communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-026-10030-4

Cell-type-aware transcriptome-wide association studies identify 91 independent risk genes for Alzheimer's disease dementia

Qiang Liu[1,2], Randy L. Parrish[1,2], Shizhen Tang[1,2], Shinya Tasaki[3], David A. Bennett[3], Nicholas T. Seyfried[4], Philip L. De Jager[5], Vilas Menon[5], Aron S. Buchman[3] & Jingjing Yang[1]

Most existing transcriptome wide association studies (TWASs) of Alzheimer's Disease (AD) dementia only use bulk RNA-seq data and a single statistical method. Here, we utilize an omnibus TWAS (TWAS-O) pipeline that leverages multiple complementary statistical methods to integrate the snRNA-seq dataset ($n = 415$) of the dorsolateral prefrontal cortex (DLPFC) and the latest GWAS data of AD dementia. We fine-map TWAS risk genes by gene-based conditional analysis and conducted validation analyses by the analogous omnibus proteome-wide association studies (PWAS-O) using bulk proteomics data of DLPFC ($n = 716$). We identify 223 unique cell-type-aware TWAS risk genes from 350 associations across six major brain cell-types, including 91 fine-mapped independent associations, 11 of which are novel. By PWAS-O, we identify 21 significant PWAS risk genes, including 13 independent associations, which validated 31.9% independent cell-type-aware TWAS associations. By protein-protein interaction network analyses, our novel cell-type-aware TWAS findings are linked to established AD risk genes such as *APOE, BIN1*, and *MAPT*.

Alzheimer's disease (AD) is the most common cause of dementia without effective cure, affecting millions of individuals worldwide[1]. Despite extensive research, the molecular mechanisms underlying AD remain largely unknown, highlighting the potential utility of integrative analyses of genetic, transcriptomic, and proteomic data to fill this critical knowledge gap. Recently, transcriptome-wide association studies[2] (TWASs) and proteome-wide association studies[3] (PWASs) have emerged as powerful tools for integrating transcriptomic and proteomic data with summary-level data generated by genome-wide association studies[4] (GWASs). The standard two-stage TWAS/PWAS (i.e., xWAS) approach[5–8] (see "Methods") first estimates the cis- genetic effect sizes (referred to cis-xQTL weights) for each quantitative gene expression or protein abundance trait, and then takes these effect sizes as weights to conduct gene-based association test for each gene (or the corresponding protein coding gene). Genes with significant xWAS p-values have their genetic effects on the phenotype of interest in the GWAS data potentially mediated through their genetically regulated gene expressions or protein abundances.

Although recent TWASs[9–11] of AD detected >100 risk genes, including both known GWAS risk genes and novel findings, these analyses primarily studied bulk RNA sequencing (RNA-seq) data of AD-related tissues such as dorsolateral prefrontal cortex (DLPFC) and often only used one statistical method. Recent studies utilizing single-nucleus RNA sequencing (snRNA-seq, including pre-mRNA and nuclear-localized RNA) have revealed critical insights into cell-type-aware mechanisms underlying AD[12,13]. For example, genetic variants associated with AD were shown with cell-type-specific patterns of enhancer enrichment, particularly in microglia cells[14], and expression quantitative trait loci (eQTL) in microglia cells are shown most enriched with GWAS risk loci of AD[15,16]. Also, recent snRNA-seq studies identified novel biomarkers, uncovered disease-associated pathways, and elucidate the important roles of non-neuronal cells such as microglia and astrocytes in AD pathologies[17]. These findings collectively highlight the

[1]Center for Computational and Quantitative Genetics, Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA. [2]Department of Biostatistics and Bioinformatics, Emory University School of Public Health, Atlanta, GA, USA. [3]Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA. [4]Department of Biochemistry, Emory University School of Medicine, Atlanta, GA, USA. [5]Center for Translational and Computational Neuroimmunology, Department of Neurology and Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University Irving Medical Center, New York, NY, USA. e-mail: jingjing.yang@emory.edu

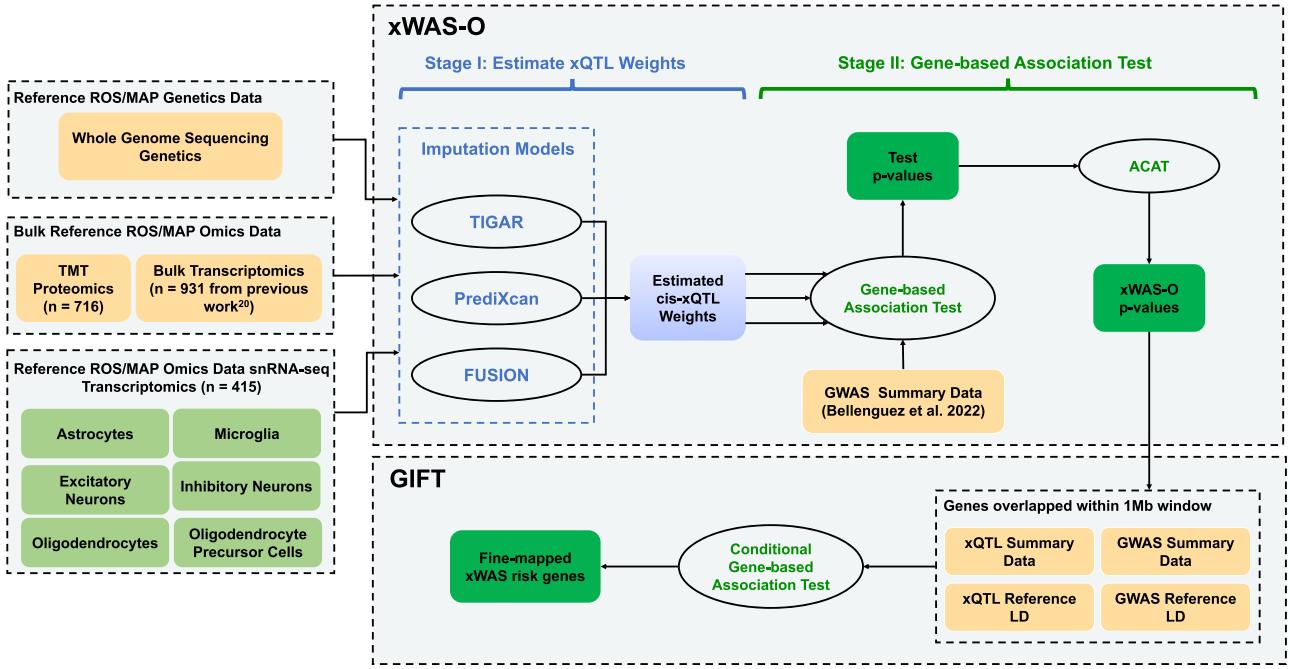

Fig. 1 | Overview of the design of this study, including inputs and outputs used in the xWAS-O pipeline. The schematic summarizes the inputs, analytical workflow, and outputs of the study. Cell-type-aware TWAS-O integrates dorsolateral prefrontal cortex snRNA-seq data from ROS/MAP ($n = 415$) with Alzheimer's disease dementia GWAS summary statistics. Bulk TWAS-O from previous work and bulk PWAS-O using dorsolateral prefrontal cortex proteomics data ($n = 716$) that was performed in this study, are included for comparison with the cell-type-aware TWAS-O findings.

importance of studies exploring cell-type-aware genomic mechanisms underlying AD.

Here, we applied our recently developed omnibus xWAS-O pipeline[18] to integrate the latest large-scale snRNA-seq data of DLPFC ($n = 415$ independent samples) from the Religious Orders Study and Rush Memory and Aging Project (ROS/MAP) Study[19] and the latest GWAS summary data of AD (Bellenguez et al.[4]; $N = \sim 789$ K independent samples)[4]. We studied six major brain cell types: astrocytes, microglia, excitatory neurons, inhibitory neurons, oligodendrocytes, and oligodendrocyte precursor cells (OPCs).

In contrast to recent cell-type-aware TWAS analyses[19] that used only one statistical method[8] to model the genetic effect sizes on gene expressions, the xWAS-O pipeline[18] aggregates TWAS findings based on multiple complementary statistical regression models (see "Methods"). Our previous publications have demonstrated that this approach improved xWAS power with calibrated false positive rates in both simulation and real studies[18,20].

To further calibrate for TWAS false positive rates, we applied the recently developed Gene-based Integrative Fine-mapping Tool[21] (GIFT) to fine-map independent risk genes. Among 223 unique significant cell-type-aware risk genes from 350 associations identified by TWAS-O, we fine-mapped 91 independent risk genes of which 11 were novel. Additionally, our analogous PWAS-O analysis using proteomics data from DLPFC ($n = 716$ independent samples) detected 21 significant PWAS risk genes, including 13 fine-mapped independent significant genes. Our PWAS-O findings validated 29 of the 91 (31.9%) independent cell-type-aware TWAS risk genes in locus level. Further, protein-protein interaction (PPI) network analyses demonstrated that many of these cell-type-aware TWAS risk genes, including novel findings, were interconnected with known AD risk genes such as *APOE*[22], *BIN1*[23], and *MAPT*[24]. Our work provides refined lists of cell-type-aware TWAS risk genes of AD dementia, along with validations by bulk PWAS in the same DLPFC tissue.

Results

Overview of the analytical procedure

As shown in Fig. 1, the xWAS-O pipeline[18] is designed to integrate transcriptomic/proteomic data with GWAS summary data using multiple complementary statistical tools (see "Methods"). Specifically, cis-xQTL weights are estimated in Stage I using the nonparametric Bayesian regression Dirichlet Process Regression (DPR) model implemented in TIGAR[5,6], penalized regression model with Elastic-Net (EN) penalty implemented in PrediXcan[7], and the best-performing model selected by FUSION[8] based on 5-fold cross-validation (CV) R^2 ("Methods"). As suggested by TIGAR-V2 paper[6], only genes with CV $R^2 > 0.5\%$ are considered as having valid prediction models with cis-xQTL genotype predictors, and then underwent further testing for the association of their genetically regulated components with AD.

We applied the xWAS-O pipeline to integrate the ROS/MAP snRNA-seq and TMT proteomics data of DLPFC and the latest GWAS summary data of AD dementia ($N = \sim 789$ K independent samples)[4] (see "Methods" about data descriptions). For each cell type, genes with omnibus TWAS p-values < 2.5E-06 (obtained by ACAT[25]) are defined as significant TWAS-O findings, which reflects a conservative genome-wide significance level corresponding to Bonferroni correction for ~20 K independent genome-wide genes. Since PWAS-O tests a smaller set of protein-coding genes (~5 K) and exhibited mild global inflation with genomic control factor $\lambda = 1.47$, we first adjusted the PWAS-O p-values according to the genomic control factor[26], and then derived q-values to control for false discovery rates (FDRs) by the Benjamini–Hochberg procedure[27]. Genes with PWAS-O q-values < 0.05 are defined as significant PWAS-O findings.

Since conventional TWAS/PWAS findings are often subject to inflated false positive rates[9–11], we employed the recently proposed tool GIFT[21] to fine-map independent significant risk genes detected by xWAS-O ("Methods"). For each genetic region containing multiple nearby significant risk genes, we used GIFT to perform the analogous gene-level conditional association testing within a ±1 Mb window around the top significant gene. Genes with independent significant associations are then reported.

We compared our cell-type-aware findings with two bulk-level analyses using reference omics data from the same ROS/MAP cohort and the same GWAS summary data of AD dementia[4]. Bulk TWAS-O results

Table 1 | The total number of test genes for studying AD dementia by three individual tools

Model	Omics data	Number of test genes	Median CV R^2
TIGAR/DPR	Astrocytes	13,482	0.013
	Excitatory neurons	13,649	0.014
	Inhibitory neurons	13,389	0.013
	Microglia	13,458	0.013
	Oligodendrocytes	13,462	0.013
	Oligodendrocyte precursor cells	13,282	0.012
	Proteomic	4069	0.013
PrediXcan/EN	Astrocytes	4536	0.012
	Excitatory neurons	4929	0.016
	Inhibitory neurons	4329	0.011
	Microglia	4126	0.011
	Oligodendrocytes	4155	0.012
	Oligodendrocyte precursor cells	3898	0.011
	Proteomic	2003	0.041
FUSION/ BestModel	Astrocytes	5478	0.0102
	Excitatory neurons	5478	0.0102
	Inhibitory neurons	5176	0.0099
	Microglia	5107	0.0095
	Oligodendrocytes	5138	0.0099
	Oligodendrocyte precursor cells	4710	0.0100
	Proteomic	4710	0.0192

Models of TIGAR/DPR, PrediXcan/EN, and FUSION/BestModel (first column) were applied to the ROS/MAP reference omics data (second column). Genes (or protein coding genes) with CV $R^2 > 0.5\%$ for the corresponding gene expression (or protein abundance) imputation models were selected as test genes in PWAS/TWAS (third column). Median CV R^2 of all test genes are presented in the fourth column.

