

ARTICLE

Open Access

# Integrative system biology analyses of CRISPR-edited iPSC-derived neurons and human brains reveal deficiencies of presynaptic signaling in FTL and PSP

Shan Jiang<sup>1,2</sup>, Natalie Wen<sup>1,2</sup>, Zeran Li<sup>1,2</sup>, Umer Dube<sup>1,2</sup>, Jorge Del Aguila<sup>1,2</sup>, John Budde<sup>1,2</sup>, Rita Martinez<sup>1,2</sup>, Simon Hsu<sup>1,2</sup>, Maria V. Fernandez<sup>1,2</sup>, Nigel J. Cairns<sup>3</sup>, Dominantly Inherited Alzheimer Network (DIAN), International FTD-Genomics Consortium (IFGC), Oscar Harari<sup>1,2</sup>, Carlos Cruchaga<sup>1,2</sup> and Celeste M. Karch<sup>1,2</sup>

## Abstract

Mutations in the microtubule-associated protein tau (*MAPT*) gene cause autosomal dominant frontotemporal lobar degeneration with tau inclusions (FTLD-tau). *MAPT* p.R406W carriers present clinically with progressive memory loss and neuropathologically with neuronal and glial tauopathy. However, the pathogenic events triggered by the expression of the mutant tau protein remain poorly understood. To identify the genes and pathways that are dysregulated in FTLD-tau, we performed transcriptomic analyses in induced pluripotent stem cell (iPSC)-derived neurons carrying *MAPT* p.R406W and CRISPR/Cas9-corrected isogenic controls. We found that the expression of the *MAPT* p.R406W mutation was sufficient to create a significantly different transcriptomic profile compared with that of the isogenic controls and to cause the differential expression of 328 genes. Sixty-one of these genes were also differentially expressed in the same direction between *MAPT* p.R406W carriers and pathology-free human control brains. We found that genes differentially expressed in the stem cell models and human brains were enriched for pathways involving gamma-aminobutyric acid (GABA) receptors and pre-synaptic function. The expression of GABA receptor genes, including *GABRB2* and *GABRG2*, were consistently reduced in iPSC-derived neurons and brains from *MAPT* p.R406W carriers. Interestingly, we found that GABA receptor genes, including *GABRB2* and *GABRG2*, are significantly lower in symptomatic mouse models of tauopathy, as well as in brains with progressive supranuclear palsy. Genome wide association analyses reveal that common variants within *GABRB2* are associated with increased risk for frontotemporal dementia ( $P < 1 \times 10^{-3}$ ). Thus, our systems biology approach, which leverages molecular data from stem cells, animal models, and human brain tissue can reveal novel disease mechanisms. Here, we demonstrate that *MAPT* p.R406W is sufficient to induce changes in GABA-mediated signaling and synaptic function, which may contribute to the pathogenesis of FTLD-tau and other primary tauopathies.

## Introduction

Frontotemporal lobar degeneration (FTLD) encompasses several disease entities that are distinguished by the molecular pathology of neuronal and glial inclusion bodies. FTLD with tau-immunoreactive inclusion bodies (FTLD-tau) make up about 50% of cases. These primary tauopathies include progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), and Pick disease<sup>1,2</sup>.

Correspondence: Oscar Harari ([harario@wustl.edu](mailto:harario@wustl.edu)) or Carlos Cruchaga ([cruchagac@wustl.edu](mailto:cruchagac@wustl.edu)) or Celeste M. Karch ([karchc@wustl.edu](mailto:karchc@wustl.edu))

<sup>1</sup>Department of Psychiatry, Washington University School of Medicine, 660S. Euclid Ave. Campus Box 8134, St. Louis, MO 63110, USA

<sup>2</sup>Hope Center for Neurological Disorders, Washington University School of Medicine, 660S. Euclid Ave. Campus Box 8111, St. Louis, MO 63110, USA  
Full list of author information is available at the end of the article.

These authors contributed equally: Oscar Harari, Carlos Cruchaga, Celeste M. Karch.

The IFGC members are listed in Supplementary Acknowledgments.

© The Author(s) 2018



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Macroscopically, FTLD-tau is characterized by frontal and temporal lobar atrophy in contrast to Alzheimer disease (AD) where atrophy is more generalized. The insular cortex and basal ganglia are frequently affected. Microscopically, FTLD-tau, as with all FTLD entities, displays neuronal loss and gliosis in affected areas. FTLD-tau is clinically heterogeneous and may present with a broad spectrum of phenotypes including behavioral, language, memory and motor disturbances<sup>3</sup>. In a subset of FTLD-tau cases, rare mutations in the microtubule associated protein tau (*MAPT*) gene are sufficient to cause disease<sup>4,5</sup>. More than 50 mutations in *MAPT* have been reported to cause FTLD-tau and are located primarily in exons 9–13<sup>6</sup>. Despite the clear association of *MAPT* mutations with FTLD-tau, we have little understanding of the downstream events that lead to neurodegeneration and dementia.

The mechanism(s) by which *MAPT* mutations disrupt tau metabolism and contribute to disease are poorly understood in part due to the complexities of the *MAPT* gene. The *MAPT* gene in the central nervous system is alternatively spliced to produce 6 isoforms that differ based on the presence of the N-terminal insert (0N, 1N, 2N) and the number of microtubule binding repeats (MTBR; 3R, 4R). Several mechanisms have been proposed to explain how *MAPT* mutations cause disease: abnormal *MAPT* splicing, altered microtubule binding kinetics, impaired degradation, tau accumulation and aggregation, among others<sup>7</sup>.

*MAPT* p.R406W heterozygous carriers are notable in presenting more commonly with AD-like progressive memory decline and a protracted clinical course that may last decades<sup>8–10</sup>. Neurofibrillary tangles of *MAPT* p.R406W contain 3R and 4R tau and are virtually indistinguishable from those of AD; they are present largely in the frontal and temporal neocortices and hippocampus<sup>11</sup>. In a transgenic mouse model, expression of human *MAPT* p.R406W leads to age-dependent accumulation of tau aggregates, reduction of tau in axonal compartments and neurodegeneration<sup>12,13</sup>. *MAPT* p.R406W does not alter tau-mediated microtubule dynamics nor does it induce filament formation in vitro<sup>14–16</sup>. *MAPT* p.R406W may block the interaction of tau with neuronal membranes and impede neurite outgrowth<sup>17,18</sup>. However, the pathogenic events triggered by the expression of the mutant tau protein remain poorly understood.

Human induced pluripotent stem cell (iPSC)-derived neurons have emerged as a powerful cellular system to model the complexities of pathological gene expression, particularly in the early stage of disease, in the context of a non-neoplastic human genome. When coupled with genome editing by CRISPR/Cas9, this system allows for molecular events to be pinpointed specifically and to impute causality for different molecular phenotypes.

However, in the context of neurodegenerative diseases, such as FTLD-tau, it remains unclear whether the stem cell models capture events that are critical and proximal to disease pathogenesis. Here, we propose a novel integrative approach to identify molecular drivers of neurodegenerative disease using a bioinformatic pipeline that reduces the risk of false-positives arising from small sample sizes and improves our ability to identify disease-relevant pathways.

To begin to understand the molecular events caused by the *MAPT* p.R406W mutation, we sought to leverage stem cell models and human brain tissue to identify molecular changes that are triggered by *MAPT* p.R406W and that lead to human disease. We found that genes differentially expressed in the stem cell models and brains were enriched for pathways involving gamma-aminobutyric acid (GABA) receptors and pre-synaptic function. The expression of GABA receptor genes, including *GABRB2* and *GABRG2*, were consistently reduced in iPSC-derived neurons and brains from *MAPT* p.R406W carriers, PSP brains, and mouse models of tauopathy. Here, we demonstrate that *MAPT* p.R406W is sufficient to induce changes in GABA-mediated signaling and synaptic function, which contribute to the pathogenesis of FTLD-tau.

## Materials and methods

### Patient consent

The Washington University School of Medicine Institutional Review Board reviewed the protocol of the Knight Alzheimer Disease Research Center (ADRC) Neuropathology Core, from which clinically and neuropathologically well-characterized brain tissues were obtained. As tissue was obtained postmortem it was exempt from IRB approval. Research participants provided autopsy consent limited to removal of the brain. All data were analyzed anonymously.

Skin punches were performed following written informed consent from the donor. The informed consent was approved by the Washington University School of Medicine Institutional Review Board and Ethics Committee (IRB 201104178 and 201306108). The consent allowed for use of tissue by all parties, commercial and academic, for the purposes of research but not for use in human therapy.

### iPSC generation and genome engineering

Dermal fibroblasts from a *MAPT* p.R406W carrier (F11362) were transduced with non-integrating Sendai virus carrying OCT3/4, SOX2, KLF4, and cMYC (Life Technologies)<sup>19,20</sup>. Cells that showed morphological evidence of reprogramming were selected manually. iPSC lines were characterized using standard methods<sup>19</sup>. Each resulting clone was analyzed for pluripotency markers by

immunocytochemistry (ICC) and quantitative PCR (qPCR); for spontaneous differentiation into the three germ layers by ICC and qPCR; and for chromosomal abnormalities by karyotyping (see below for details). Mutation status was confirmed in newly generated iPSC by Sanger sequencing (Supplemental Fig. 1). Cell lines were confirmed to be free of mycoplasma.

iPSC that were heterozygous for *MAPT* p.R406W were edited using CRISPR/Cas9 as previously reported<sup>21</sup>. Briefly, allele-specific guideRNAs (gRNAs) were designed for the mutant allele. gRNAs were designed to have at least 3 bp of mismatch to any other gene in the human genome and validated for activity using the T7E1 assay. In this assay, the T7E1 enzyme recognizes and cleaves non-perfectly matched DNA. To prepare cells for editing, iPSC colonies were dissociated into single cells via incubation in Accutase for 10 min. Single cell iPSC cultures were maintained on Matrigel supplemented with fibronectin and cultured in mTesR. To edit cells, human iPSCs were nucleofected with 1 µg pMaxGFP (used to assess nucleofection efficiency), 1 µg gRNA, 3 µg Cas9, and 300 µM single stranded oligodeoxynucleotides (ssODN) and the P3 Primary Cell 4D reaction mix (Lonza). At least 96 clones were screened for editing. We routinely selected 2–5 edited clones for expansion and characterization. In addition to edited clones, we selected 1–2 unedited clones that were exposed to the genome-editing pipeline but remained unmodified. Characterization of edited and unedited clones included qPCR and ICC for pluripotency markers, karyotyping and Sanger sequencing of on and predicted off-target sites (Supplemental Fig. 1).

### Immunocytochemistry

Immunocytochemistry was performed to confirm pluripotency of the cultures (Supplemental Fig. 1). Culture media was aspirated, and cells were washed and fixed with 4% paraformaldehyde (Sigma). Cells were washed and permeabilized with permeabilization buffer (0.1% Triton X-100 in PBS). Cells were then blocked in 0.1% bovine serum albumin (BSA; Sigma) and treated with primary and secondary antibodies diluted in 0.1% BSA. Immunostained cells were then imaged (Nikon Eclipse 80i fluorescent microscope). The following antibodies were used: SOX2, Tra1-60 (Abcam, San Francisco, CA, USA); DAPI (Life Technologies); Alexa488 anti-rabbit (Life Technologies); and Alexa594 anti-mouse (Life Technologies).

