. A meta-analysis that jointly analyzed ALS and FTLD-TDP cases identified significant associations in *UNC13A(20)* which had previously only been identified as a susceptibility SNP for ALS and has since been implicated in an FTLD-TDP GWAS.*(21)* Another study used a genetic enrichment approach that found an up to 300-fold enrichment of the ALS phenotype as a function of the increased association with FTLD-TDP, indicating these phenotypes are genetically related.*(22)* However, conjunction analyses performed to identify which SNPs were jointly associated with increased risk for ALS and FTLD-TDP cases only revealed a SNP that is used as a surrogate marker for *C9orf72* repeat expansions, which can cause FLTD-TDP and/or ALS even within the same family.*(23, 24)*

Despite mounting evidence for shared genetic risk, considerably less effort has focused on the disparate, phenotype-specific genetic features across TDP-43 proteinopathies and whether these genotypes contribute to distinct clinical presentations. Rather than focusing only on overlapping risk factors, identifying disparate risk alleles may reveal specific neuroanatomic and/or cellular vulnerabilities that contribute to our understanding of disease-specific drivers. Using existing GWAS summary statistics with novel bioinformatics strategies, we aimed to identify the shared and disparate genomic features that drive individual-level presentations of these syndromes.

We compared the relative associations of case-control GWAS summary statistics for ALS, FTLD-TDP, and HS-aging to identify data-driven modules of highly correlated SNPs that may contribute to the phenotypic heterogeneity across these neurodegenerative diseases. We identified the biological pathways associated with these modules. We then related these module SNPs to independently evaluated individual-level features, including clinical presentation and the anatomic distribution of pathology across brain regions, in a well-characterized sample of more than 250 individuals with ALS and/or FTLD-TDP, through the construction of module-specific polygenic risk scores (PRS). We identified modules of genotypic contributions to selective vulnerability in the anatomic distribution of TDP-43 associated disease, and, for the first time, established potential genetic contributions to risk for either FTLD-TDP or ALS within individuals with *C9orf72* repeat expansions.

RESULTS

Data-driven modules of correlated SNPs identified using summary statistics

We first aimed to identify the shared and disparate genomic features that drive individual-level cognitive/behavioral and/or neuromuscular presentations of TDP-43 proteinopathies. To accomplish this we compared the relative associations of existing GWAS summary statistics for published case-control studies of TDP-43 proteinopathies (including FTLD-TDP*(11)* and ALS,*(9)* as well as HS-Aging,*(12)* which is enriched for TDP-43*(2–4)*) to derive modules of associated SNPs using weighted correlation analysis. Following alignment (see Materials and Methods), we identified 494,417 SNPs commonly genotyped across these GWAS summary statistics. Beta coefficients for each of the three studies were transformed into z-scores to reduce the effects of sample size differences, and the top 1% of SNPs (see Materials and Methods) with the strongest associations to any GWAS phenotype (by |z|, n= 4,945) were selected for analysis (Fig. S1).

Weighted correlation network analysis, performed using the WGCNA package in R*(25)*, identified 5 modules of correlated SNPs associated with either a single GWAS phenotype (accordingly referred to as M1 ALS, M2 FTLD-TDP, and M3 HS-Aging from now on) or phenotypic pairs (accordingly referred to as M4 ALS-FTLD and M5 FTLD-HS from now on, Fig. 1). This approach uses the patterns of associations of SNPs with the respective GWAS phenotypes to identify correlations between SNPs and construct modules based on those correlations. Module assignment was robust, particularly for SNPs associated with a single GWAS phenotype (Fig. S2). After modules were identified, SNPs were LD pruned within modules to reduce redundancy. After pruning, M1 ALS contained 1,025 SNPs, M2 FTLD-TDP contained 1,021 SNPs, M3 HS-Aging contained 646 SNPs, M4 ALS-FTLD contained 80 SNPs, and FTLD-HS contained 72 SNPs (Fig. S3). These data-driven modules of highly correlated SNPs may contribute to phenotypic heterogeneity across TDP-43 proteinopathies, and in the remaining sections we identify the biological pathways associated with these module SNPs and test for associations between module-specific polygenic risk and phenotypic heterogeneity.

Pathway analysis revealed biologically plausible associations for module SNPs

Beyond individual SNPs, it is likely that biological pathways are involved in the development of complex neurodegenerative diseases. To investigate the biological pathway associations of each of our identified modules, we performed a gene ontology enrichment analysis for each module using gprofiler2.*(26)* Enrichment was determined by overrepresentation test, which calculates an expected value by matching the proportion of genes in a particular annotation data category in the human genome reference list to the number of genes provided. Module-specific overrepresented pathways suggest the identified modules are biologically meaningful (FDR $p<.05$, Fig. 2 and Data file S1). Notably, M1 ALS uniquely included pathways related to motor neurons and spinal cord, while neurogenesis-related terms were among the top overrepresented pathways in M3 HS-Aging. Additionally, many pathways that were common across modules related to nervous system processes that may be broadly implicated across neurodegenerative diseases (Fig. S4). Together, these pathway associations support the presence of differentiated and biologically meaningful genetic contributions within the identified modules. Further, they reinforce the potential for shared susceptibility across diseases, while pointing to potential disparate drivers of ALS, FTLD-TDP, and HS-Aging.

Module-specific polygenic risk associated with individual-level phenotypic heterogeneity

While the identified data-driven modules suggest phenotype-specific SNPs which may contribute to our understanding of phenotypic heterogeneity across TDP-43 proteinopathies, these summary-level results alone cannot address individual-level risk. To that end, we focused on these phenotype-specific SNPs and hypothesized that they would contribute to individual-level differences in both clinical presentation and anatomic distribution of pathology within the ALS-FTD clinicopathologic spectrum, where individuals have heterogeneous pathological, clinical, and genetic features.

To ensure a TDP-43 proteinopathy, only individuals who met neuropathological criteria for FTLD-TDP or ALS or had a known TDP-43 mutation were included for analysis (see Materials and Methods, Table 1). Within known TDP-43 proteinopathy cases, individuals were characterized based on clinical features of evidence for neuromuscular impairment (MND-only), cognitive and/or behavioral impairment (Cog-only), or any evidence of a combination of impairment (Cog+MND) at any point in the natural history of disease.

Module-specific polygenic risk scores (PRS) were constructed for M1 ALS and M2 FTLD-TDP as the weighted sum of alleles per SNP per module for each individual, centered and mean scaled (see Materials and Methods). The difference between an individual's M2 FTLD-TDP risk and M1 ALS risk was evaluated to assess relative risk.

Polygenic risk associates with the anatomic distribution of pathology

TDP-43 inclusions are observed in characteristic anatomic regions correlating with phenotype. To assess whether relative polygenic risk was associated with regional TDP-43 deposition across the ALS-FTD spectrum, the difference between M2 FTLD-TDP risk and M1 ALS risk was tested for association with ordinal measures of TDP-43 burden. First, this was done in characteristic FTLD-TDP regions, specifically the cingulate gyrus, angular gyrus, middle frontal cortex, and superior/middle temporal cortex (Fig. S5). One-way ANOVAs revealed that higher M2 FTLD-TDP risk, relative to M1 ALS risk, was associated with higher TDP-43 burden in each FTLD-TDP characteristic region (all FDR adjusted $p < .05$, Fig. 3), controlling for age at death, sex, and inclusion in the prior ALS or FTLD-TDP GWAS (see Materials and Methods). Second, we tested whether relative polygenic risk was associated with ordinal measures of TDP-43 burden in characteristic ALS regions, motor cortex and spinal cord. However, no significant differences in relative PRS by TDP-43 burden were observed in either ALS characteristic region. These observations suggest that these data-driven module-specific polygenic risk scores relate to characteristic individual-level differences in the anatomic distribution of TDP-43 within the ALS-FTD clinicopathologic spectrum.

Polygenic risk associates with phenotypic heterogeneity within mutation carriers at risk for MND and/or Cognitive Impairment

As our sample contained both sporadic and familial cases, module-specific polygenic risk was subsequently specifically tested for association with phenotypic heterogeneity within known *C9orf72*, *GRN*, *TARDBP*, *TBK1*, and *VCP* mutation carriers, where autosomal dominant inherited mutations may primarily drive disease more than underlying polygenic risk. One-way ANOVAs revealed that the M1 ALS PRS was higher in mutation carriers with MND (MND-only or Cog+MND, compared to Cog-only; $F(1,125)=[6.87]$, $p=0.01$), while the M2 FTLD-TDP PRS was higher in mutation carriers with cognitive and/or behavioral impairment (Cog-only or Cog+MND, compared to MND-only; ($F(1,125)=[12.47]$, $p<0.001$), controlling for last contact age, sex, and inclusion in the prior ALS or FTLD-TDP GWAS, respectively (Fig. S6). This suggests that even in the context of autosomal dominant disease, genetic variation associates with risk of a specific impairment including M1 ALS risk for motor disease and M2 FTLD-TDP risk for cognitive disease.

Further limiting the sample to *C9orf72* cases which can develop ALS, FTLD-TDP, or both, a one-way ANOVA revealed that the M1 ALS PRS was higher in *C9orf72* carriers with MND (MND-only or Cog+MND, compared to Cog-only ($F(1,104)=[4.2]$, $p=0.04$), while the M2 FTLD-TDP PRS was higher in *C9orf72* carriers with cognitive and/or behavioral impairment (Cog-only or Cog+MND, compared to MND-only; $F(1,104)=[8.62]$, $p=0.004$), Fig. 4), controlling for last contact age, sex, and inclusion in the prior ALS or FTLD-TDP GWAS, respectively. This provides individual-level evidence that each module independently contributes to risk for a motor or cognitive syndrome in the context of a *C9orf72* expansion.

DISCUSSION

In this study, we hypothesized that both shared and disparate genetic variants would be identified across TDP-43 proteinopathies which contribute to the heterogenous individual-level clinical and pathological features of ALS and FTLD-TDP. We identified data-driven modules of correlated SNPs that were either phenotype specific (M1 ALS, M2 FTLD-TDP, and M3 HS-Aging) or shared across TDP-43 proteinopathy phenotypes (M4 ALS-FTLD and M5 FTLD-HS). Leveraging commonly genotyped SNPs across existing independent GWAS allowed for the identification of patterns of correlations between SNPs and the respective GWAS phenotypes. Though performed at the summary statistic level, this analysis suggested phenotype-specific module SNPs which were then shown to associate with phenotypic heterogeneity at the individual level, including pathological burden in phenotype-specific anatomic regions and individual risk of a motor or cognitive clinical presentation.

Gene ontology enrichment analysis of module SNPs provided plausible reinforcement for the presence of both shared and phenotype-specific biological pathways across TDP-43 proteinopathies. Terms relevant to nervous system development, such as neuron projection development and morphogenesis, as well as those related to synapse structure, organization, activity, and assembly were common across M1 ALS, M2 FTLD-TDP, and M3 HS-Aging, as would be expected in neurodegenerative diseases. However, top overrepresented pathways also support disease-specific drivers. Several of the top overrepresented pathways in M1 ALS related to the selective vulnerability of motor neurons and spinal cord in ALS. Among these were terms related to motor neuron migration, including spinal cord motor neuron migration and lateral motor column neuron migration. This analysis also identified several terms in M1 ALS related to axonogenesis, including central nervous system projection neuron axonogenesis and specifically spinal cord ventral commissure morphogenesis, as well as anterior/posterior axon guidance and negative regulation of collateral sprouting. Further, detailed molecular studies have identified inflammatory responses following neurodegeneration in ALS,*(27, 28)* and have implicated inflammatory pathways as contributors to further neurodegeneration.*(29, 30)* Gene expression studies in ALS have identified an upregulation of immune and inflammatory genes in disease relevant regions*(31–34)*. Here, we observed pathways related to the response to cytokines in M1 ALS, including gene ontology terms such as response to stem cell factor, cellular response to stem cell factor stimulus, cytokine-mediated signaling pathway, Kit signaling pathway, response to chemokine, and the regulation of chemokine-mediated signaling pathway. The top overrepresented pathway in M1 ALS was positive regulation of retrograde transport, endosome to Golgi. Golgi changes are well documented in ALS and one of the most consistent pathologic findings in motor neuron disease.*(35–37)* Finally, M1 ALS included terms related to the regulation of metabolic process, including regulation of kinase activity, regulation of hexokinase activity, and regulation of glucokinase activity. Several lines of evidence support the role of kinases in ALS pathogenesis, including the identification of kinase-encoding genes such as *TBK1*, mutations in which can cause ALS.*(38)* In M2 FTLD-TDP, several of the top overrepresented pathways related to cytoskeleton organization, including actin filament network formation, actin filament bundle organization, and actin filament bundle assembly. Differential expression of genes relating to the structure and function of the cytoskeleton have been observed in FTLD-TDP, and the cytoskeleton has been implicated more broadly in aging and neurodegeneration.*(39–42)* For M3 HS-Aging, we observed several overrepresented terms related to neurogenesis and neuron development, including neuron recognition, which was

among the top overrepresented pathways for this module. As the hippocampus is the site of adult neurogenesis, the inclusion of these pathways align with the selective vulnerability of the hippocampus in HS-aging.*(43, 44)*

Within the ALS-FTD spectrum, the genetic architecture is complex. In addition to shared causal mutations, such as the *C9orf72* repeat expansion, shared significant associations have been identified in *UNC13A*.*(20, 23, 24)* Similarly, *TMEM106B* has been implicated for risk in both FTLD-TDP and HS-Aging.*(3, 11, 12)* Combined studies across TDP-43 proteinopathies have largely focused on identifying shared genetic risk factors that would explain susceptibility across phenotypes through a common underlying biological mechanism. However, the identification of phenotype-specific module SNPs allowed us to demonstrate the contribution of several genetic variants to individual-level pathological and clinical features through the construction of module-specific PRS. Specifically, PRS analyses showed that higher M2 FTLD-TDP risk, relative to M1 ALS risk, was associated with higher TDP-43 burden in FTLD-TDP characteristic regions including cingulate gyrus, angular gyrus, middle frontal cortex, and superior/middle temporal cortex. Orbito-frontal regions are affected early in FTLD-TDP, even when the overall burden of pathology is lower, though more widespread cortical pathology is observed in advanced cases.*(45)* In contrast, these regions are either spared or affected later in the course of ALS, where pathology spreads from motor neurons.*(46)* While no significant differences in relative PRS by TDP-43 burden were observed in either ALS characteristic region evaluated, it is possible that this is confounded by neuron loss, particularly for the spinal cord where neuronal dropout can lower TDP-43 burden.

It is currently unknown why some individuals with TDP-43 pathology develop a motor disorder while others develop a cognitive disorder. Both the M1 ALS PRS and the M2 FTLD-TDP PRS showed associations with a motor or cognitive clinical presentation, respectively, even within known mutation carriers, where disease may be less influenced by underlying polygenic risk. Strikingly, even within individuals with a *C9orf72* repeat expansion, for which there is no current way to predict the development of ALS, FTLD-TDP, or both, the M1 ALS PRS and M2 FTLD-TDP PRS associated with individual-level clinical presentations. While further validation is required, this observation has important implications for predictive genetic testing and management of clinical care. Moreover, current clinical trials focus on either ALS or FTLD-TDP populations in the absence of knowledge about individual risk. These novel PRS may therefore be important for prognostication and stratification in the context of interpreting outcomes in a combined trial. *(47–50)*

This study is limited by the complexity of diagnosis across TDP-43 proteinopathies. A more complete clinical evaluation of motor dysfunction in suspected FTLD-TDP cases and cognitive/behavioral impairment in ALS cases is needed to better characterize these individuals. Additional challenges to complete, accurate diagnoses in the FTD-ALS spectrum include a lack of TDP-43 biomarkers and often rapid progression which may preclude evaluation of the full clinical picture. Thus, there is a reliance on autopsy. However, the regional distribution of pathology alone cannot determine the presence of a secondary motor or cognitive/behavioral syndrome within FTLD-TDP or ALS. Therefore, our individual-level analyses focused on either ordinal measures of TDP-43 or clinical motor and/or cognitive/behavioral features, rather than separate diagnostic groups.