(Supplementary Data 1) were obtained from our previously published work[20] using bulk DLPFC RNA-seq data ($n = 931$ independent samples). The total number of significant bulk TWAS-O risk genes shown in Supplementary Data 1 is less than the one reported in the previous publication (70 vs. 113), for using a more stringent genome-wide significance threshold p-values < 2.5E-6 versus the FDR < 0.05 use in the previous work. Bulk PWAS-O was newly performed in this study ($n = 716$) as shown in Fig. 1, with results provided in Supplementary Data 2.

These bulk analyses were used exclusively for downstream comparison and validation and were not incorporated into model training for any cell-type-aware TWAS-O analyses. For each cell type, its cell-type-aware TWAS-O risk genes are considered as overlapping with bulk TWAS-O findings when their test genetic regions overlapped. Independent cell-type-aware TWAS risk genes are considered as validated in PWAS-O findings, when their test genetic regions overlapped with any of the protein-coding genes with PWAS-O q-values < 0.05.

Cell-type-aware TWAS-O results

As shown in Table 1, for each cell type, ~13 K valid gene imputation models with CV $R^2 > 0.5\%$ were obtained by TIGAR/DPR with median CV $R^2 = \sim 0.013$; ~4 K valid gene imputation models were obtained by PrediXcan/EN with median CV $R^2 = \sim 0.012$; ~5 K valid gene imputation models were obtained by FUSION/BestModel with median CV $R^2 = \sim 0.01$. Scatter plots and histograms of CV R^2 for the three tools are presented in Supplementary Figs. 1–3. The substantially larger number of valid expression models from TIGAR/DPR is only because the nonparametric Bayesian DPR model best models the genetic architecture underlying most gene expression quantitative traits. The PrediXcan/EN method obtained much fewer valid prediction models for expression traits, which is mainly due to the limitation of EN method with equal proportion of L1 and L2 penalties that can only select cis-eQTL with relatively large effect sizes.

Quantile–Quantile plots of TWAS results by these three individual tools and TWAS-O are shown in Supplementary Figs. 5–10. The Manhattan plots of TWAS-O p-values of all six cell types are shown in Fig. 2. Independent TWAS-O risk genes that either have fine-mapped GIFT p-values < 0.05 or contain no overlapped test cis- genetic variants with other significant TWAS-O genes are identified, colored in red, and labeled in the Manhattan plots. Independent significant cell-type-aware TWAS-O risk genes are presented in Tables 2–4. Example GIFT locus plots from microglia are shown in Fig. 3, where all genomic regions have at least one gene prioritized by GIFT. Because of joint modeling all nearby genes in the same model by GIFT, the prioritized genes are often not the genes with most significant marginal TWAS-O p-values in their test genomic regions.

For astrocytes, we identified 65 significant cell-type-aware TWAS-O risk genes (Fig. 2A and Supplementary Data 3), including 20 that were previously linked to AD clinical or pathological traits in GWAS Catalog[28]. We fine-mapped 15 independent significant risk genes by GIFT (Table 2 and Supplementary Fig. 12), including 2 novel findings that have not been reported in GWAS Catalog nor bulk TWAS-O[20] — *CYP19A1* is involved in estrogen biosynthesis and previously reported to be associated with AD risk and neuroprotection[29]; *KNOP1* is linked to chromatin condensation and likely to contribute to AD pathology through transcription factor regulation[30].

For microglia, we identified 64 TWAS-O significant cell-type-aware risk genes (Fig. 2B and Supplementary Data 4), including 23 that were previously reported to be associated with AD-related traits in the GWAS Catalog[28]. We fine-mapped 23 independent significant genes by GIFT (Table 2 and Fig. 3), including 2 novel findings (*MED18* and *TP53INP1*). *MED18* encodes a subunit of the Mediator complex, a key regulator of transcription, which has been implicated in neurodegenerative diseases[31]. Other related mediator subunits, such as *MED12* and *MED23*, that have been found to be deregulated in AD. This suggests a potential role for *MED18* in transcriptional dysregulation and neuronal dysfunction in AD pathology[32]. *TP53INP1* is involved in oxidative stress responses, regulating p53 activity, that has been implicated in AD-related neurodegeneration[33].

In neurons, we identified 77 significant cell-type-aware TWAS-O risk genes in excitatory neurons (Supplementary Data 5) and 49 in inhibitory neurons (Supplementary Data 6), including 24 and 13 that were previously associated with AD-related traits in GWAS Catalog[28]. We fine-mapped 22 and 15 independent significant genes by GIFT, respectively in excitatory and inhibitory neurons (Table 3 and Supplementary Figs. 13 and 14). Two novel findings were identified in excitatory neurons (*PDE8B* and *BLID*). *PDE8B*, which encodes a cAMP-specific phosphodiesterase, is enriched in neurons and microglia within cognition-critical regions such as the cerebral cortex and hippocampus. Additionally, its age-dependent up-regulation in 3 × Tg-AD mice links impaired cAMP turnover to AD progression[34]. One novel finding was identified in inhibitory neurons: *KDM6B*, which encodes histone *H3K27* demethylase. This gene is essential for synaptic plasticity and cognitive function, and its Tau-dependent dysregulation has been linked to neuronal activity changes, cognitive deficits, and AD pathogenesis[35].

In oligodendrocytes and OPCs, we identified 50 and 45 significant TWAS-O risk genes, respectively (Supplementary Data 7 and 8), including 18 and 14 that were previously associated with AD-related traits in GWAS Catalog. We fine-mapped 17 and 14 independent significant genes by GIFT, respectively (Table 4 and Supplementary Figs. 15 and 16). We identified three novel independent risk genes (*KNOP1*, *PDPR* and *ITGA3*) in oligodendrocytes, and two novel independent risk genes (*GRHL3* and *PLA2G2F*) in OPCs.

To further evaluate whether the significant TWAS-O associations reflect causal genetic effects mediated through gene expression, we applied

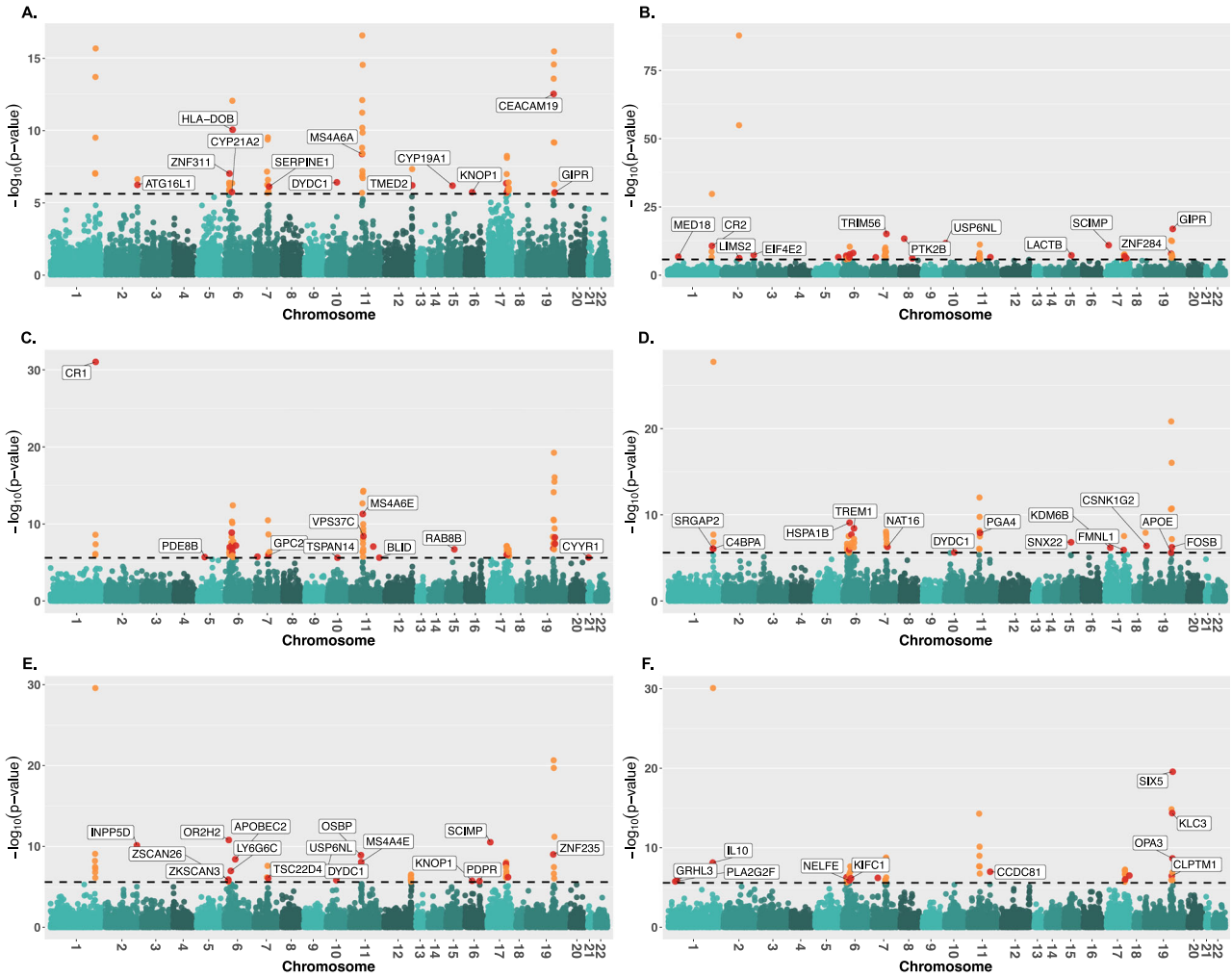

Fig. 2 | Cell-type-aware TWAS-O association signals for AD dementia across six cell types. Manhattan plots of −log10(cell-type-aware TWAS-O *p*-values) of AD dementia in astrocytes (**A**), microglia (**B**), excitatory neurons (**C**), inhibitory neurons (**D**), oligodendrocytes (**E**), and oligodendrocyte precursor cells (**F**). Transcriptome-wide significant threshold −log10(2.5E-06) was plotted as the dashed horizontal line. Independent significant genes are colored in red, including the ones detected by GIFT and the ones having no overlapped test regions with other significant TWAS-O risk genes. Most independent genes were labeled, with a few clustered with other independent genes that were not labeled due to the limitation of the "max.overlaps" parameter in the *R* plot function "ggplot". Other non-independent TWAS-O risk genes are colored in orange.

the probabilistic Mendelian randomization method PMR-Egger[36] to all cell-type-aware significant genes. Significant PMR-Egger FDRs are presented in Supplementary Data 3–8. We found that 34% of the of the total 350 significant TWAS-O associations in six cell types show significant PMR-Egger FDR < 0.05. A substantial proportion of TWAS-O independent signals showed significant mediated causal genetic effects by PMR-Egger: 9 of 15 in astrocytes, 16 of 23 in microglia, 15 of 22 in excitatory neurons, 10 of 15 in inhibitory neurons, 10 of 17 in oligodendrocytes, and 8 of 14 in OPCs. These results provide a computational validation of our identified cell-type-aware risk genes of AD dementia.

