

Gene expression imputation provides insight into the genetic architecture of frontotemporal dementia

A transcriptome-wide analysis on frontotemporal dementia

Lianne M. Reus¹, Bogdan Pasaniuc^{2,3,4}, Danielle Posthuma⁵, Toni Boltz², International FTD-Genomics Consortium (IFGC), Yolande A.L. Pijnenburg¹, Roel A Ophoff^{2,6,7}

¹ Alzheimer Center Amsterdam, Department of Neurology, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands.

² Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California.

³ Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California.

⁴ Department of Computational Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California.

⁵ Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive research, VU University Amsterdam, The Netherlands.

⁶ Center for Neurobehavioral Genetics, University of California Los Angeles, Los Angeles, California.

⁷ Department of Psychiatry, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands.

* Correspondence to Lianne M. Reus, l.reus@amsterdamumc.nl, phone: +310 20 44 40 685, fax: +310 020 444 8529.

Postal address: De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands

Word count: 335/3,659 (abstract/main text)

Figures: 3, Tables: 2, Supplementary figures: 13, Supplementary tables: 25

Abstract

The etiology of genetically sporadic frontotemporal dementia is poorly understood. Although genome-wide association studies for frontotemporal dementia have identified a small number of candidate risk regions, most of the risk genes remain largely unknown. To identify candidate genes with predicted expression levels associated with frontotemporal dementia, we integrated genome-wide summary statistics with external reference gene expression data, using a transcriptome-wide association studies approach.

FUSION software was used to leverage summary statistics on frontotemporal dementia (n=2,340 cases, n=7,252 controls) and clinical subtypes (behavioral variant frontotemporal dementia n=1,337 cases/2,754 controls; semantic dementia n=308 cases/616 controls; progressive non-fluent aphasia n=269 cases/538 controls, frontotemporal dementia with motor neuron disease n=200 cases/400 controls) from the International Frontotemporal Dementia Genomics Consortium with 53 expression quantitative loci tissue type panels (n=12,205 from five consortia). Significance was assessed using a 5% false discovery rate threshold. We identified 73 significant gene-tissue associations for frontotemporal dementia, representing 44 unique genes in 34 tissue types. Most significant findings were derived from dorsolateral prefrontal cortex splicing data (n=19 genes, 26%). Furthermore, the 17q21.31 inversion locus contained 23 significant associations, representing six unique genes whose predicted expression associated with frontotemporal dementia. Other top hits included *SEC22B* on chromosome 1, a gene involved in vesicle trafficking, *TRGV5* on chromosome 17 and *ZNF302* on chromosome 19. A single gene finding was observed for behavioral variant frontotemporal dementia (i.e., *RAB38* on chromosome 11) with evidence from multiple tissue types. For the other clinical subtypes no significant associations were observed.

We used transcriptome-wide association studies to prioritize candidate genes for frontotemporal dementia and identified a number of specific genes, including potential novel candidate genes (such as *SEC22B*) and previously reported risk regions (e.g., 17q.21.31). Most significant associations were observed in the dorsolateral prefrontal cortex, despite the modest sample size of the gene expression reference panel of this tissue type. This suggests that our findings are specific to frontotemporal dementia and are likely to be biologically relevant highlights of genes at different frontotemporal dementia risk loci that are contributing to the disease pathology.

Keywords: frontotemporal dementia, transcriptome-wide association study

List of abbreviations

bvFTD	behavioral variant of frontotemporal dementia
CADD	Combined Annotation Dependent Depletion
CMC	CommonMind Consortium
DLPFC	dorsolateral prefrontal cortex
eQTL	expression quantitative trait loci
FDR	false discovery rate
FTD	frontotemporal dementia
FTD-MND	frontotemporal dementia with motor neuron diseases
FTLD	frontotemporal lobar degeneration
FTLD-tau	frontotemporal lobar degeneration with tau pathology
FTLD-TDP	frontotemporal lobar degeneration with transactive response DNA binding protein pathology
FUMA	Functional Mapping and Annotation
GO	Gene Ontology
GTEX	Genotype-Tissue Expression project
GWAS	genome-wide association studies
IFGC	International Frontotemporal Dementia Genomics Consortium
LD	linkage disequilibrium
MESC	Mediated Expression Score Regression
METSIM	Metabolic Syndrome in Men Study
MHC	major histocompatibility
NTR	Netherlands Twin Registry
PNFA	progressive non-fluent aphasia
PP	posterior probability
ROS	reactive oxygen-containing species
SD	semantic dementia
SMR	summary data-based Mendelian randomization
SNPs	single nucleotide polymorphisms
TDP	transactive response DNA binding protein
TWAS	transcriptome-wide association study
TWAS-GSEA	TWAS-based gene set enrichment analysis
YFS	The Cardiovascular Risk in Young Finns Study

Introduction

Frontotemporal dementia (FTD) is a heterogeneous neurodegenerative disorder, characterized by frontal and/or temporal patterns of atrophy. Clinically, FTD patients present with the behavioral variant of FTD (bvFTD) or language variants, such as semantic dementia (SD) and progressive non-fluent aphasia (PNFA) (Rascovsky *et al.*, 2011). In 10% of all cases, FTD co-occurs with motor neuron diseases (FTD-MND) (Seelaar *et al.*, 2011).

Where FTD is mostly sporadic (80%), approximately 20% of all FTD cases are familial, with the most common Mendelian mutations including the hexanucleotide repeat expansion at the *C9ORF72* locus on chromosome 9, and mutations in microtubule-associated protein tau (*MAPT*) and progranulin (*GRN*) genes in and near the chromosome 17q21 inversion locus (Hutton *et al.*, 1998; Baker *et al.*, 2006; DeJesus-Hernandez *et al.*, 2011; Renton *et al.*, 2011; Greaves and Rohrer, 2019). Genome-wide association studies (GWAS) in FTD have also identified genetic risk variants, each having small associations with disease risk (Van Deerlin *et al.*, 2010; Diekstra *et al.*, 2014; Ferrari *et al.*, 2014; Pottier *et al.*, 2019). The number of known FTD disease susceptibility loci remains small due to limited power for discovery in the relatively small sample sizes of the GWAS studies thus far with $n_{\text{cases}} < 5,000$. At this time, it is poorly understood how genetic risk variants for FTD exert effects on etiology, while such knowledge is essential for understanding disease pathology and the development of therapeutic interventions.

Genetic risk variants identified in GWAS are often located in noncoding regions with and without regulatory motifs, outside the protein encoding sequences (Maurano *et al.*, 2012). These risk variants are likely to predispose individuals to disease susceptibility by modulating mRNA expression levels, through local (*cis*) or distal (*trans*) expression quantitative trait loci (eQTL) (Nicolae *et al.*, 2010). The FTD risk variant rs302652 nearby *RAB38* is a local eQTL, decreasing *RAB38* gene expression in monocytes (Ferrari *et al.*, 2014) and potentially influencing bvFTD disease risk by modulating *RAB38* gene expression levels in specific brain areas. However, the joint effects of genetic risk loci for FTD on (differential) gene expression across multiple tissue types is unclear.