Further, even though the vast majority of HS-Aging cases have TDP-43 pathology, the pathological definition does not require the presence TDP-43. The more recently defined Limbic-predominant age-related TDP-43 encephalopathy (LATE) relies only on the presence of TDP-43 in the amygdala and subsequently hippocampus and middle frontal gyrus.*(51)* As hippocampal sclerosis is neither sufficient nor necessary for a LATE diagnosis, future work including a LATE GWAS is needed to disentangle the genotype associations with HS-aging from limbic TDP-43. In the present study we were unable to evaluate the relationship between M3 HS-Aging SNPs and individual-level differences in either clinical presentation or anatomic distribution of pathology due to a lack of directionality information in the published GWAS summary statistics.

The threshold for inclusion of SNPs initially selected for analysis from the GWAS summary statistics was arbitrary. This threshold (the top 1% of SNPs with the strongest associations to any GWAS phenotype) is lenient by GWAS standards, including SNPs with summary statistic p-values as high as 0.003. However, strict GWAS correction would leave too few SNPs to compare across phenotypes, and sub-threshold associations are still likely informative. However, relaxing the threshold beyond a certain point would introduce non-meaningful SNPs which have no associations with any of the phenotypes. Importantly, we show that the SNPs chosen at this threshold are biologically meaningful with converging evidence from our summary statistic and individual level results.

Finally, similar to limitations of the vast majority of PRS studies, the module-specific polygenic scores were both predominantly derived from and tested on individuals of European ancestry. Therefore, future work is required to validate the broader applicability of these findings in a diverse population.

Overall, examining GWAS summary statistics revealed modules of both shared and disparate correlated SNPs across phenotypes associated with TDP-43 proteinopathies. These modules have biologically plausible pathway associations, and module SNPs relate to individual-level clinical and neuropathological features providing suggestive evidence of a potentially predictive and prognostic genetic marker of cognitive and/or motor risk. We suggest that genotypic variation across TDP-43 proteinopathies may contribute to individual-level presentations of these syndromes.

MATERIALS AND METHODS

Module identification and annotation

Summary Statistic Data

Summary statistics for published case-control FTLD-TDP,*(11)* ALS,*(9)* and HS-Aging*(12)* GWAS were used to derive modules of associated SNPs using weighted correlation analysis. The FTLD-TDP GWAS included 515 FTLD-TDP individuals of European descent with dementia (with or without MND) and either a pathological diagnosis of FTLD-TDP (n=499) or, in the case of living individuals, a pathogenic GRN mutation (n=16). The ALS GWAS included a combined discovery sample of 20,806 sporadic and familial ALS cases. 12,577 cases were included from

summary statistics downloaded from van Rheenen et al*(13)*. The remaining 8,229 cases were of non-Hispanic white race/ethnicity and were diagnosed with ALS according to the El Escorial criteria*(52)*. The HS-Aging GWAS evaluated 310 cases of HS-Aging as defined by presence vs. absence at pathological evaluation.

Only SNPs that were common across GWAS were considered for analysis. FTLD-TDP summary statistics were transformed from NCBI36/hg18 into GRCh37/hg19 using liftOver*(53)* for comparison with the ALS and HS-Aging summary statistics. GWAS statistics for each of the three studies were transformed into z-scores to reduce the effects of sample size differences, and only the absolute value of the z-score was used for analysis. For the ALS GWAS, z-scores were calculated as the effect divided by the standard error. For the FTLD-TDP and HS-Aging GWAS, z-scores were calculated as $|z|=|\Phi-1(p/2)|$ where $\Phi-1$ is the inverse cumulative distribution function of the normal distribution, and p is the p-value. The top 1% of SNPs with the strongest associations to any GWAS phenotype (by $|z|$) were selected for analysis (Fig. S1).

Weighted Correlation Analysis

The WGCNA package in R*(25)* was used to perform WGCNA. This approach uses the patterns of associations of SNPs with the respective GWAS phenotypes to identify correlations between SNPs and construct modules based on those correlations. First, an adjacency matrix is constructed using the strength of the association between each SNP and GWAS phenotype. From the adjacency matrix values, a topological overlap matrix is created to alleviate the effect of noise, and this is converted into a distance matrix which is used as the basis for clustering. Modules are identified with unsupervised clustering (i.e., without the use of a priori defined gene or SNP sets), specifically hierarchical clustering, along with dynamic tree cut.

As only SNPs that were common across GWAS were considered for analysis, there were 0 SNPs removed due to missingness. Modules were constructed with a minimum module size of 30 SNPs. Modules whose distance was less than 0.25 were merged. To test for module robustness, we created 100 samples, each containing a random 80% of the SNPs selected for weighted correlation analysis, without replacement (Fig. S2). For each resample, WGCNA was performed as in the full sample.

After modules were identified, multialleleic SNPs and SNPs with > 10% sample missingness in the individual level genotype data (see Genotype Data) were removed, and SNPs were pruned within modules using an R^2 threshold of 0.3 with LDlinkR.*(54)*

Module Annotation

To annotate, module SNPs were mapped to overlapping genes using the biomaRt package in R*(55)*, specifically ENSEMBL_MART_SNP. A gene ontology enrichment analysis for biological processes was performed separately for each module using gprofiler2. *(26)*

Enrichment was determined by overrepresentation test, which calculates an expected value by matching the proportion of genes in a particular annotation data category in the human genome reference list to the number of genes provided and calculates a fold enrichment value by comparing the true number of genes in a particular annotation data category to the expected value. A p-value was calculated for each term with the hypergeometric test followed by

correction for multiple testing using the False Discovery Rate (FDR), calculated by the Benjamini-Hochberg procedure. Pathways that met FDR p<.05 were examined for each module.

Individual- level data

Case selection

To ensure a TDP-43 proteinopathy, only individuals who met neuropathological criteria for FTLD-TDP or ALS or had a known TDP-43 mutation were included for analysis, described in detail below. *SOD1* ALS cases and *VCP* tauopathy cases were excluded to focus on TDP-43 proteinopathy. Within known TDP-43 proteinopathy cases, individuals were characterized based on neuromuscular impairment (MND-only), cognitive and/or behavioral impairment (Cog-only), and any evidence of a combination of impairment (Cog+MND) at any point in the natural history of disease.

Genotype Data

Individual level imputed genotype data, derived from a variety of prior GWAS array genotyping chips, was available in the Penn Integrated Neurodegenerative Disease Database*(56)*. To harmonize data, we performed imputation of allele dosages on samples passing QC using the Michigan Imputation Server*(57)* with the TOPMed reference panel based on >97K sequenced samples.*(58)* Prior to genotype analyses or imputation, all genotype data were QCed using standard metrics including but not limited to sex discrepancy, sample relatedness, heterozygosity, and Hardy-Weinberg equilibrium.*(59)*

Genetic Sequencing

DNA was extracted from peripheral blood or frozen brain tissue following the manufacturer's protocols (QuickGene DNA whole blood kit (Autogen) for blood, and QIAamp DNA Mini Kit (Qiagen) for brain tissue). All individuals were tested for *C9orf72* hexanucleotide repeat expansions using a modified repeat-primed PCR as previously described *(60)* and screened for mutations associated with FTD and/or ALS using whole exome/genome sequencing, and/or a custom targeted multi neurodegenerative disease sequencing panel *(61)* which included *SOD1(62)*, *TBK1(63)*, *TARDBP (64)*, *VCP (65)*, and *GRN(66, 67)*). The sequencing data was analyzed using Geneticist Assistant software (Soft Genetics, State College, PA).

Neuropathological Data

Detailed neuropathological assessments were performed using established and uniform methods of fixation, tissue processing, IHC with well-characterized antibodies, and current neuropathological criteria, as has been described in detail elsewhere.*(61, 68, 69)* Briefly, 15 brain regions from one hemisphere, alternating right and left at random, are routinely sampled at autopsy, formalin-fixed, and processed for immunohistochemical staining using 1D3 for phosphorylated TDP-43. Each brain region was semi-quantitatively scored for the severity of TDP-43 (0, absent; 0.5, rare; 1, mild; 2, moderate; 3, severe). ALS was diagnosed based on the

presence of upper and lower motor neuron degeneration, and TDP-43 subtypes were determined based on TDP-43 immunohistochemistry as previously described.*(70)*

Polygenic risk score calculation

To assess whether module-specific polygenic risk was associated with phenotypic heterogeneity, we limited our analysis to modules associated with a single GWAS phenotype (M1 ALS, M2 FTLD-TDP, and M3 HS-Aging), rather than phenotypic pairs (M4 ALS-FTLD and M5 FTLD-HS), in an effort to focus on phenotype-specific drivers. Due to a lack of directionality information in the provided summary statistics, we were unable to calculate a PRS for M3 HS-Aging.

Module-specific PRS were constructed for M1 ALS and M2 FTLD-TDP as the weighted sum of alleles per SNP per module for each individual, using z-score transformed associations with the phenotype from the respective GWAS, ALS or FTLD-TDP. The effect allele was labelled as 'Allele1' in the ALS summary statistics and was the minor allele in the FTLD-TDP summary statistics. Here, we assigned minor allele using the European 1,000 Genomes sample allele frequencies. The direction of the weight for each SNP was determined by whether the dose of the effect (direction consistent with summary statistic) or non-effect (direction flipped) allele was provided in the individual-level data.

Each PRS was centered using the European 1,000 Genomes sample allele frequencies and mean scaled. Missing allele dosages were replaced with the European 1,000 Genomes sample average. SNPs without allele frequency information were removed. The M1 ALS PRS included 1021 SNPs and the M2 FTLD-TDP PRS included 1013 SNPs. Individuals with >10% SNP missingness for either PRS were removed. This resulted in the removal of 130 individuals.

PRS were tested for association with presentation and regional postmortem measures of TDP-43 burden using one-way ANOVAs controlling for age at last contact and sex. Some individuals in the sample were included in a prior ALS or FTLD-TDP GWAS for which summary statistics were used to generate the modules (Table 1). Therefore, inclusion in a prior ALS or FTLD-TDP GWAS was additionally controlled for in all ANOVAs.

List of Supplementary Materials

Fig. S1 to S6

Data file S1

References

1. M. Neumann, D. M. Sampathu, L. K. Kwong, A. C. Truax, M. C. Micsenyi, T. T. Chou, J. Bruce, T. Schuck, M. Grossman, C. M. Clark, L. F. McCluskey, B. L. Miller, E. Masliah, I. R. Mackenzie, H. Feldman, W. Feiden, H. A. Kretzschmar, J. Q. Trojanowski, V. M.-Y. Lee, Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science* **314**, 130–133 (2006).

2. C. Amador-Ortiz, Z. Ahmed, C. Zehr, D. W. Dickson, Hippocampal sclerosis dementia differs from hippocampal sclerosis in frontal lobe degeneration. *Acta Neuropathol* **113**, 245–252 (2007).

3. P. T. Nelson, S. Estus, E. L. Abner, I. Parikh, M. Malik, J. H. Neltner, E. Ighodaro, W.-X. Wang, B. R. Wilfred, L.-S. Wang, W. A. Kukull, K. Nandakumar, M. L. Farman, W. W. Poon, M. M. Corrada, C. H. Kawas, D. H. Cribbs, D. A. Bennett, J. A. Schneider, E. B. Larson, P. K. Crane, O. Valladares, F. A. Schmitt, R. J. Kryscio, G. A. Jicha, C. D. Smith, S. W. Scheff, J. A. Sonnen, J. L. Haines, M. A. Pericak-Vance, R. Mayeux, L. A. Farrer, L. J. Van Eldik, C. Horbinski, R. C. Green, M. Gearing, L. W. Poon, P. L. Kramer, R. L. Woltjer, T. J. Montine, A. B. Partch, A. J. Rajic, K. Richmire, S. E. Monsell, G. D. Schellenberg, D. W. Fardo, ABCC9 gene polymorphism is associated with hippocampal sclerosis of aging pathology. *Acta Neuropathol* **127**, 825–843 (2014).

4. S. Nag, L. Yu, A. W. Capuano, R. S. Wilson, S. E. Leurgans, D. A. Bennett, J. A. Schneider, Hippocampal sclerosis and TDP-43 pathology in aging and Alzheimer disease. *Ann Neurol* **77**, 942–952 (2015).

5. M. J. Strong, S. Abrahams, L. H. Goldstein, S. Woolley, P. Mclaughlin, J. Snowden, E. Mioshi, A. Roberts-South, M. Benatar, T. HortobáGyi, J. Rosenfeld, V. Silani, P. G. Ince, M. R. Turner, Amyotrophic lateral sclerosis - frontotemporal spectrum disorder (ALS-FTSD): Revised diagnostic criteria. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* **18**, 153–174 (2017).

6. D. W. Dickson, P. Davies, C. Bevona, K. H. Van Hoeven, S. M. Factor, E. Grober, M. K. Aronson, H. A. Crystal, Hippocampal sclerosis: a common pathological feature of dementia in very old (> or = 80 years of age) humans. *Acta Neuropathol* **88**, 212–221 (1994).

7. J. Corey-Bloom, M. N. Sabbagh, M. W. Bondi, L. Hansen, M. F. Alford, E. Masliah, L. J. Thal, Hippocampal sclerosis contributes to dementia in the elderly. *Neurology* **48**, 154–160 (1997).

8. J. B. Leverenz, C. M. Agustin, D. Tsuang, E. R. Peskind, S. D. Edland, D. Nochlin, L. DiGiacomo, J. D. Bowen, W. C. McCormick, L. Teri, M. A. Raskind, W. A. Kukull, E. B. Larson, Clinical and neuropathological characteristics of hippocampal sclerosis: a community-based study. *Arch Neurol* **59**, 1099–1106 (2002).

9. A. Nicolas, K. P. Kenna, A. E. Renton, N. Ticozzi, F. Faghri, R. Chia, J. A. Dominov, B. J. Kenna, M. A. Nalls, P. Keagle, A. M. Rivera, W. van Rheenen, N. A. Murphy, J. J. F. A. van Vugt, J. T. Geiger, R. A. Van der Spek, H. A. Pliner, Shankaracharya, B. N. Smith, G. Marangi, S. D. Topp, Y. Abramzon, A. S. Gkazi, J. D. Eicher, A. Kenna, G. Mora, A. Calvo, L. Mazzini, N. Riva, J. Mandrioli, C. Caponnetto, S. Battistini, P. Volanti, V. La Bella, F. L. Conforti, G. Borghero, S. Messina, I. L. Simone, F. Trojsi, F. Salvi, F. O. Logullo, S. D'Alfonso, L. Corrado, M. Capasso, L. Ferrucci, C. de A. M. Moreno, S. Kamalakaran, D. B. Goldstein, A. D. Gitler, T. Harris, R. M. Myers, H. Phatnani, R. L. Musunuri, U. S. Evani, A. Abhyankar, M. C. Zody, J. Kaye, S. Finkbeiner, S. K. Wyman, A. LeNail, L. Lima, E. Fraenkel, C. N. Svendsen, L. M. Thompson, J. E. Van Eyk, J. D. Berry, T. M. Miller, S. J. Kolb, M. Cudkowicz, E. Baxi, M. Benatar, J. P. Taylor, E. Rampersaud, G. Wu, J. Wuu, G. Lauria, F. Verde, I. Fogh, C. Tiloca, G. P. Comi, G. Sorarù, C. Cereda, P. Corcia, H. Laaksovirta, L. Myllykangas, L. Jansson, M. Valori, J. Ealing, H. Hamdalla, S. Rollinson, S. Pickering-Brown, R. W. Orrell, K. C. Sidle, A. Malaspina, J. Hardy, A. B. Singleton, J. O. Johnson, S. Arepalli, P. C. Sapp, D. McKenna-Yasek, M. Polak, S. Asress, S. Al-Sarraj, A. King, C.