Bulk PWAS-O findings

As shown in Table 1, 4069 valid protein abundance imputation models with CV $R^2 > 0.5\%$ were obtained by TIGAR/DPR with median CV $R^2 = 0.013$; 2003 valid protein abundance imputation models were obtained by PrediXcan/EN with median CV $R^2 = 0.041$; 4710 valid protein abundance imputation models were obtained by FUSION/BestModel with median CV $R^2 = 0.0192$. Scatter plots and histograms of the CV R^2 of these three tools are presented in Supplementary Fig. 4. Because a sparse cis-genetic architecture with low heritability is observed for most quantitative protein traits, both TIGAR/DPR and FUSION/BestModel obtained similar number of valid prediction models. The much smaller number of valid protein abundance imputation models obtained by PrediXcan/EN is still due to the limitation of the Elastic-Net method with equal proportion of L1 and L2 penalties that can only select cis-eQTL with relatively large effect sizes.

Quantile-Quantile plots of PWAS-O and PWAS results by three tools are presented in Supplementary Fig. 11. We identified 21 PWAS-O significant risk genes of AD dementia, including 14 risk genes for AD-related traits previously listed in the GWAS Catalog[28], 14 overlapped with previous bulk TWAS-O findings[20], and 3 novel findings that were not associated with AD in GWAS Catalog, nor detected by our previous bulk TWAS-O analyses[20] (Fig. 4 and Supplementary Data 2). We fine-mapped 13 independent significant PWAS risk genes for AD dementia by GIFT (Supplementary Fig. 17).

We further compared significant bulk PWAS-O, bulk TWAS-O, and cell-type-aware TWAS-O findings by calculating the pair-wise Jaccard Similarity Indexes[37] from lists of unique gene names (Supplementary Fig. 18). The Jaccard Similarity Index quantifies the proportion of overlaps among the union of the compared pair. We found that cell-type-aware TWAS-O findings have 9 ~ 20% overlaps with each other, with top overlaps found between astrocytes and excitatory neurons (20%), astrocytes and oligodendrocytes (17%), microglia and inhibitory neurons (16%).

Table 2 | Independent significant cell-type-aware TWAS-O risk genes in astrocytes (Ast) and microglia (Mic)

GeneName	CHR	Cell type	TWAS-O *p*-values	GIFT *p*-values	GeneName	CHR	Cell type	TWAS-O *p*-values	GIFT *p*-values
CR2[b]	1	Mic	2.58e-11	2.75e-02	*DYDC1*[b]	10	Ast	4.04e-07	–
MED18[d,e]	1	Mic	2.38e-07	–	*USP6NL*[a]	10	Mic	2.46e-12	–
ATG16L1[c,e]	2	Ast	6.09e-07	4.34e-04	*MS4A6A*[a,b,e]	11	Ast	4.50e-09	2.94e-02
EIF4E2[c,e]	2	Mic	5.55e-08	–	*OR4D9*[b,e]	11	Mic	1.08e-06	8.92e-06
LIMS2[a,e]	2	Mic	7.59e-07	6.11e-44	*PICALM*[a,b,e]	11	Mic	3.55e-07	–
RASGEF1C[a,e]	5	Mic	3.60e-07	–	*TMED2*[b,e]	12	Ast	6.53e-07	2.75e-04
CYP21A2[b,e]	6	Ast	1.91e-06	1.28e-40	*CYP19A1*[d,e]	15	Ast	6.83e-07	–
GPSM3[b,e]	6	Mic	3.80e-07	2.56e-02	*LACTB*[b,c,e]	15	Mic	7.68e-08	–
HLA-DOB[b]	6	Ast	9.31e-11	9.19e-14	*KNOP1*[d]	16	Ast	1.95e-06	–
MCCD1[b,e]	6	Mic	1.04e-06	4.12e-02	*C1QL1*[b,c]	17	Ast	4.72e-07	1.69e-02
TAP2[b]	6	Mic	3.21e-08	8.46e-03	*RPRML*[b,c]	17	Mic	1.78e-07	2.32e-05
TREM1[a,b]	6	Mic	1.07e-08	–	*SAMD14*[d]	17	Mic	7.66e-07	–
ZNF311[b]	6	Ast	1.00e-07	1.92e-03	*SCIMP*[a,b]	17	Mic	1.43e-11	–
ZNF311[b]	6	Mic	9.85e-08	2.60e-06	*SNF8*[b,c,e]	17	Ast	1.21e-06	2.75e-06
JAZF1[a,e]	7	Mic	3.90e-07	–	*WNT9B*[b,c]	17	Ast	1.87e-06	6.74e-05
SERPINE1[a,b,e]	7	Ast	8.03e-07	1.00e-02	*CEACAM19*[a,b,c]	19	Ast	3.13e-13	1.31e-02
TRIM56[a,b,e]	7	Mic	1.07e-15	2.04e-02	*GIPR*[a,b,c]	19	Ast	2.00e-06	1.04e-04
PTK2B[a,b,c,e]	8	Mic	5.69e-14	–	*GIPR*[a,b,c]	19	Mic	1.61e-17	5.07e-04
TP53INP1[d,e]	8	Mic	9.60e-07	–	*ZNF284*[b,c,e]	19	Mic	2.00e-08	4.56e-02

All significant genes here are either not overlapping with other genes, or independent significant genes identified by GIFT. NAs are shown as "–" in the table.
[a]Known GWAS risk genes of AD.
[b]Shared TWAS risk genes of AD in bulk TWAS-O analyses, or with test regions overlapped with shared bulk TWAS-O risk genes.
[c]Shared PWAS risk genes of AD in PWAS-O analyses, or with test regions overlapped with shared PWAS-O risk genes.
[d]Novel TWAS risk genes of AD.
[e]TWAS risk genes of AD detected only in this cell type.

Validating cell-type-aware TWAS-O findings through bulk transcriptomics and proteomics

For each of the independent cell-type-aware TWAS-O risk genes, it is considered as locus-level validated through bulk transcriptomics/proteomics if its test gene region is overlapped with the ones of any significant bulk TWAS-O/PWAS-O risk genes. Among 15 independent cell-type-aware TWAS-O risk gene regions in astrocytes, we identified 12 (80%) overlapping with bulk TWAS-O results, and 6 overlapping with PWAS-O findings (40%) (Table 2). We found that 4 of these 6 genes validated in proteomics (*C1QL1, SNF8, CEACAM19*, and *ATG16L1*) are specifically identified in astrocytes. In particular, *C1QL1*, a member of the C1q complement family, is implicated in AD and other neurodegenerative disorders due to its role in synapse formation, maintenance, and clearance of abnormal protein aggregates[38]. *CEACAM19*, a carcinoembryonic antigen-related cell adhesion molecule that has been linked to mitochondrial pathology, may play a role in the mechanisms underlying AD[39]. *ATG16L1*, encodes autophagy related 16 like 1 protein, whose activation has been studied to reduce AD-like pathology[40].

In microglia, among 23 independent cell-type-aware TWAS-O risk regions, we identified 15 (65%) overlapping with bulk TWAS-O and 6 (26%) overlapping with PWAS-O findings (Table 2). We found that 4 of 6 genes validated in proteomics are specifically identified in microglia (*EIF4E2, PTK2B, LACTB*, and *ZNF284*). In particular, *EIF4E2*, encodes a cap-binding translation factor; its phosphorylation rises sharply in late-stage AD and tracks with hyper-phosphorylated Tau. These findings link EIF4E2-mediated translation to neurofibrillary tangle formation[41]. *PTK2B*, a synaptic tyrosine kinase implicated in NMDA receptor regulation, dendritic spine plasticity, and tau phosphorylation pathways, has been shown to colocalize with both Aβ and tau pathology in the entorhinal cortex (EC) of AD. *PTK2B* expression has been found in microglia in the EC of AD, suggesting a potential role of neuroinflammatory mechanisms[42]. *LACTB*, whose reduced expression in myeloid cells—and its association with higher, AD-protective succinyl-carnitine—correlates with lower risk of AD[43].

In excitatory neurons, 18 of 22 (82%) independent cell-type-aware TWAS-O risk regions overlapped with bulk TWAS-O and 7 of 22 (38%) with PWAS-O; in inhibitory neurons, 14 of 15 (93%) overlapped with bulk TWAS-O and 3 of 15 (20%) with PWAS-O (Table 3). Among risk genes validated in proteomics, 6 out of 7 excitatory neurons genes are specifically identified in excitatory neurons (*ARL4A, RAB8B, HEXIM1, ZNF225, SNRPD2*, and *FBXO46*), and all 3 inhibitory neuron genes are found specific to inhibitory neurons (*SNX22*, *FMNL1*, and *FOSB*). In particular, *RAB8B* encodes a vesicle-trafficking Rab GTPase; experimental modulation of *RAB8B* activity eases AD-like pathology, marking it as a potential therapeutic target[44]. *HEXIM1* encodes a P-TEFb-binding modulator that gates release of paused RNA polymerase II, whose elevated expression in AD cortex perturbs activity-dependent transcription crucial for synaptic plasticity and memory[45]. *SNRPD2*, whose down-regulation in MCI and AD relaxes nuclear retention of lncRNAs/mRNAs, may hasten Alzheimer's progression[46]. The dysfunction of *SNX22* was found as a susceptible factor contributing to the development or progression of AD[47].

In oligodendrocytes, 12 of 17 (70%) cell-type-aware TWAS-O risk regions overlapped with bulk TWAS-O and 3 of 17 (18%) overlapped with PWAS-O findings; in OPCs, 11 of 14 (78%) cell-type-aware TWAS-O risk regions overlapped with bulk TWAS-O and 6 of 14 (43%) overlapped with PWAS-O findings (Table 4). Among risk genes validated in proteomics, all 3 oligodendrocytes genes are found specific to oligodendrocytes (*INPP5D, ZNF235*, and *GFAP*), and 5 out of 6 OPC genes are found specific to OPCs (*LIMD2*, *CLPTM1*, *KLC3*, *OPA3*, and *SIX5*). Particularly, *INPP5D* encodes the microglial phosphatase SHIP1, whose expression increases with amyloid burden and localizes to plaque-associated microglia in AD[48]. *GFAP* encodes glial fibrillary acidic protein; higher plasma *GFAP* mirrors tau accumulation in temporal cortex and predicts faster cognitive decline, underscoring astrocytic neuroinflammation in AD progression[49]. *CLPTM1*, a transmembrane protein, has been implicated in AD, with studies showing its association with both AD and epilepsy[30]. *SIX5*, a specific transcription factor