Transcriptome-wide association studies (TWAS) have emerged as a way to identify associations between traits and gene expression. The most common TWAS methods include PrediXcan, summary data-based Mendelian randomization (SMR) and FUSION (Gamazon *et al.*, 2015; Gusev *et al.*, 2016; Zhu *et al.*, 2016). TWAS leverage the combined effects of

multiple SNPs, either on individual-level (PrediXcan, SMR) or summary-level (s-PrediXcan, FUSION), on gene expression, thereby increasing power to find novel associations over a traditional GWAS when gene expression mediates risk (Gamazon *et al.*, 2015; Gusev *et al.*, 2016; Zhu *et al.*, 2016). Imputation of the genetic control of gene expression is now widely used to decipher how GWAS identified alleles may contribute to disease risk and to identify specific candidate genes through which this effect is regulated. In this study, we performed a multi-tissue TWAS on sporadic FTD and its clinical subtypes, to identify genes whose changes in expression plays a role in FTD and to identify tissue types relevant to FTD. As a secondary aim of the study, we performed a TWAS-based enrichment analysis and explored whether FTD shows overlap in differential expression with psychiatric disorders that show clinical overlap with FTD.

Materials and methods

GWAS summary statistics

GWAS summary statistics from the International Frontotemporal Dementia Genomics Consortium (IFGC) (<https://ifgcsite.wordpress.com/>) on frontotemporal dementia (FTD; n=2,154 cases/4,308 controls) and FTD clinical subtypes, behavioral variant FTD (bvFTD; n=1,377 cases/2,754 controls), semantic dementia (SD; n=306 cases/616 controls), progressive non-fluent dementia (PNFA; n=269 cases/ 538 controls) and FTD with motor neuron disease (FTD-MND; n=200 cases/400 controls), were used (Table S1). Written informed consent was obtained from all participants according to the Declaration of Helsinki. For all study sites, the study was approved by the Medical Ethics Committee. Preprocessing and quality check procedures have been described previously (Ferrari *et al.*, 2014). Single nucleotide polymorphisms (SNPs) were converted from chr:bp to rsID coordinates using Phase 3 1000 Genomes Project data (Genomes Project *et al.*, 2015). Summary statistics were converted to LD-score format using the munge_stats.py utility from LDSC, leaving 1,068,995 SNPs for final analysis for all phenotypes (Bulik-Sullivan *et al.*, 2015).

Expression quantitative trait loci (eQTL) reference panels

Local eQTL datasets from five different cohorts (n=12,205) on 53 tissue types were downloaded from the FUSION website (<http://gusevlab.org/projects/fusion>) (Table 1). The five cohorts included the CommonMind Consortium (CMC, n=452) (Fromer *et al.*, 2016), Netherlands Twin Registry (NTR, n=1,247) (Wright *et al.*, 2014), The Cardiovascular Risk in

Young Finns Study (YFS, n=1,264) (Laaksonen *et al.*, 2017), Metabolic Syndrome in Men Study (METSIM, n=562) (Laakso *et al.*, 2017) and the Genotype Tissue Expression project (GTEx) v7 (<https://gtexportal.org/home/datasets>, n=752). Local eQTLs were calculated by leveraging gene expression with genetic variation data (i.e., SNPs within ± 1 Mb of the transcriptional start site (TSS) of the gene). More detailed information on genotyping and gene expression analyses for these datasets have been described previously: CMC (Gusev *et al.*, 2018), NTR, YFS, METSIM (Gusev *et al.*, 2016) and GTEx (Consortium., 2015).

Local eQTL datasets from tissue types less relevant to FTD (e.g., blood) were included in this study, as local eQTLs are highly conserved across tissues (GTEx Consortium *et al.*, 2017) and eQTL datasets with non-brain tissues consist of substantially larger sample sizes, thereby maximizing power to detect significant associations between local gene expression and FTD GWAS SNPs.

FUMA

To examine the proportion of noncoding variants amongst FTD-risk SNPs, we annotated SNPs from the IFGC GWAS on FTD using Functional Mapping and Annotation (FUMA, <https://fuma.ctglab.nl/>) The most significant ($p < 5 \times 10^{-6}$) SNPs and SNPs in linkage disequilibrium (LD, $r^2 \geq 0.6$) with these were used for further inspection, using 1000 Genomes Project data (Genomes Project *et al.*, 2015). Lead SNPs were defined as being independent from each other at $r^2 > 0.1$. LD blocks of independent SNPs were merged into a genomic locus if they were closely located to each other (i.e., less than 250kb).

Lead and correlated SNPs were annotated for potential regulatory functions (RegulomeDB, RDB) (Boyle *et al.*, 2012), 15-core chromatin state predicted by ChromHMM (Ernst and Kellis, 2012), functional consequences on gene functions annotated by ANNOVAR (Wang *et al.*, 2010) and deleteriousness score (Combined Annotation Dependent Depletion, CADD) (Rentzsch *et al.*, 2019).

Statistical analysis

TWAS analysis

To identify genes whose local-regulated expression is associated with FTD and its clinical subtypes (i.e., bvFTD, SD, PNFA and FTD-MND), we performed TWAS analyses using FUSION software (<http://gusevlab.org/projects/fusion/>) (Gusev *et al.*, 2016). FUSION estimates the genetic correlation between local gene expression and FTD, by integrating

GWAS summary statistics with external gene expression reference panel data while accounting for LD structure among SNPs.

To study whether GWAS SNPs colocalized with eQTLs, we performed a Bayesian colocalization analysis for all associations with $p_{TWAS \text{ uncorrected}} < 0.05$ using the COLOC package in R (<https://cran.r-project.org/web/packages/coloc/>) (Plagnol *et al.*, 2009) implemented in FUSION. A joint analysis was performed to identify which genes are conditionally independent.

TWAS results are presented including the major histocompatibility (MHC) locus, as the FTD GWAS included genome-wide significant loci within the MHC region (Ferrari *et al.*, 2014). Results on gene-tissue associations per phenotype (i.e., FTD, bvFTD, SD, PNFA and FTD-MND) were corrected for multiple comparisons using a 5% false discovery rate (FDR) significance threshold. Significant TWAS loci were identified as novel if the strongest FTD associated SNP was not nominal significant ($P > 0.05$) in the IFGC GWAS (Ferrari *et al.*, 2014) within ± 1 Mb of the TSS of the gene's region.

MESC analysis

In order to estimate the proportion of disease heritability mediated by local gene expression, we performed a Mediated Expression Score Regression (MESC) analysis per tissue type (<https://github.com/douglasyao/mesc>) (Yao *et al.*, 2020). Here, we define h^2_{med} as heritability mediated by local gene expression, h^2_g as disease heritability and h^2_{med}/h^2_g as the proportion of heritability mediated by local gene expression. First, for each gene, local heritability scores were estimated while accounting for LD structure. Genes were partitioned into bins according to their local heritability, as this has shown to provide unbiased h^2_{med}/h^2_g estimates. Second, we estimated h^2_{med}/h^2_g from expression scores estimated in the previous step and GWAS summary statistics on FTD. As MESC produces biased estimates for eQTL reference panels with small sample sizes, only eQTL datasets with sample size $n > 300$ ($n = 17$) were included.