Troakes, C. Vance, J. de Belleroche, F. Baas, A. L. M. A. ten Asbroek, J. L. Muñoz-Blanco, D. G. Hernandez, J. Ding, J. R. Gibbs, S. W. Scholz, M. K. Floeter, R. H. Campbell, F. Landi, R. Bowser, S. M. Pulst, J. M. Ravits, D. J. L. MacGowan, J. Kirby, E. P. Pioro, R. Pamphlett, J. Broach, G. Gerhard, T. L. Dunckley, C. B. Brady, N. W. Kowall, J. C. Troncoso, I. Le Ber, K. Mouzat, S. Lumbroso, T. D. Heiman-Patterson, F. Kamel, L. Van Den Bosch, R. H. Baloh, T. M. Strom, T. Meitinger, A. Shatunov, K. R. Van Eijk, M. de Carvalho, M. Kooyman, B. Middelkoop, M. Moisse, R. L. McLaughlin, M. A. Van Es, M. Weber, K. B. Boylan, M. Van Blitterswijk, R. Rademakers, K. E. Morrison, A. N. Basak, J. S. Mora, V. E. Drory, P. J. Shaw, M. R. Turner, K. Talbot, O. Hardiman, K. L. Williams, J. A. Fifita, G. A. Nicholson, I. P. Blair, G. A. Rouleau, J. Esteban-Pérez, A. García-Redondo, A. Al-Chalabi, E. Rogaeva, L. Zinman, L. W. Ostrow, N. J. Maragakis, J. D. Rothstein, Z. Simmons, J. Cooper-Knock, A. Brice, S. A. Goutman, E. L. Feldman, S. B. Gibson, F. Taroni, A. Ratti, C. Gellera, P. Van Damme, W. Robberecht, P. Fratta, M. Sabatelli, C. Lunetta, A. C. Ludolph, P. M. Andersen, J. H. Weishaupt, W. Camu, J. Q. Trojanowski, V. M. Van Deerlin, R. H. Brown, L. H. van den Berg, J. H. Veldink, M. B. Harms, J. D. Glass, D. J. Stone, P. Tienari, V. Silani, A. Chiò, C. E. Shaw, B. J. Traynor, J. E. Landers, F. O. Logullo, I. Simone, G. Logroscino, F. Salvi, I. Bartolomei, G. Borghero, M. R. Murru, E. Costantino, C. Pani, R. Puddu, C. Caredda, V. Piras, S. Tranquilli, S. Cuccu, D. Corongiu, M. Melis, A. Milia, F. Marrosu, M. G. Marrosu, G. Floris, A. Cannas, S. Tranquilli, M. Capasso, C. Caponnetto, G. Mancardi, P. Origone, P. Mandich, F. L. Conforti, S. Cavallaro, G. Mora, K. Marinou, R. Sideri, S. Penco, L. Mosca, C. Lunetta, G. L. Pinter, M. Corbo, N. Riva, P. Carrera, P. Volanti, J. Mandrioli, N. Fini, A. Fasano, L. Tremolizzo, A. Arosio, C. Ferrarese, F. Trojsi, G. Tedeschi, M. R. Monsurrò, G. Piccirillo, C. Femiano, A. Ticca, E. Ortu, V. La Bella, R. Spataro, T. Colletti, M. Sabatelli, M. Zollino, A. Conte, M. Luigetti, S. Lattante, G. Marangi, M. Santarelli, A. Petrucci, M. Pugliatti, A. Pirisi, L. D. Parish, P. Occhineri, F. Giannini, S. Battistini, C. Ricci, M. Benigni, T. B. Cau, D. Loi, A. Calvo, C. Moglia, M. Brunetti, M. Barberis, G. Restagno, F. Casale, G. Marrali, G. Fuda, I. Ossola, S. Cammarosano, A. Canosa, A. Ilardi, U. Manera, M. Grassano, R. Tanel, F. Pisano, M. B. Harms, D. B. Goldstein, N. A. Shneider, S. Goutman, Z. Simmons, T. M. Miller, S. Chandran, S. Pal, G. Manousakis, S. H. Appel, E. Simpson, L. Wang, R. H. Baloh, S. Gibson, R. Bedlack, D. Lacomis, D. Sareen, A. Sherman, L. Bruijn, M. Penny, A. S. Allen, S. Appel, R. H. Baloh, R. S. Bedlack, B. E. Boone, R. Brown, J. P. Carulli, A. Chesi, W. K. Chung, E. T. Cirulli, G. M. Cooper, J. Couthouis, A. G. Day-Williams, P. A. Dion, S. Gibson, A. D. Gitler, J. D. Glass, D. B. Goldstein, Y. Han, M. B. Harms, T. Harris, S. D. Hayes, A. L. Jones, J. Keebler, B. J. Krueger, B. N. Lasseigne, S. E. Levy, Y.-F. Lu, T. Maniatis, D. McKenna-Yasek, T. M. Miller, R. M. Myers, S. Petrovski, S. M. Pulst, A. R. Raphael, J. M. Ravits, Z. Ren, G. A. Rouleau, P. C. Sapp, N. A. Shneider, E. Simpson, K. B. Sims, J. F. Staropoli, L. L. Waite, Q. Wang, J. R. Wimbish, W. W. Xin, H. Phatnani, J. Kwan, D. Sareen, J. R. Broach, Z. Simmons, X. Arcila-Londono, E. B. Lee, V. M. Van Deerlin, N. A. Shneider, E. Fraenkel, L. W. Ostrow, F. Baas, N. Zaitlen, J. D. Berry, A. Malaspina, P. Fratta, G. A. Cox, L. M. Thompson, S. Finkbeiner, E. Dardiotis, T. M. Miller, S. Chandran, S. Pal, E. Hornstein, D. J. MacGowan, T. Heiman-Patterson, M. G. Hammell, N. A. Patsopoulos, J. Dubnau, A. Nath, J. Kaye, S. Finkbeiner, S. Wyman, A. LeNail, L. Lima, E. Fraenkel, J. D. Rothstein, C. N. Svendsen, L. M. Thompson, J. Van Eyk, N. J. Maragakis, J. D. Berry, J. D. Glass, T. M. Miller, S. J. Kolb, R. H. Baloh, M. Cudkowicz, E. Baxi, M. Benatar, J. P. Taylor, G. Wu, E. Rampersaud, J. Wuu, R. Rademakers, S. Züchner, R. Schule, J. McCauley, S. Hussain, A. Cooley, M. Wallace, C. Clayman, R. Barohn, J. Statland, J. Ravits, A. Swenson, C. Jackson, J. Trivedi, S. Khan, J. Katz, L. Jenkins, T. Burns, K. Gwathmey, J. Caress, C. McMillan, L. Elman, E. Pioro, J. Heckmann, Y. So, D. Walk, S. Maiser, J. Zhang, V. Silani, N. Ticozzi, C. Gellera, A. Ratti, F. Taroni, G. Lauria, F. Verde, I. Fogh, C. Tiloca, G. P. Comi, G. Sorarù, C. Cereda, S. D'Alfonso, L. Corrado, F. De Marchi, S. Corti, M. Ceroni, L. Mazzini, G. Siciliano, M. Filosto, M. Inghilleri, S. Peverelli, C. Colombrita, B. Poletti, L. Maderna, R. Del Bo, S. Gagliardi, G. Querin, C. Bertolin, V. Pensato, B. Castellotti, W. Camu, K. Mouzat, S. Lumbroso, P. Corcia, V. Meininger, G. Besson, E. Lagrange, P. Clavelou, N. Guy, P. Couratier, P. Vourch, V. Danel, E. Bernard, G. Lemasson, A. Al Kheifat, A. Al-Chalabi, P. Andersen, A. N. Basak, I. P. Blair, A. Chio, J. Cooper-Knock, P. Corcia, P. Couratier, M. de Carvalho, A. Dekker, V. Drory, A. G. Redondo, M. Gotkine, O. Hardiman, W. Hide, A. Iacoangeli, J. Glass, K. Kenna, M. Kiernan, M. Kooyman, J. Landers, R. McLaughlin, B. Middelkoop, J. Mill, M. M. Neto, M. Moisse, J. M. Pardina, K. Morrison, S. Newhouse, S. Pinto, S. Pulit, W. Robberecht, A. Shatunov, P. Shaw, C. Shaw, V. Silani, W. Sproviero, G. Tazelaar, N. Ticozzi, P. van Damme, L. van den Berg, R. van der Spek, K. van Eijk, M. van Es, W. van Rheenen, J. van Vugt, J. Veldink, M. Weber, K. L. Williams, M. Zatz, D. C. Bauer, N. A. Twine, Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron* **97**, 1268-1283.e6 (2018).

10. M. A. van Es, J. H. Veldink, C. G. J. Saris, H. M. Blauw, P. W. J. van Vught, A. Birve, R. Lemmens, H. J. Schelhaas, E. J. N. Groen, M. H. B. Huisman, A. J. van der Kooi, M. de Visser, C. Dahlberg, K. Estrada, F. Rivadeneira, A. Hofman, M. J. Zwarts, P. T. C. van Doormaal, D. Rujescu, E. Strengman, I. Giegling, P. Muglia, B. Tomik, A. Slowik, A. G. Uitterlinden, C. Hendrich, S. Waibel, T. Meyer, A. C. Ludolph, J. D. Glass, S. Purcell, S.

Cichon, M. M. Nöthen, H.-E. Wichmann, S. Schreiber, S. H. H. M. Vermeulen, L. A. Kiemeney, J. H. J. Wokke, S. Cronin, R. L. McLaughlin, O. Hardiman, K. Fumoto, R. J. Pasterkamp, V. Meininger, J. Melki, P. N. Leigh, C. E. Shaw, J. E. Landers, A. Al-Chalabi, R. H. Brown, W. Robberecht, P. M. Andersen, R. A. Ophoff, L. H. van den Berg, Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet* **41**, 1083–1087 (2009).

11. V. M. Van Deerlin, P. M. A. Sleiman, M. Martinez-Lage, A. Chen-Plotkin, L.-S. Wang, N. R. Graff-Radford, D. W. Dickson, R. Rademakers, B. F. Boeve, M. Grossman, S. E. Arnold, D. M. A. Mann, S. M. Pickering-Brown, H. Seelaar, P. Heutink, J. C. van Swieten, J. R. Murrell, B. Ghetti, S. Spina, J. Grafman, J. Hodges, M. G. Spillantini, S. Gilman, A. P. Lieberman, J. A. Kaye, R. L. Woltjer, E. H. Bigio, M. Mesulam, S. al-Sarraj, C. Troakes, R. N. Rosenberg, C. L. White, I. Ferrer, A. Lladó, M. Neumann, H. A. Kretzschmar, C. M. Hulette, K. A. Welsh-Bohmer, B. L. Miller, A. Alzualde, A. L. de Munain, A. C. McKee, M. Gearing, A. I. Levey, J. J. Lah, J. Hardy, J. D. Rohrer, T. Lashley, I. R. A. Mackenzie, H. H. Feldman, R. L. Hamilton, S. T. Dekosky, J. van der Zee, S. Kumar-Singh, C. Van Broeckhoven, R. Mayeux, J. P. G. Vonsattel, J. C. Troncoso, J. J. Kril, J. B. J. Kwok, G. M. Halliday, T. D. Bird, P. G. Ince, P. J. Shaw, N. J. Cairns, J. C. Morris, C. A. McLean, C. DeCarli, W. G. Ellis, S. H. Freeman, M. P. Frosch, J. H. Growdon, D. P. Perl, M. Sano, D. A. Bennett, J. A. Schneider, T. G. Beach, E. M. Reiman, B. K. Woodruff, J. Cummings, H. V. Vinters, C. A. Miller, H. C. Chui, I. Alafuzoff, P. Hartikainen, D. Seilhean, D. Galasko, E. Masliah, C. W. Cotman, M. T. Tuñón, M. C. C. Martínez, D. G. Munoz, S. L. Carroll, D. Marson, P. F. Riederer, N. Bogdanovic, G. D. Schellenberg, H. Hakonarson, J. Q. Trojanowski, V. M.-Y. Lee, Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet* **42**, 234–239 (2010).

12. G. W. Beecham, K. Hamilton, A. C. Naj, E. R. Martin, M. Huentelman, A. J. Myers, J. J. Corneveaux, J. Hardy, J.-P. Vonsattel, S. G. Younkin, D. A. Bennett, P. L. De Jager, E. B. Larson, P. K. Crane, M. I. Kamboh, J. K. Kofler, D. C. Mash, L. Duque, J. R. Gilbert, H. Gwirtsman, J. D. Buxbaum, P. Kramer, D. W. Dickson, L. A. Farrer, M. P. Frosch, B. Ghetti, J. L. Haines, B. T. Hyman, W. A. Kukull, R. P. Mayeux, M. A. Pericak-Vance, J. A. Schneider, J. Q. Trojanowski, E. M. Reiman, the Alzheimer's Disease Genetics Consortium (ADGC), G. D. Schellenberg, T. J. Montine, G. Gibson, Ed. Genome-Wide Association Meta-analysis of Neuropathologic Features of Alzheimer's Disease and Related Dementias. *PLoS Genet* **10**, e1004606 (2014).