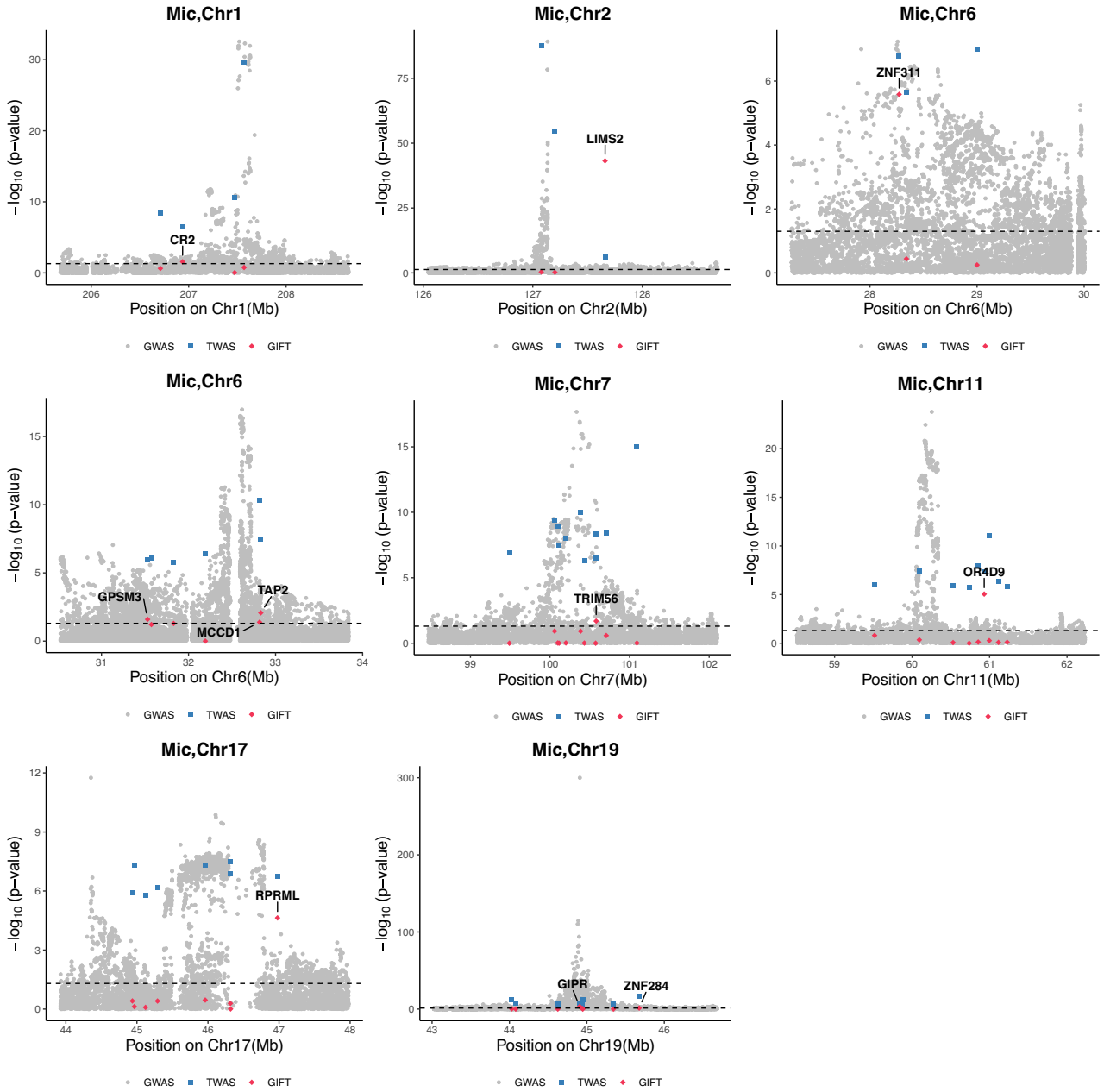

Fig. 3 | Fine-mapping results of cell-type-aware TWAS-O results of AD dementia in microglia by GIFT. Each panel represents the fine-mapped results by GIFT within one genomic region (±1 Mb around the top significant TWAS-O risk genes) that contains multiple significant TWAS-O risk genes. *X*-axis: chromosomal position (Mb); *Y*-axis: −log10(*p*-values) of SNPs by GWAS (gray dots), and of genes by TWAS-O (blue squares) and GIFT (red diamonds); Dashed horizontal line: the significance threshold of 0.05 for GIFT *p*-value. Genes with significant GIFT *p*-values < 0.05 are labeled.

(TF), plays a role in retinal function and is also involved in AD and age-related macular degeneration[50].

PPI network and enrichment analyses results

Using the STRING webtool[51] ("Methods"), we conducted PPI network analyses using the lists of PWAS-O risk genes (Fig. 4B) and cell-type-aware TWAS-O risk genes of all six brain cell types (Fig. 5). The edges in the PPI network represent existing PPI evidence and edge thickness indicates the strength of data support. In addition, pathways significantly enriched with our PWAS-O and cell-type-aware TWAS-O findings are found by the pathDIP tool[52] (Figs. 4B and 6 and Supplementary Data 9). We found 8 out of 21 (38.1%) PWAS-O risk genes interconnected in the PPI network (Fig. 4B). In particular, the major PPI network of 6 PWAS-O risk genes include gene *CLU*, a well-known AD risk gene encoding a secretory protein predominantly synthesized in brain astrocytes that would inhibit Aβ degradation and promote Aβ generation[53].

Pathway enrichment analyses with PWAS-O risk genes revealed AD-relevant significant pathways. Specifically, the pathway *Cadmium induces DNA synthesis and proliferation in macrophages* was found significant (Fig. 4C, FDR = 1.06E-02) with leading risk genes of *CLU*, *DIS3L2*, and *PPP1R37*. *PPP1R37* encodes protein phosphatase 1 regulatory subunit 37, which has been reported to be associated with AD[54]. The other significant pathway *Fc Epsilon Receptor I Signaling in Mast Cells* pathway (Fig. 4C, FDR = 1.68E-02; with leading risk genes of *CD2AP*, *FCER1G*, and *PPP1R37*), involves Mast Cells that were associated with AD progression[55].

Across all six brain cell types, 42.2–66.2% of our detected cell-type-aware TWAS-O risk genes formed interconnected PPI networks (Fig. 5), featuring well-established AD risk genes such as *APOE*[22], *BIN1*[23], and

Table 3 | Independent significant cell-type-aware TWAS-O risk genes identified in excitatory neurons (Ex) and inhibitory neurons (In)

GeneName	CHR	Cell type	TWAS-O *p*-values	GIFT *p*-values	GeneName	CHR	Cell type	TWAS-O *p*-values	GIFT *p*-values
C4BPA[b]	1	In	8.88e-07	8.04e-04	*BLID*[d,e]	11	Ex	2.38e-06	–
CR1[a,b]	1	Ex	9.25e-32	5.99e-03	*FZD4*[b,e]	11	Ex	8.49e-08	–
SRGAP2[b,e]	1	In	8.29e-07	8.60e-07	*MS4A6E*[a,b,e]	11	Ex	5.14e-12	1.34e-02
PDE8B[d,e]	5	Ex	1.97e-06	–	*PGA4*[b]	11	In	1.26e-08	1.40e-02
ABHD16A[b,e]	6	Ex	1.32e-09	9.39e-19	*VPS37C*[b,e]	11	Ex	3.93e-09	2.33e-02
APOBEC2[b]	6	Ex	6.30e-08	–	*RAB8B*[b,c,e]	15	Ex	1.96e-07	–
HSPA1B[b]	6	Ex	5.92e-08	4.03e-02	*SNX22*[b,c,e]	15	In	1.53e-07	–
HSPA1B[b]	6	In	8.15e-10	6.34e-03	*FMNL1*[b,c,e]	17	In	1.21e-06	2.59e-07
LY6G6C[b]	6	In	1.10e-06	1.24e-02	*HEXIM1*[b,c]	17	Ex	9.71e-07	1.16e-04
NELFE[b]	6	Ex	4.96e-07	3.48e-05	*KDM6B*[d,e]	17	In	6.62e-07	–
TNXB[a,b,e]	6	Ex	1.32e-07	3.32e-02	*WNT9B*[b,c]	17	Ex	9.36e-07	4.87e-22
TREM1[a,b]	6	In	3.76e-09	3.22e-03	*APOE*[a,b,c]	19	In	2.45e-06	2.91e-02
ZBTB22[b,e]	6	In	2.02e-08	7.62e-17	*CSNK1G2*[b,e]	19	In	4.08e-07	7.43e-43
ZNF311[b]	6	Ex	1.00e-07	1.02e-10	*FBXO46*[b,c,e]	19	Ex	3.76e-08	4.61e-02
ARL4A[c,e]	7	Ex	1.74e-06	–	*FOSB*[b,c]	19	In	5.76e-07	3.45e-02
GPC2[b]	7	Ex	1.32e-06	8.64e-05	*SNRPD2*[b,c,e]	19	Ex	5.51e-09	2.38e-03
NAT16[b,e]	7	In	5.29e-07	1.96e-03	*ZNF225*[a,b,c,e]	19	Ex	1.51e-07	1.37e-03
DYDC1[b]	10	In	2.31e-06	–	*CYYR1*[a,e]	21	Ex	2.24e-06	–
TSPAN14[a,b,e]	10	Ex	2.40e-06	–					

All significant genes here are either not overlapping with other genes, or independent significant genes identified by GIFT. NAs are shown as "–" in the table.
[a]Known GWAS risk genes of AD.
[b]Shared TWAS risk genes of AD in bulk TWAS-O analyses, or with test regions overlapped with shared bulk TWAS-O risk genes.
[c]Shared PWAS risk genes of AD in PWAS-O analyses, or with test regions overlapped with shared PWAS-O risk genes.
[d]Novel TWAS risk genes of AD.
[e]TWAS risk genes of AD only detected in this cell type.

Table 4 | Independent significant cell-type-aware TWAS-O risk genes identified in oligodendrocytes (Oli) and oligodendrocyte precursor cells (OPCs)

GeneName	CHR	Cell type	TWAS-O *p*-values	GIFT *p*-values	GeneName	CHR	Cell type	TWAS-O *p*-values	GIFT *p*-values
GRHL3[d,e]	1	OPC	1.19e-06	–	*CCDC81*[b]	11	OPC	1.01e-07	–
IL10[b]	1	OPC	7.24e-09	3.18e-16	*MS4A4E*[a,b]	11	Oli	8.62e-09	7.10e-06
PLA2G2F[d,e]	1	OPC	1.66e-06	–	*OSBP*[b]	11	Oli	1.18e-09	1.23e-04
INPP5D[a,c,e]	2	Oli	7.26e-11	–	*KNOP1*[d]	16	Oli	1.74e-06	–
APOBEC2[b]	6	Oli	3.79e-09	–	*PDPR*[d,e]	16	Oli	1.88e-06	–
KIFC1[b,e]	6	OPC	6.65e-07	4.86e-08	*GFAP*[b,c,e]	17	Oli	1.37e-08	4.94e-02
LY6G6C[b]	6	Oli	1.05e-07	–	*ITGA3*[d]	17	Oli	6.21e-07	–
NELFE[b]	6	OPC	1.28e-06	4.08e-03	*LIMD2*[b,c,e]	17	OPC	3.02e-07	–
OR2H2[b,e]	6	Oli	1.58e-11	8.31e-04	*RPRML*[b,c]	17	OPC	9.07e-07	1.33e-05
ZKSCAN3[b]	6	Oli	1.70e-06	1.08e-06	*SCIMP*[a,b]	17	Oli	2.96e-11	–
ZSCAN16[b,e]	6	OPC	5.04e-07	5.48e-08	*CLPTM1*[a,b,c,e]	19	OPC	3.31e-07	1.64e-02
ZSCAN26[b]	6	Oli	1.22e-06	2.51e-02	*KLC3*[a,b,c]	19	OPC	4.33e-15	4.30e-03
GPR141[a,e]	7	OPC	5.91e-07	–	*OPA3*[a,b,c,e]	19	OPC	2.11e-09	2.56e-02
TSC22D4[b,e]	7	Oli	8.48e-07	8.53e-03	*SIX5*[b,c,e]	19	OPC	2.71e-20	5.54e-03
DYDC1[b]	10	Oli	1.48e-06	–	*ZNF235*[b,c,e]	19	Oli	9.52e-10	1.86e-10
USP6NL[a]	10	Oli	4.87e-08	–					

All significant genes here are either not overlapping with other genes, or independent significant genes identified by GIFT. NAs are shown as "-" in the table.
[a]Known GWAS risk genes of AD.
[b]Shared TWAS risk genes of AD in bulk TWAS-O analyses, or with test regions overlapped with shared bulk TWAS-O risk genes.
[c]Shared PWAS risk genes of AD in PWAS-O analyses, or with test regions overlapped with shared PWAS-O risk genes.
[d]Novel TWAS risk genes of AD.
[e]TWAS risk genes of AD only detected in this cell type.