### Quantitative PCR

To determine whether iPSC clones express endogenous markers of pluripotency, we performed qPCR (Supplemental Fig. 1). RNA was extracted from cell pellets with an RNeasy kit (Qiagen), following the manufacturer's protocol. Extracted RNA (10 µg) was converted to cDNA

by PCR using the High-Capacity cDNA Reverse Transcriptase kit (Life Technologies). Gene expression was measured in iPSCs using qPCR as previously described (*SOX2*, *POU5F1*, *LIN28A*, *NANOG*, *B3GALT5*, *PODXL*, *SEV*)<sup>22</sup>. Primers specific to *GAPDH* were used as a loading control.

### Karyotyping

Chromosomal abnormalities were assessed by G-band karyotyping after clonal isolation of iPSC and after genome editing (Supplemental Fig. 1).

### Differentiation of iPSC into cortical neurons

One *MAPT* p.R406W iPSC donor and one isogenic control iPSC line were selected for evaluation by RNA-seq. To generate neurons, iPSCs were harvested for neural aggregate formation upon reaching 75–85% confluency. iPSCs were dissociated with Accutase (MP Biomedicals) and pelleted by centrifugation. Cells were counted and resuspended in mTesR1 supplemented with the ROCK inhibitor, Y-27632 (10 µM), to achieve 450,000–650,000 cells/mL. Then, 100 µL/well of iPSC suspension was plated into a v-bottom 96-well plate. The v-bottom plate was centrifuged at 750 rpm for 3 min to pellet the iPSCs and promote formation of spheres. Cells were then incubated at 37 °C, 5% CO<sub>2</sub> for 24 h. Neurospheres were cultured for 5 additional days with daily media change in Neural Induction Medium (NIM; Stemcell Technologies).

Per manufacturer's guidelines, on Day 5 of neural aggregate formation, spheres were washed with 100 µL of DMEM/F12. Then, 100 µL of NIM was added to each well. Spheres were transferred from the conical wells into a 6 well plate, pre-coated with poly-Ornithin (PLO)/Laminin (Millipore-Sigma), at a density of 32 spheres per well. Daily media changes followed with NIM. Spheres were monitored daily for formation of neural rosette structures. Neural rosettes were harvested when spheres had completely flattened and clusters were clearly visible (3–7 days after plating; line dependent). Neural rosettes were harvested by aspirating spent medium, washing with DMEM/F12, and then adding 1 mL of Neural Rosette Selection reagent (Stemcell Technologies) to each well for 1 h at 37 °C. Rosette clusters were then detached by pipetting with DMEM/F12. Rosette clusters were transferred into a conical tube, centrifuged at 750 rpm for 3 min, and then either cryopreserved or plated for neural progenitor cell expansion. NPCs were cultured on PLO and Laminin-coated plates and terminal differentiation was initiated with the addition of cortical maturation media (Neurobasal Media (Life Technologies) supplemented with B27 (GIBCO), BDNF (Peprotech), GDNF (Peprotech), cAMP (Sigma), and L-glutamate (Sigma)). Neurons were maintained for 6 weeks.

**Table 1 Characteristics of brain samples analyzed in this study**

Sample group	Neuropathologic diagnosis	Sample size	Brain region	Age (years) <sup>a</sup>	Female (%)	DV200 <sup>a</sup>	Reference
<i>MAPT</i> p.R406W	FTLD-Tau	2	Insular cortex	69.46 ± 3.88	50%	85.33 ± 11.55	<sup>23</sup>
Controls	None	2	Insular cortex	71.81 ± 2.14	50%	88.5 ± 2.12	<sup>23</sup>
Sporadic FTLD	FTLD-TDP	11	Parietal cortex	74.09 ± 17.00	45.5%	89.45 ± 4.16	<sup>23</sup>
Late-onset AD	AD	80	Parietal cortex	85.83 ± 6.88	61.25%	89.39 ± 3.81	<sup>23</sup>
ADAD <sup>b</sup>	AD	18	Parietal cortex	50.17 ± 12.03	37.5%	86.06 ± 4.22	<sup>23</sup>
Controls	None	11	Parietal cortex	85.01 ± 9.06	45.5%	90.09 ± 1.87	<sup>23</sup>
PSP	PSP	84	Temporal cortex	73.96 ± 6.54	39.29%	NA	<sup>24</sup>
PSP	PSP	84	Cerebellum	73.95 ± 6.54	39.29%	NA	<sup>24</sup>
Controls	None	80	Temporal cortex	82.55 ± 8.80	48.75%	NA	<sup>24</sup>
Controls	None	79	Cerebellum	82.67 ± 8.11	49.37%	NA	<sup>24</sup>

<sup>a</sup>Mean ± standard deviation

<sup>b</sup>*PSEN1* A79V, N135S, I143T, H163R, S170F, G206A, G217R, L226R, I229F, S290C, T245P, C410Y, A431E

NA not available

### Human brain tissue

Clinically and neuropathologically well-characterized brain tissue was obtained from the Knight ADRC and the Dominantly Inherited Alzheimer Network (DIAN), both at Washington University in St. Louis. Genotyping was performed to determine whether brain tissue carried a pathogenic mutation as described previously<sup>23</sup>. Brain tissue from the insular cortex was obtained from *MAPT* p.R406W carriers ( $n = 2$ ) and pathology-free control cases ( $n = 2$ ). We also used brain tissue from the parietal cortex from sporadic FTLD-TDP ( $n = 11$ ), sporadic late onset AD (LOAD;  $n = 80$ ), *PSEN1* (p.A79V, p.N135S, p.I143T, p.H163R, p.S170F, p.G206A, p.G217R, p.L226R, p.I229F, p.S290C, p.T245P, p.C410Y, p.A431E) with AD ( $n = 18$ ), and pathology-free control cases ( $n = 11$ ). In one case, we evaluated parietal and insular cortices from the same case (control no. 10 and 13, respectively; Supplementary Table 1). In general, control brains were matched based on age, sex and DV200, defined as the percentage of RNA fragments >200 nucleotides in size. Specifically, *MAPT* p.R406W and control brains were matched based on sex and DV200. ANNOVA was performed to demonstrate that there was no statistically significant difference between covariates ( $P$ -value = 0.988 for sex and  $P$ -value = 0.659 for DV200).

To evaluate whether genes differentially expressed in iPSC-derived neurons and brains from *MAPT* p.R406W compared with controls were also differentially expressed in primary tauopathies, we examined previously generated gene expression profiles from the temporal cortex and cerebellum of 80 control and 84 progressive supranuclear palsy (PSP) brains (syn5550404)<sup>24</sup>. Brain samples used in this study are summarized in Table 1 and Supplemental Table 1.

### RNA extraction, library construction, and sequencing

RNA was extracted from frozen brain tissues and iPSC-derived neuronal cell pellets using Tissue Lyser LT and RNeasy Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. RNA yields of individual samples were quantified by the Quant-iT RNA Assay (Life Technologies) on the Qubit Fluorometer (Fisher Scientific). Quality of yielded RNA of individual samples was assessed by the DV200 metric. DV200 of individual samples was measured with the RNA 6000 Pico Assay using the Bioanalyzer 2100 (Agilent Technologies). cDNA libraries of individual samples were constructed using the TruSeq Stranded Total RNA Sample Prep with Ribo-Zero Gold kit (Illumina) and then sequenced by the HiSeq 4000 (Illumina) at the McDonnell Genome Institute at Washington University in St. Louis with a mean of  $58.14 \pm 8.62$  million reads as previously described<sup>23</sup>.

### Quality control, alignment and quantification

Quality control, alignment and quantification was performed on all newly generated RNA-seq samples, as well as publically available sequencing reads from PSP and control brains (syn5550404)<sup>23,24</sup>. After converting the sequencing reads to FASTQ format, FastQC was used to check quality of the RNA-seq data. STAR (v. 2.5.2b) was then used to align the FASTQ files to human GRCh37 primary assembly<sup>25</sup>. We used the primary assembly and aligned reads to the assembled chromosomes, unlocalized and unplaced scaffolds, and discarded alternative haploid sequences. Sequencing metrics, including coverage, distribution of reads in the genome, ribosomal and mitochondrial contents, and alignment quality, were further obtained by applying Picard CollectRna-SeqMetrics (ver 2.8.2) to detect sample deviation.

Aligned and sorted bam files were loaded into IGV to visually identify target variants<sup>26</sup>. Samples carrying unexpected or missing expected variants were labeled as potentially swapped samples. In addition, variants were called from RNA-seq following BWA/GATK pipeline<sup>27,28</sup>. The identities of the samples were later verified by performing identity-by-descent analysis against genomic typing from genome-wide association study chipsets.

After alignment, Salmon (v. 0.7.2) was used to quantify expression levels of individual genes included in the GENCODE reference genome (GRCh37.75)<sup>29</sup>. We applied quality control, alignment and quantification as previously described<sup>23</sup>.

### Differential expression analysis

We tested for differential expression using the R (v.3.4.2) package DESeq2 (v.1.18.1)<sup>30</sup>, that uses the negative binomial distribution to calculate the statistical significance. DESeq2 reports *P* values and False Discovery Rate (FDR)—corrected *P* values for individual genes—determined using a Wald test. It has previously been demonstrated that there are many fewer spurious differentially expressed genes (DEGs) for a given false discovery rate (FDR) when isogenic controls are used with paired mutant lines rather than populations controls, even with smaller sample sizes<sup>31</sup>. A paired analysis (mutant vs. isogenic controls) is a highly sensitive and specific approach provided a more stringent FDR threshold is adopted<sup>31</sup>. Thus, we applied the more stringent Benjamini-Yekutieli FDR (B-Y FDR) correction to guarantee the specificity of the differentially expressed genes<sup>32</sup>. B-Y extends Benjamini-Hochberg (B-H) by modeling and penalizing for positive dependency, such as the comparison of multiple treatments with a single control, correlated tests or effects of latent variables on endpoints of same brain. Genes with B-Y FDR-corrected *P* values less than 0.05 were considered differentially expressed. Volcano plots were generated using built-in R function plot. Heat maps of differentially expressed genes were generated using R package pheatmap (v.1.0.8) (<https://cran.r-project.org/web/packages/pheatmap/index.html>).

### Digital deconvolution of bulk RNA-seq

To determine whether the cellular composition of iPSC-derived neurons from *MAPT* p.R406W and isogenic controls were similar, we applied a machine learning approach to digitally deconvolve bulk RNA-seq data to reveal information as to the percentage of neurons, astrocytes, microglia and oligodendrocytes in the cultures<sup>23,33</sup>. To infer cellular composition of the iPSC-neurons, we assembled and validated a reference panel composed of cell type specific samples and genes. Then we applied a digital deconvolution algorithm termed

semi-supervised non-negative matrix factorization algorithm (ssNMF<sup>34</sup>) to infer cellular composition structure.