Enrichment analysis

Competitive enrichment analysis on FTD TWAS results was performed using TWAS-based gene set enrichment analysis (TWAS-GSEA) (<https://github.com/opain/TWAS-GSEA>) (Pain *et al.*, 2019). TWAS-GSEA is an adapted method of GWAS-based enrichment analysis implemented in software MAGMA (de Leeuw *et al.*, 2015). In brief, this method examines whether TWAS results are enriched for specific pathways while accounting for LD structure. Per phenotype, TWAS-GSEA was performed simultaneously for all 53 eQTL datasets. If

genes were present in multiple local eQTL datasets, the gene with the best prediction of expression (as estimated by cross-validated R^2 , MODELCV.R2) was used in the GSEA. Gene identifiers in TWAS result files were converted to Entrez ID format using the biomaRt package in R, resulting in 15,004 (14,813 non-MHC) unique Entrez IDs for FTD and all clinical FTD subtypes. TWAS results were tested for enrichment across 6,778 Gene Ontology (GO) biological processes gene sets. Per phenotype, results were corrected for the number of gene sets using a 5% FDR significance threshold.

Data availability

Data can be made available upon request.

Results

Most risk variants for FTD are located in noncoding regions

For FTD, FUMA annotated 3,103 SNPs from thirteen independent lead SNPs located in ten genomic risk loci (Watanabe *et al.*, 2017). Of all SNPs, 50.2% were located in intronic, 24.3% in non-coding RNA, 19.3% in intergenic, whereas only 1.4% was located in exonic regions. Most SNPs (93.1%) were located in open chromatin regions (range minimum chromatin state across 127 tissue/cell types=1-7) and 11.4% SNPs had potential regulatory elements, as indicated by a RDB score below 2 (Figure S1).

Predicted gene expression levels show 73 associations with FTD

Predicted gene expression levels in 53 tissue types (range of genes per tissue type=1,505-9,229) were tested for association with FTD. We identified 73 significant gene-tissue associations for FTD, representing 44 (40 non-MHC) unique genes in 34 tissue types (Table 1, Table S2, Figure 1, Figure 2). The strongest genic FTD TWAS associations with supporting evidence from colocalization analyses included *ARL17B* on chromosome 17 (brain cerebellar hemisphere $P_{FDR}=9.02\times 10^{-22}$), *ZNF302* on chromosome 19 (DLPFC splicing data $P_{FDR}=5.80\times 10^{-8}$), *LRRC37A* (lung $P_{FDR}=1.58\times 10^{-5}$), *SEC22B* on chromosome 1 (thyroid $P_{FDR}=2.28\times 10^{-3}$) and *TRGV5P* on chromosome 17 (cells transformed fibroblasts $P_{FDR}=2.39\times 10^{-3}$) (Table 2, Table S3). Of all transcriptome-wide significant genes with supporting colocalization evidence, only the association of *SEC22B* with FTD was novel, showing no evidence for association in the FTD GWAS (minimal P within ± 1 Mb of the gene's region= 6.14×10^{-2}) (Ferrari *et al.*, 2014) (Table S4).

One region of interest is 17q21.31 on chromosome 17, which contained 23 significant associations, representing six unique genes (i.e., *ARL17B*, *KANSL1-AS1*, *LRRC37A*, *MAPT*,

MAPT-AS1 and *NSFP1*). This locus is an inversion polymorphism that has been associated previously with neurodegenerative tauopathies, but also with psychiatric disorders, such as autism spectrum disorders (Li *et al.*, 2014; Pain *et al.*, 2019). Gene expression of most gene-tissue pairs were highly correlated, except for *KANSL1-AS1*, *MAPT* and *MAPT-AS1* (Figure S2). For the majority of significant associations in 17q21.31 (n=16, 69.6%), colocalization supported model 4 (i.e., shared variant between gene expression and FTD, range PP3=0.001-0.22, range PP4=0.02-0.85).

Another region was 7p14.1, for which predicted gene expression of *TRGV5* and its pseudogene *TRGV5P* achieved transcriptome-wide significance in four different tissue types. Colocalization supported model 4 (range PP3=0.002-0.08, range PP4=0.76-0.98), suggesting that FTD and 17q21.31 gene expression share a single causal association.

Most TWAS associations were detected in dorsolateral prefrontal cortex splicing data

The brain-derived reference panels contributed the most to the significant associations between gene expression and FTD (43.8%, 32 variants), with the majority derived from the dorsolateral prefrontal cortex (DLPFC) splicing data (19 variants, 13 unique, all outside MHC). The sample size and number of measured genes of the eQTL reference panel has shown to correlate with the number of significant hits (Gamazon *et al.*, 2019). Despite the modest sample size (n_{sample}=452) and number of measured genes (n_{genes unique}=3,221, n_{genes total}=7,514), the DLPFC splicing data accounted for 26% of all transcriptome-wide hits, thereby exceeding the number of significant hits compared to eQTL tissue types with larger sample sizes (e.g., 0% for YFS whole blood, n_{sample}=1,264) and more measured genes (e.g., 3% for thyroid, n_{genes unique}=9,225, n_{genes total}=9,229) (Figure S3, Figure S4).

MESC analysis showed that a substantial proportion of FTD heritability was mediated by the local component of gene expression levels (mean h²_{med}=35(4.7)%). The tibial nerve had the highest heritability mediated by local gene expression levels (h²_{med}=59.5(2.2)%), potentially reflecting a genetic component underlying the comorbidity underlying FTD and motor neuron diseases. For DLPFC splicing data the h²_{med} was 43.8(8.5)%, whereas for the eQTL panel with the largest sample size (YFS whole blood data) this was 12.6(7.4)%. A full overview of local mediated heritability is presented in Figure S5.

Predicted gene expression levels on clinical subtypes separately show association with bvFTD only

Predicted gene expression levels in 53 tissue types (range of genes per tissue type=1,505-9,229) were tested for association with bvFTD, SD, PNFA and FTD-MND. Gene expression

of *RAB38* on chromosome 11 was significantly associated with bvFTD risk in 8 out of 25 tissue panels (colon sigmoid $P_{FDR}=4.02\times 10^{-4}$, range significant gene-tissue associations $P_{FDR}=4.02\times 10^{-4}$ - 4.37×10^{-2}) (Figure 3, Figure S6, Table S5). Colocalization supported model 4 with a range posterior probability PP4 of 0.64-1.0 (range PP3=0.003-0.04) (Table S6). The former GWAS on bvFTD showed nominal evidence for the association of *RAB38* with FTD (rs302668 odd ratio(OR)=0.81(0.71-0.91); $p_{GWAS}=2.44\times 10^{-7}$) (Ferrari *et al.*, 2014). For SD, PNFA and FTD-MND, no significant transcriptome-wide associations were observed (Figure S7, S8, S9, Table S7-S12).