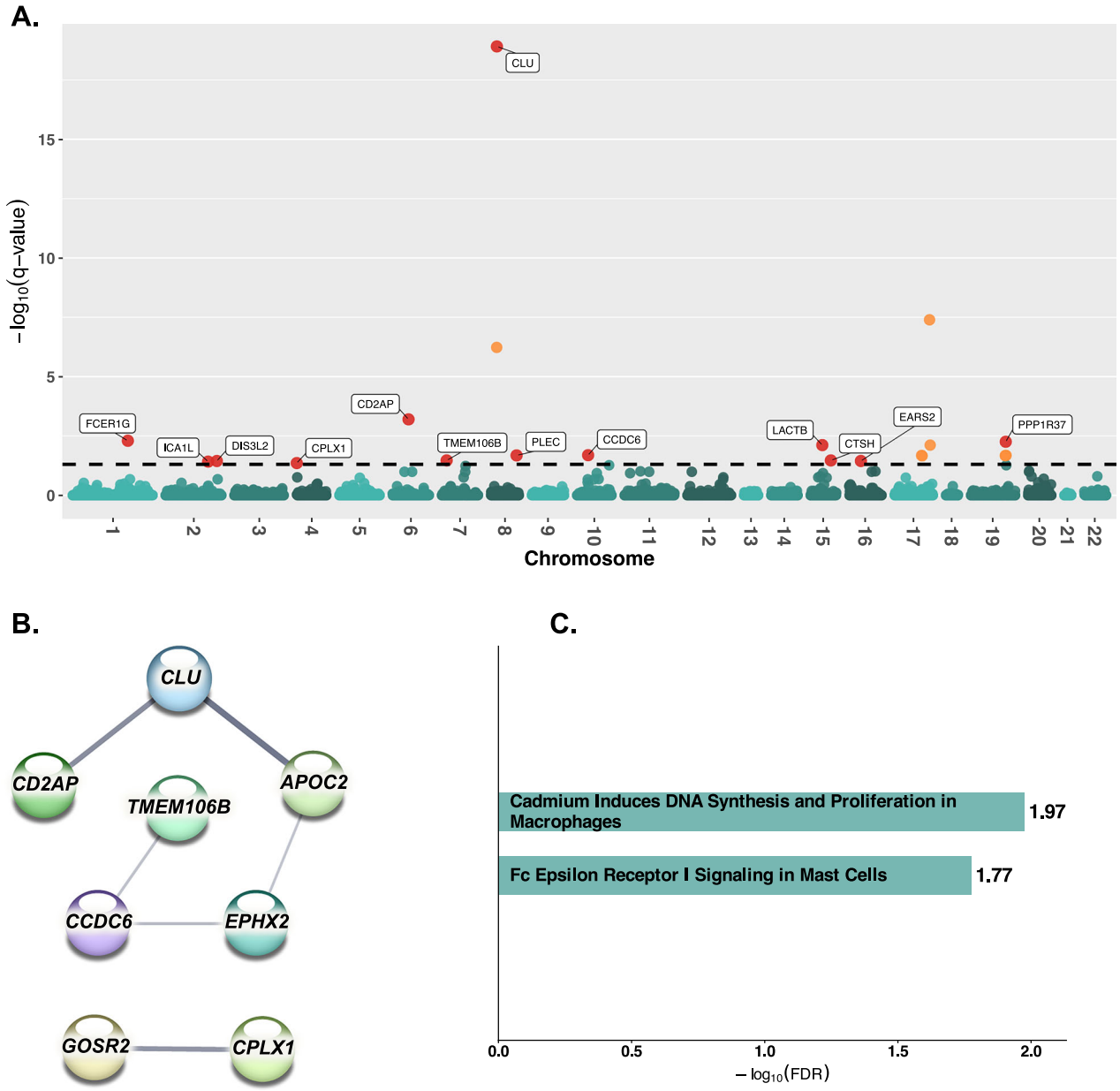

Fig. 4 | PWAS-O results of AD dementia. A. Manhattan plot of -log10(PWAS-O *q*-values): Transcriptome-wide significance threshold −log10(0.05) was plotted as the dashed horizontal line. Independent significant PWAS-O risk genes are colored in red and labelled. Other non-independent ones are colored in orange. **B.** PPI network of significant PWAS-O risk genes: Edges in the PPI network plot represent PPI links, with edge thickness indicating the strength of data support (i.e., STRING confidence score). **C.** Pathways significantly enriched with PWAS-O risk genes are plotted, with the −log10(Enrichment FDRs) in the *x*-axis.

MAPT[24] as hubs in multiple cell types. Importantly, *APOE* encodes apolipoprotein E, the brain's chief lipid-transport protein; its E4 isoform accelerates and amplifies amyloid-β plaque seeding, making *APOE* a major genetic driver of Alzheimer's risk[22]. *BIN1*, with *BIN1*-dependent alterations in neuronal properties, has been associated with AD pathophysiology[23]. *MAPT* encodes the predominant component of neurofibrillary tangles tau, which are neuropathological biomarker of AD pathology[24].

In astrocytes, 32 of 65 (49.2%) TWAS-O significant genes formed PPI networks (Fig. 5A), with known AD risk genes *MS4A1* and *MS4A6A* as hubs. *MS4A1*, which is known to serve as a functional component of a Ca2+-permeable cation channel, has been shown involved in AD pathogenesis[56]. *MS4A6A* is linked to cerebrospinal fluid soluble TREM2 levels regulation and associated with AD progression[56,57]. Specifically, the *Parkinson disease* pathway (Fig. 6A, FDR = 3.08E-02; with leading risk genes of *MS4A1*, *HEXIM1*, *ATG16L1*, and *PFKFB2*) has known neuropathological overlap with AD[58]. Interestingly, the recurrent inclusion of *BCL3* and *PPP1R9B* across multiple significant pathways underscores a pro-inflammatory, GPCR-driven astrocytic response that could modulate APP processing and neurovascular tone in AD[59,60].

In microglia, 27 of 64 (42.2%) TWAS-O risk genes were interconnected in PPI networks (Fig. 5B), with known AD risk genes *CD5*, *MAPT*, *BIN1*, and *APOE* as hubs. Importantly, *BIN1*, *MAPT*, and *MAPKAPK2*, among others, were enriched in *B cell activation* (Fig. 6B, FDR = 8.99E-07) and *CCKR signaling map* (Fig. 6B, FDR = 8.33E-04), reinforcing the role of microglial immune response in AD. The *T cell activation* pathway (Fig. 6B, FDR = 1.23E-04; with leading risk genes of *CD5*, *MAPKAPK2, MS4A2*, and *PTK2B*) is associated with T cells that are often dysfunctional in AD and are involved in AD pathology[61]. Additionally, the *VEGF signaling* pathway (Fig. 6B, FDR = 8.64E-04; with leading risk genes *CD5*, *MAPKAPK2, MAPT, and PTK2B*), is a known pathway implicated in AD progression[62].

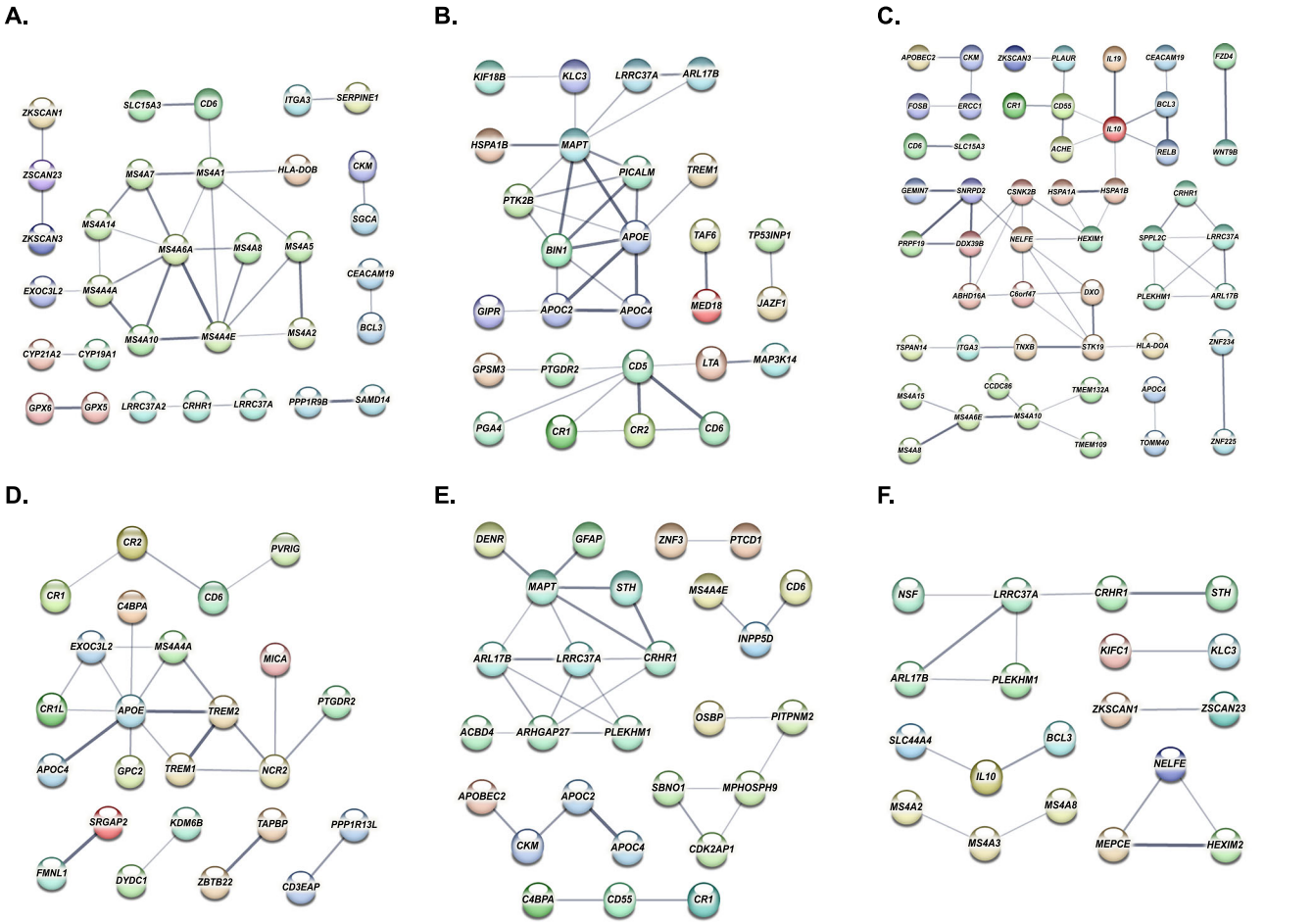

Fig. 5 | Protein-protein interaction networks of cell-type-aware TWAS-O risk genes for AD dementia. PPI network analyses results with cell-type-aware TWAS-O risk genes of AD dementia in astrocytes (**A**), microglia (**B**), excitatory neurons (**C**), inhibitory neurons (**D**), oligodendrocytes (**E**), and OPCs (**F**). Edges in the networks represent PPI links, with edge thickness indicating the strength of data support (i.e., STRING confidence score). Cell-type-aware TWAS-O risk genes not connected in the networks are not included in the network plots.

Interestingly, the convergence of B- and T-cell activation, VEGF, and cytokine pathways suggests that microglial immune signaling and angiogenic cross-talk form an integrated axis driving AD-related neuroinflammation.

Fifty-one of 77 (66.2%) excitatory neuron genes were interconnected in PPI networks (Fig. 5C), with known AD risk genes *SPPL2C, MS4A6E*, and *IL10*[63] as hubs. Enrichment analysis revealed that excitatory neuron risk genes were associated with *CCKR signaling map* (Fig. 6C, FDR = 7.20E-04) and *AD-presenilin pathway* (Fig. 6C, FDR = 1.88E-02), the most common cause of familial AD[64]. Moreover, 24 of 49 (49.0%) inhibitory neuron genes were interconnected in PPI networks (Fig. 5D), with known AD risk genes *APOE* and *TREM2* as hubs. Particularly, inhibitory neuron risk genes were enriched in *immunoregulatory interactions between lymphoid and non-lymphoid cells* (Fig. 6D, FDR = 1.48E-02).

In oligodendrocytes, 27 of 50 (54.0%) formed PPI networks (Fig. 5E), with known AD risk genes *MAPT* as hubs. Oligodendrocyte risk genes were significantly enriched in *CCKR signaling map* (Fig. 6E, FDR = 6.81E-04) and *T cell activation* (Fig. 6E, FDR = 2.15E-03); in OPCs, 19 of 45 (42.2%) risk genes formed PPI networks (Fig. 5F), with known AD risk genes *LRRC37A*[65] and *IL10*[63] as hubs. Risk genes in OPCs were also enriched in *VEGF signaling pathway* (Fig. 6F, FDR = 3.36E-02) and *T cell activation* (Fig. 6F, FDR = 3.8E-02). Notably, the *Interleukin signaling* pathway (Fig. 6F, FDR = 2.73E-02; with leading risk genes of *BCL3*, *IL10*, and *MS4A2*), is a crucial inflammatory pathway associated with AD pathology[66].