### Functional annotation of differentially expressed genes

Functional annotation was performed on genes that were significantly differentially expressed in the same direction in both iPSC-derived neurons and brains from *MAPT* p.R406W carriers and controls. Gene relationships, including physical interaction, co-localization, pathway, shared protein domain, and genetic interaction, with or without co-expression among these verified differentially expressed genes were annotated using the Cytoscape plugin GeneMANIA (v. 3.4.1)<sup>35</sup>. Gene co-expression data were derived from brain tissues of donors free of neuropathology<sup>36</sup>. Gene topological properties were calculated using Cytoscape built-in tool NetworkAnalyzer. To identify enrichments in gene ontologic features, biological pathways and drug targets, we used Webgestalt (<http://www.webgestalt.org/>)<sup>36</sup>. In Webgestalt, the 61 differentially expressed genes were input as a target gene set and all sequenced genes were input as the reference gene set. Enrichments with B-H FDR-corrected *P* values less than 0.05 were considered significantly enriched.

### Mouseac

To determine whether genes differentially expressed in iPSC-derived neurons and brains from *MAPT* p.R406W carriers and controls were similarly altered in a mouse model of tauopathy, we evaluated gene expression in the *MAPT* p.P301L transgenic mouse model ([www.mouseac.org](http://www.mouseac.org))<sup>37</sup>. We focused on the GABA receptor genes that were differentially expressed in iPSC-derived neurons, FTLD-tau brains and PSP brains compared with controls. Briefly, microarray gene expression data was evaluated from two brain regions (cortex and hippocampus) from non-transgenic C57Bl/6J (2, 4, 8, 18 months; *n* = 4 per time point) and *MAPT* p.P301L transgenic mice under the control of a CaMKII promoter (2 months, *n* = 11, 4 months, *n* = 12; 8 months, *n* = 10; 18 months, *n* = 7). Gene expression levels were log2 transformed and expressed as a function of age. The normality of distribution was assessed using the Kolmogorov-Smirnov test. Between-group differences of normally distributed data were assessed using an unpaired *t*-test. The presence of neurofibrillary tangle (NFT) pathology was evaluated and scored as previously reported by immunohistochemistry<sup>37</sup>.

### Developmental expression of differentially expressed genes

To determine whether genes differentially expressed in iPSC-derived neurons from *MAPT* p.R406W carriers and

isogenic controls (328 genes) change during development, we evaluated human brain expression data from Brain-Span ([www.brainspan.org](http://www.brainspan.org)) across 11 developmental stages to generate a heat map<sup>38</sup>. Samples were grouped into developmental stages: embryonic (age 4–7 post-conception weeks (pcw)); early prenatal A (age 8–9 pcw); early prenatal B (age 10–12 pcw); early mid-prenatal A (age 13–15 pcw); early mid-prenatal B (age 16–18 pcw); late mid-prenatal (age 19–24 pcw); late prenatal (age 25–38 pcw); early infancy (age birth–5 months); late infancy (age 6–18 months); early childhood (age 19 months–5 years); late childhood (age 6–11 years); adolescence (age 12–19 years); and adulthood (20–40 years). Samples from each developmental stage were compared to samples from adulthood stage, which were used as the reference group to assess the developmental expression difference of differentially expressed genes. Sex and brain region were included as covariates in the analysis.

#### Gene enrichment in FTD GWAS loci

To determine whether genes differentially expressed in iPSC-derived neurons and brains from *MAPT* p.R406W carriers and controls occur within loci that increase frontotemporal dementia (FTD) risk, we used summary statistics from a genome wide association study (GWAS) of FTD<sup>39</sup>. The International FTD-Genomics Consortium (IFGC) provided phase 1 FTD-GWAS summary statistic data, which consisted of 2154 FTD cases and 4308 controls with genotypes or imputed data at 6,026,384 SNPs. The FTD dataset included multiple clinical subtypes within the FTD spectrum: behavioral variant FTD, semantic dementia, progressive non-fluent aphasia and FTD with motor neuron disease<sup>39</sup>. The criteria for annotating an FTD-GWAS SNP to a given differentially expressed gene was that the GWAS SNP was located within 5 kb of the gene boundaries.

To determine whether the differentially expressed genes in *MAPT* p.R406W carriers were significantly enriched for FTD risk, improved gene-set enrichment analysis for GWAS (*i*-GSEA4GWAS) was employed<sup>40,41</sup>. The maximum  $-\log(P\text{-value})$  of the SNPs located within 5 kb of a gene boundary was assigned to represent the gene. Instead of the commonly used phenotype label permutation, SNP label permutations were implemented to generate the distribution of the enrichment score (*ES*). A gene set significance proportion based enrichment score (*SPES*) was calculated based on gene rank using the following equation:  $SPES = k/K \times ES$ , where *k* is the proportion of significant genes of the gene set and *K* is the proportion of significant genes of the total number of genes in the GWAS. Gene-length bias was corrected by applying adaptive permutation in PLINK before *i*-GSEA4GWAS. Then, a comparative quantile-quantile (QQ) plot was used to demonstrate the differential association with FTD

between all genes and the genes differentially expressed in *MAPT* p.R406W carriers.

To demonstrate the specificity of the enrichment of genes differentially expressed in iPSC-derived neurons and brains from *MAPT* p.R406W carriers compared to controls, *i*-GSEA4GWAS was also employed on genes differentially expressed in AD, schizophrenia (SCZ), bipolar disorder (BD) and autism spectrum disorder (ASD) for FTD risk. Genes differentially expressed in SCZ, BD, and ASD were from obtained from a recently published study<sup>42</sup>.

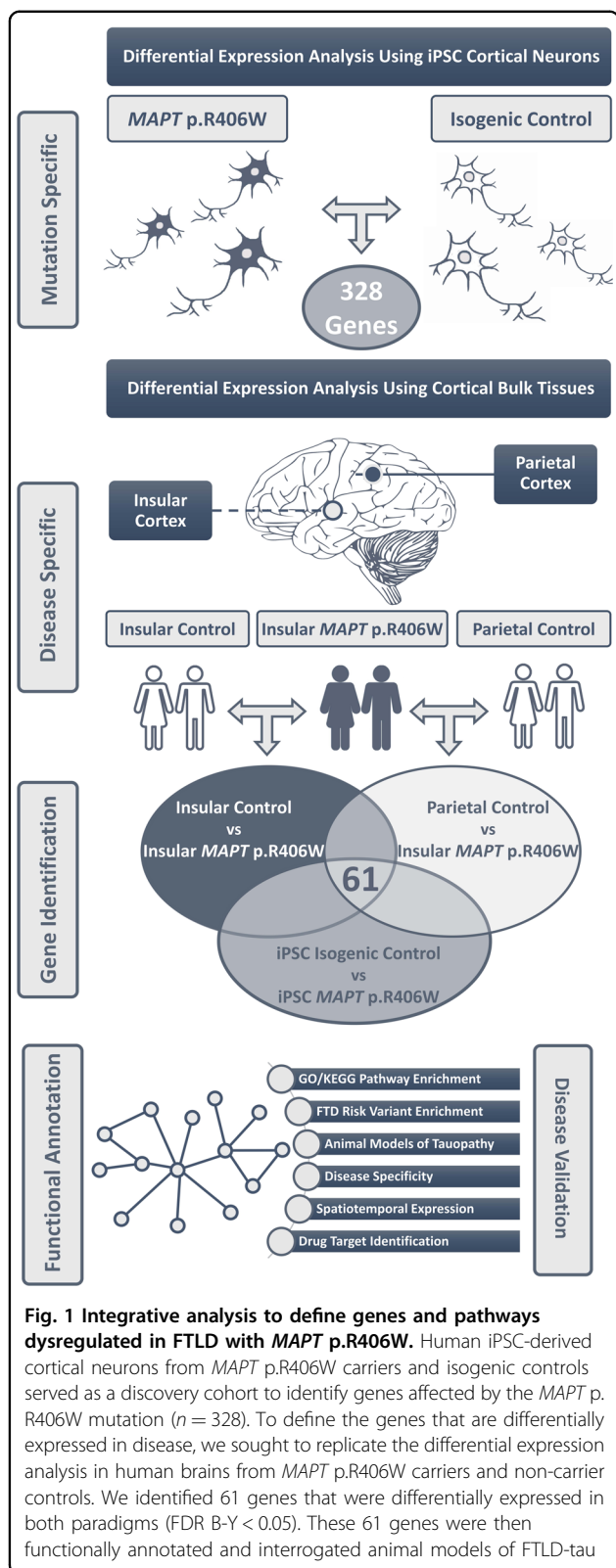
## Results

### Systems biology approach to identify molecular drivers of FTLT-tau

The goal of our study was to identify genes and networks that are critical drivers of disease pathogenesis in FTLT-tau downstream of *MAPT* mutations. As human brain tissue from FTLT with *MAPT* p.R406W mutation is rare, identifying and replicating genetic and molecular networks involved in the etiology of disease has been challenging, resulting in little focus outside of the tau protein. Thus, we sought to develop a novel integrative approach in which we coupled iPSC-derived neurons from *MAPT* p.R406W carriers and CRISPR/Cas9-generated isogenic controls as a discovery cohort with validation using human brain tissue to identify consistently differential expressed genes (Fig. 1). We also compared the differentially expressed genes in the iPSC-derived neurons vs. isogenic cell lines with those genes differentially expressed in autosomal dominant AD (*PSEN1* mutation carriers), FTLT-TDP, and PSP in order to determine if those genes were related to *MAPT* p.R406W, tau pathology or neurodegeneration in general. Finally, we performed pathway analysis and functional annotation to identify the main networks that lead to disease in *MAPT* p.R406W carriers and to identify novel potential therapeutically targets. Together, this approach led us to conclude that GABA-related molecular pathways are involved in FTLT with *MAPT* p.R406W mutation and PSP and may constitute novel targets in these diseases.