13. W. van Rheenen, A. Shatunov, A. M. Dekker, R. L. McLaughlin, F. P. Diekstra, S. L. Pulit, R. A. A. van der Spek, U. Võsa, S. de Jong, M. R. Robinson, J. Yang, I. Fogh, P. T. van Doormaal, G. H. P. Tazelaar, M. Koppers, A. M. Blokhuis, W. Sproviero, A. R. Jones, K. P. Kenna, K. R. van Eijk, O. Harschnitz, R. D. Schellevis, W. J. Brands, J. Medic, A. Menelaou, A. Vajda, N. Ticozzi, K. Lin, B. Rogelj, K. Vrabec, M. Ravnik-Glavač, B. Koritnik, J. Zidar, L. Leonardis, L. D. Grošelj, S. Millecamps, F. Salachas, V. Meininger, M. de Carvalho, S. Pinto, J. S. Mora, R. Rojas-García, M. Polak, S. Chandran, S. Colville, R. Swingler, K. E. Morrison, P. J. Shaw, J. Hardy, R. W. Orrell, A. Pittman, K. Sidle, P. Fratta, A. Malaspina, S. Topp, S. Petri, S. Abdulla, C. Drepper, M. Sendtner, T. Meyer, R. A. Ophoff, K. A. Staats, M. Wiedau-Pazos, C. Lomen-Hoerth, V. M. Van Deerlin, J. Q. Trojanowski, L. Elman, L. McCluskey, A. N. Basak, C. Tunca, H. Hamzeiy, Y. Parman, T. Meitinger, P. Lichtner, M. Radivojkov-Blagojevic, C. R. Andres, C. Maurel, G. Bensimon, B. Landwehrmeyer, A. Brice, C. A. M. Payan, S. Saker-Delye, A. Dürr, N. W. Wood, L. Tittmann, W. Lieb, A. Franke, M. Rietschel, S. Cichon, M. M. Nöthen, P. Amouyel, C. Tzourio, J.-F. Dartigues, A. G. Uitterlinden, F. Rivadeneira, K. Estrada, A. Hofman, C. Curtis, H. M. Blauw, A. J. van der Kooi, M. de Visser, A. Goris, M. Weber, C. E. Shaw, B. N. Smith, O. Pansarasa, C. Cereda, R. Del Bo, G. P. Comi, S. D'Alfonso, C. Bertolin, G. Sorarù, L. Mazzini, V. Pensato, C. Gellera, C. Tiloca, A. Ratti, A. Calvo, C. Moglia, M. Brunetti, S. Arcuti, R. Capozzo, C. Zecca, C. Lunetta, S. Penco, N. Riva, A. Padovani, M. Filosto, B. Muller, R. J. Stuit, PARALS Registry, SLALOM Group, SLAP Registry, FALS Sequencing Consortium, SLAGEN Consortium, NNIPPS Study Group, I. Blair, K. Zhang, E. P. McCann, J. A. Fifita, G. A. Nicholson, D. B. Rowe, R. Pamphlett, M. C. Kiernan, J. Grosskreutz, O. W. Witte, T. Ringer, T. Prell, B. Stubendorff, I. Kurth, C. A. Hübner, P. N. Leigh, F. Casale, A. Chio, E. Beghi, E. Pupillo, R. Tortelli, G. Logroscino, J. Powell, A. C. Ludolph, J. H. Weishaupt, W. Robberecht, P. Van Damme, L. Franke, T. H. Pers, R. H. Brown, J. D. Glass, J. E. Landers, O. Hardiman, P. M. Andersen, P. Corcia, P. Vourc'h, V. Silani, N. R. Wray, P. M. Visscher, P. I. W. de Bakker, M. A. van Es, R. J. Pasterkamp, C. M. Lewis, G. Breen, A. Al-Chalabi, L. H. van den Berg, J. H. Veldink, Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat Genet* **48**, 1043–1048 (2016).

14. N. Finch, M. M. Carrasquillo, M. Baker, N. J. Rutherford, G. Coppola, M. Dejesus-Hernandez, R. Crook, T. Hunter, R. Ghidoni, L. Benussi, J. Crook, E. Finger, K. J. Hantanpaa, A. M. Karydas, P. Sengdy, J. Gonzalez, W. W. Seeley, N. Johnson, T. G. Beach, M. Mesulam, G. Forloni, A. Kertesz, D. S. Knopman, R. Uitti, C. L. White, R. Caselli, C. Lippa, E. H. Bigio, Z. K. Wszolek, G. Binetti, I. R. Mackenzie, B. L. Miller, B. F. Boeve, S. G. Younkin, D. W. Dickson, R. C. Petersen, N. R. Graff-Radford, D. H. Geschwind, R. Rademakers, TMEM106B regulates progranulin levels and the penetrance of FTLD in GRN mutation carriers. *Neurology* **76**, 467–474 (2011).

15. M. D. Gallagher, E. Suh, M. Grossman, L. Elman, L. McCluskey, J. C. Van Swieten, S. Al-Sarraj, M. Neumann, E. Gelpi, B. Ghetti, J. D. Rohrer, G. Halliday, C. Van Broeckhoven, D. Seilhean, P. J. Shaw, M. P. Frosch, I. Alafuzoff, A. Antonell, N. Bogdanovic, W. Brooks, N. J. Cairns, J. Cooper-Knock, C. Cotman, P. Cras, M. Cruts, P. P. De Deyn, C. DeCarli, C. Dobson-Stone, S. Engelborghs, N. Fox, D. Galasko, M. Gearing, I. Gijselinck, J. Grafman, P. Hartikainen, K. J. Hatanpaa, J. R. Highley, J. Hodges, C. Hulette, P. G. Ince, L.-W. Jin, J. Kirby, J. Kofler, J. Kril, J. B. J. Kwok, A. Levey, A. Lieberman, A. Llado, J.-J. Martin, E. Masliah, C. J. McDermott, A. McKee, C. McLean, S. Mead, C. A. Miller, J. Miller, D. G. Munoz, J. Murrell, H. Paulson, O. Piguet, M. Rossor, R. Sanchez-Valle, M. Sano, J. Schneider, L. C. Silbert, S. Spina, J. van der Zee, T. Van Langenhove, J. Warren, S. B. Wharton, C. L. White, R. L. Woltjer, J. Q. Trojanowski, V. M. Y. Lee, V. Van Deerlin, A. S. Chen-Plotkin, TMEM106B is a genetic modifier of frontotemporal lobar degeneration with C9orf72 hexanucleotide repeat expansions. *Acta Neuropathol* **127**, 407–418 (2014).

16. A. M. Nicholson, N. A. Finch, A. Wojtas, M. C. Baker, R. B. Perkerson, M. Castanedes-Casey, L. Rousseau, L. Benussi, G. Binetti, R. Ghidoni, G.-Y. R. Hsiung, I. R. Mackenzie, E. Finger, B. F. Boeve, N. Ertekin-Taner, N. R. Graff-Radford, D. W. Dickson, R. Rademakers, TMEM106B p.T185S regulates TMEM106B protein levels: implications for frontotemporal dementia. *J Neurochem* **126**, 781–791 (2013).

17. X. Zhou, A. M. Nicholson, Y. Ren, M. Brooks, P. Jiang, A. Zuberi, H. N. Phuoc, R. B. Perkerson, B. Matchett, T. M. Parsons, N. A. Finch, W. Lin, W. Qiao, M. Castanedes-Casey, V. Phillips, A. L. Librero, Y. Asmann, G. Bu, M. E. Murray, C. Lutz, D. W. Dickson, R. Rademakers, Loss of TMEM106B leads to myelination deficits: implications for frontotemporal dementia treatment strategies. *Brain* **143**, 1905–1919 (2020).

18. M. van Blitterswijk, B. Mullen, A. M. Nicholson, K. F. Bieniek, M. G. Heckman, M. C. Baker, M. DeJesus-Hernandez, N. A. Finch, P. H. Brown, M. E. Murray, G.-Y. R. Hsiung, H. Stewart, A. M. Karydas, E. Finger, A. Kertesz, E. H. Bigio, S. Weintraub, M. Mesulam, K. J. Hatanpaa, C. L. White, M. J. Strong, T. G. Beach, Z. K. Wszolek, C. Lippa, R. Caselli, L. Petrucelli, K. A. Josephs, J. E. Parisi, D. S. Knopman, R. C. Petersen, I. R. Mackenzie, W. W. Seeley, L. T. Grinberg, B. L. Miller, K. B. Boylan, N. R. Graff-Radford, B. F. Boeve, D. W. Dickson, R. Rademakers, TMEM106B protects C9ORF72 expansion carriers against frontotemporal dementia. *Acta Neuropathol* **127**, 397–406 (2014).

19. J. van der Zee, T. Van Langenhove, G. Kleinberger, K. Sleegers, S. Engelborghs, R. Vandenberghe, P. Santens, M. Van den Broeck, G. Joris, J. Brys, M. Mattheijssens, K. Peeters, P. Cras, P. P. De Deyn, M. Cruts, C. Van Broeckhoven, TMEM106B is associated with frontotemporal lobar degeneration in a clinically diagnosed patient cohort. *Brain* **134**, 808–815 (2011).

20. F. P. Diekstra, V. M. Van Deerlin, J. C. van Swieten, A. Al-Chalabi, A. C. Ludolph, J. H. Weishaupt, O. Hardiman, J. E. Landers, R. H. Brown, M. A. van Es, R. J. Pasterkamp, M. Koppers, P. M. Andersen, K. Estrada, F. Rivadeneira, A. Hofman, A. G. Uitterlinden, P. van Damme, J. Melki, V. Meininger, A. Shatunov, C. E. Shaw, P. N. Leigh, P. J. Shaw, K. E. Morrison, I. Fogh, A. Chiò, B. J. Traynor, D. Czell, M. Weber, P. Heutink, P. I. W. de Bakker, V. Silani, W. Robberecht, L. H. van den Berg, J. H. Veldink, C9orf72 and UNC13A are shared risk loci for amyotrophic lateral sclerosis and frontotemporal dementia: a genome-wide meta-analysis. *Ann Neurol* **76**, 120–133 (2014).

21. C. Pottier, Y. Ren, R. B. Perkerson, M. Baker, G. D. Jenkins, M. van Blitterswijk, M. DeJesus-Hernandez, J. G. J. van Rooij, M. E. Murray, E. Christopher, S. K. McDonnell, Z. Fogarty, A. Batzler, S. Tian, C. T. Vicente, B. Matchett, A. M. Karydas, G.-Y. Robin Hsiung, H. Seelaar, M. O. Mol, E. C. Finger, C. Graff, L. Öijerstedt, M. Neumann, P. Heutink, M. Synofzik, C. Wilke, J. Prudlo, P. Rizzu, J. Simon-Sanchez, D. Edbauer, S. Roeber, J. Diehl-Schmid, B. M. Evers, A. King, M.-M. Mesulam, S. Weintraub, C. Geula, K. F. Bieniek, L. Petrucelli, G. L. Ahern, E. M. Reiman, B. K. Woodruff, R. J. Caselli, E. D. Huey, M. R. Farlow, J. Grafman, S. Mead, L. T.

Grinberg, S. Spina, M. Grossman, D. J. Irwin, E. B. Lee, E. Suh, J. Snowden, D. Mann, N. Ertekin-Taner, R. J. Uitti, Z. K. Wszolek, K. A. Josephs, J. E. Parisi, D. S. Knopman, R. C. Petersen, J. R. Hodges, O. Piguet, E. G. Geier, J. S. Yokoyama, R. A. Rissman, E. Rogaeva, J. Keith, L. Zinman, M. C. Tartaglia, N. J. Cairns, C. Cruchaga, B. Ghetti, J. Kofler, O. L. Lopez, T. G. Beach, T. Arzberger, J. Herms, L. S. Honig, J. P. Vonsattel, G. M. Halliday, J. B. Kwok, C. L. White, M. Gearing, J. Glass, S. Rollinson, S. Pickering-Brown, J. D. Rohrer, J. Q. Trojanowski, V. Van Deerlin, E. H. Bigio, C. Troakes, S. Al-Sarraj, Y. Asmann, B. L. Miller, N. R. Graff-Radford, B. F. Boeve, W. W. Seeley, I. R. A. Mackenzie, J. C. van Swieten, D. W. Dickson, J. M. Biernacka, R. Rademakers, Genome-wide analyses as part of the international FTLD-TDP whole genome sequencing consortium reveals novel disease risk factors and increases support for immune dysfunction in FTLD. *Acta Neuropathol* **137**, 879–899 (2019).

22. C. M. Karch, N. Wen, C. C. Fan, J. S. Yokoyama, N. Kouri, O. A. Ross, G. Höglinger, U. Müller, R. Ferrari, J. Hardy, G. D. Schellenberg, P. M. Sleiman, P. Momeni, C. P. Hess, B. L. Miller, M. Sharma, V. Van Deerlin, O. B. Smeland, O. A. Andreassen, A. M. Dale, R. S. Desikan, for the International Frontotemporal Dementia (FTD)–Genomics Consortium, International Collaboration for Frontotemporal Dementia, Progressive Supranuclear Palsy (PSP) Genetics Consortium, and International Parkinson's Disease Genomics Consortium, Selective Genetic Overlap Between Amyotrophic Lateral Sclerosis and Diseases of the Frontotemporal Dementia Spectrum. *JAMA Neurol* **75**, 860 (2018).

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

24. M. DeJesus-Hernandez, I. R. Mackenzie, B. F. Boeve, A. L. Boxer, M. Baker, N. J. Rutherford, A. M. Nicholson, N. A. Finch, H. F. Gilmer, J. Adamson, N. Kouri, A. Wojtas, P. Sengdy, G.-Y. R. Hsiung, A. Karydas, W. W. Seeley, K. A. Josephs, G. Coppola, D. H. Geschwind, Z. K. Wszolek, H. Feldman, D. Knopman, R. Petersen, B. L. Miller, D. Dickson, K. Boylan, N. Graff-Radford, R. Rademakers, Expanded GGGGCC hexanucleotide repeat in non-coding region of C9ORF72 causes chromosome 9p-linked frontotemporal dementia and amyotrophic lateral sclerosis. *Neuron* **72**, 245–256 (2011).

25. P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).

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

27. T. Town, V. Nikolic, J. Tan, The microglial "activation" continuum: from innate to adaptive responses. *Journal of Neuroinflammation* **2**, 24 (2005).

28. S. A. Liddelow, K. A. Guttenplan, L. E. Clarke, F. C. Bennett, C. J. Bohlen, L. Schirmer, M. L. Bennett, A. E. Münch, W.-S. Chung, T. C. Peterson, D. K. Wilton, A. Frouin, B. A. Napier, N. Panicker, M. Kumar, M. S. Buckwalter, D. H. Rowitch, V. L. Dawson, T. M. Dawson, B. Stevens, B. A. Barres, Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* **541**, 481–487 (2017).

29. W. Zhao, W. Xie, W. Le, D. R. Beers, Y. He, J. S. Henkel, E. P. Simpson, A. A. Yen, Q. Xiao, S. H. Appel, Activated microglia initiate motor neuron injury by a nitric oxide and glutamate-mediated mechanism. *J Neuropathol Exp Neurol* **63**, 964–977 (2004).

30. A. M. Haidet-Phillips, M. E. Hester, C. J. Miranda, K. Meyer, L. Braun, A. Frakes, S. Song, S. Likhite, M. J. Murtha, K. D. Foust, M. Rao, A. Eagle, A. Kammesheidt, A. Christensen, J. R. Mendell, A. H. M. Burghes, B. K. Kaspar, Astrocytes from familial and sporadic ALS patients are toxic to motor neurons. *Nat Biotechnol* **29**, 824–828 (2011).

31. A. M. D'Erchia, A. Gallo, C. Manzari, S. Raho, D. S. Horner, M. Chiara, A. Valletti, I. Aiello, F. Mastropasqua, L. Ciaccia, F. Locatelli, F. Pisani, G. P. Nicchia, M. Svelto, G. Pesole, E. Picardi, Massive transcriptome sequencing of human spinal cord tissues provides new insights into motor neuron degeneration in ALS. *Sci Rep* **7**, 10046 (2017).

32. D. G. Brohawn, L. C. O'Brien, J. P. Bennett, RNAseq Analyses Identify Tumor Necrosis Factor-Mediated Inflammation as a Major Abnormality in ALS Spinal Cord. *PLoS One* **11**, e0160520 (2016).

33. O. Dols-Icardo, V. Montal, S. Sirisi, G. López-Pernas, L. Cervera-Carles, M. Querol-Vilaseca, L. Muñoz, O. Belbin, D. Alcolea, L. Molina-Porcel, J. Pegueroles, J. Turón-Sans, R. Blesa, A. Lleó, J. Fortea, R. Rojas-García, J. Clarimón, Motor cortex transcriptome reveals microglial key events in amyotrophic lateral sclerosis. *Neurol Neuroimmunol Neuroinflamm* **7**, e829 (2020).

34. P. Andrés-Benito, J. Moreno, E. Aso, M. Povedano, I. Ferrer, Amyotrophic lateral sclerosis, gene deregulation in the anterior horn of the spinal cord and frontal cortex area 8: implications in frontotemporal lobar degeneration. *Aging (Albany NY)* **9**, 823–851 (2017).