Implicated genes highlight involvement of amino acid transport in FTD pathogenesis

Full competitive results for the enrichment analysis on FTD and its clinical subtypes are presented in Table S13-S22. TWAS results for FTD were significantly enriched for ‘Sulfur amino acid transport’ (with MHC $P_{FDR}=0.04$, without MHC $P_{FDR}=0.03$) (Figure S11, S12). For all other gene sets and traits, no gene sets were significant after FDR correction.

No genetic correlation between gene expression FTD and primary psychiatric disorders

Given the clinical similarities between FTD and primary psychiatric disorders, we also explored the genetic correlation between the predicted gene expression for FTD and schizophrenia, autism spectrum disorder and major depressive disorder, using RHOGE (Mancuso *et al.*, 2017) (see Supplementary method section). No significant correlations were observed after FDR correction (Table S23,24, Figure S13).

Discussion

In this study, we aimed to better understand the genetic etiology of sporadic FTD by identifying genes whose expression plays a role in FTD, using a TWAS approach with increased power of detecting loci compared to a traditional GWAS. We identified 73 significant gene-tissue associations for FTD, representing 44 unique genes in 34 tissue types. The 17q.21.31 inversion region was replicated as risk region for FTD. *SEC22B* was identified as likely novel risk gene for FTD. Interestingly, most associations were derived from DLPFC (a brain region relevant to FTD) splicing data, thereby providing some biological validation to the multi-tissue TWAS approach in FTD, and highlighting the importance of splicing events for disease risk (Li *et al.*, 2016). Our results indicate that a large proportion of FTD risk loci modulate gene expression levels, and we highlight these genes as potential candidates for functional follow-up studies.

The majority of FTD risk variants were located in noncoding regions, demonstrating that these variants likely have regulatory functions. Forty-four genes were identified as differentially expressed in FTD. We replicated the 17q21.31 locus as risk factor for FTD. This region contained 23 significant associations from six different genes, including *ARL17B*, *KANSL1-AS1*, *LRRK37A*, *MAPT*, *MAPT-AS1* and *NSFP1*. The 17q21.31 region contains a common inversion polymorphism and has been associated with several neurodegenerative disorders (e.g., progressive supranuclear palsy, corticobasal degeneration, Alzheimer's disease and FTD), but also with psychiatric disorders such as autism spectrum disorder (Myers *et al.*, 2005; Webb *et al.*, 2008; Li *et al.*, 2014; Mishra *et al.*, 2017; Gandal *et al.*, 2018; Pain *et al.*, 2019). Previous research has shown that different haplotypes of the 17q21.31 inversion affects expression of 17q21.31 genes in blood and different brain regions (de Jong *et al.*, 2012). Here, we highlight the role of differential gene expression of 17q21.31 genes in the pathogenesis of FTD.

Another implicated gene was *SEC22B* on chromosome 1, which showed evidence for differential gene expression in FTD without achieving genome-wide significance in the corresponding FTD GWAS ($P > 0.05$ within $\pm 1\text{Mb}$ of *SEC22B*). *SEC22B* is a vesicle trafficking protein, playing an important role in trafficking, autophagy and membrane fusion. The latter is essential for the development of the nervous system including axonal and dendritic growth (Petkovic *et al.*, 2014). Little is known about the precise role of *SEC22B* in neurodegeneration, but differential expression of this gene in the brain has been associated with normal aging and Alzheimer's disease (Berchtold *et al.*, 2013; Zhao *et al.*, 2016).

We found increased *C4A* gene expression to be significantly associated with FTD. The *C4* gene has two functionally different isoforms (i.e., C4A and C4B, both can vary in structure and copy number) and is located on the major histocompatibility (MHC) locus, a locus strongly associated with immune-related processes. Structural variation in *C4A/B* has been associated with schizophrenia, probably affecting synaptic pruning (Sekar *et al.*, 2016; Kamitaki *et al.*, 2020). The potential role of *C4* (structure) in the etiology of FTD has not been fully understood yet. Human post-mortem and mice model studies on FTD demonstrate an association between upregulated *C4A* gene expression and aggregation of transactive response (TAR) DNA binding protein, 43 kDa (TDP), one of the most common pathological subtypes of frontotemporal lobar degeneration (FTLD) (Chen-Plotkin *et al.*, 2008; Wu *et al.*, 2019). Upregulated *C4A* gene expression is probably generic to neurological disorders rather

than specific for FTD, as it has been observed in other neurological and psychiatric disorders (McCarthy *et al.*, 2019).

Proteins differentially expressed in FTD showed enrichment for the transport of sulfur amino acids. Sulfur amino acids (e.g., methionine and cysteine) are sensitive to oxidative modifications by reactive oxygen-containing species (ROS). The transport of sulfur amino acids is essential for the synthesis of antioxidants. For example, transport of L-cystine (i.e., oxidized form of cysteine) is needed for the production of antioxidant glutathione in the brain (McBean and Flynn, 2001). A balance between the production of ROS and antioxidants protects cells against invaders. However, an imbalance leads to increased oxidative stress, which is particularly damaging to cells in high demand of oxygen, such as neuronal cells (Haque *et al.*, 2019). Increased oxidative stress has been associated with aging, and has been observed in several disorders, including FTD (Stadtman *et al.*, 2005; Palluzzi *et al.*, 2017; Haque *et al.*, 2019).

The DLPFC contributed toward a large number of significant transcriptome-wide findings, thereby highlighting the topology-specific neurodegenerative nature of FTD. MESC analysis, an approach to examine the genome-wide distribution of heritability, showed that the tibial nerve had the largest proportion of heritability mediated by local gene expression, which may reflect the comorbidity of FTD with motor neuron diseases. We also observed various associations outside the brain, potentially highlighting the importance of other organ systems in FTD. In line with this, other organ systems, such as the gastrointestinal and musculoskeletal system, have been associated with FTD (Ikegami *et al.*, 2000; Ahmed *et al.*, 2016). On the other hand, we included local eQTL data from many tissue types - also those that seemingly are less disease-relevant - to increase power and to include as many genes in this exploratory study. As a result we may not have detected the true mechanism of disease due to a shared cross-tissue regulatory architecture of eQTLs between the tissue types related and non-related to FTD (GTEx Consortium *et al.*, 2017; Wainberg *et al.*, 2019). This is illustrated by our finding on bvFTD, for which we only identified differential regulatory gene expression of *RAB38* in tissue types outside the brain. As *RAB38* is expressed throughout the brain (<https://gtexportal.org/home/gene/RAB38>) but not available in the brain tissue panels we used, we hypothesize that differential expression of *RAB38* in the brain contributes to bvFTD disease risk too.

This study is a starting point for bridging the gap between genetic variation and disease pathogenesis involving specific genes in FTD. Nevertheless, several limitations should be taken into account. First, where TWAS increases power over a traditional GWAS study, the small sample size of the current FTD GWAS study still reduces the power to find novel transcriptome-wide associations. A second major limitation is that this study does not address the pathological heterogeneity in FTD. The most common pathological subtypes of FTD include abnormal aggregation of tau (FTLD-tau) and FTLD-TDP (Mackenzie *et al.*, 2010). As we performed a TWAS on the clinical entity of FTD, this study provides only insights into generic mechanisms underlying FTD but not into specific mechanisms underlying pathological subtypes. Additional studies are required to gain more insight into distinct mechanisms underlying pathological subtypes of FTD. Finally, it should be noted that TWAS or colocalization analysis cannot be used for causal inference (Wainberg *et al.*, 2019). It is therefore essential that our efforts will be extended to functional validation, to further understand the relationship between FTD and genes reported in this study.