### *MAPT* p.R406W is sufficient to induce transcriptome-wide changes in iPSC-derived neurons

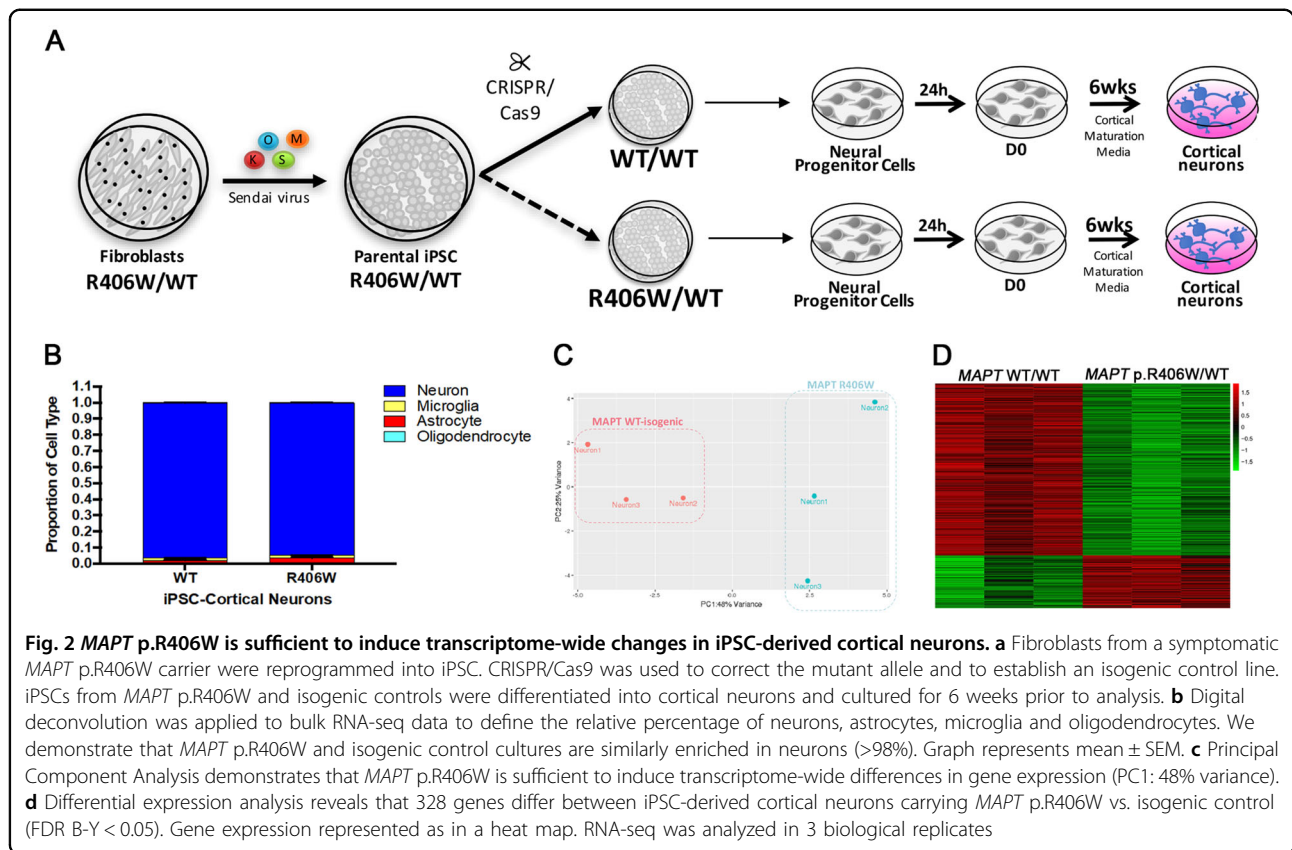
To determine whether *MAPT* p.R406W is sufficient to induce molecular changes in human neurons that may lead to neuronal dysfunction, we leveraged human somatic cells isolated from *MAPT* p.R406W carriers<sup>8</sup>. Fibroblasts were isolated from a skin punch biopsy and reprogrammed into pluripotent cells using non-integrating Sendai virus expressing *KLF4*, *OCT4*, *SOX2*, and *cMYC* (Fig. 2a; Supplemental Fig. 1). Clonal lines were picked and expanded. Only those clones expressing endogenous pluripotency markers by



immunohistochemistry and qPCR and that also displayed a normal karyotype were used in this study (Supplemental Fig. 1). Genetic background of individual donors is a large contributor to phenotypic variability in iPSCs;<sup>43</sup> thus, cell lines from non-mutation carriers are not ideal controls. To define phenotypes driven specifically by the mutant allele, we used CRISPR/Cas9 genome editing to establish isogenic controls for donor iPSC line (Fig. 2a). All cell lines used in this study were fully characterized for pluripotency and chromosomal stability (Supplemental Fig. 1). One clone from the MAPT p.R406W and the isogenic, corrected control (MAPT WT) were then differentiated into cortical neurons using a growth factor-based approach (see Methods). After 6 weeks in cortical maturation medium, we collected RNA and evaluated the whole transcriptome by RNA-seq. At 6 weeks in culture, iPSC-derived neurons produce spontaneous action potentials, form functional synapses, and display a profile of tau isoform expression that is similar to tau found in the central nervous system<sup>44</sup>.

To begin to evaluate the impact of MAPT p.R406W on the overall transcriptome profile, we first sought to determine whether the MAPT p.R406W and isogenic controls exhibited a similar capacity to form neurons. We applied a digital deconvolution approach to predict the proportion of neurons, astrocytes, microglia and oligodendrocytes in the cultures<sup>23</sup>. We found that MAPT p.R406W and isogenic controls were similarly enriched for neuronal-specific genes (Fig. 2b). Additionally, we examined the expression of individual genes specific to cortical layer markers to determine whether the MAPT p.R406W mutation influenced survival of neuronal subtypes in the culture. Deep cortical layer markers (TBR1, PCP4, TLE4); upper cortical layer markers (TLE1, SATB2, CUX1, POU3F2); layer 4 marker (RORB); layer 5 marker (BCL11B); and subplate markers (CTGF, NR4A2)<sup>45</sup> were not significantly differentially expressed between the MAPT p.R406W neurons and isogenic controls (Supplemental Table 2). Thus, we would predict that any transcriptome-wide changes between the MAPT p.R406W and isogenic controls are driven by molecular dysregulation within neurons rather than being a reflection of differences in cellular composition between cultures.

To determine whether the MAPT p.R406W mutation is sufficient to induce changes in gene expression in human neurons, we performed differential expression analyses. We found that the expression of the MAPT p.R406W mutation leads to a unique transcriptome profile that is significantly different compared to that of the isogenic controls. Principal component analysis revealed that MAPT mutation explains 48% of the transcriptome-wide variance between MAPT p.R406W and isogenic controls (Fig. 2c), and the MAPT p.R406W was sufficient to



discriminate the two groups: the mutation carriers from the isogenic controls. These results indicate that the *MAPT* p.R406W is sufficient to shift the normal transcriptome state of human neurons and some of these changes could be involved in disease pathogenesis.

Next, to define the differentially expressed genes that lead to the major changes in the transcriptome analyses, we applied the strict B-Y FDR on the 18,537 genes that were measured across all the cells. We found that 328 genes are differentially expressed between neurons expressing *MAPT* p.R406W and isogenic controls (Fig. 2d and Fig. 3a; Supplemental Table 3).

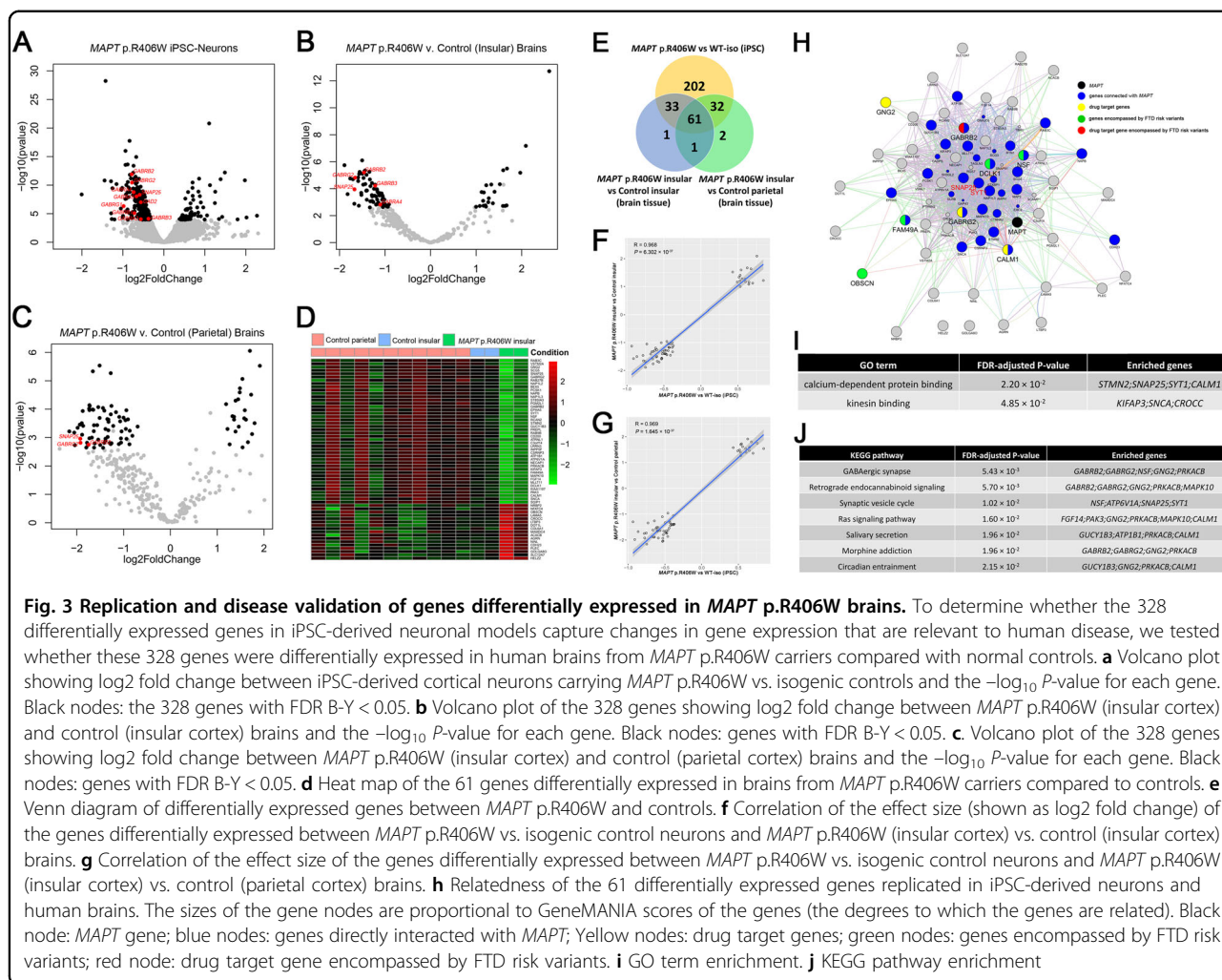
#### Gene expression patterns in brain tissue from *MAPT* p.R406W carriers overlaps with those observed in human stem cell models

Given that FTL with *MAPT* p.R406W is a neurodegenerative disease and stem cell systems are thought to represent relatively immature states, we were uncertain as to whether iPSC-derived neurons could capture molecular phenotypes relevant to a disease that develops clinical symptoms several decades after birth or whether the differentially expressed genes only capture different states of development. Combining the differential

expression analyses from iPSC-derived neurons and brains should allow us to identify genes altered as a function of the *MAPT* mutation and involved in pathology rather than those genes implicated in neurodevelopment (iPSC-only artifacts) or in general neurodegeneration and cell death-related processes (autopsied brain-only artifacts). To determine whether any of the 328 genes differentially expressed between *MAPT* p.R406W and isogenic controls are related to disease pathogenesis, we examined brain tissue from cases of FTL with *MAPT* p.R406W and non-mutation carriers free of pathology (Fig. 3). We performed RNA-seq in the insular cortex, a region affected early in disease<sup>46–48</sup>. Due to the limited amount of tissue available from *MAPT* p.R406W carriers, we expanded our dataset by including additional control brain samples (parietal cortex) from non-carriers free of pathology (Fig. 3).