35. K. Y. Soo, M. Halloran, V. Sundaramoorthy, S. Parakh, R. P. Toth, K. A. Southam, C. A. McLean, P. Lock, A. King, M. A. Farg, J. D. Atkin, Rab1-dependent ER–Golgi transport dysfunction is a common pathogenic mechanism in SOD1, TDP-43 and FUS-associated ALS. *Acta Neuropathol* **130**, 679–697 (2015).

36. V. Sundaramoorthy, J. M. Sultana, J. D. Atkin, Golgi fragmentation in amyotrophic lateral sclerosis, an overview of possible triggers and consequences. *Front Neurosci* **9**, 400 (2015).

37. K. Burk, R. J. Pasterkamp, Disrupted neuronal trafficking in amyotrophic lateral sclerosis. *Acta Neuropathol* **137**, 859–877 (2019).

38. W. Guo, T. Vandoorne, J. Steyaert, K. A. Staats, L. Van Den Bosch, The multifaceted role of kinases in amyotrophic lateral sclerosis: genetic, pathological and therapeutic implications. *Brain* **143**, 1651–1673 (2020).

39. A. S. Chen-Plotkin, F. Geser, J. B. Plotkin, C. M. Clark, L. K. Kwong, W. Yuan, M. Grossman, V. M. Van Deerlin, J. Q. Trojanowski, V. M.-Y. Lee, Variations in the progranulin gene affect global gene expression in frontotemporal lobar degeneration. *Hum Mol Genet* **17**, 1349–1362 (2008).

40. P. Andrés-Benito, E. Gelpi, M. Povedano, G. Santpere, I. Ferrer, Gene Expression Profile in Frontal Cortex in Sporadic Frontotemporal Lobar Degeneration-TDP. *Journal of Neuropathology & Experimental Neurology* **77**, 608–627 (2018).

41. K. Kounakis, N. Tavernarakis, The Cytoskeleton as a Modulator of Aging and Neurodegeneration. *Adv Exp Med Biol* **1178**, 227–245 (2019).

42. M. O. Mol, S. S. M. Miedema, J. C. van Swieten, J. G. J. van Rooij, E. G. P. Dopper, Molecular Pathways Involved in Frontotemporal Lobar Degeneration with TDP-43 Proteinopathy: What Can We Learn from Proteomics? *Int J Mol Sci* **22**, 10298 (2021).

43. J. Altman, Are new neurons formed in the brains of adult mammals? *Science* **135**, 1127–1128 (1962).

44. P. S. Eriksson, E. Perfilieva, T. Björk-Eriksson, A. M. Alborn, C. Nordborg, D. A. Peterson, F. H. Gage, Neurogenesis in the adult human hippocampus. *Nat Med* **4**, 1313–1317 (1998).

45. J. Brettschneider, K. Del Tredici, D. J. Irwin, M. Grossman, J. L. Robinson, J. B. Toledo, L. Fang, V. M. Van Deerlin, A. C. Ludolph, V. M.-Y. Lee, H. Braak, J. Q. Trojanowski, Sequential distribution of pTDP-43 pathology in behavioral variant frontotemporal dementia (bvFTD). *Acta Neuropathol* **127**, 423–439 (2014).

46. J. Brettschneider, K. Del Tredici, J. B. Toledo, J. L. Robinson, D. J. Irwin, M. Grossman, E. Suh, V. M. Van Deerlin, E. M. Wood, Y. Baek, L. Kwong, E. B. Lee, L. Elman, L. McCluskey, L. Fang, S. Feldengut, A. C. Ludolph, V. M.-Y. Lee, H. Braak, J. Q. Trojanowski, Stages of pTDP-43 pathology in amyotrophic lateral sclerosis: ALS Stages. *Ann Neurol.* **74**, 20–38 (2013).

47. B. E. Spencer, L. A. Digma, R. G. Jennings, J. B. Brewer, Gene- and age-informed enrichment method for preclinical Alzheimer's disease trials: Approach and estimated cost savings. *Alzheimer's & Dementia* **16**, e040021 (2020).

48. I. J. Broce, P. A. Castruita, J. S. Yokoyama, Moving Toward Patient-Tailored Treatment in ALS and FTD: The Potential of Genomic Assessment as a Tool for Biological Discovery and Trial Recruitment. *Frontiers in Neuroscience* **15** (2021) (available at https://www.frontiersin.org/articles/10.3389/fnins.2021.639078).

49. X. Lu, Z. Liu, Q. Cui, F. Liu, J. Li, X. Niu, C. Shen, D. Hu, K. Huang, J. Chen, X. Xing, Y. Zhao, F. Lu, X. Liu, J. Cao, S. Chen, H. Ma, L. Yu, X. Wu, X. Wu, Y. Li, H. Zhang, X. Mo, L. Zhao, J. Huang, L. Wang, W. Wen, X.-O. Shu, F. Takeuchi, W.-P. Koh, E. S. Tai, C.-Y. Cheng, T. yin Wong, X. Chang, M. Y.-Y. Chan, W. Gao, H. Zheng, K. Chen, J. Chen, J. He, C. S. Tang, K. S. L. Lam, H. Tse, C. Y. Y. Cheung, A. Takahashi, M. Kubo, N. Kato, C. Terao, Y. Kamatani, P. C. Sham, C.-K. Heng, Z. Hu, Y. E. Chen, T. Wu, H. Shen, C. J. Willer, D. Gu, A polygenic risk score improves risk stratification of coronary artery disease: a large-scale prospective Chinese cohort study. *European Heart Journal* **43**, 1702–1711 (2022).

50. C. Huntley, B. Torr, A. Sud, C. F. Rowlands, R. Way, K. Snape, H. Hanson, C. Swanton, J. Broggio, A. Lucassen, M. McCartney, R. S. Houlston, A. D. Hingorani, M. E. Jones, C. Turnbull, Utility of polygenic risk scores in UK cancer screening: a modelling analysis. *Lancet Oncol* **24**, 658–668 (2023).

51. P. T. Nelson, D. W. Dickson, J. Q. Trojanowski, C. R. Jack, P. A. Boyle, K. Arfanakis, R. Rademakers, I. Alafuzoff, J. Attems, C. Brayne, I. T. S. Coyle-Gilchrist, H. C. Chui, D. W. Fardo, M. E. Flanagan, G. Halliday, S. R. K. Hokkanen, S. Hunter, G. A. Jicha, Y. Katsumata, C. H. Kawas, C. D. Keene, G. G. Kovacs, W. A. Kukull, A. I. Levey, N. Makkinejad, T. J. Montine, S. Murayama, M. E. Murray, S. Nag, R. A. Rissman, W. W. Seeley, R. A. Sperling, C. L. White Iii, L. Yu, J. A. Schneider, Limbic-predominant age-related TDP-43 encephalopathy (LATE): consensus working group report. *Brain* **142**, 1503–1527 (2019).

52. B. R. Brooks, El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors. *J Neurol Sci* **124 Suppl**, 96–107 (1994).

53. A. S. Hinrichs, D. Karolchik, R. Baertsch, G. P. Barber, G. Bejerano, H. Clawson, M. Diekhans, T. S. Furey, R. A. Harte, F. Hsu, J. Hillman-Jackson, R. M. Kuhn, J. S. Pedersen, A. Pohl, B. J. Raney, K. R. Rosenbloom, A. Siepel, K. E. Smith, C. W. Sugnet, A. Sultan-Qurraie, D. J. Thomas, H. Trumbower, R. J. Weber, M. Weirauch, A. S. Zweig, D. Haussler, W. J. Kent, The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res* **34**, D590-598 (2006).

54. T. A. Myers, S. J. Chanock, M. J. Machiela, LDlinkR: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations. *Frontiers in Genetics* **11** (2020) (available at https://www.frontiersin.org/articles/10.3389/fgene.2020.00157).

55. S. Durinck, Y. Moreau, A. Kasprzyk, S. Davis, B. De Moor, A. Brazma, W. Huber, BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**, 3439–3440 (2005).

56. S. X. Xie, Y. Baek, M. Grossman, S. E. Arnold, J. Karlawish, A. Siderowf, H. Hurtig, L. Elman, L. McCluskey, V. Van Deerlin, V. M.-Y. Lee, J. Q. Trojanowski, Building An Integrated Neurodegenerative Disease Database At An Academic Health Center. *Alzheimers Dement* **7**, e84–e93 (2011).

57. S. Das, L. Forer, S. Schönherr, C. Sidore, A. E. Locke, A. Kwong, S. I. Vrieze, E. Y. Chew, S. Levy, M. McGue, D. Schlessinger, D. Stambolian, P.-R. Loh, W. G. Iacono, A. Swaroop, L. J. Scott, F. Cucca, F. Kronenberg, M. Boehnke, G. R. Abecasis, C. Fuchsberger, Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).

58. D. Taliun, D. N. Harris, M. D. Kessler, J. Carlson, Z. A. Szpiech, R. Torres, S. A. G. Taliun, A. Corvelo, S. M. Gogarten, H. M. Kang, A. N. Pitsillides, J. LeFaive, S. Lee, X. Tian, B. L. Browning, S. Das, A.-K. Emde, W. E. Clarke, D. P. Loesch, A. C. Shetty, T. W. Blackwell, A. V. Smith, Q. Wong, X. Liu, M. P. Conomos, D. M. Bobo, F. Aguet, C. Albert, A. Alonso, K. G. Ardlie, D. E. Arking, S. Aslibekyan, P. L. Auer, J. Barnard, R. G. Barr, L. Barwick, L. C. Becker, R. L. Beer, E. J. Benjamin, L. F. Bielak, J. Blangero, M. Boehnke, D. W. Bowden, J. A. Brody, E. G. Burchard, B. E. Cade, J. F. Casella, B. Chalazan, D. I. Chasman, Y.-D. I. Chen, M. H. Cho, S. H. Choi, M. K. Chung, C. B. Clish, A. Correa, J. E. Curran, B. Custer, D. Darbar, M. Daya, M. de Andrade, D. L. DeMeo, S. K. Dutcher, P. T. Ellinor, L. S. Emery, C. Eng, D. Fatkin, T. Fingerlin, L. Forer, M. Fornage, N. Franceschini, C. Fuchsberger, S. M. Fullerton, S. Germer, M. T. Gladwin, D. J. Gottlieb, X. Guo, M. E. Hall, J. He, N. L. Heard-Costa, S. R. Heckbert, M. R. Irvin, J. M. Johnsen, A. D. Johnson, R. Kaplan, S. L. R. Kardia, T. Kelly, S. Kelly, E. E. Kenny, D. P. Kiel, R. Klemmer, B. A. Konkle, C. Kooperberg, A. Köttgen, L. A. Lange, J. Lasky-Su, D. Levy, X. Lin, K.-H. Lin, C. Liu, R. J. F. Loos, L. Garman, R. Gerszten, S. A. Lubitz, K. L. Lunetta, A. C. Y. Mak, A. Manichaikul, A. K. Manning, R. A. Mathias, D. D. McManus, S. T. McGarvey, J. B. Meigs, D. A. Meyers, J. L. Mikulla, M. A. Minear, B. D. Mitchell, S. Mohanty, M. E. Montasser, C. Montgomery, A. C. Morrison, J. M. Murabito, A. Natale, P. Natarajan, S. C. Nelson, K. E. North, J. R. O'Connell, N. D. Palmer, N. Pankratz, G. M. Peloso, P. A. Peyser, J. Pleiness, W. S. Post, B. M. Psaty, D. C. Rao, S. Redline, A. P. Reiner, D. Roden, J. I. Rotter, I. Ruczinski, C. Sarnowski, S. Schoenherr, D. A. Schwartz, J.-S. Seo, S. Seshadri, V. A. Sheehan, W. H. Sheu, M. B. Shoemaker, N. L. Smith, J. A. Smith, N. Sotoodehnia, A. M. Stilp, W. Tang, K. D. Taylor, M. Telen, T. A. Thornton, R. P. Tracy, D. J. Van Den Berg, R. S. Vasan, K. A. Viaud-Martinez, S. Vrieze, D. E. Weeks, B. S. Weir, S. T. Weiss, L.-C. Weng, C. J. Willer, Y. Zhang, X. Zhao, D. K. Arnett, A. E. Ashley-Koch, K. C. Barnes, E. Boerwinkle, S. Gabriel, R. Gibbs, K. M. Rice, S. S. Rich, E. K. Silverman, P. Qasba, W. Gan, G. J. Papanicolaou, D. A. Nickerson, S. R. Browning, M. C. Zody, S. Zöllner, J. G. Wilson, L. A. Cupples, C. C. Laurie, C. E. Jaquish, R. D. Hernandez, T. D. O'Connor, G. R. Abecasis, Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* **590**, 290–299 (2021).

59. S. Zhao, W. Jing, D. C. Samuels, Q. Sheng, Y. Shyr, Y. Guo, Strategies for processing and quality control of Illumina genotyping arrays. *Brief Bioinform* **19**, 765–775 (2018).

60. E. Suh, E. B. Lee, D. Neal, E. M. Wood, J. B. Toledo, L. Rennert, D. J. Irwin, C. T. McMillan, B. Krock, L. B. Elman, L. F. McCluskey, M. Grossman, S. X. Xie, J. Q. Trojanowski, V. M. V. Deerlin, Semi-automated quantification of C9orf72 expansion size reveals inverse correlation between hexanucleotide repeat number and disease duration in frontotemporal degeneration. *Acta neuropathologica* **130**, 363 (2015).

61. J. B. Toledo, V. M. Van Deerlin, E. B. Lee, E. Suh, Y. Baek, J. L. Robinson, S. X. Xie, J. McBride, E. M. Wood, T. Schuck, D. J. Irwin, R. G. Gross, H. Hurtig, L. McCluskey, L. Elman, J. Karlawish, G. Schellenberg, A. Chen-Plotkin, D. Wolk, M. Grossman, S. E. Arnold, L. M. Shaw, V. M.-Y. Lee, J. Q. Trojanowski, A platform for discovery: The University of Pennsylvania Integrated Neurodegenerative Disease Biobank. *Alzheimers Dement* **10**, 477-484.e1 (2014).

62. D. R. Rosen, Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **364**, 362 (1993).

63. A. Freischmidt, T. Wieland, B. Richter, W. Ruf, V. Schaeffer, K. Müller, N. Marroquin, F. Nordin, A. Hübers, P. Weydt, S. Pinto, R. Press, S. Millecamps, N. Molko, E. Bernard, C. Desnuelle, M.-H. Soriani, J. Dorst, E. Graf, U. Nordström, M. S. Feiler, S. Putz, T. M. Boeckers, T. Meyer, A. S. Winkler, J. Winkelman, M. de Carvalho, D. R. Thal, M. Otto, T. Brännström, A. E. Volk, P. Kursula, K. M. Danzer, P. Lichtner, I. Dikic, T. Meitinger, A. C.

Ludolph, T. M. Strom, P. M. Andersen, J. H. Weishaupt, Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. *Nat Neurosci* **18**, 631–636 (2015).

64. E. Kabashi, P. N. Valdmanis, P. Dion, D. Spiegelman, B. J. McConkey, C. Vande Velde, J.-P. Bouchard, L. Lacomblez, K. Pochigaeva, F. Salachas, P.-F. Pradat, W. Camu, V. Meininger, N. Dupre, G. A. Rouleau, TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat Genet* **40**, 572–574 (2008).