Results presented in this study could be used as a point of reference in future genetic association studies on FTD. We provide evidence for the contribution of many genes to the pathogenesis of FTD, including potential novel (i.e., *SEC22B*) and previously reported FTD risk loci (e.g., 17q21.31 inversion region). Most associations were detected in DLPFC splicing data, but tissues outside the brain may be involved in FTD as well. However, functional validation is needed as TWAS is sensitive to detecting associations not relevant for disease if the disease-relevant tissue is not well-represented across reference panels. Identifying which biological processes are genetically influenced by FTD is important for understanding the disease etiology, and eventually for the development of treatments.

Figure(s) captions

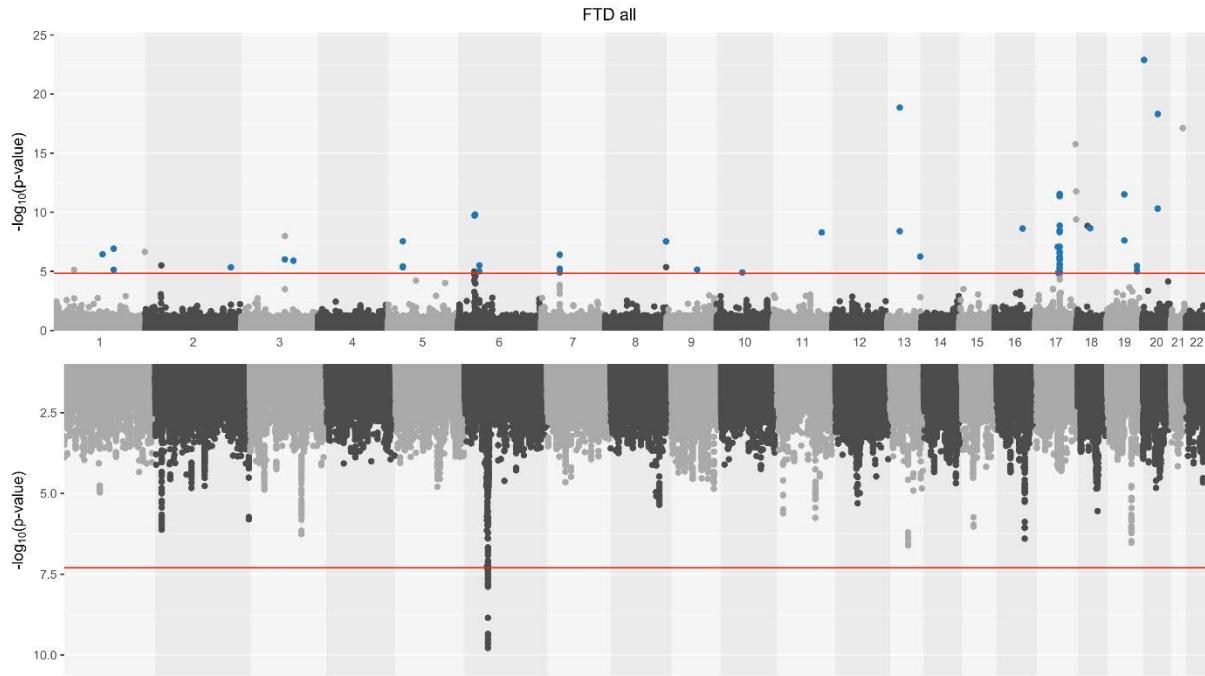


Figure 1. Manhattan plot on FTD TWAS (top) and GWAS (bottom). 44 unique genes were associated with FTD across 34 tissue types. Each point depicts a distinct gene-tissue association. TWAS hits with supporting evidence from colocalization analysis are highlighted blue. The red line depicts the significance threshold; $P_{\text{FDR}} < 0.05$ for TWAS and $P < 5\text{e-}8$ for GWAS.

FTD; frontotemporal dementia, TWAS; transcriptome-wide association study, GWAS: genome-wide association study.