Principal component analysis of the whole transcriptome illustrates that brains from FTL with *MAPT* p.R406W also cluster separately from non-mutation carriers free of pathology (Supplemental Fig. 2). We also demonstrate that control brains from the insular and parietal cortices cluster similarly, suggesting there are few overall differences in gene expression between these two





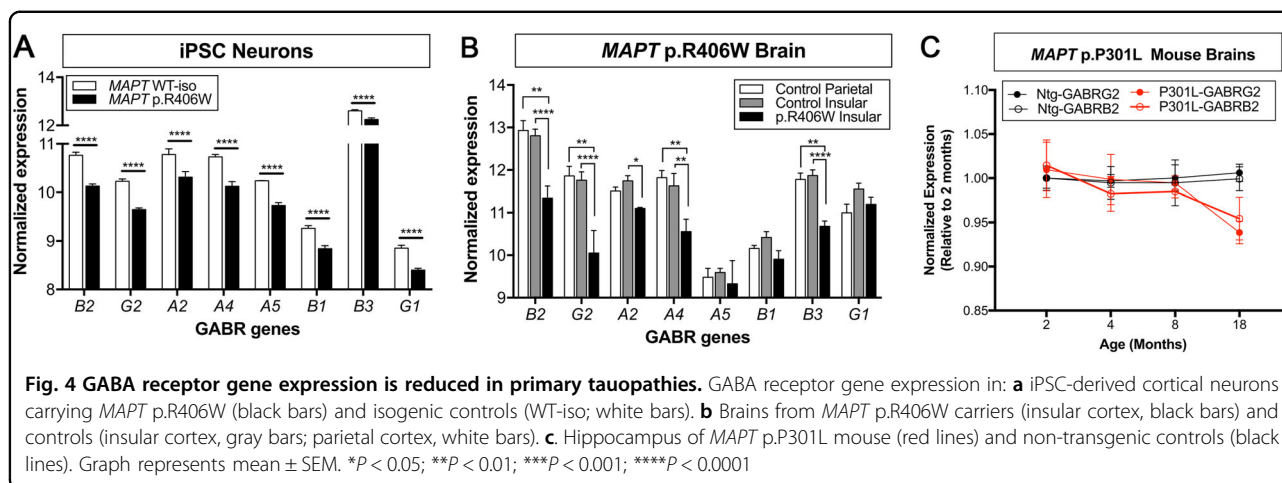
brain regions and analyses including both regions should not bias our results (Supplemental Fig. 2). Of the 328 genes differentially expressed between *MAPT* p.R406W and isogenic control neurons, we found that 61 of these genes were also differentially expressed in the same direction with a B-Y FDR < 0.05 between the insular cortex of FTLTD with *MAPT* p.R406W and controls and between the insular cortex and partial lobe of mutation carriers and controls (Fig. 3; Supplemental Table 4).

We found that these 61 differentially expressed genes exhibited similar effect sizes in the neurons and brains (Fig. 3f, g; Pearson correlation  $R = 0.969$  and  $P = 1.645 \times 10^{-37}$  for neurons vs. parietal cortices and Pearson correlation  $R = 0.968$  and  $P = 6.302 \times 10^{-37}$  for neurons vs. insular cortices). Interestingly, only three genes were differentially expressed but in the opposite direction in the neurons and brains (Supplemental Tables 3 and 4). Importantly, this suggests that our iPSC-derived neuronal system is capable of capturing gene expression changes similar to those present in human brains at autopsy. Thus,

these 61 differentially genes represent genes that are altered by the *MAPT* p.R406W mutation and are likely to be relevant to disease pathogenesis.

### Genes dysregulated by *MAPT* p.R406W are enriched for pathways involving GABA receptors and pre-synaptic function

We next sought to determine whether the 61 genes differentially expressed between *MAPT* p.R406W carriers and controls in iPSC-derived neurons and human brains were enriched in any specific pathways or whether they interacted with one another. GeneMANIA analysis revealed that more than half of the 61 genes directly interacted with *MAPT* (Fig. 3h). The GeneMANIA interactions of the 61 genes span over multiple relationships: co-expression (70.52%), co-localization (16.07%), physical interaction (7.89%), co-pathway (2.9%), computationally predicted (2.26%), genetic interaction (0.33%) and shared protein domain (0.03%). Importantly, the GABA-associated genes *SNAP25* and *SYT1* are hubs in



the interaction network (Supplemental Table 5). Hub genes are defined by the total number of nodes directly connected to a given node; rankings of hub genes also factor in betweenness and closeness centrality. Reconstructing the GeneMANIA network without co-expression revealed that the top two types of gene relationships are co-localization (55.28%) and physical interaction (26.22%). Twenty-six genes were directly connected to *MAPT*, and GABA-associated genes *SYT1* and *SNAP25* remained the hub genes of the network (Supplemental Fig. 3).

Next, we determined if these 61 genes were enriched in specific biological pathways. We found that the 61 genes are enriched for functions related to calcium-dependent presynaptic function (FDR-adjusted  $P = 2.2 \times 10^{-2}$ ; *STMN2*, *SNAP25*, *SYT1*, and *CALM1*), kinesin binding (FDR-adjusted  $P = 4.85 \times 10^{-2}$ ; *KIFAP3*, *SNCA*, *CROCC*; Fig. 3i) and pathways involving GABAergic signaling (FDR-adjusted  $P = 5.43 \times 10^{-3}$ ; *GABRB2*, *GABRG2*, *NSF*, *NGN2*, *PRKACB*; Fig. 3j). Importantly, these enriched pathways appear to be functionally related, as *SNAP25* is calcium-dependent protein binding that is critical for evoked GABA release and is expressed in the presynaptic terminals of mature GABAergic neurons<sup>49</sup>.

In addition to evaluating gene ontologies and biological pathways that are enriched by the 61 differentially expressed genes, we tested whether these genes are enriched for known drugs using Webgestalt. These analyses identified 32 GABA agonists and antagonists that target *GABRB2* and *GABRG2* (Supplemental Table 6). This served as an important proof of concept that we can leverage our molecular and biological findings with publicly available data on drug interactions to identify and prioritize putative FDA-approved drug targets and repurpose these molecules as potential treatments for FTLD with *MAPT* p.R406W.

### GABA receptors

Given the enrichment in GABAergic signaling among the 61 genes differentially expressed between *MAPT* p.R406W and controls in iPSC-derived neurons and human brains, we examined expression of all genes encoding GABA receptor subunits (Fig. 4; Supplemental Table 7). In iPSC-derived neurons expressing *MAPT* p.R406W, we observed broad downregulation of GABA receptor genes: *GABRB2*, *GABRG2*, *GABRA2*, *GABRA4*, *GABRA5*, *GABRB1*, *GABRB3*, *GABRG1* (Fig. 4a). Similarly, most of these GABA receptor genes were also downregulated in human brains from *MAPT* p.R406W carriers, with the exception of *GABRA5*, *GABRB1*, *GABRG1* (Fig. 4b), although some of these genes did not pass the stringent multiple correction we applied in the initial analyses. It is important to note that the GABA<sub>B</sub> receptors form functional heterodimers composed of GABA B1 and GABA B2 subunits. The GABA B2 subunit, which is encoded by *GABRB2*, is essential for the functional expression of the receptor dimer at the cell surface. Interestingly, the observed change in receptor expression was specific, as genes encoding subunits for AMPA and NMDA receptors were not differentially expressed (Supplemental Table 8).

### Mouse models of tauopathy display altered GABA receptor expression profiles

Our stem cell models and human brain tissue samples capture a single time point in the progression of the disease; thus, we next sought to resolve whether GABA receptor gene expression is impaired early in disease or as a consequence of pathologic events. To begin to address this question, we examined temporal expression of *Gabbr1*, *Gabrb2*, *Gabrb3*, and *Gabrg2* in a mouse model of tauopathy that overexpresses human *MAPT* p.P301L<sup>37</sup>. We observed robust and consistent down regulation of *Gabbr1*, *Gabrb2*, *Gabrb3*, and *Gabrg2* at 18 months of age in *MAPT* p.P301L mice compared with non-transgenic

**Table 2** Enrichment of genes differentially expressed by *MAPT* p.R406W among FTD risk SNPs

SNP	Chromosome	Allele 1	Allele 2	$\beta$	P value	Nearest gene	Location in Gene
rs9531160	13	T	C	-0.2723	3.54E-04	<i>DCLK1</i>	Intron
rs6743717	2	A	G	0.1804	3.63E-04	<i>FAM49A</i>	Intron
rs17521304	5	T	C	0.225	5.64E-04	<i>GABRB2</i>	Intron
rs199533	17	A	G	-0.1948	5.94E-05	<i>NSF</i>	Exon; K720K
rs3738684	1	A	T	0.2893	8.49E-05	<i>OBSCN</i>	Intron

controls but not prior (Fig. 4c; Supplemental Fig. 4). This time point coincides with the presence of overt tau aggregation and neurodegeneration in this model<sup>37</sup>. Thus, we hypothesize that *MAPT* mutations, including p.P301L and p.R406W are sufficient to induce pathologic changes that lead to altered GABAergic signaling.

#### Common variants in GABA receptor genes may influence FTD risk

To determine whether the 61 differentially expressed genes are involved more broadly in FTD, we asked whether variants present in the 61 differentially expressed genes confer risk for FTD. We used summary statistics from the largest GWAS to date of FTD ( $n = 6462$ )<sup>39</sup>. We extracted all SNPs that occur within 5 kb of the 61 differentially expressed genes. Five genes, including the GABA receptor gene, *GABRB2*, were found to be marginally associated with FTD risk (Table 2; Supplemental Table 9).