65. J. O. Johnson, J. Mandrioli, M. Benatar, Y. Abramzon, V. M. Van Deerlin, J. Q. Trojanowski, J. R. Gibbs, M. Brunetti, S. Gronka, J. Wuu, J. Ding, L. McCluskey, M. Martinez-Lage, D. Falcone, D. G. Hernandez, S. Arepalli, S. Chong, J. C. Schymick, J. Rothstein, F. Landi, Y.-D. Wang, A. Calvo, G. Mora, M. Sabatelli, M. R. Monsurrò, S. Battistini, F. Salvi, R. Spataro, P. Sola, G. Borghero, ITALSGEN Consortium, G. Galassi, S. W. Scholz, J. P. Taylor, G. Restagno, A. Chiò, B. J. Traynor, Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* **68**, 857–864 (2010).

66. M. Cruts, I. Gijselinck, J. van der Zee, S. Engelborghs, H. Wils, D. Pirici, R. Rademakers, R. Vandenberghe, B. Dermaut, J.-J. Martin, C. van Duijn, K. Peeters, R. Sciot, P. Santens, T. De Pooter, M. Mattheijssens, M. Van den Broeck, I. Cuijt, K. Vennekens, P. P. De Deyn, S. Kumar-Singh, C. Van Broeckhoven, Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21. *Nature* **442**, 920–924 (2006).

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

68. I. R. A. Mackenzie, M. Neumann, A. Baborie, D. M. Sampathu, D. D. Plessis, E. Jaros, R. H. Perry, J. Q. Trojanowski, D. M. A. Mann, V. M. Y. Lee, A harmonized classification system for FTLD-TDP pathology. *Acta Neuropathol* **122**, 111–113 (2011).

69. I. R. A. Mackenzie, M. Neumann, E. H. Bigio, N. J. Cairns, I. Alafuzoff, J. Kril, G. G. Kovacs, B. Ghetti, G. Halliday, I. E. Holm, P. G. Ince, W. Kamphorst, T. Revesz, A. J. M. Rozemuller, S. Kumar-Singh, H. Akiyama, A. Baborie, S. Spina, D. W. Dickson, J. Q. Trojanowski, D. M. A. Mann, Nomenclature and nosology for neuropathologic subtypes of frontotemporal lobar degeneration: an update. *Acta Neuropathol* **119**, 1–4 (2010).

70. E. B. Lee, S. Porta, G. Michael Baer, Y. Xu, E. Suh, L. K. Kwong, L. Elman, M. Grossman, V. M.-Y. Lee, D. J. Irwin, V. M. Van Deerlin, J. Q. Trojanowski, Expansion of the classification of FTLD-TDP: distinct pathology associated with rapidly progressive frontotemporal degeneration. *Acta Neuropathol* **134**, 65–78 (2017).

71. B. Benyamin, J. He, Q. Zhao, J. Gratten, F. Garton, P. J. Leo, Z. Liu, M. Mangelsdorf, A. Al-Chalabi, L. Anderson, T. J. Butler, L. Chen, X.-D. Chen, K. Cremin, H.-W. Deng, M. Devine, J. Edson, J. A. Fifita, S. Furlong, Y.-Y. Han, J. Harris, A. K. Henders, R. L. Jeffree, Z.-B. Jin, Z. Li, T. Li, M. Li, Y. Lin, X. Liu, M. Marshall, E. P. McCann, B. J. Mowry, S. T. Ngo, R. Pamphlett, S. Ran, D. C. Reutens, D. B. Rowe, P. Sachdev, S. Shah, S. Song, L.-J. Tan, L. Tang, L. H. van den Berg, W. van Rheenen, J. H. Veldink, R. H. Wallace, L. Wheeler, K. L. Williams, J. Wu, X. Wu, J. Yang, W. Yue, Z.-H. Zhang, D. Zhang, P. G. Noakes, I. P. Blair, R. D. Henderson, P. A. McCombe, P. M. Visscher, H. Xu, P. F. Bartlett, M. A. Brown, N. R. Wray, D. Fan, Cross-ethnic meta-analysis identifies association of the GPX3-TNIP1 locus with amyotrophic lateral sclerosis. *Nat Commun* **8**, 611 (2017).

Acknowledgments:

Funding:

National Institutes of Health grant F32AG079618 (BES), P01AG066597 (CTM), NS109260 (DJI), Penn Institute on Aging, and The DeCrane Family Fund for Primary Progressive Aphasia.

Author contributions:

Conceptualization: BES, CTM, DAW

Methodology: BES, CTM, EBL

Investigation: BES, DJI, LEB, CCQ, DA, MG

Visualization: BES, CTM

Funding acquisition: BES, CTM, DJI

Project administration: CTM, DJI

Supervision: CTM, DJI

Writing—original draft: BES, CTM

Writing—review & editing: BES, DJI, VVD, EBL, LBE, CCQ, DA, MG, DAW, CTM

Competing interests: All authors declare they have no relevant competing interests.

Data and materials availability: All summary statistic data was accessed from prior GWAS*(9, 11, 12)* Genotype, diagnosis, and neuropathological data may be requested and upon approval of reasonable requests may be shared with individual investigators. Data requests can be completed through the Penn Neurodegenerative Data Sharing Committee webform: https://www.pennbindlab.com/data-sharing

Figures

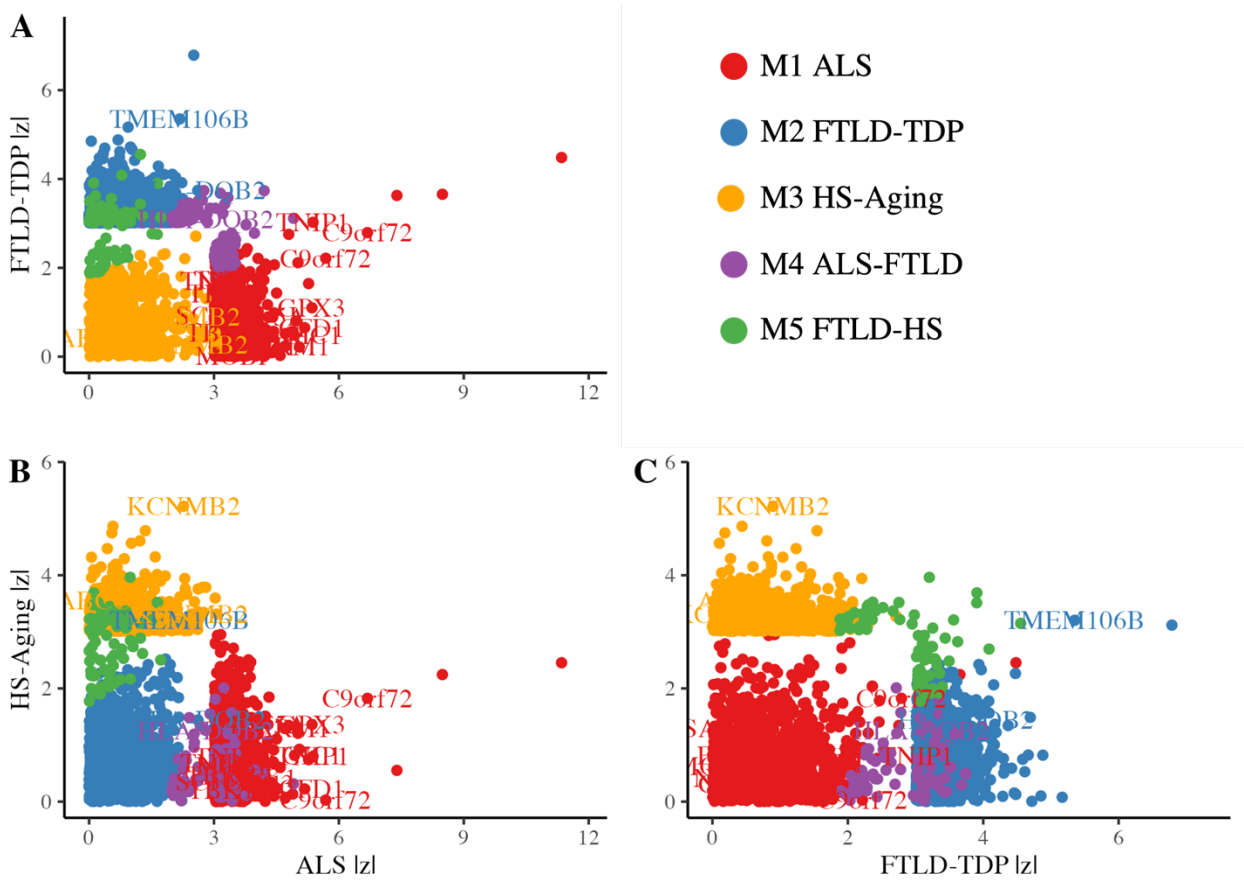

Fig. 1. WGCNA revealed modules of correlated SNPs across phenotypes associated with TDP-43 proteinopathies. Module SNPs are plotted by relative associations (by |z|) to (**A**) FTLD-TDP and ALS, (**B**) HS-Aging and ALS, or (**C**) HS-Aging and FTLD-TDP GWAS. SNPs mapping to GWAS significant genes*(3, 9, 11–13, 21, 71)* are labelled.

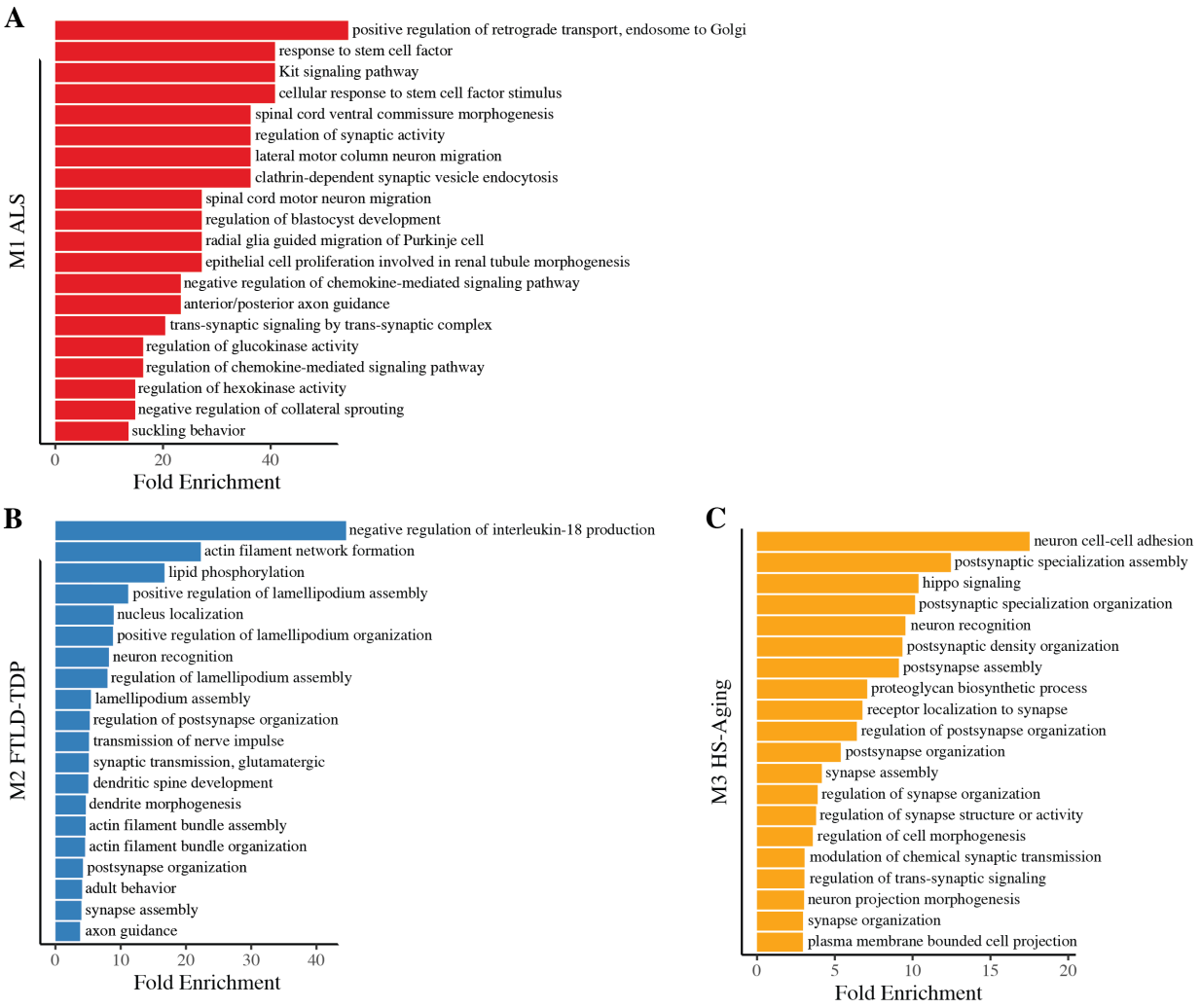

Fig. 2. WGCNA-derived modules have biologically plausible pathway associations. The top 20 overrepresented pathways by fold enrichment that met FDR-adjusted p<.05 are displayed for (**A**) M1 ALS, (**B**) M2 FTLD-TDP, and (**C**) M3 HS-Aging.

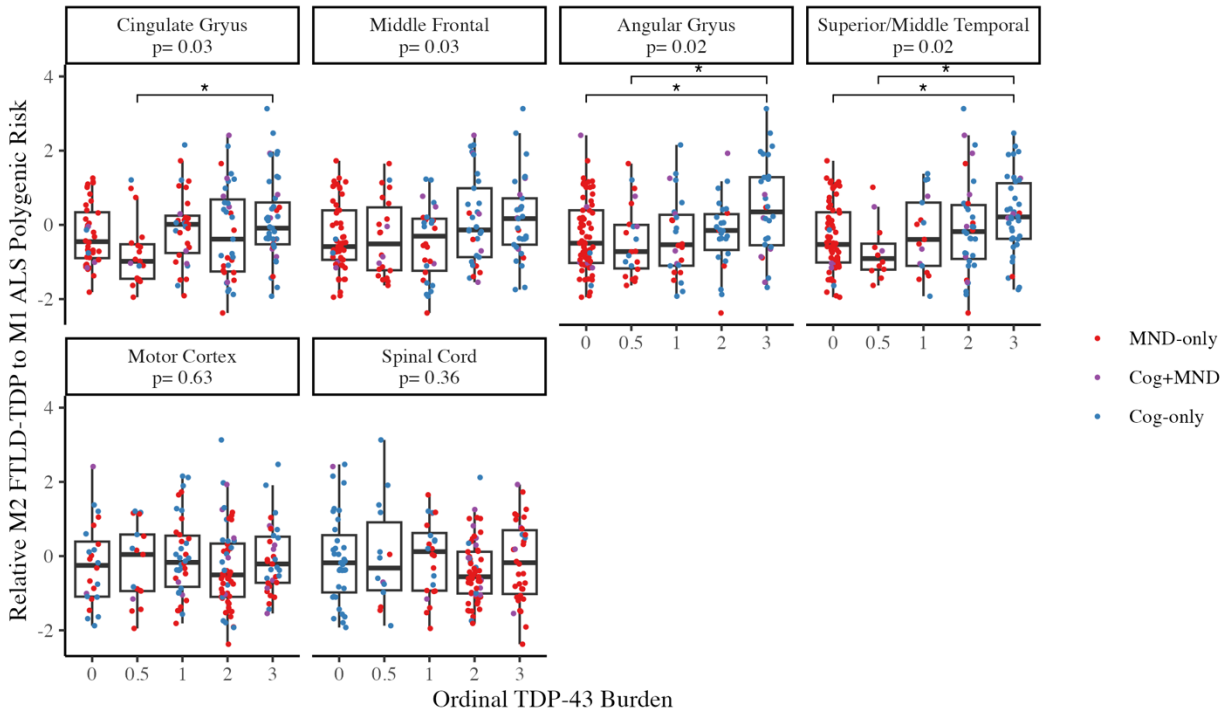

Fig. 3. Relative polygenic risk associates with regional TDP-43 burden. The difference between an individual's M2 FTLD-TDP risk and M1 ALS risk was tested for association with TDP-43 burden in FTLD-TDP characteristic regions (top row: cingulate gyrus, angular gyrus, middle frontal cortex, and superior/middle temporal cortex), and ALS characteristic regions (bottom row: motor cortex and spinal cord). The FDR adjusted p-values for each ANOVA are displayed by region. Asterisks represent FDR-adjusted significant pairwise t-tests within region.