Several pathways were significant in multiple cell types, e.g., CCKR signaling map, T cell activation, B cell activation, VEGF signaling pathway, inflammation mediated by chemokine and cytokine signaling pathway, and Parkinson disease. These findings suggest that similar immune, vascular, and neurodegeneration processes affect several major brain cell types in AD. These overlapped pathways suggest shared immune and neurotransmission process in major brain cell types underlying AD pathology.

Discussion

Leveraging multiple complementary statistical methods[5–8] by xWAS-O, we identified 223 unique cell-type-aware significant risk genes across six brain cell types, which were further fine-mapped to 91 independent genes by GIFT[21], including 11 novel findings. We also performed gene-level causal inference using PMR-Egger, which tests if causal genetic effects are mediated through omics data while adjusting for horizontal pleiotropy. Comparing to the recently published cell-type-aware cis-eQTL mapping[67] derived from an independent snRNA-seq cohort of European ancestry, we found that 6–17% of our TWAS-O risk genes across all six brain cell types were identified as eGenes containing at least one significant cis-eQTL. This is expected as all cis-SNPs with non-zero effect sizes on the target gene expression, often not statistically significant cis-eQTL, are tested by TWAS-O. Across six brain cell types, PMR-Egger detected significant causal genetic effects in more than half of our fine-mapped independent cell-type-aware TWAS-O risk genes, providing a computational validation of our xWAS-O signals.

Compared to the previous cell-type-aware TWAS analyses by FUSION[8] using the same ROS/MAP snRNA-seq and GWAS summary data[19], we detected more significant TWAS risk genes, even by using the more stringent genome-wide significance threshold of p-values < 2.5E-06

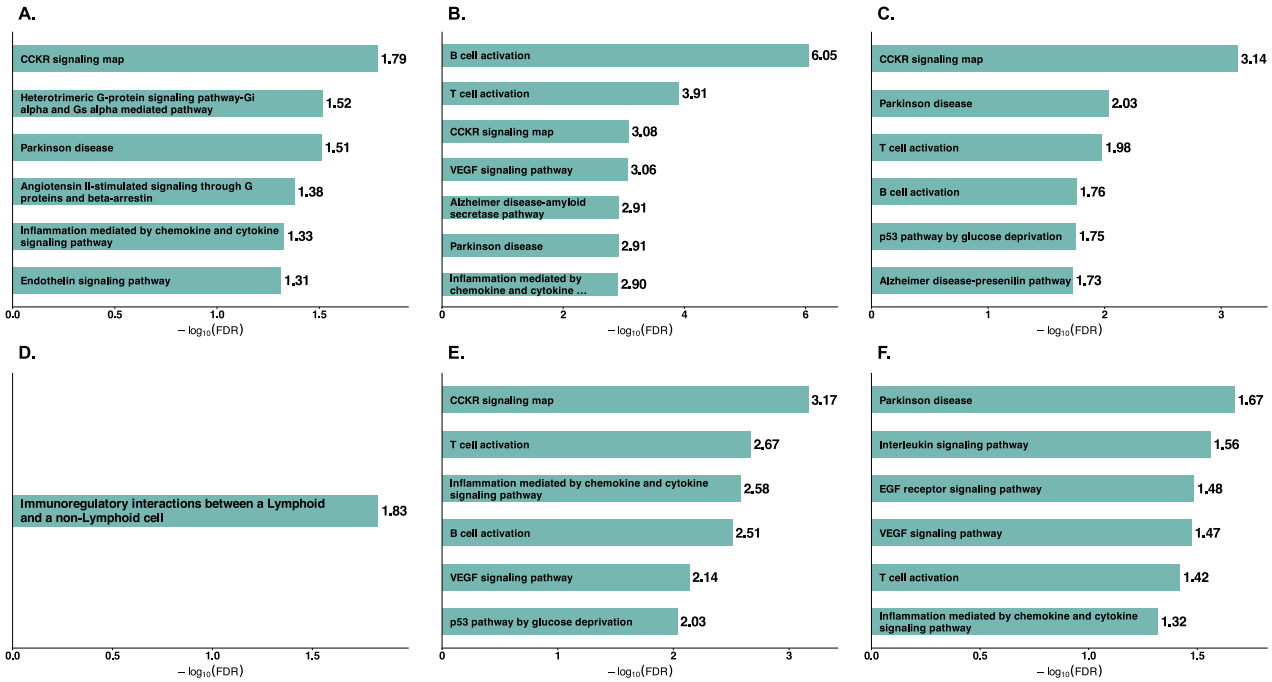

Fig. 6 | Pathway enrichment of cell-type-aware TWAS-O risk genes for AD dementia. Enrichment analyses results with cell-type-aware TWAS-O risk genes of AD dementia in astrocytes (**A**), microglia (**B**), excitatory neurons (**C**), inhibitory neurons (**D**), oligodendrocytes (**E**), and OPCs (**F**). Pathways significantly enriched with cell-type-aware TWAS-O risk genes are plotted, with the −log10(Enrichment FDRs) in the *x*-axis.

(vs. FDR < 0.05), 223 vs. 165 in the same 6 major brain cell types. This difference is likely attributable to the larger set of tested genes with expression prediction models (with CV $R^2 > 0.5\%$) trained by TIGAR/DPR, and the power gained by aggregating TWAS p-values based on three complementary statistical models (by the ACAT method). Also, 40 out of our 91 independent cell-type-aware TWAS-O findings are shown overlapped with findings by this previous study by standard FUSION using the same ROS/MAP data.

Another novel perspective of this study is that we conducted analogous PWAS-O analyses using large-scale proteomics data[68] of DLPFC in the same ROS/MAP cohorts and the same GWAS summary data[4,69]. We identified 21 significant AD risk genes, including 13 fine-mapped independent ones. Among these 21 PWAS-O risk genes, 14 were overlapped with previous PWAS findings[18,69] using a subset of the ROS/MAP proteomics data ($n = 355$ independent samples) and a different GWAS summary dataset of AD dementia[70,71]. Comparing our cell-type-aware TWAS-O findings to these PWAS-O findings, we validated 29 out of 91 (31.9%) independent cell-type-aware TWAS-O risk genes in proteomics, which were overlapped with at least one of these 21 PWAS-O significant genes.

Further PPI network analyses of lists of our cell-type-aware TWAS-O and PWAS-O risk genes detected interconnected PPI networks with known AD risk genes, such as *SPPL2C*[72], *APOE*[22], *BIN1*[23], and *MAPT*[24], as hub genes. Our detected cell-type-aware TWAS-O risk genes were found enriched in several shared AD-related pathways, such as CCKR signaling map, B cell activation, T cell activation, and p53 pathway by glucose deprivation (Fig. 6). Cell-type specific enriched pathways were also detected, such as Endothelin signaling pathway and Angiotensin II-stimulated signaling through G proteins and beta-arrestin for astrocytes, and Alzheimer disease-amyloid secretase pathway for microglia, and Alzheimer disease-presenilin pathway for excitatory neurons. Our detected PWAS-O risk genes were found enriched with known risk genes of Cadmium Induces DNA Synthesis and Proliferation in Macrophages and Fc Epsilon Receptor I Signaling in Mast Cells (Fig. 4B). These results suggest that important cell-type-specific TWAS risk genes could help illustrate cell-type-specific mechanisms underlying AD, and that the PWAS analyses would complement TWAS analyses by identifying post-transcriptionally regulated risk genes that are enriched in lipoprotein measurements.

Especially, our results reinforce the crucial role of microglia and astrocytes in AD. TWAS-O analysis in microglia showed the highest number of independent risk genes (23), and 27 out of 64 (42.2%) TWAS-O risk genes in microglia exhibited robust PPI connectivity. Well-established AD genes like *APOE*[22], *MAPT*[24], and *BIN1*[23] served as key hubs in the PPI networks. Genes such as *CD5*[73] and *MAPKAPK2*[74] were shown enrichment in T cell activation and VEGF signaling, supporting microglial involvement in neuroinflammation and cognitive aging. Importantly, novel microglia-aware risk genes (*MED18*[31,32] and *TP53INP1*[33]) were identified, suggesting transcriptional regulation and oxidative stress pathways may be underappreciated contributors to microglial-mediated AD risk.

In astrocytes, 32 out of 65 (49.2%) TWAS-O genes were interconnected in PPI networks, with hubs genes *MS4A1*[56] and *MS4A6A*[57] that were previously linked to amyloid processing and mitochondrial function. The discovery of novel astrocyte-aware risk genes, including *CYP19A1*[29] and *KNOP1*[30]—highlights emerging mechanisms involving estrogen synthesis and chromatin condensation, which may intersect with astrocyte-mediated neuroinflammatory and homeostatic functions. In both neuronal and oligodendrocyte lineage cells, over 40% of TWAS-O risk genes formed interconnected PPI networks, highlighting coordinated molecular mechanisms relevant to AD, suggesting strong functional organization within each cell type.

Beyond identifying risk genes in individual cell types, our results provide mechanistic insight into how genetic susceptibility to AD is distributed across the major neural and glial lineages. The 11 novel, GIFT fine-mapped independent genes discovered in this study highlight biological processes that may contribute to AD pathogenesis, including transcriptional regulation (*MED18*, *KNOP1*), chromatin remodeling and neuronal plasticity (*KDM6B*), oxidative stress responses (*TP53INP1*), cAMP signaling (*PDE8B*), estrogen biosynthesis (*CYP19A1*), and lipid-associated glial pathways (*PLA2G2F*, *GRHL3*). Importantly, these signals were not detectable in GWAS or bulk TWAS analyses, underscoring the added value of resolving genetically regulated expression at single-cell resolution. By

integrating the results of TWAS-O, PWAS-O, fine-mapping (GIFT), our study delivers a prioritized set of cell-type-level candidate causal genes that extend current AD biology, and provide experimentally actionable entry points for mechanistic studies. Collectively, these findings refine the etiological map of AD risk by linking genetic effects to specific cellular contexts where pathogenic processes may be initiated or amplified.

Despite these important findings of cell-type-aware TWAS-O and PWAS-O risk genes of AD in this study, several limitations should be acknowledged. First, inflation is observed in our cell-type-aware TWAS-O and PWAS-O results. This is expected due to the inflation in GWAS summary data due to LD, overlapped test cis-xQTL of nearby genes, and correlated genetically regulated gene expression or protein abundances of nearby genes. We expect to calibrate the false positive inflation rates by fine-mapping these results for independent signals, as implemented in this study.

Second, the publicly available GWAS summary statistics used in this study excluded variants with significant *p*-values < 1E-300 in the *APOE* locus, including the well-known risk variant *APOE E2 rs7412* and *E4 rs429358*. Because TWAS/PWAS framework would test all available variants within ±1 MB region around the test gene, the remaining genetic effects were still tested for gene *APOE* and its nearby genes. However, the *APOE* region is characterized by extremely strong linkage disequilibrium with the dominant *APOE E2/E4* GWAS loci, making the nearby TWAS/PWAS risk genes difficult to interpret without explicitly conditioning on *APOE E2/E4*. To avoid over-interpretation, we therefore excluded the *APOE* locus and nearby TWAS/PWAS risk genes from downstream interpretations.

Third, we only studied cis-xQTL in this study, which might miss risk genes whose gene expression or protein abundances are affected by trans-xQTL. Further studies are ongoing about exploring both cis- and trans-xQTL in xWAS analyses of AD dementia. Finally, although 31.9% of cell-type-aware TWAS-O risk genes were also validated in the PWAS-O analyses in locus level, follow-up biological validations are needed for understanding the underlying biological mechanisms of these detected risk genes.