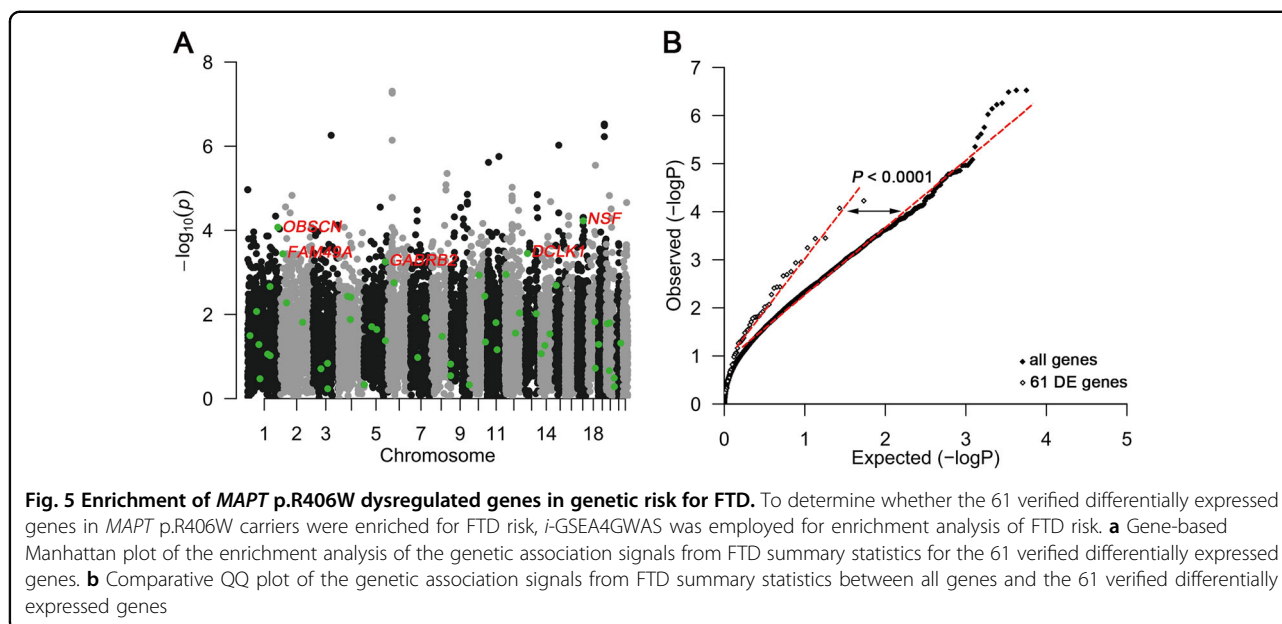
To determine whether the 61 differentially expressed genes were significantly enriched for FTD risk variants, we performed *i*-GSEA4GWAS. The 61 differentially expressed genes were defined as a gene set and enrichment analysis was performed to determine whether the FTD risk variants were significantly enriched within the gene set. The 61 genes were significantly enriched for FTD risk variants (Fig. 5a;  $P$ -value  $< 0.0001$ ). Furthermore, we demonstrate that the SNPs within the 61 differentially expressed genes are more significantly associated with FTD risk than expected by chance (Fig. 5b). To demonstrate specificity of this genetic enrichment, we sought to determine whether genes differentially expressed in AD, SCZ, BP, and ASD (Supplemental Table 10) were also significantly enriched among FTD risk variants<sup>42</sup>. We observed no significant enrichment (Supplemental Fig. 5). These analyses provide further evidence to support an important role for these 61 genes in FTD.

#### GABA receptor involvement in other neurodegenerative diseases

There is pathologic overlap between FTL with *MAPT* p.R406W and other primary tauopathies, including PSP

and CBD<sup>50</sup>. Thus, we next sought to determine whether the 328 genes differentially expressed in iPSC-derived neurons from *MAPT* p.R406W carriers and isogenic controls capture aspects of disease pathogenesis that are more broadly related to another primary tauopathy, PSP. We performed differential expression analyses in two brain regions affected in disease, temporal neocortex and cerebellum, from PSP cases and pathology-free controls. We found that 23 genes were differentially expressed in the same direction in the stem cell model and in the temporal cortex of PSP cases compared to controls (Fig. 6; Supplemental Table 11). We also found that 94 genes were differentially expressed in the same direction in the stem cell model and in the cerebellum of PSP cases compared to controls (Fig. 6b, c; Supplemental Table 11). The effect size of the expression of these genes were highly correlated in temporal cortex (Fig. 6d; Pearson correlation  $R = 0.849$  and  $P = 3.04 \times 10^{-7}$ ) and cerebellum (Fig. 6e; Pearson correlation  $R = 0.843$  and  $P = 1.454 \times 10^{-26}$ ). In total, 12 of the 328 genes (*NALCN*, *CX3CR1*, *DSP*, *MGP*, *CELSR1*, *KIAA1107*, *EMB*, *GREBIL*, *GAD2*, *LGI2*, *SLC6A1*, and *TONSL*) were differentially expressed in the same direction in the stem cell models and both PSP brain regions (Fig. 6c). Only a single gene, *KIAA1107*, was significantly reduced in stem cell models and brains from *MAPT* p.R406W carriers and PSP brains (Supplemental Fig. 6C; Supplemental Table 11). *KIAA1107* is an uncharacterized protein that has been implicated in white matter lesions in multiple sclerosis and has been implicated in risk for FTD, PSP, and CBD<sup>39,51-53</sup>.

To determine whether genes involved in GABAergic signaling are similarly differentially expressed in PSP brains, we performed pathway and differential expression analysis. We found that the GABAergic synapse pathway was significantly enriched among the differentially expressed genes in the temporal cortex and cerebellum of PSP cases (Fig. 6g). We also found that *GABRB1* and *GABRB2* were significantly lower in the temporal cortex of PSP brains and that *GABRA2*, *GABRA4*, *GABRA5*, and *GABRB1* were significantly lower in the cerebellum of PSP brains compared with pathology-free controls (Fig. 6f). Of



note, *GAD2*, which was among the 12 differentially expressed genes in both PSP brain regions and in the stem cell model, encodes an enzyme that catalyzes the decarboxylation of glutamate to GABA and  $\text{CO}_2$ .

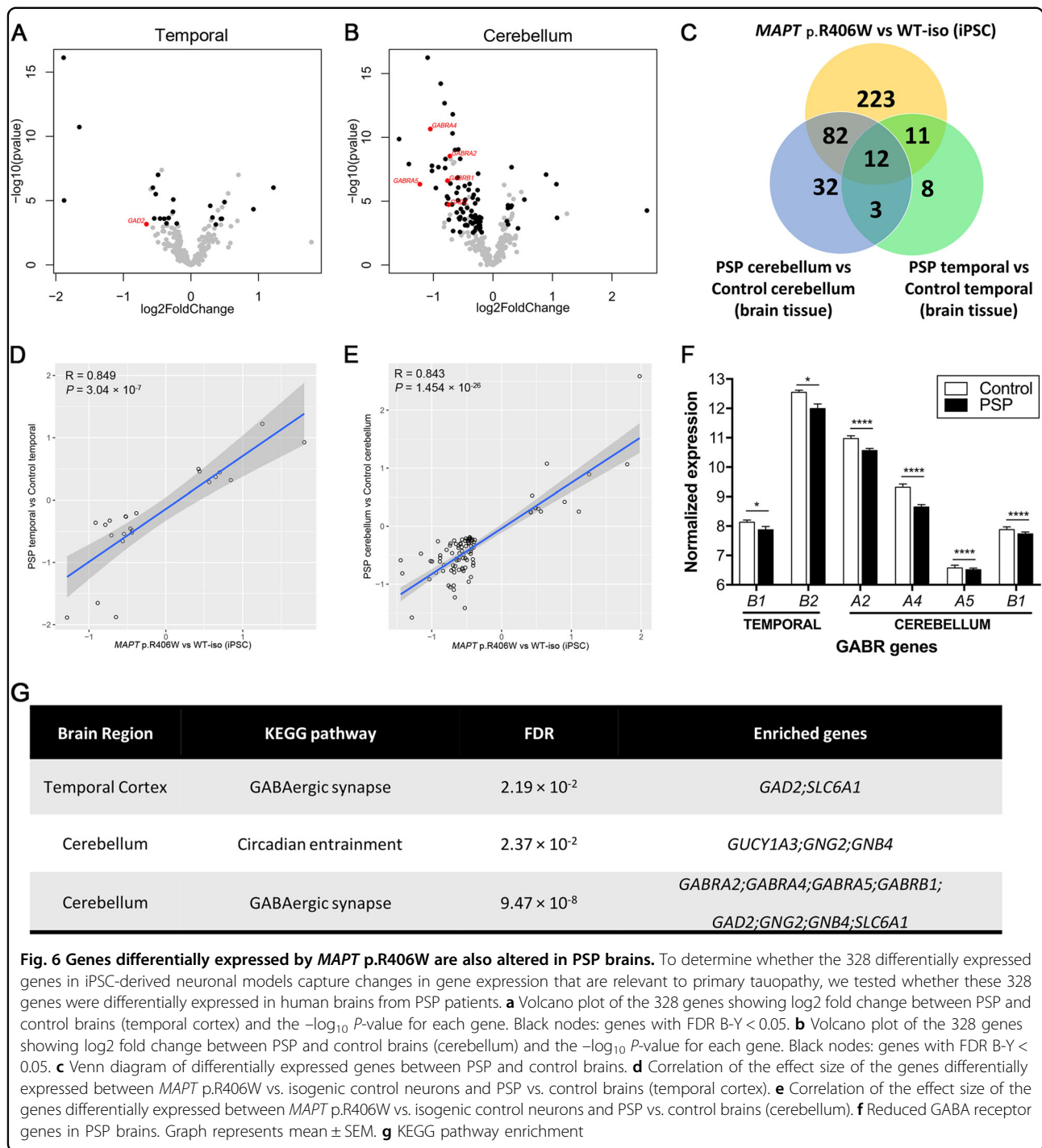
To determine whether the 328 differentially expressed genes in the stem cell model of *MAPT* p.R406W represent genes that are broadly related to neurodegenerative disease, we examined RNA-seq data from the brains of autosomal dominant AD carrying *PSEN1* mutations and from brains with FTLN-TDP that are sporadic in nature. We failed to observe any of the 328 differentially expressed genes in either dataset (Supplemental Fig. 6A-B; Supplemental Table 12). Additionally, we found that the 328 genes are not significantly differentially expressed in the human developing brain (Supplemental Fig. 7). Thus, these 328 genes do not capture disease processes related to AD, where tauopathy is an event secondary to  $\beta$ -amyloidosis; disease processes related to FTLN associated with TDP-43 inclusions; nor developmental processes. Furthermore, these results indicate that the genes shared between the stem cell model and human brains from *MAPT* p.R406W carriers did not occur by chance.

## Discussion

Here, we use a systems biology approach to identify molecular drivers of FTLN with *MAPT* p.R406W. By leveraging transcriptomics data from human stem cell models and human brain tissue and bioinformatic approaches, we demonstrate that genes enriched in pre-synaptic function are altered in FTLN with *MAPT* p.R406W. Furthermore, we demonstrate that GABA receptor expression is altered as a function of disease.

The goal of using patient-derived cell culture models is to develop tractable, human platforms that recapitulate disease-specific and patient-specific phenotypes. However, reagent costs and experimental variables preclude us and the scientific community in general from studying exceedingly large numbers of patient lines in parallel. Here, we demonstrate that even in a relative small number of patient stem cell lines and paired human brain tissues have sufficient power to detect molecular drivers of disease.

