

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	No software were used for data collection
Data analysis	Mayo Genome GPS v4.0, Burrows–Wheeler Aligner, Genome Analysis Toolkit, FREEMIX, PREST, R 3.6.0, ANNOVAR (version2016Feb01), PLINKv.00a23LM2, Metal, FUMA

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The processed data (summary statistics) are available on GWAS catalog under access code GCP001210 (GCST90558311, GCST90558312, GCST90558313, GCST90558314, https://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/). The data (project VCF) that was generated in this study have been deposited in the dbGAP platform under accession code phs003309. Genetic data is not available for participants of sequencing Phase I because of data sharing constraints. The

gene-based data generated in this study are provided in the Supplementary Information.

Datasets and molecular QTLs used in the gene prioritization are publicly available (see also Supplementary Data 3):
 eQTLs and eTWAS reference panels in AD-relevant bulk brain regions from AMP-AD cohorts and in LCLs from the EADB Belgian cohort, as analyzed by Bellenguez et al.14: <https://doi.org/10.5281/zenodo.5745927>;
 sQTLs and sTWAS reference panels in AD-relevant bulk brain regions from AMP-AD cohorts and in LCLs from the EADB Belgian cohort, as analyzed by Bellenguez et al.14.: <https://doi.org/10.5281/zenodo.5745929>;
 Bryois et al.83 ct-eQTL catalogues (<https://doi.org/10.5281/zenodo.5543734>);
 eQTL Catalogue database (<https://www.ebi.ac.uk/eqtl/>);
 Brain xQTL serve mQTL and haQTL catalogues issued from Ng et al., 2017 (PMID:28869584);
 GTEx v8 eQTL and sQTL catalogues (<https://www.gtexportal.org/>);
 GTEx v8 expression and splicing prediction models for eTWAS/sTWAS (<https://predictdb.org/post/2021/07/21/gtex-v8-models-on-eqtl-and-sqtl/#mashr-based-models>);
 MiGA eQTLs (<https://doi.org/10.5281/zenodo.4118605>);
 MiGA sQTLs (<https://doi.org/10.5281/zenodo.4118403>);
 MiGA meta-analysis (<https://doi.org/10.5281/zenodo.4118676>); and
 Wingo et al.86 pQTLs v2 (<https://www.synapse.org/#!Synapse:syn23627957>).

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	Sex was included as covariate in regression models. Sex-based analyses were not performed as they are beyond the scope of the current manuscript. We have sequencing data for a total of 4138 individuals, including 2224 females (53.75%). Consent for sharing individual data was collected for a subset of the samples which will be available on dbGAP upon publication.
Reporting on race, ethnicity, or other socially relevant groupings	Only sequencing data from self declared non-Hispanic white was included in this study, which represents the largest group of individuals with FTLD and sequencing data available at this point. Other race/ethnic groups show very low statistical power to identify new genes associated with FTLD.
Population characteristics	Patient/control status, FTLD diagnosis subtype, sex and genetic principal component were used in our study and statistical models.
Recruitment	No patients were recruited for this particular study. Autopsy confirmed participants were diagnosed by experts neuropathologists in multiple Brain Banks. Clinical participants were diagnosed by expert neurologists in multiple specialized centers.
Ethics oversight	This study was approved by the appropriate Mayo Clinic Institutional Review Board

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For determining the sample sizes in our study using whole genome sequencing as part of an international consortium, we followed a systematic approach. Firstly, we reached out to several centers to request DNA samples. The choice of sample size was primarily driven by the availability of DNA that was already accessible or extracted from samples collected for other purposes. The rationale for choosing these sample sizes is based on the goal of achieving a comprehensive dataset for studying frontotemporal lobar degeneration. By using the DNA that was already available or extracted from samples collected during autopsies, we were able to create the largest autopsy-enriched cohort known for FTLD. This larger cohort size provides statistical power and increases the likelihood of capturing a wide range of genetic variations pertaining to FTLD. Overall, the choice of sample sizes in our study was based on practical considerations, the availability of DNA samples, and the aim of creating a large cohort to enhance the statistical power and robustness of our findings in the study of FTLD. We excluded samples and variants based on standard quality controls which are described in the methods section of the manuscript (relatedness, sex mismatch, missingness, Hardy-Weinberg equilibrium, batch effect, low quality variants)
Data exclusions	Individuals of non-European descent and Hispanic populations were excluded from our study due to limitations in statistical power within these groups. The decision to exclude these populations was based on the need to ensure sufficient statistical power to detect meaningful associations and accurately interpret the data.

Replication	To ensure the reproducibility of our experimental findings, we undertook measures to replicate our research using publicly available datasets. Specifically, we utilized the DementiasEQ dataset for the purpose of replication. However, it is important to note that not all of our findings could be replicated in this cohort due to the lack of granularity in phenotypic information. Where feasible, we made efforts to replicate our findings to assess the robustness and generalizability of our results. We are confident in the overall validity of our study, acknowledging the importance of future studies with more granular phenotypic data to confirm and expand upon our discoveries.
Randomization	No randomization of sample selection was performed as we did not collect samples for the purpose of this study. Principal component were used in all regression models we used to control for potential population structure.
Blinding	Sequencing was done blindly without knowing the case/control status of the participants. Blinding is not relevant as we are providing case/control status to the program used for association studies (PLINK or SKAT).

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	<i>Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.</i>
Novel plant genotypes	<i>Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.</i>
Authentication	<i>Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.</i>