In summary, we demonstrated that by leveraging large-scale snRNA-seq and proteomics data of DLPFC, xWAS-O findings provide a more granular characterization of AD risk genes. In particular, by leveraging multiple complementary statistical method as implemented in three tools of TIGAR[5,6], PrediXcan[7], and FUSION[8], our analyses show higher power than using only one tool. From a translational perspective, our results provide a prioritized list of cell-type-aware candidate TWAS risk genes for functional validation and therapeutic targeting. Our findings emphasize that risk genes of AD dementia might have genetic effects mediated through transcriptomic data of multiple cell types, and that our detected cell-type-aware TWAS risk genes provides valuable insights into disease etiology of AD. Our findings in this study underscore the importance of cell-type-aware TWAS and PWAS analyses in unraveling the molecular mechanisms underlying AD.

Methods

xWAS-O analytical framework

As shown in our recently published PWAS-O paper[18], prediction models for gene expression quantitative traits ($\mathbf{T_g}$) are first trained in Stage I using individual-level genetic and transcriptomic/proteomic data from reference panels. Here, we considered cell-type-aware gene expression profiled by snRNA-seq and protein abundance quantified by TMT proteomics data. Cis- genotype data ($\mathbf{G}$, ± 1 Mb of the transcription start and termination sites of the target gene) are used as predictors in the assumed multiple linear regression model (Eq. (1)), with effect sizes ($\mathbf{w}$) estimated by multiple complementary statistical methods.

$$\boldsymbol{T_g = Gw + \epsilon_g; \epsilon \sim N(0, \sigma_\epsilon^2 I)}. \quad (1)$$

Covariates such as age, sex, top 20 Probabilistic Estimation of Expression Residuals (PEER) factors[75], top 5 genotype principal components (PC), and batch effects were regressed out from the raw log2-transformed gene expression (or protein abundance) quantitative traits. Residuals after adjusting for these covariates are taken as response variables $\mathbf{T_g}$.

xWAS-O employs three imputation models: (i) nonparametric Bayesian DPR as implemented by TIGAR tool[5,6], (ii) penalized regression with Elastic-Net penalty (initially proposed in the PrediXcan[7] tool, also implemented in the TIGAR tool), (iii) best model selected by FUSION tool[8] according to CV R^2, out of single best eQTL (Top-QTL), penalized regression with LASSO penalty, penalized regression with Elastic-Net penalty, best linear unbiased predictor computed from all SNPs (BLUP). For each quantitative trait of a test gene, three sets of xQTL weights ($\hat{\mathbf{w}}$) will be estimated by TIGAR/DPR, PrediXcan/EN, and FUSION/BestModel.

As suggested by TIGAR-V2 paper[6], we consider the prediction model fitted for a gene with cross-validated prediction $R^2 > 0.5\%$ is a valid model. Only genes with valid prediction models are subject to the follow-up association test. Importantly, as shown by the TIGAR-V2 paper[6], a relaxed threshold does not inflate type I error, because TWAS *Z*-scores are linear combinations of GWAS statistics weighted by estimated eQTL effects, imprecise weights reduce power but do not introduce bias.

In Stage II, xWAS-O first uses these three sets of xQTL weights ($\hat{\mathbf{w}}$) in Eq. (1) from Stage I to predict the genetically regulated genetic component of expression (GReX) in the GWAS test data by $G_{gwas}w$, and then test the association between the phenotype *Y* and the GReX component as follows:

$$\boldsymbol{f(E[Y]) = \alpha\left(G_{gwas}\hat{w}\right); H_o : \alpha = 0, H_a : \alpha \neq 0}. \quad (2)$$

When individual genotype data of the GWAS cohort ($\mathbf{G_{gwas}}$) are not available, it is equivalent to take $\hat{\mathbf{w}}$ as variant weights to implement gene-based burden test as shown in the S-PrediXcan paper[76]. A gene-based association *Z*-score test statistic can be calculated using the summary-level GWAS test data as follows:

$$\widetilde{\boldsymbol{Z}}_{\boldsymbol{S\_PrediXcan}} = \frac{\sum_{l=1}^{m}(\hat{w}_l\hat{\sigma}_l Z_l)}{\sqrt{\hat{w}' V \hat{w}}}; \hat{\sigma_l^2} = \boldsymbol{Var}(\boldsymbol{G}_{0,l}); \boldsymbol{V = Cov}(\boldsymbol{G}_0). \quad (3)$$

Here, $\mathbf{Z_l}$ denotes the single-variant *Z*-score test statistics in GWAS summary data for the l^{th} genetic variant, $\hat{\sigma_l^2}$ denotes genotype variance of the l^{th} genetic variant, and $\mathbf{V}$ denotes the genotype variance-covariance matrix. Both $\hat{\sigma_l^2}$ and $\mathbf{V}$ could be approximated from individual-level genotype data of an external reference panel of the same ancestry as the test GWAS data (with genotype matrix $\mathbf{G_0}$).

Reference LD derived from the WGS genotype data of the ROS/MAP cohorts[68] was used for deriving xWAS *Z*-score test statistics in this study. Gene-based association test (two-sided) *p*-values can be easily derived from the *Z*-score statistic in Eq. (3) that has a standard normal distribution under the null hypothesis of no association exists between genetically regulated molecular traits and the phenotype of interest.

As shown by our previous studies[18,77], there is no single statistical method that can achieve optimal performance of modeling the underlying genetic architecture of genome-wide genes. Thus, to enhance the robustness and statistical power of gene-based association testing, xWAS-O utilizes the Aggregated Cauchy Association Test[25] (ACAT) to combine *p*-values from all three imputation tools based on complementary statistical methods as:

$$T_{ACAT} = \frac{1}{k}\sum_{k=1}^{K=3}\tan[\frac{\pi}{2}(0.5 - p_k)]; p \in [0, 1] \quad (4)$$

Here, *K* denotes the total number of xWAS *p*-values per gene, and p_k denotes the *p*-value obtained from a gene-based association test by each of these three tools. ACAT generates an omnibus test *p*-value (two-sided) for each test molecular trait of a test gene by aggregating *p*-values derived from three complementary statistical models, which is then used to identify significant xWAS risk genes[18]. By comprehensive simulation studies and real

studies of complex traits, previous studies[18,77] have shown that improved power and calibrated type I error rates were obtained using ACAT omnibus xWAS p-values.

Motivation for including TIGAR/DPR in xWAS-O

In addition to the PrediXcan/EN and FUSION/BestModel tools, xWAS-O framework incorporates the TIGAR tool that implements the nonparametric Bayesian DPR model. DPR assumes a nonparametric prior on cis-QTL effect sizes, allowing the model to capture unknown architectures. As demonstrated in previous TIGAR paper[5,6], DPR tends to yield a larger number of expression predictions models with CV $R^2 > 0.5\%$ (a criterion used to select genes/proteins for the Stage II gene-based association tests), particularly for genes with relatively low expression heritability and less sparse genetic architecture. In contrast, the widely used PrediXcan tool implements the penalized linear regression model with the Elastic-Net penalty of equal L1 and L2 penalties, which can only capture relatively large eQTL effect sizes for genes with relatively high heritability and sparse genetic architectures in practice. Because all methods in our pipeline use the same cis-window, genotype QC protocols, and filtering criteria (CV $R^2 > 0.5\%$), differences in the number of valid models reflect the modeling assumptions of each statistical model rather than technical discrepancies.

Fine-map xWAS-O risk genes by GIFT

Since conventional TWAS/PWAS findings are often subject to inflated false positive rates[9–11], we applied the Gene-based Integrative Fine-mapping through conditional TWAS[21] tool, referred to as GIFT, to fine-map genetic regions containing multiple overlapped significant xWAS-O risk genes. For each target genetic region, we first extended each significant gene's start and end positions by 1 Mb, thereby defining the flanking regions. Overlapping flanking regions among xWAS-O risk genes were then merged into a single genomic region. Second, GIFT jointly models genetically regulated gene expression or protein abundance components of all significant xWAS-O risk genes in the target genetic region, while explicitly accounting for correlations among the genetically regulated gene molecular traits and LD among cis-SNPs. As a result, GIFT can effectively pinpoint conditionally independent signals. When individual-level data are unavailable, GIFT leverages summary xQTL and GWAS data and LD derived from reference panels, enabling robust fine-mapping of xWAS-O signals.

Here, we briefly describe the underlying model assumptions by GIFT for using summary xQTL and GWAS data. For each gene i, the marginal association Z-scores from the xQTL data can be modeled as:

$$\hat{\boldsymbol{Z}}_{T_i} = \sqrt{(n_1 - 1)}\boldsymbol{\Sigma}_{1i}\boldsymbol{w}_i + \boldsymbol{\epsilon}_{T_i}, \boldsymbol{i} = \boldsymbol{1}, \ldots, \boldsymbol{k}. \quad (5)$$

Here, n_1 is the sample size for xQTL summary data, $\boldsymbol{\Sigma}_{1i}$ is the LD correlation matrix of cis-SNPs for gene i, $\mathbf{w}_i$ represents the vector of cis-SNP effect sizes on gene expression or protein abundance as in Eq. (1), and $\boldsymbol{\epsilon}_{T_i}$ captures residual errors.

The marginal Z-score statistics of all cis-SNPs in the target genetic region from the GWAS summary data can be modeled as:

$$\hat{\boldsymbol{Z}}_y = \sqrt{(n_2 - 1)}\boldsymbol{\Sigma}_2\left(\boldsymbol{\alpha}_1\boldsymbol{w}_1^T, \ldots, \boldsymbol{\alpha}_k\boldsymbol{w}_k^T\right)^T + \boldsymbol{\epsilon}_y. \quad (6)$$

Here, n_2 is the sample size for GWAS summary data; $\boldsymbol{\Sigma}_2$ is the LD correlation matrix of all cis-SNPs; $\mathbf{w}_i$ are xQTL effect sizes defined in Eq. (5); α_i denotes the conditional effect of the GReX or genetically regulated protein abundance of gene i on the phenotype, while accounting for the genetic components of all neighboring genes (α_{-i}) in the same multiple linear regression model; and $\boldsymbol{\epsilon}_y$ represents residual errors. A gene with significant conditional effect α_i will be the independent signal prioritized by GIFT. A gene with most significant marginal TWAS-O p-value may not always be the one prioritized by GIFT in a joint model of all significant genes in the same target genetic region as in Eq. (6).

Since our omnibus framework does not provide a combined xQTL weights, we implemented GIFT with xQTL and GWAS summary data. We generated xQTL summary data ($\hat{\boldsymbol{Z}}_{T_i}$) and the corresponding LD ($\boldsymbol{\Sigma}_{1i}$) from the individual-level snRNA-seq, TMT proteomics, and whole genome sequencing data of ROS/MAP samples[68]. We used GWAS LD ($\boldsymbol{\Sigma}_2$) derived from the UK Biobank[78] reference panel. Reference LD derived from different reference panels are recommended by GIFT developers for resolving potential computational issues.

Causal inference using PMR-Egger

To provide an alternative computational validation, we applied the probabilistic Mendelian Randomization tool PMR-Egger[36] to all significant TWAS-O genes. The PMR-Egger paper[36] first recognizes the same statistical framework assumed by Mendelian Randomization and TWAS. The only difference between PMR-Egger (Eq. (7)) and standard two-stage TWAS (Eqs. (1) and (2)) is whether the horizontal pleiotropy effect (γ) between the molecular trait and the phenotype is accounted for when testing for the causal genetic effect (α) mediated through the molecular trait. The PMR-Egger method jointly models cis-eQTL effect sizes ($\mathbf{w}$) and the mediated causal genetic effect size ($\alpha\mathbf{w}$), while explicitly accounting for the horizontal pleiotropy effect (γ).

$$\boldsymbol{T}_g = \boldsymbol{G}\boldsymbol{w} + \boldsymbol{\epsilon}_g; \boldsymbol{f}(\boldsymbol{E}[\boldsymbol{Y}]) = \boldsymbol{\alpha}\left(\boldsymbol{G}_{new}\boldsymbol{w}\right) + \boldsymbol{G}_{new}\boldsymbol{\gamma}. \quad (7)$$

PMR-Egger can also be implemented to test for the causal genetic effects mediated through the molecular trait by using only xQTL and GWAS summary data, along with reference LDs. However, the PMR-Egger tool suffers computation burden. Thus, we only applied the PMR-Egger to validate our cell-type-aware TWAS-O risk genes.