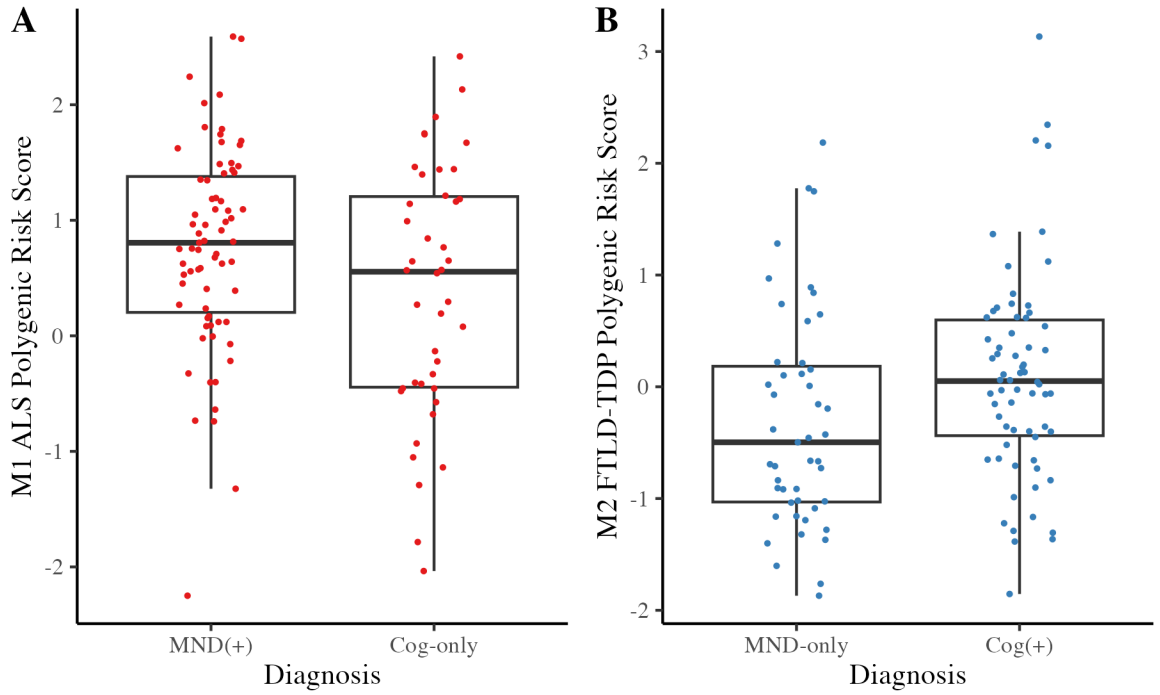

Fig. 4. M1 ALS and M2 FTLD-TDP associate with individual presentation in C9orf72 cases. The (**A**) M1 ALS and (**B**) M2 FTLD-TDP PRS were tested for association with neuromuscular or cognitive and/or behavioral impairment or in C9orf72 mutation carriers.

	MND-only	Cog+MND	Cog-only
n	125	31	104
Sex (Male, %)	72 (57.6)	17 (54.8)	59 (56.7)
Education (mean (SD))	14.88 (3.09)	14.43 (2.66)	15.67 (3.36)
Deceased (%)	121 (96.8)	31 (100.0)	86 (82.7)
Age at last contact (mean (SD))	62.16 (10.21)	60.42 (8.55)	68.54 (8.47)
Race (%)			
Asian	0 (0.0)	0 (0.0)	1 (1.0)
Black or African American	3 (2.4)	1 (3.2)	1 (1.0)
More than one race	0 (0.0)	0 (0.0)	1 (1.0)
White	121 (96.8)	30 (96.8)	99 (95.2)
Unknown or not reported	1 (0.8)	0 (0.0)	2 (1.9)
Ethnicity (%)			
Hispanic or Latino	2 (1.6)	0 (0.0)	1 (1.0)
Not Hispanic or Latino	120 (96.0)	31 (100.0)	101 (97.1)
Unknown or not reported	3 (2.4)	0 (0.0)	2 (1.9)
Mutation (%)			
No known mutation	77 (61.6)	7 (22.6)	41 (39.4)
C9orf72	47 (37.6)	24 (77.4)	42 (40.4)
GRN	0 (0.0)	0 (0.0)	17 (16.3)
TARDBP	1 (0.8)	0 (0.0)	0 (0.0)
TBK1	0 (0.0)	0 (0.0)	2 (1.9)
VCP	0 (0.0)	0 (0.0)	2 (1.9)
Prior GWAS Case* (%)			
No prior GWAS inclusion	64 (51.2)	15 (48.4)	82 (78.8)
ALS GWAS	61 (48.8)	12 (38.7)	0 (0.0)
FTLD-TDP GWAS	0 (0.0)	4 (12.9)	22 (21.2)
TDP Type (%)			
A	4 (4.4)	2 (11.8)	26 (39.4)
B	6 (6.6)	9 (52.9)	12 (18.2)
C	0 (0.0)	0 (0.0)	23 (34.8)
D	0 (0.0)	0 (0.0)	2 (3.0)
E	5 (5.5)	1 (5.9)	2 (3.0)
Not specified	76 (83.5)	5 (29.4)	1 (1.5)

Table 1. Sample Characteristics. Demographic, genetic, and pathological data for individuals with neuromuscular impairment (MND-only), cognitive and/or behavioral impairment (Cog-only), or any evidence of a combination of impairment (Cog+MND) within the sample.

Inclusion in either the ALS(9)* or FTLD-TDP*(11)* GWAS for which summary statistics were used to generate the modules.

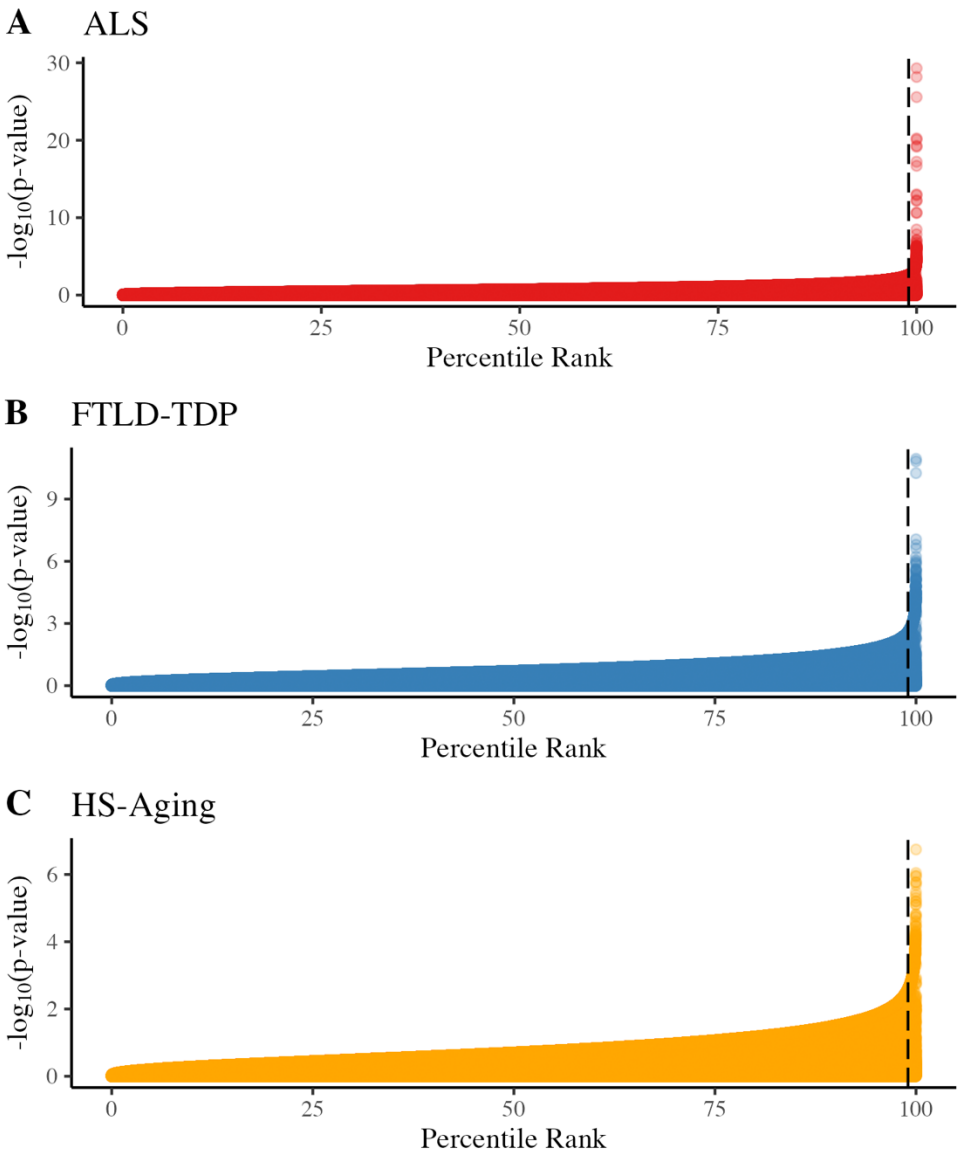

Fig. S1. The top 1% of SNPs with the strongest associations to any GWAS phenotype were selected for analysis. 494,417 unique SNPs commonly genotyped across the ALS, FTLD-TDP, and HS-Aging GWAS summary statistics were ranked according to each SNP's strongest association to any GWAS phenotype (by |z|). For each GWAS, (A) ALS, (B) FTLD-TDP, (C) HS-Aging, all SNPs are plotted by -$\log_{10}$(p-value) from the original GWAS and the percentile rank. The dashed line indicates the cut point for selection.

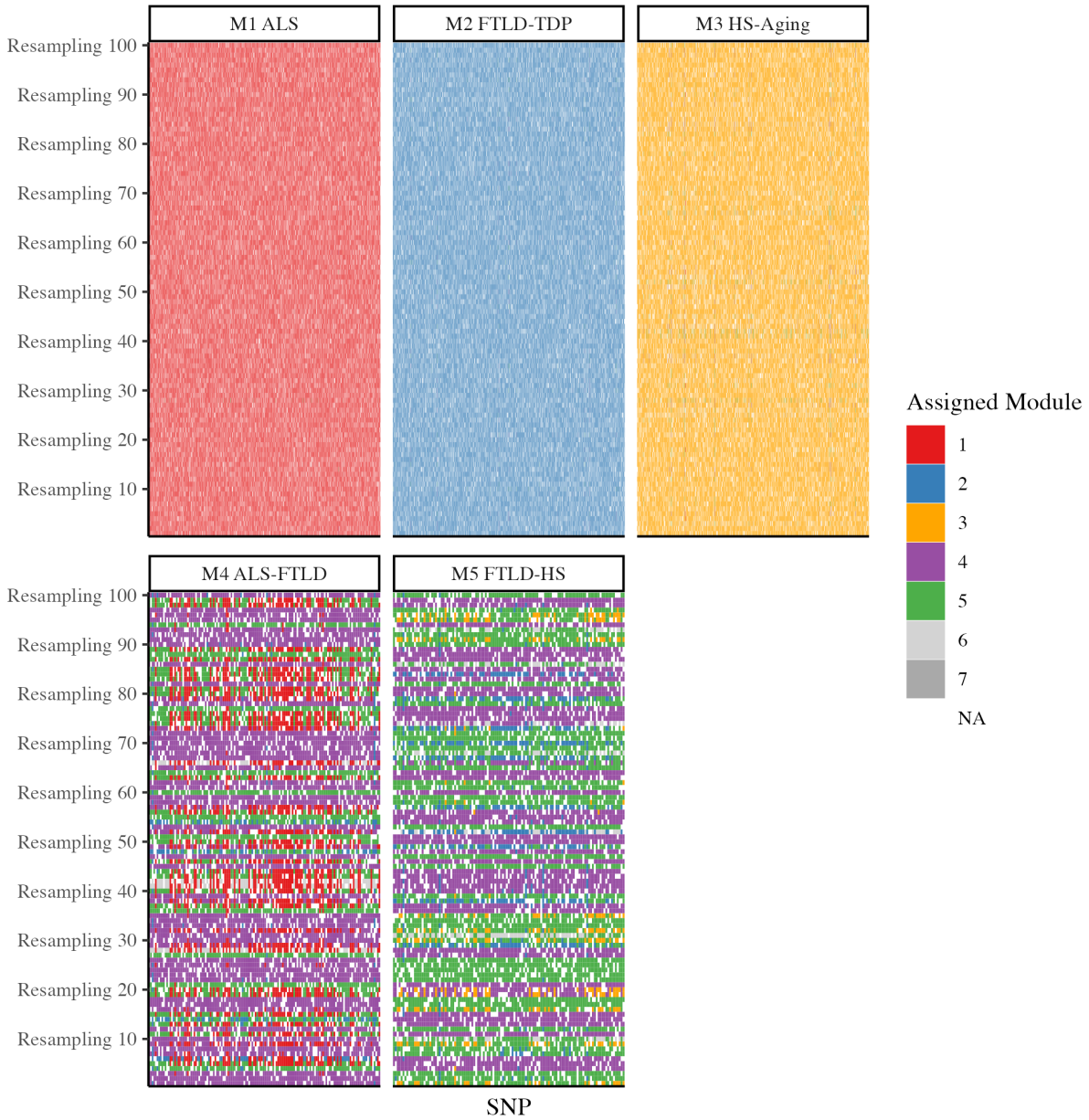

Fig. S2. Module assignment is robust for SNPs associated with a single GWAS phenotype. To test for module robustness, we created 100 samples, each containing a random 80% of the 4,945 SNPs selected for weighted correlation analysis, without replacement. For each resample, WGCNA was performed as in the full sample. Each of the 4,945 SNPs is represented on the x-axis, faceted by module assignment in the full sample. Each row on the y-axis is a resample, and the colors indicate which modules the SNPs have been assigned in that resample. All resamples have 20% NAs by definition (white). Light and dark gray represent an assigned module 6 or 7, assigned in 31 and 1 resamples respectively, which does not exist in the full sample. In the majority of instances (n= 16 resamples), module 6 contained SNPs from only a single module, indicating the split of a single module into 2 modules. As is visually apparent, modules associated with a single GWAS phenotype, M1 ALS, M2 FTLD-TDP, and M3 HS-Aging, are particularly robust, with 99.76%, 99.63%, and 97.9% SNPs, respectively, assigned to the same module as in the full sample across 100 resamples. In each resample, module numbers are

assigned by size of the module. Importantly, while M4 ALS-FTLD and M5 FTLD-HS SNPs are assigned to both modules 4 and 5 across the resamples, there is only one instance of a resample in which M5 FTLD-HS SNPs are assigned to both module 4 and 5. Thus, the separation of M4 ALS-FTLD from M5 FTLD-HS SNPs is maintained. 73.98% of M4 ALS-FTLD and 85.31% of M5 FTLD-HS SNPs were assigned to either the module 4 or 5 across all resamples. The largest differences in M4 ALS-FTLD and M5 FTLD-HS SNP assignments in the resamples compared to the full sample assignment came from 21.41% of M4 ALS-FTLD SNPs assigned to module 1 and 7.77% of M5 FTLD-HS SNPs assigned to module 2. The reduced robustness in the modules representing phenotypic pairs (M4 ALS-FTLD and M5 FTLD-HS) compared to a single GWAS phenotype (M1 ALS, M2 FTLD-TDP, and M3 HS-Aging), is likely influenced by their size. Single GWAS phenotype modules are an order of magnitude larger than phenotypic pair modules (1042-1910 vs 121-144 SNPs per module, respectively).