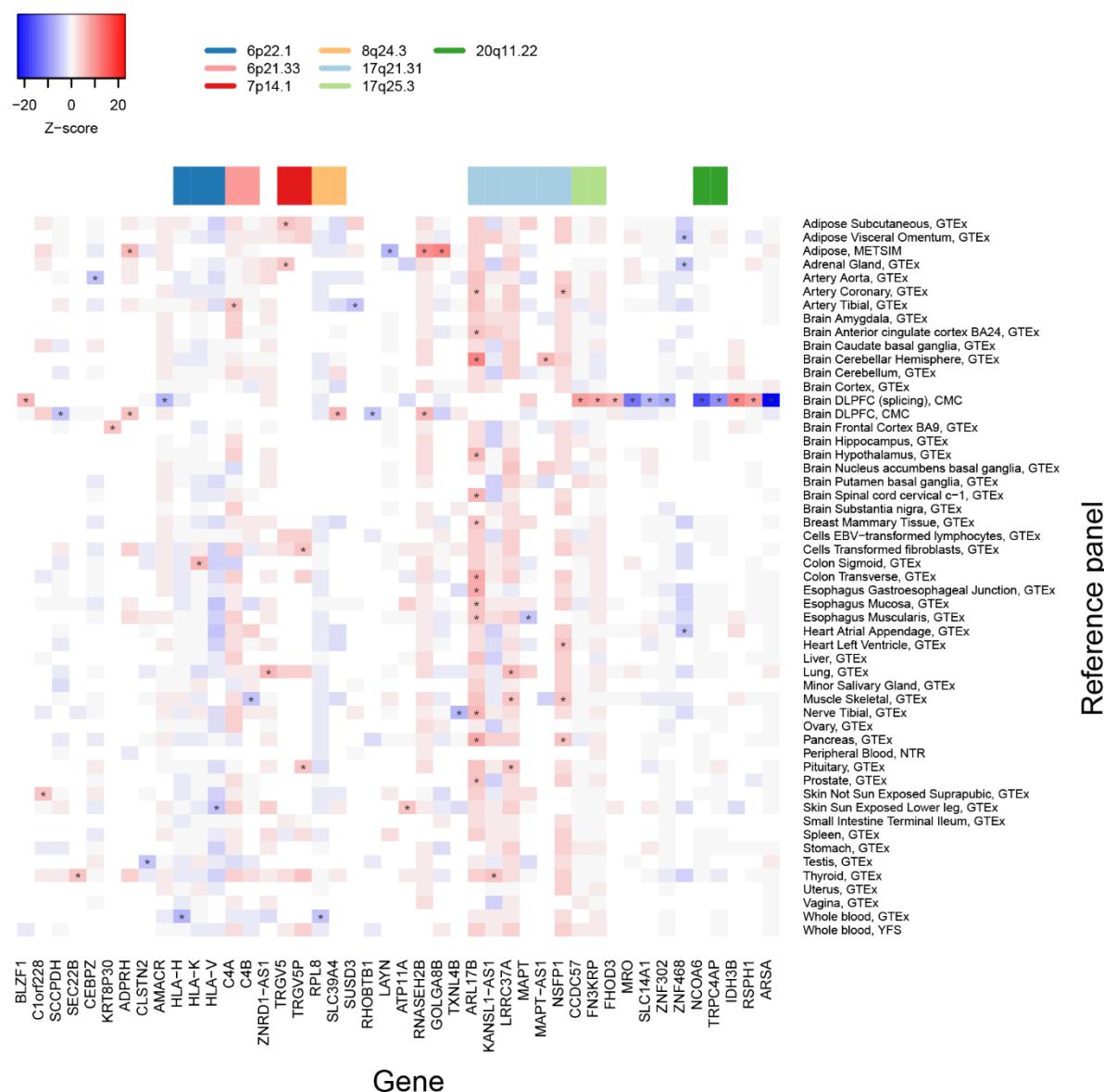


Figure 2. Heatmap of Z scores of genes with at least one transcriptome-wide significant association with FTD. Transcriptome-wide significant associations ($P_{FDR} < 0.05$) are depicted with an asterisk. Blank squares indicate that gene weights were not available in the reference panel.

FTD; frontotemporal dementia, TWAS; transcriptome-wide association study, FDR; false-discovery rate.

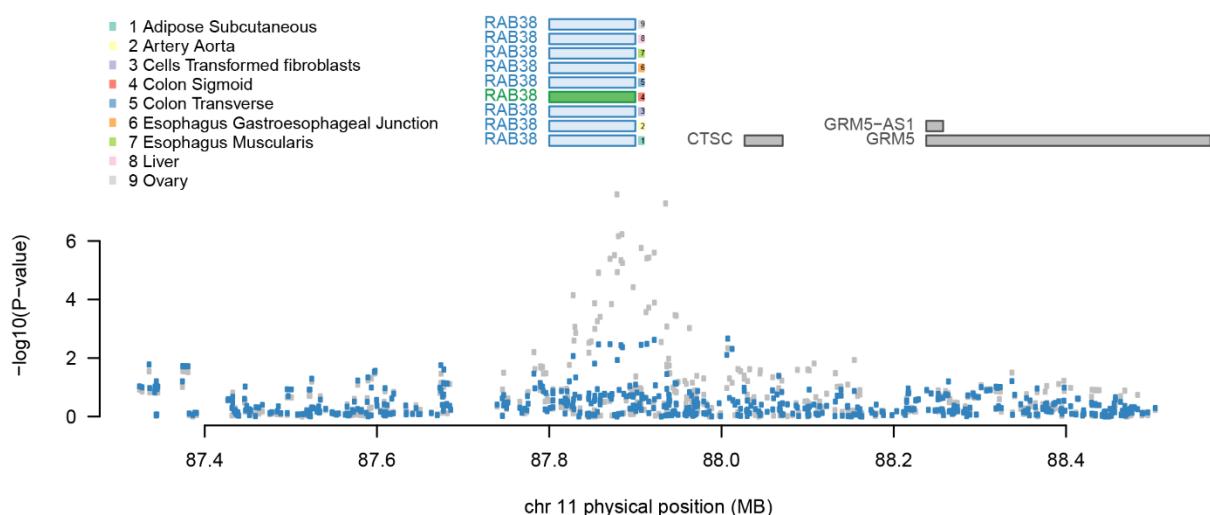


Figure 3. Regional association plot of *RAB38* for behavioral variant FTD. The top panel shows all of the genes in the locus. The marginally TWAS associated genes are highlighted in blue, and those that are jointly significant (i.e., *RAB38* in Colon Transverse) highlighted in green. The bottom panel shows a Manhattan plot of the GWAS data before (gray) and after (blue) conditioning on the green genes. This locus goes from being genome-wide significant to non-significant after conditioning on the predicted expression of *RAB38*.

FTD; frontotemporal dementia.

Tables

Table 1. Descriptive statistics for tissue reference panels and TWAS results.

Study	Reference Panel	Subjects, N a	Genes, N b	FTD TWAS significant	behavioral FTD TWAS significant
CMC	Brain - Dorsolateral prefrontal cortex	452	5244	5	0
CMC	Brain - Dorsolateral prefrontal cortex (splicing)	452	7514 (3221)	19 (13)	0
GETx	Adipose - Subcutaneous	385	7669 (7668)	1	1
GETx	Adipose - Visceral (Omentum)	313	5765 (5763)	1	0
GETx	Adrenal Gland	175	4252 (4251)	2	0
GETx	Artery - Aorta	267	6040	1	0
GETx	Artery - Coronary	152	3026	2	0
GETx	Artery - Tibial	388	7732 (7730)	2	0
GETx	Brain - Amygdala	88	1710	0	0
GETx	Brain - Anterior cingulate cortex (BA24)	109	2523	1	0
GETx	Brain - Caudate (basal ganglia)	144	3418	0	0
GETx	Brain - Cerebellar Hemisphere	125	4131 (4130)	2	0
GETx	Brain - Cerebellum	154	5513	0	0
GETx	Brain - Cortex	136	3761	0	0
GETx	Brain - Frontal Cortex (BA9)	118	2934	1	0
GETx	Brain - Hippocampus	111	2129	0	0
GETx	Brain - Hypothalamus	108	2147	1	0

GETx	Brain - Nucleus accumbens (basal ganglia)	130	3032	0	0
GETx	Brain - Putamen (basal ganglia)	111	2638	0	0
GETx	Brain - Spinal cord (cervical c-1)	83	1892	1	0
GETx	Brain - Substantia nigra	80	1505	0	0
GETx	Breast - Mammary Tissue	251	4701 (4700)	1	0
	Blood - EBV-transformed lymphocytes	117	2558 (2557)	0	0
GETx	Transformed fibroblasts	300	6957 (6956)	1	1
GETx	Colon - Sigmoid	203	4559 (4558)	1	1
GETx	Colon - Transverse	246	4935 (4934)	1	1
	Esophagus - Gastroesophageal Junction	213	4563 (4562)	1	1
GETx	Esophagus - Mucosa	358	7551 (7549)	1	0
GETx	Esophagus - Muscularis	335	7287 (7286)	2	1
GETx	Heart - Atrial Appendage	264	5316	1	0
GETx	Heart - Left Ventricle	272	4750	1	0
GETx	Liver	153	2711	0	1
GETx	Lung	383	7270 (7268)	2	0
GETx	Minor Salivary Gland	85	1681	0	0
GETx	Muscle - Skeletal	491	6990 (6989)	3	0
GETx	Nerve - Tibial	361	9064 (9062)	2	0
GETx	Ovary	122	2620 (2619)	0	1
GETx	Pancreas	220	4768 (4767)	2	0
GETx	Pituitary	157	4122 (4121)	2	0
GETx	Prostate	132	2600	1	0
	Skin - Not Sun Exposed (Suprapubic)	335	6984 (6983)	1	0
GETx	Skin - Sun Exposed (Lower leg)	414	8343 (8342)	2	0

GETx	Small Intestine - Terminal Ileum	122	2664	0	0
GETx	Spleen	146	4161	0	0
GETx	Stomach	237	4145 (4143)	0	0
GETx	Testis	225	8685 (8682)	1	0
GETx	Thyroid	399	9229 (9225)	2	0
GETx	Uterus	101	1972	0	0
GETx	Vagina	106	1852	0	0
GETx	Whole Blood	369	1898	2	0
METSIM	Adipose	563	4458	4	0
NTR	Peripheral Blood	1247	2356	0	0
YFS	Whole blood	1264	5568 (5567) 246,320 (241,893 non-MHC, 233,420 unique)	0	0
Total	-	-	73 (67)	8	

No significant gene-tissue interactions were observed for semantic dementia (SD), progressive non-fluent aphasia (PNFA) and FTD with motor neuron diseases (FTD-MND).