In this study, we demonstrate that GABA receptor genes are reduced in FTLN with *MAPT* p.R406W. GABA receptors occur in the pre-synaptic and post-synaptic regions where they function to bind GABA, the major inhibitory neurotransmitter in the brain. During development, GABA is the major excitatory neurotransmitter in the brain prior to the maturation of glutamatergic synapses, where it plays critical roles in neural progenitor cell proliferation, elongation of neurites and formation of synapses. There are two major classes of GABA receptors:  $\text{GABA}_A$  and  $\text{GABA}_B$ .  $\text{GABA}_A$  receptors compose the ligand-gated ion channel complex, while  $\text{GABA}_B$  receptors are G protein-coupled neuromodulatory receptors. While we observed a marginal effect of *MAPT* p.R406W on  $\text{GABA}_A$  receptor subunit genes (Supplemental Table 2 and 3), in the iPSC-derived neurons, human brains and mouse model,  $\text{GABA}_B$  receptors were significantly reduced in *MAPT* mutation carriers. Importantly, *GABRB2* was robustly reduced across the model systems. *GABRB2* encodes the GABA B2 subunit, which is the major regulator of the functional expression of the receptor dimer on the cell surface<sup>54</sup>. Thus, reduction of this gene could have a significant effect on the number of



functional GABA receptors at the synapse. Beyond the eight GABA receptor genes that were identified in the stem cell model, we also observed differential expression of genes related to GABA signaling. *SNAP25* is calcium-dependent protein binding that is critical for evoked GABA release and is expressed in the presynaptic terminals of mature GABAergic neurons<sup>49</sup>. *GAD2* encodes an enzyme that catalyzes the decarboxylation of glutamate to

GABA and CO<sub>2</sub><sup>55</sup>. Interestingly, *SYT1*, an integral membrane protein involved in synaptic vesicles, co-localizes with *KIAA1107*, the gene differentially expressed in stem cell models and brains from *MAPT* p.R406W carriers and PSP brains, in neurons<sup>56</sup>.

Defects in GABAergic signaling have been previously implicated in tauopathies. A mouse model of tauopathy expressing *MAPT* p.P301L exhibited a reduction in

hippocampal GABAergic interneurons<sup>57</sup>. Enhanced late-phase long-term potentiation was observed in these mice that could be rescued with zolpidem, a GABA<sub>A</sub>R agonist<sup>57</sup>. Several prior studies have reported qualitative evidence for the role of GABAergic neurons in FTLD-tau. *Caenorhabditis elegans* and stem cell models of an FTLD-tau risk variant in *MAPT* (*MAPT* p.A152T) exhibit abnormal morphology and degeneration of GABA-positive cells<sup>58,59</sup>. GABA<sub>A</sub> receptor activity may regulate tau phosphorylation<sup>60</sup>. GABAergic interneurons are reduced in some FTLD-tau brains, including PSP<sup>61,62</sup>. PSP brains also exhibit GABA<sub>A</sub> receptor loss by flumazenil-PET<sup>63,64</sup>. However, how these defects fit within molecular cascade that leads to disease remains unresolved. Importantly, these studies did not reveal whether changes in GABAergic neurons and GABA receptors were a consequence of early molecular dysfunction or end-stage neurodegeneration. Our findings suggest that *MAPT* p.R406W is necessary and sufficient to induce changes in GABA receptor genes and other genes that regulate GABA function. Furthermore, we demonstrate that this is maintained in human tissue, suggesting that the event triggered by the *MAPT* mutation is maintained throughout the disease course.

Drug enrichment analysis revealed a number of GABA receptor agonists and antagonists that act on the genes differentially expressed in our stem cell model and human brains. The clinical syndrome underlying FTLD-tau pathology in the brain presents with behavioral abnormalities, language and executive dysfunction, memory loss, and other changes<sup>65,66</sup>. GABA receptor agonists have been considered as a possible treatment for these clinical phenotypes<sup>67</sup>. However, our findings suggest that drugs that increase GABA receptor expression or enhance their function could be a therapeutic strategy that targets central mechanism of disease.

iGSEA4GWAS analysis showed that the 61 verified differentially expressed genes in *MAPT* p.R406W carriers are generally enriched with more FTD risk variants than average genes, indicating a genetic basis for the perturbed expression profile. Interestingly, we found that GABA-associated genes, *GABRB2* and *NSF*, are within susceptibility loci for FTLD. Two marginal risk variants ( $P$ -value =  $7.24 \times 10^{-3}$  for rs13178594 and  $P$ -value =  $9.14 \times 10^{-3}$  for rs194072) were also mapped to *GABRB2* in the GWAS for PSP risk<sup>68</sup>. Of note, the haplotype block constituted by genetic susceptibility loci of FTLD mapped to *NSF*, or *MAPT* H1 haplotype<sup>69</sup>, is a significant GWAS hit shared with FTD, PSP, CBD, and Parkinson disease<sup>53,68,70</sup>.

Our results also indicate that the iPSC model of *MAPT* p.R406W has some correlation with PSP, implicating the ability of the iPSC model of *MAPT* p.R406W to capture disease mechanisms related to other primary tauopathies.

The reduced GABA receptor genes, GABAergic pathway, and drugs targeting GABA receptors were also identified in PSP, suggesting the loss-of-function of GABAergic neurons is a common mechanism of tauopathy<sup>57,71,72</sup>. However, the iPSC model of *MAPT* p.R406W is uninformative with regards to *PSEN1* mutations in AD or FTLD-TDP, further supporting its specificity for primary tauopathy.

Performing transcriptome analyses (differential expression and network analyses) in human brain tissue has been highly informative in identifying potential co-expression networks. Constructing gene regulatory networks from 1647 post-mortem brain tissues of AD and pathology-free controls revealed dysregulation of genes involved in immune and microglia networks and identified *TYROBP* as a key regulator of these networks<sup>73</sup>. More recently, multiscale networks in human AD brains suggest that human herpesvirus subtypes may impact AD risk via regulation of genes involved in APP processing<sup>74</sup>. However, it remains unclear whether these networks are implicated in disease pathogenesis or an unrelated event in the neurodegenerative process. At the same time, performing similar analyses in iPSC-derived neural cells presents huge potential, but it is not clear whether this cellular system recapitulates the events that lead to disease late in life. The results of this study indicate that iPSC-derived neurons are highly informative and recapitulate the pathogenic events leading to disease. While, of course, not all the transcriptome changes in these models will be important to disease, when coupled with human brain tissue, it is clear that the differentially expressed genes shared between the two models are implicated in disease and are unlikely to be a mere artifact. These results also confirm that using human brain tissue from deceased individuals can be informative, but when paired with additional models, our data indicate that it may not be necessary to study hundreds of samples to identify reliable and replicable signals. In this study, we paired the human brain tissue with iPSC-derived neurons because the *MAPT* gene is mainly expressed in neurons, but we recognize that we may be missing other relevant pathways involved in non-neuronal cell types, including oligodendrocytes, astrocytes, and microglia cells. Thus, future studies using iPSC-derived glia can lead to the identification of additional pathways implicated in disease. This study indicates that pairing iPSC-derived neurons with human tissue is a highly powerful and informative approach that helps to identify novel pathways involved in disease. We predict that similar study designs may be helpful in identifying pathogenic pathways leading to neurodegeneration in other diseases and in other genes. This is especially important for studies trying to characterize different events that lead to disease in rare

(Mendelian mutations causing AD, PD or FTD) or low frequency variants (i.e., *TREM2* p.R47H) with a large effect size conditions, where it is difficult to identify a large number of brain samples that carry these variants.

#### Availability of data and materials

The datasets analyzed during the current study are available from these sources: Mayo Clinic Brain Bank: <https://www.synapse.org/#!Synapse:syn5550404>; Mount Sinai Brain Bank: <https://www.synapse.org/#!Synapse:syn3157743>; The Knight-ADRC: <https://www.synapse.org/#!Synapse:syn12181323>. According to the data request terms, DIAN data are available upon request: <http://dian.wustl.edu>.

#### Acknowledgements

This work was supported by grants from the National Institutes of Health (K01 AG046374, R01 AG056923, R01 AG056293, R01 AG044546, P01 AG003991, RF1 AG053303, R01 AG058501, U01 AG05241102, and U01 AG058922), the Tau Consortium, the Alzheimer Association (NIRG-11-200110, BAND-14-338165, and BFG-15-362540). The recruitment and clinical characterization of research participants at Washington University were supported by NIH P50 AG05681, P01 AG03991, and P01 AG026276. This work was supported by access to equipment made possible by the Hope Center for Neurological Disorders and the Departments of Neurology and Psychiatry at Washington University School of Medicine. Data collection and sharing for this project was supported by The Dominantly Inherited Alzheimer's Network (DIAN, UF1AG032438) funded by the National Institute on Aging (NIA), the German Center for Neurodegenerative Diseases (DZNE), Raul Carrea Institute for Neurological Research (FLENI), Partial support by the Research and Development Grants for Dementia from Japan Agency for Medical Research and Development, AMED, and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI). This manuscript has been reviewed by DIAN Study investigators for scientific content and consistency of data interpretation with previous DIAN Study publications. We acknowledge the altruism of the participants and their families and contributions of the DIAN research and support staff at each of the participating sites for their contributions to this study. The authors thank the International FTD-Genomics Consortium (IFGC) for providing summary statistics data for these analyses. The list of IFGC members and further acknowledgments for the IFGC can be found in the online supplementary material.

#### Author details

<sup>1</sup>Department of Psychiatry, Washington University School of Medicine, 660S. Euclid Ave. Campus Box 8134, St. Louis, MO 63110, USA. <sup>2</sup>Hope Center for Neurological Disorders, Washington University School of Medicine, 660S. Euclid Ave. Campus Box 8111, St. Louis, MO 63110, USA. <sup>3</sup>Department of Pathology and Immunology, Washington University in St. Louis, School of Medicine, 660S. Euclid Ave, Campus Box 8118, Saint Louis, MO 63110, USA

#### Conflict of interest

The authors declare that they have no conflict of interest.

#### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41398-018-0319-z>).