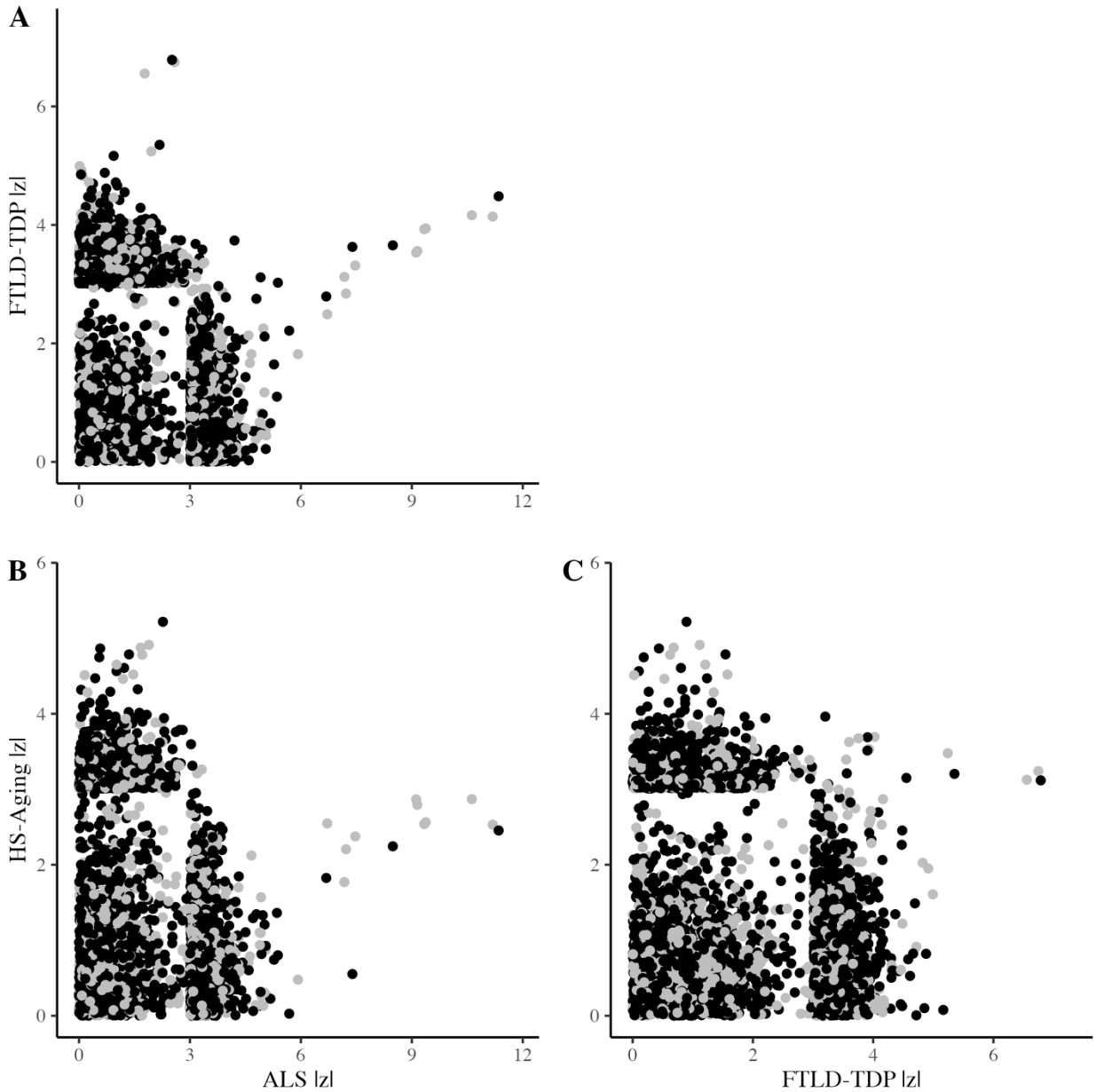

Fig. S3. WGCNA module SNPs were pruned by LD. Module SNPs are plotted by relative associations (by |z|) to (**A**) FTLD-TDP and ALS, (**B**) HS-Aging and ALS, or (**C**) HS-Aging and FTLD-TDP GWAS. SNPs were LD pruned within modules using an R^2 threshold of 0.3. SNPs shown in black were carried forward for analysis, while those in gray were removed.

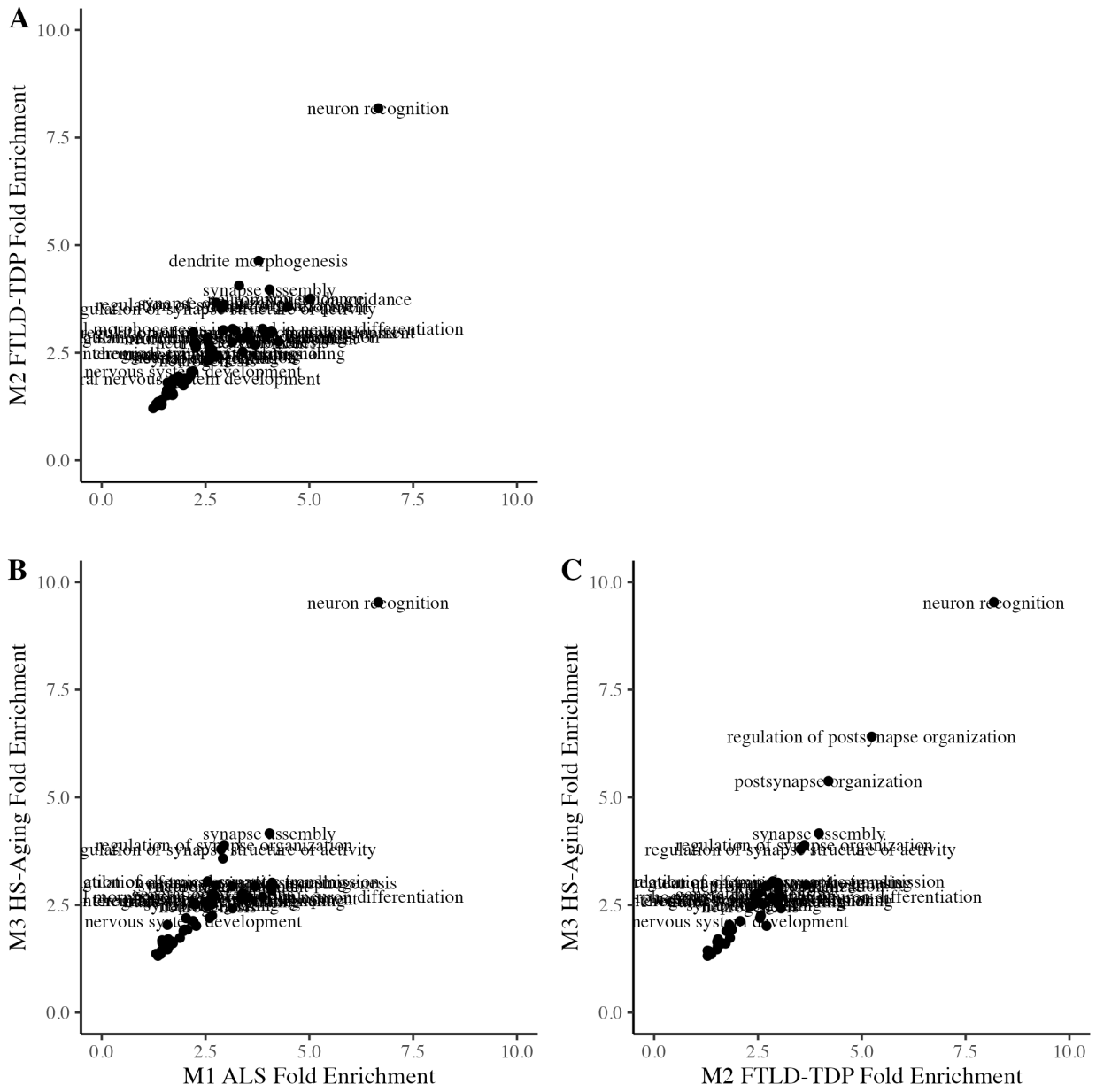

Fig. S4. Common pathways between modules are relevant to nervous system processes. Of 382 unique pathways significant across modules, 76 were common to M1 ALS and M2 FTLD-TDP, 57 were common to M1 ALS and M3 HS-Aging, and 52 were common to M2 FTLD-TDP and M3 HS-Aging. All common pathways are plotted by fold enrichment within (**A**) M1 ALS and M2 FTLD-TDP, (**B**) M1 ALS and M3 HS-Aging, or (**C**) M2 FTLD-TDP and M3 HS-Aging. Labelled terms include those containing neruo- (neuron, neurogenesis, neurotransmitter), nervous system, synap- (synapse, synaptic), dendrit- (dendrite, dendritic), or axon.

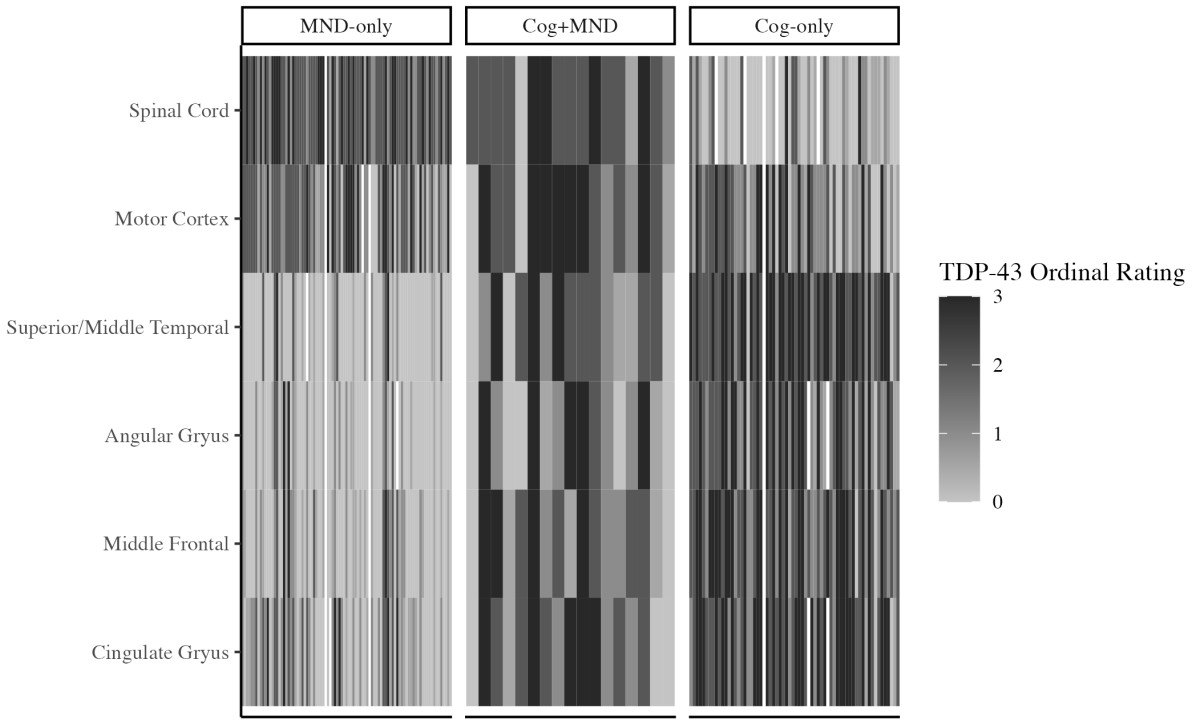

Fig. S5. Regional burden of TDP-43 in characteristic FTLD-TDP and ALS regions. Ordinal ratings of TDP-43 burden are displayed for each individual, grouped based on clinical evidence for neuromuscular impairment (MND-only), cognitive and/or behavioral impairment (Cog-only), or any evidence of a combination of impairment (Cog+MND) at any point in the natural history of disease.

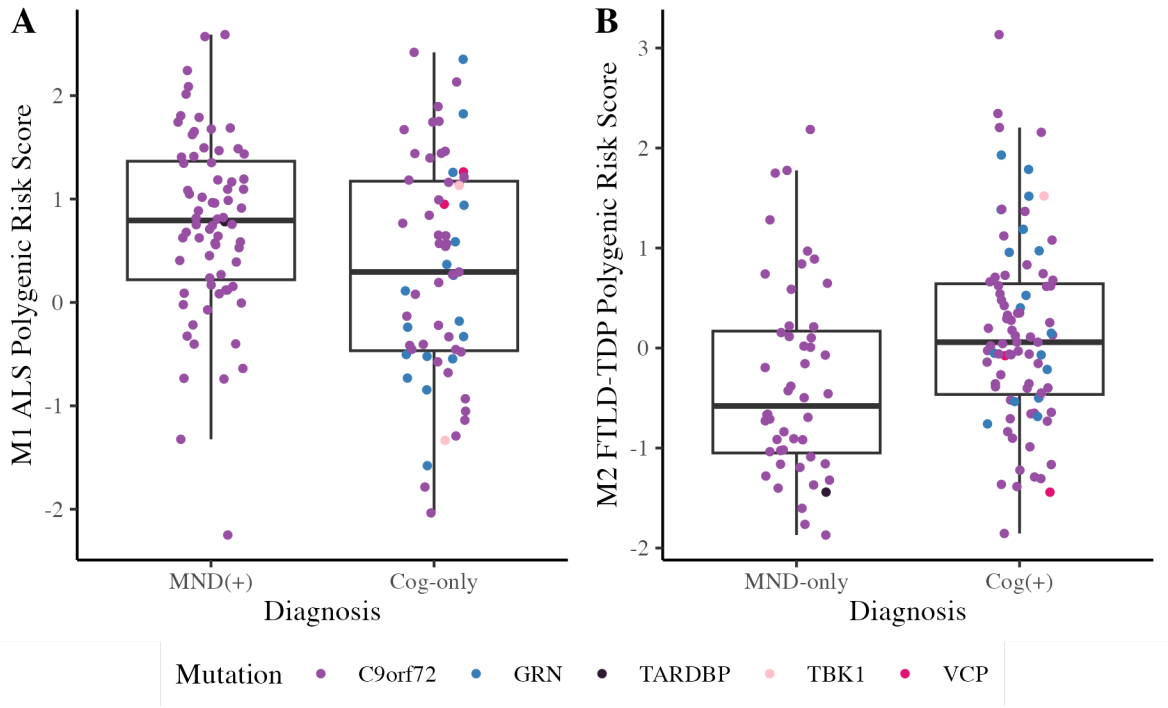

Fig. S6. M1 ALS and M2 FTLD-TDP associate with individual presentation in known mutation carriers. The (**A**) M1 ALS and (**B**) M2 FTLD-TDP PRS were tested for association with neuromuscular or cognitive and/or behavioral impairment in *C9orf72, GRN, TARDBP, TBK1*, and *VCP* mutation carriers.

Data file S1. Gene ontology enrichment analysis results by module. Pathways that met FDR-adjusted $p<.05$ are listed for each module.