FTD; frontotemporal dementia, TWAS; transcriptome-wide association study, CMC; CommonMind Consortium, GTEx; Genotype-Tissue Expression project, METSIM; Metabolic Syndrome in Men Study, NTR; Netherlands Twin Registry, YFS; Young Finns Study.

a. Number of subjects included in reference panel study.

b. Number of genes that could be estimated

Table 2. Transcriptome-wide significant associations with supporting evidence from colocalization analysis.

Location	Min p (TWAS)	Min p (GWAS)	Jointly significant	Marginally significant
17.q21.31	1.83e-26	5.94e-05	<i>ARL17B, LRRC37A, NSFP1</i>	<i>ARL17B KANSL1-AS1 LRRC37A NSFP1</i>
13.q34	5.56e-07	2.30e-03	<i>ATP11A</i>	
6.p21.33	1.00e-05	6.49e-04	<i>C4A</i>	
3.q23	1.26e-06	3.07e-03	<i>CLSTN2</i>	
2.q35	4.57e-06	9.48e-04	<i>KRT8P30</i>	
10.q21.2	1.24e-05	2.69e-02	<i>RHOBTB1</i>	
1.p12	3.61e-07	8.88e-02	<i>SEC22B</i>	
9.q22.31	7.21e-06	6.14e-02	<i>SUSD3</i>	
7.p14.1	3.89e-07	2.24e-02	<i>TRGV5</i>	<i>TRGV5, TRGV5P</i>
19.q13.11	3.06e-12	2.63e-02	<i>ZNF302</i>	<i>ZNF302</i>
19.q13.41	9.09e-06	6.04e-02	<i>ZNF468</i>	

Min p (GWAS) represents the p value for the top SNP association within ± 1 Mb of transcriptional site of the gene's region.

TWAS; transcriptome-wide association study, GWAS; genome-wide association study.

Acknowledgement(s)

We thank and acknowledge the International FTD-Genomics Consortium (IFGC).

Funding

This project has been supported by a personal Alzheimer Nederland fellowship for Lianne Reus, called 'Genetic and functional overlap between behavioural variant frontotemporal dementia and psychiatric disorders' (WE.15-2018-11). Yolande Pijnenburg received a personal fellowship from the Dutch brain foundation.

Research of the Alzheimer center Amsterdam is part of the neurodegeneration research program of Amsterdam Neuroscience. The Alzheimer Center Amsterdam is supported by Stichting Alzheimer Nederland and Stichting VUmc fonds. Analyses were supported by the EU-PRISM Project (www.prism-project.eu), which received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 115916. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and EFPIA. Funding sources had no role in design and conduct of the study, data collection, data analysis, data interpretation, or in writing or approval of this report.

Competing interests

We declare that we have no conflict of interest.

References

- Ahmed RM, Irish M, Piguet O, Halliday GM, Ittner LM, Farooqi S, *et al.* Amyotrophic lateral sclerosis and frontotemporal dementia: distinct and overlapping changes in eating behaviour and metabolism. *Lancet Neurol* 2016; 15(3): 332-42.
- Baker M, Mackenzie IR, Pickering-Brown SM, Gass J, Rademakers R, Lindholm C, *et al.* Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. *Nature* 2006; 442(7105): 916-9.
- Berchtold NC, Coleman PD, Cribbs DH, Rogers J, Gillen DL, Cotman CW. Synaptic genes are extensively downregulated across multiple brain regions in normal human aging and Alzheimer's disease. *Neurobiol Aging* 2013; 34(6): 1653-61.
- Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; 22(9): 1790-7.
- Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; 47(3): 291-5.
- Chen-Plotkin AS, Geser F, Plotkin JB, Clark CM, Kwong LK, Yuan W, *et al.* Variations in the progranulin gene affect global gene expression in frontotemporal lobar degeneration. *Hum Mol Genet* 2008; 17(10): 1349-62.

Consortium. G. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015; 348(6235): 648-60.

de Jong S, Chepelev I, Janson E, Strengman E, van den Berg LH, Veldink JH, *et al.* Common inversion polymorphism at 17q21.31 affects expression of multiple genes in tissue-specific manner. *Bmc Genomics* 2012; 13.

de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* 2015; 11(4): e1004219.

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, *et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 2011; 72(2): 245-56.

Diekstra FP, Van Deerlin VM, van Swieten JC, Al-Chalabi A, Ludolph AC, Weishaupt JH, *et al.* C9orf72 and UNC13A are shared risk loci for amyotrophic lateral sclerosis and frontotemporal dementia: a genome-wide meta-analysis. *Ann Neurol* 2014; 76(1): 120-33.

Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 2012; 9(3): 215-6.

Ferrari R, Hernandez DG, Nalls MA, Rohrer JD, Ramasamy A, Kwok JBJ, *et al.* Frontotemporal dementia and its subtypes: a genome-wide association study. *The Lancet Neurology* 2014; 13(7): 686-99.

Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, *et al.* Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat Neurosci* 2016; 19(11): 1442-53.

Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet* 2015; 47(9): 1091-8.

Gamazon ER, Zwinderman AH, Cox NJ, Denys D, Derkx EM. Multi-tissue transcriptome analyses identify genetic mechanisms underlying neuropsychiatric traits. *Nat Genet* 2019; 51(6): 933-40.

Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, *et al.* Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* 2018; 362(6420).

Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, *et al.* A global reference for human genetic variation. *Nature* 2015; 526(7571): 68-74.

Greaves CV, Rohrer JD. An update on genetic frontotemporal dementia. *J Neurol* 2019; 266(8): 2075-86.

GTEx Consortium, Laboratory DA, Coordinating Center -Analysis Working G, Statistical Methods groups-Analysis Working G, Enhancing Gg, Fund NIHC, *et al.* Genetic effects on gene expression across human tissues. *Nature* 2017; 550(7675): 204-13.

Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet* 2016; 48(3): 245-52.

Gusev A, Mancuso N, Won H, Kousi M, Finucane HK, Reshef Y, *et al.* Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. *Nat Genet* 2018; 50(4): 538-48.

Haque MM, Murale DP, Kim YK, Lee JS. Crosstalk between Oxidative Stress and Tauopathy. *Int J Mol Sci* 2019; 20(8).