Received: 5 October 2018 Accepted: 13 November 2018

Published online: 13 December 2018

#### References

- Cairns, N. J. et al. Neuropathologic diagnostic and nosologic criteria for frontotemporal lobar degeneration: consensus of the Consortium for Frontotemporal Lobar Degeneration. *Acta Neuropathol.* **114**, 5–22 (2007).
- Mackenzie, I. R. et al. Nomenclature and nosology for neuropathologic subtypes of frontotemporal lobar degeneration: an update. *Acta Neuropathol.* **119**, 1–4 (2010).
- Yasuda, M. et al. A mutation in the microtubule-associated protein tau in pallido-nigro-luysian degeneration. *Neurology* **53**, 864–868 (1999).
- Ghetti, B. et al. Invited review: Frontotemporal dementia caused by microtubule-associated protein tau gene (MAPT) mutations: a chameleon for neuropathology and neuroimaging. *Neuropathol. Appl. Neurobiol.* **41**, 24–46 (2015).
- Forrest, S. L. et al. Retiring the term FTDP-17 as MAPT mutations are genetic forms of sporadic frontotemporal tauopathies. *Brain* **141**, 521–534 (2018).
- Cruts, M., Theuns, J. & Van Broeckhoven, C. Locus-specific mutation databases for neurodegenerative brain diseases. *Hum. Mutat.* **33**, 1340–1344 (2012).
- van Swieten, J. & Spillantini, M. G. Hereditary frontotemporal dementia caused by Tau gene mutations. *Brain. Pathol.* **17**, 63–73 (2007).
- Reed, L. A. et al. Autosomal dominant dementia with widespread neurofibrillary tangles. *Ann. Neurol.* **42**, 564–572 (1997).
- van Swieten, J. C. et al. Phenotypic variation in hereditary frontotemporal dementia with tau mutations. *Ann. Neurol.* **46**, 617–626 (1999).
- Lindquist, S. G. et al. Alzheimer disease-like clinical phenotype in a family with FTDP-17 caused by a MAPT R406W mutation. *Eur. J. Neurol.* **15**, 377–385 (2008).
- Miyasaka, T. et al. Molecular analysis of mutant and wild-type tau deposited in the brain affected by the FTDP-17 R406W mutation. *Am. J. Pathol.* **158**, 373–379 (2001).
- Zhang, B. et al. Retarded axonal transport of R406W mutant tau in transgenic mice with a neurodegenerative tauopathy. *J. Neurosci.* **24**, 4657–4667 (2004).
- Tatebayashi, Y. et al. Tau filament formation and associative memory deficit in aged mice expressing mutant (R406W) human tau. *Proc. Natl Acad. Sci. USA* **99**, 13896–13901 (2002).
- Bunker, J. M., Kamath, K., Wilson, L., Jordan, M. A. & Feinstein, S. C. FTDP-17 mutations compromise the ability of tau to regulate microtubule dynamics in cells. *J. Biol. Chem.* **281**, 11856–11863 (2006).
- Aoyagi, H., Hasegawa, M. & Tamaoka, A. Fibrillogenic nuclei composed of P301L mutant tau induce elongation of P301L tau but not wild-type tau. *J. Biol. Chem.* **282**, 20309–20318 (2007).
- Chang, E., Kim, S., Yin, H., Nagaraja, H. N. & Kuret, J. Pathogenic missense MAPT mutations differentially modulate tau aggregation propensity at nucleation and extension steps. *J. Neurochem.* **107**, 1113–1123 (2008).
- Brandt, R., Leger, J. & Lee, G. Interaction of tau with the neural plasma membrane mediated by tau's amino-terminal projection domain. *J. Cell Biol.* **131**, 1327–1340 (1995).
- Gauthier-Kemper, A. et al. The frontotemporal dementia mutation R406W blocks tau's interaction with the membrane in an annexin A2-dependent manner. *J. Cell Biol.* **192**, 647–661 (2011).
- Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
- Ban, H. et al. Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors. *Proc. Natl Acad. Sci. USA* **108**, 14234–14239 (2011).
- Budde, J. et al. Precision genome-editing with CRISPR/Cas9 in human induced pluripotent stem cells. *bioRxiv* 187377 (2017).
- Karch, C. M. et al. Expression of novel Alzheimer's disease risk genes in control and Alzheimer's disease brains. *PLoS ONE* **7**, e50976 (2012).
- Li, Z. et al. Genetic variants associated with Alzheimer's disease confer different cerebral cortex cell-type population structure. *Genome Med.* **10**, 43 (2018).
- Allen, M. et al. Human whole genome genotype and transcriptome data for Alzheimer's and other neurodegenerative diseases. *Sci. Data* **3**, 160089 (2016).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
- Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
- McKenna, A. et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
- Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

30. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
31. Germain, P. L. & Testa, G. Taming human genetic variability: transcriptomic meta-analysis guides the experimental design and interpretation of iPSC-based disease modeling. *Stem Cell Rep.* **8**, 1784–1796 (2017).
32. Kim, K. I. & van de Wiel, M. A. Effects of dependence in high-dimensional multiple testing problems. *BMC Bioinforma.* **9**, 114 (2008).
33. Gaujoux, R. & Seoighe, C. CellMix: a comprehensive toolbox for gene expression deconvolution. *Bioinformatics* **29**, 2211–2212 (2013).
34. Gaujoux, R. & Seoighe, C. Semi-supervised Nonnegative Matrix Factorization for gene expression deconvolution: a case study. *Infect. Genet. Evol.* **12**, 913–921 (2012).
35. Montojo, J. et al. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics* **26**, 2927–2928 (2010).
36. Wang, J., Vasaikar, S., Shi, Z., Greer, M. & Zhang, B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* **45**(W1), W130–W137 (2017).
37. Matarin, M. et al. A genome-wide gene-expression analysis and database in transgenic mice during development of amyloid or tau pathology. *Cell Rep.* **10**, 633–644 (2015).
38. Kang, H. J. et al. Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).
39. Ferrari, R. et al. Frontotemporal dementia and its subtypes: a genome-wide association study. *Lancet Neurol.* **13**, 686–699 (2014).
40. Zhang, K., Cui, S., Chang, S., Zhang, L. & Wang, J. i-GSEA4GWAS: a web server for identification of pathways/gene sets associated with traits by applying an improved gene set enrichment analysis to genome-wide association study. *Nucleic Acids Res.* **38**(Web Server issue), W90–W95 (2010).
41. Jiang, S. et al. A systems view of the differences between APOE epsilon4 carriers and non-carriers in Alzheimer's disease. *Front. Aging Neurosci.* **8**, 171 (2016).
42. Gandal, M. J. et al. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359**, 693–697 (2018).
43. Kilpinen, H. et al. Common genetic variation drives molecular heterogeneity in human iPSCs. *Nature* **546**, 370–375 (2017).
44. Sato, C. et al. Tau kinetics in neurons and the human central nervous system. *Neuron* **97**, 1284–1298 e1287 (2018).
45. Hevner, R. F. Layer-specific markers as probes for neuron type identity in human neocortex and malformations of cortical development. *J. Neuropathol. Exp. Neurol.* **66**, 101–109 (2007).
46. Rohrer, J. D. Structural brain imaging in frontotemporal dementia. *Biochim. Biophys. Acta* **1822**, 325–332 (2012).
47. Seeley, W. W. Selective functional, regional, and neuronal vulnerability in frontotemporal dementia. *Curr. Opin. Neurol.* **21**, 701–707 (2008).
48. Seeley, W. W. Anterior insula degeneration in frontotemporal dementia. *Brain Struct. Funct.* **214**, 465–475 (2010).
49. Tafoya, L. C. et al. Expression and function of SNAP-25 as a universal SNARE component in GABAergic neurons. *J. Neurosci.* **26**, 7826–7838 (2006).
50. Dickson, D. W. Neuropathologic differentiation of progressive supranuclear palsy and corticobasal degeneration. *J. Neurol.* **246**(Suppl 2), II6–II15 (1999).
51. Höglinger, G. U. et al. Identification of common variants influencing risk of the tauopathy progressive supranuclear palsy. *Nat. Genet.* **43**, 699 (2011).
52. Kouri, N. et al. Genome-wide association study of corticobasal degeneration identifies risk variants shared with progressive supranuclear palsy. *Nat. Commun.* **6**, 7247 (2015).
53. Yokoyama, J. S. et al. Shared genetic risk between corticobasal degeneration, progressive supranuclear palsy, and frontotemporal dementia. *Acta Neuropathol.* **133**, 825–837 (2017).
54. Lujan, R., Shigemoto, R. & Lopez-Bendito, G. Glutamate and GABA receptor signalling in the developing brain. *Neuroscience* **130**, 567–580 (2005).
55. Schwab, C., Yu, S., Wong, W., McGeer, E. G. & McGeer, P. L. GAD65, GAD67, and GABAT immunostaining in human brain and apparent GAD65 loss in Alzheimer's disease. *J. Alzheimers Dis.* **33**, 1073–1088 (2013).
56. Piccini, A. et al. APACHE Is an AP2-Interacting Protein Involved in Synaptic Vesicle Trafficking and Neuronal Development. *Cell Rep.* **21**, 3596–3611 (2017).
57. Levenga, J. et al. Tau pathology induces loss of GABAergic interneurons leading to altered synaptic plasticity and behavioral impairments. *Acta Neuropathol. Commun.* **1**, 34 (2013).
58. Fong, H. et al. Genetic correction of tauopathy phenotypes in neurons derived from human induced pluripotent stem cells. *Stem Cell Rep.* **1**, 226–234 (2013).
59. Pir, G. J., Choudhary, B., Mandelkow, E. & Mandelkow, E. M. Tau mutant A152T, a risk factor for FTD/PSP, induces neuronal dysfunction and reduced lifespan independently of aggregation in a *C. elegans* Tauopathy model. *Mol. Neurodegener.* **11**, 33 (2016).
60. Nykanen, N. P., Kysenius, K., Sakha, P., Tammela, P. & Huttunen, H. J. gamma-Aminobutyric acid type A (GABAA) receptor activation modulates tau phosphorylation. *J. Biol. Chem.* **287**, 6743–6752 (2012).
61. Ferrer, I. Neurons and their dendrites in frontotemporal dementia. *Dement. Geriatr. Cogn. Disord.* **10**(Suppl 1), 55–60 (1999).
62. Levy, R. et al. Alterations of GABAergic neurons in the basal ganglia of patients with progressive supranuclear palsy: an in situ hybridization study of GAD67 messenger RNA. *Neurology* **45**, 127–134 (1995).
63. Foster, N. L. et al. PET measures of benzodiazepine receptors in progressive supranuclear palsy. *Neurology* **54**, 1768–1773 (2000).
64. Murley, A. G. & Rowe, J. B. Neurotransmitter deficits from frontotemporal lobar degeneration. *Brain* **141**, 1263–1285 (2018).
65. Neary, D. et al. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology* **51**, 1546–1554 (1998).
66. Boxer, A. L. & Miller, B. L. Clinical features of frontotemporal dementia. *Alzheimer Dis. Assoc. Disord.* **19**(Suppl 1), S3–S6 (2005).
67. Solas, M., Puerta, E. & Ramirez, M. J. Treatment options in Alzheimer's disease: the GABA story. *Curr. Pharm. Des.* **21**, 4960–4971 (2015).
68. Höglinger, G. U. et al. Identification of common variants influencing risk of the tauopathy progressive supranuclear palsy. *Nat. Genet.* **43**, 699–705 (2011).
69. Stefansson, H. et al. A common inversion under selection in Europeans. *Nat. Genet.* **37**, 129–137 (2005).
70. Simon-Sanchez, J. et al. Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat. Genet.* **41**, 1308–1312 (2009).
71. Soler, H. et al. The GABAergic septohippocampal connection is impaired in a mouse model of tauopathy. *Neurobiol. Aging* **49**, 40–51 (2017).
72. Loreth, D. et al. Selective degeneration of septal and hippocampal GABAergic neurons in a mouse model of amyloidosis and tauopathy. *Neurobiol. Dis.* **47**, 1–12 (2012).
73. Zhang, B. et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* **153**, 707–720 (2013).
74. Readhead, B. et al. Multiscale analysis of independent Alzheimer's cohorts finds disruption of molecular, genetic, and clinical networks by human herpesvirus. *Neuron* **99**, 64–82.e7 (2018).