Hutton M, Lendon CL, Rizzu P, Baker M, Froelich S, Houlden H, *et al.* Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* 1998; 393(6686): 702-5.

Ikegami S, Harada A, Hirokawa N. Muscle weakness, hyperactivity, and impairment in fear conditioning in tau-deficient mice. *Neurosci Lett* 2000; 279(3): 129-32.

Kamitaki N, Sekar A, Handsaker RE, De Rivera H, Tooley K, Morris DL, *et al.* Complement genes contribute sex-biased vulnerability in diverse disorders. *Nature* 2020; 1-27.

Laakso M, Kuusisto J, Stancakova A, Kuulasmaa T, Pajukanta P, Lusis AJ, *et al.* The Metabolic Syndrome in Men study: a resource for studies of metabolic and cardiovascular diseases. *J Lipid Res* 2017; 58(3): 481-93.

Laaksonen J, Taipale T, Seppala I, Raitoharju E, Mononen N, Lyytikainen LP, *et al.* Blood pathway analyses reveal differences between prediabetic subjects with or without dyslipidaemia. The Cardiovascular Risk in Young Finns Study. *Diabetes Metab Res Rev* 2017; 33(7).

Li Y, Chen JA, Sears RL, Gao F, Klein ED, Karydas A, *et al.* An epigenetic signature in peripheral blood associated with the haplotype on 17q21.31, a risk factor for neurodegenerative tauopathy. *PLoS Genet* 2014; 10(3): e1004211.

Li YI, van de Geijn B, Raj A, Knowles DA, Petti AA, Golan D, *et al.* RNA splicing is a primary link between genetic variation and disease. *Science* 2016; 352(6285): 600-4.

Mackenzie IR, Neumann M, Bigio EH, Cairns NJ, Alafuzoff I, Kril J, *et al.* Nomenclature and nosology for neuropathologic subtypes of frontotemporal lobar degeneration: an update. *Acta Neuropathol* 2010; 119(1): 1-4.

Mancuso N, Shi H, Goddard P, Kichaev G, Gusev A, Pasaniuc B. Integrating Gene Expression with Summary Association Statistics to Identify Genes Associated with 30 Complex Traits. *Am J Hum Genet* 2017; 100(3): 473-87.

Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* 2012; 337(6099): 1190-5.

McBean GJ, Flynn J. Molecular mechanisms of cystine transport. *Biochem Soc Trans* 2001; 29(Pt 6): 717-22.

McCarthy N, Laws SM, Porter T, Burnham SC, Moses EK, Jablensky A. Increased predicted C4A expression is associated with cognitive deficit in both schizophrenia and Alzheimer's disease. *European Neuropsychopharmacology* 2019; 29: S871.

Mishra A, Ferrari R, Heutink P, Hardy J, Pijnenburg Y, Posthuma D, *et al.* Gene-based association studies report genetic links for clinical subtypes of frontotemporal dementia. *Brain* 2017; 140(5): 1437-46.

Myers AJ, Kaleem M, Marlowe L, Pittman AM, Lees AJ, Fung HC, *et al.* The H1c haplotype at the MAPT locus is associated with Alzheimer's disease. *Human Molecular Genetics* 2005; 14(16): 2399-404.

Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet* 2010; 6(4): e1000888.

Pain O, Pocklington AJ, Holmans PA, Bray NJ, O'Brien HE, Hall LS, *et al.* Novel Insight Into the Etiology of Autism Spectrum Disorder Gained by Integrating Expression Data With Genome-wide Association Statistics. *Biol Psychiatry* 2019; 86(4): 265-73.

Palluzzi F, Ferrari R, Graziano F, Novelli V, Rossi G, Galimberti D, *et al.* A novel network analysis approach reveals DNA damage, oxidative stress and calcium/cAMP homeostasis-associated biomarkers in frontotemporal dementia. *PLoS One* 2017; 12(10): e0185797.

Petkovic M, Jemaiel A, Daste F, Specht CG, Izeddin I, Vorkel D, *et al.* The SNARE Sec22b has a non-fusogenic function in plasma membrane expansion. *Nat Cell Biol* 2014; 16(5): 434-U121.

Plagnol V, Smyth DJ, Todd JA, Clayton DG. Statistical independence of the colocalized association signals for type 1 diabetes and RPS26 gene expression on chromosome 12q13. *Biostatistics* 2009; 10(2): 327-34.

Pottier C, Ren Y, Perkerson RB, 3rd, Baker M, Jenkins GD, van Blitterswijk M, *et al.* 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* 2019; 137(6): 879-99.

Rascovsky K, Hedges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, *et al.* Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 2011; 134(Pt 9): 2456-77.

Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR, *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 2011; 72(2): 257-68.

Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* 2019; 47(D1): D886-D94.

Seelaar H, Rohrer JD, Pijnenburg YA, Fox NC, van Swieten JC. Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a review. *J Neurol Neurosurg Psychiatry* 2011; 82(5): 476-86.

Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, *et al.* Schizophrenia risk from complex variation of complement component 4. *Nature* 2016; 530(7589): 177-83.

Stadtman ER, Van Remmen H, Richardson A, Wehr NB, Levine RL. Methionine oxidation and aging. *Biochim Biophys Acta* 2005; 1703(2): 135-40.

Van Deerlin VM, Sleiman PM, Martinez-Lage M, Chen-Plotkin A, Wang LS, Graff-Radford NR, *et al.* Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet* 2010; 42(3): 234-9.

Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, *et al.* Opportunities and challenges for transcriptome-wide association studies. *Nat Genet* 2019; 51(4): 592-9.

Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010; 38(16): e164.

Watanabe K, Taskesen E, van Bochoven A, Posthuma D. FUMA: functional mapping and annotation of genetic associations. *Nature Communications* 2017; 8: 1-11.

Webb A, Miller B, Bonasera S, Boxer A, Karydas A, Wilhelmsen KC. Role of the Tau Gene Region Chromosome Inversion in Progressive Supranuclear Palsy, Corticobasal Degeneration, and Related Disorders. *Arch Neurol-Chicago* 2008; 65(11): 1473-8.

Wright FA, Sullivan PF, Brooks AI, Zou F, Sun W, Xia K, *et al.* Heritability and genomics of gene expression in peripheral blood. *Nat Genet* 2014; 46(5): 430-7.

Wu LS, Cheng WC, Chen CY, Wu MC, Wang YC, Tseng YH, *et al.* Transcriptopathies of pre- and post-symptomatic frontotemporal dementia-like mice with TDP-43 depletion in forebrain neurons. *Acta Neuropathol Commun* 2019; 7(1): 50.

Yao DW, O'Connor LJ, Price AL, Gusev A. Quantifying genetic effects on disease mediated by assayed gene expression levels. *Nature Genetics* 2020.

Zhao Y, Tan W, Sheng W, Li X. Identification of Biomarkers Associated With Alzheimer's Disease by Bioinformatics Analysis. *Am J Alzheimers Dis Other Demen* 2016; 31(2): 163-8.

Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet* 2016; 48(5): 481-7.