Protein-protein interaction network analyses by STRING

The STRING[51,79,80] webtool (version 12.0) integrates public data sources of protein interaction and analyzes the PPI network connectivity of proteins. STRING integrates evidence from curated databases, experiments, text mining, co-expression, gene neighborhood/co-occurrence, and homology to derive an overall confidence score for each association. Protein-protein edges represent the functional association evidence (i.e., proteins that are likely to jointly contribute to a shared biological function) and do not necessarily imply direct physical binding. Edge thickness indicates the strength of data support (i.e., confidence score) for the association. Minimum required interaction score used for generating PPI network was 0.4. We utilized the STRING webtool to conduct PPI network analyses with the lists of cell-type-aware TWAS-O and PWAS-O risk genes.

Enrichment analyses by pathDIP

PathDIP 5[52] integrates pathways from 6500-plus curated records across many databases, unifying them with a KEGG-/Reactome-based ontology so enrichment results can be filtered, consolidated, and mapped into 53 functional categories, with FDR q-values calculated by the Benjamini–Hochberg method. Specific pathway sources can be selected for gene enrichment analyses, which aim to detect pathways that are significantly enriched with the curated gene list versus a similar list of random genes. We utilized the pathDIP webtool to conduct gene pathway enrichment analyses with respect to data bases of Panther[81], PathBank[82], and Reactome[83] with the lists of cell-type-aware TWAS-O and PWAS-O risk genes.

ROS/MAP omics data

The ROS/MAP genetics, snRNA-seq, and proteomics data originate from two longitudinal, community-based clinic pathological cohort studies (ROS and MAP)[68]. Participants were recruited without known dementia at the time of enrollment and agreed to annual clinical evaluations and brain donation at the time of death. Each study received approval from the Institutional Review Board of Rush University Medical Center, and at the

time of enrollment all participants sign informed consent forms for the studies, Anatomical Gift Act agreements, and repository consents[84].

In this study, snRNA-seq data were profiled from 436 postmortem specimens from DLPFC from ROS/MAP decedents, yielding transcriptomic profiles of 1,546,794 cells[19]. Our analyses focused on six major brain cell types: astrocytes (Ast), microglia (Mic), excitatory neurons (Ex), inhibitory neurons (In), oligodendrocytes (Oli), and OPCs. For each gene, its pseudo-bulk expression quantitative trait was generated for each cell type by summarizing the raw reads counts across all cells of the same cell type. The raw read counts in single nuclei were combined into pseudo-bulk RNA-seq counts per gene per cell type, which were then converted to counts per million and log2 transformed for improved normality.

Human brain proteomic profiles were quantified by tandem mass spectrometry from DLPFC of 716 postmortem brains donated by participants of European ancestry from the ROS/MAP cohorts[68,69]. As described in a previous publication[85], protein abundances were normalized by scaling each protein abundance with respect to a sample-specific total protein abundance, and proteins with missing values in more than 50% of the participants were excluded from the analyses. Protein abundance ratios related to baseline (sample-specific total protein abundance) were log2 transformed for improved normality, referred to as protein abundance traits, with missing values imputed by the corresponding mean in the cohort. Outlier samples were removed using the iterative principal component analysis method as described in previous publication of PWAS analyses[85].

Confounding covariates were regressed out from the cell-type-aware pseudo-bulk expression quantitative traits and protein abundance traits. For expression quantitative traits, sex, age at death, study cohort (ROS or MAP), and top 20 PEER[75] factors were regressed out. For protein abundance traits, sex, age at death, postmortem interval, batches (i.e., sequencing batch for gene expressions or MS2 versus MS3 mass spectrometry reporter quantification mode for protein abundances), and top 5 genotype PC were regressed out.

Genotype data of ROS/MAP samples were profiled by whole genome sequencing (WGS)[86]. Genetic variants with Hardy–Weinberg equilibrium p-values < 10^{-5} and minor allele frequency > 1% were used for estimating cis-xQTL weights.

GWAS summary data of AD dementia by Bellenguez et al.

The most recent GWAS summary data (N = ~789 K independent samples) of AD dementia[4] were generated by meta-analysis, including clinically diagnosed AD cases and proxy AD and related dementia (proxy-ADD), resulting in a total of approximately 64,498 clinically diagnosed AD cases, 46,828 proxy-ADD cases, and 677,663 controls. It is noticed that the variants in the *APOE* gene region were capped by their GWAS p-values < 1e-300 in this public GWAS summary data. In this study, we were only able to consider 16 variants in the APOE region 19:44905791-44909393, including two genome-wide significant variants rs769448 (19: 44906322; GWAS p-value = 7.81E-14) and rs769452 (19: 44907853; GWAS p-value = 7.49 E-18).

Statistics and reproducibility

All statistical and computational analyses were performed using the methods and software described in the corresponding subsections above. For xWAS-O analyses, two-sided gene-based association p-values were derived from estimated xQTL weights and summary-level GWAS data, and omnibus p-values were obtained using ACAT to combine p-values across TIGAR/DPR, PrediXcan/EN, and FUSION/BestModel for each test gene. For pathway enrichment analyses, FDR q-values were calculated using the Benjamini–Hochberg method from PathDIP website tool.

Sample sizes were determined by data availability. Cell-type-aware TWAS-O analyses used ROS/MAP DLPFC snRNA-seq data, with an effective sample size of 415 biologically independent individuals for model training. Bulk PWAS-O analyses used proteomics data from 716 biologically independent individuals. Each sample represents a biologically independent individual in this study. AD dementia GWAS summary statistics were obtained from Bellenguez et al.[4]. Reproducibility was supported by publicly available analytical tools and pipelines, GWAS summary data and xQTL weights, controlled-access ROS/MAP omics datasets, and the availability of our analysis scripts on Github.

Ethics declarations

ROS/MAP studies received approval from an Institutional Review Board of Rush University Medical Center, with all participants signing informed consent forms, Anatomical Gift Act agreements, and repository consents. All individual-level snRNAseq and proteomics data analyzed in this work were de-identified.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

ROS/MAP resources by Bennett et al. 2018 can be requested at www.radc.rush.edu and www.synapse.org. All omics data of ROS/MAP samples are available from Synapse (https://www.synapse.org/Synapse:syn3219045), with approved access. GWAS summary data by the Bellenguez et al.[4] are available from European Bioinformatics Institute GWAS Catalog (https://www.ebi.ac.uk/gwas/) under accession no. GCST90027158. Trained cell-type-specific cis-eQTL weights, cis-pQTL weights, and TWAS/PWAS summary data are publicly available through SYNAPSE (syn74009946; https://doi.org/10.7303/syn74009946).

Code availability

Analysis scripts used in this work are available at https://github.com/Leo-LiuQiang/CTA-TWAS, and archived in Zenodo at https://doi.org/10.5281/zenodo.18994273.

Received: 9 August 2025; Accepted: 30 March 2026;
Published online: 22 April 2026

References

1. Scheltens, P. et al. Alzheimer's disease. *Lancet* **397**, 1577–1590 (2021).
2. Wainberg, M. et al. Opportunities and challenges for transcriptome-wide association studies. *Nat. Genet.* **51**, 592–599 (2019).
3. Brandes, N., Linial, N. & Linial, M. PWAS: proteome-wide association study—linking genes and phenotypes by functional variation in proteins. *Genome Biol.* **21**, 173 (2020).
4. Bellenguez, C. et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat. Genet.* **54**, 412–436 (2022).
5. Nagpal, S. et al. TIGAR: an improved bayesian tool for transcriptomic data imputation enhances gene mapping of complex traits. *Am. J. Hum. Genet.* **105**, 258–266 (2019).
6. Parrish, R. L., Gibson, G. C., Epstein, M. P. & Yang, J. TIGAR-V2: efficient TWAS tool with nonparametric Bayesian eQTL weights of 49 tissue types from GTEx V8. *Hum. Genet. Genomics Adv.* **3**, 100068 (2022).
7. Gamazon, E. R. et al. A gene-based association method for mapping traits using reference transcriptome data. *Nat. Genet.* **47**, 1091–1098 (2015).
8. Gusev, A. et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
9. Sun, Y. et al. A transcriptome-wide association study of Alzheimer's disease using prediction models of relevant tissues identifies novel candidate susceptibility genes. *Genome Med.* **13**, 141 (2021).
10. Liu, N. et al. Hippocampal transcriptome-wide association study and neurobiological pathway analysis for Alzheimer's disease. *PLoS Genet.* **17**, e1009363 (2021).

11. Guo, S. & Yang, J. Bayesian genome-wide TWAS with reference transcriptomic data of brain and blood tissues identified 141 risk genes for Alzheimer's disease dementia. *Alzheimers Res. Ther.* **16**, 120 (2024).
12. Olah, M. et al. Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. *Nat. Commun.* **11**, 6129 (2020).
13. Olayinka, O. A. et al. Single nucleus RNA sequencing reveals interaction between microglia and endothelial cells in mixed Alzheimer's disease and vascular pathology. *Alzheimer's. Dement.* **20**, e089628 (2025).
14. Nott, A. et al. Brain cell type–specific enhancer–promoter interactome maps and disease-risk association. *Science* **366**, 1134–1139 (2019).
15. Kosoy, R. et al. Genetics of the human microglia regulome refines Alzheimer's disease risk loci. *Nat. Genet.* **54**, 1145–1154 (2022).
16. Chen, J. et al. A scalable Bayesian functional GWAS method accounting for multivariate quantitative functional annotations with applications for studying Alzheimer disease. *HGG Adv.* **3**, 100143 (2022).
17. Jin, W. et al. Comprehensive review on single-cell RNA sequencing: a new frontier in Alzheimer's disease research. *Ageing Res. Rev.* **100**, 102454 (2024).
18. Hu, T. et al. Omnibus proteome-wide association study identifies 43 risk genes for Alzheimer disease dementia. *Am. J. Hum. Genet.* **111**, 1848–1863 (2024).
19. Fujita, M. et al. Cell subtype-specific effects of genetic variation in the Alzheimer's disease brain. *Nat. Genet.* **56**, 605–614 (2024).
20. Hu, T. et al. Proteome-wide association studies using summary pQTL data of brain, CSF, and plasma identify 30 risk genes of Alzheimer's disease dementia. *Alzheimer's. Res. Ther.* **17**, 135 (2025).
21. Liu, L. et al. Conditional transcriptome-wide association study for fine-mapping candidate causal genes. *Nat. Genet.* **56**, 348–356 (2024).
22. Raulin, A.-C. et al. ApoE in Alzheimer's disease: pathophysiology and therapeutic strategies. *Mol. Neurodegener.* **17**, 72 (2022).
23. Saha, O. et al. The Alzheimer's disease risk gene BIN1 regulates activity-dependent gene expression in human-induced glutamatergic neurons. *Mol. Psychiatry* **29**, 2634–2646 (2024).
24. Allen, M. et al. Association of MAPT haplotypes with Alzheimer's disease risk and MAPT brain gene expression levels. *Alzheimer's. Res. Ther.* **6**, 39 (2014).
25. Liu, Y. et al. ACAT: a fast and powerful p value combination method for rare-variant analysis in sequencing studies. *Am. J. Hum. Genet.* **104**, 410–421 (2019).
26. Devlin, B., Roeder, K. & Wasserman, L. Genomic control, a new approach to genetic-based association studies. *Theor. Popul. Biol.* **60**, 155–166 (2001).
27. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.: Ser. B* **57**, 289–300 (1995).
28. Buniello, A. et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
29. Zheng, J. et al. The CYP19A1 rs3751592 variant confers susceptibility to Alzheimer disease in the Chinese Han population. *Medicine* **95**, e4742 (2016).
30. Gockley, J. et al. Multi-tissue neocortical transcriptome-wide association study implicates 8 genes across 6 genomic loci in Alzheimer's disease. *Genome Med.* **13**, 76 (2021).
31. Karijolich, J. J. & Hampsey, M. The mediator complex. *Curr. Biol.* **22**, R1030–R1031 (2012).
32. Schiano, C., Luongo, L., Maione, S. & Napoli, C. Mediator complex in neurological disease. *Life Sci.* **329**, 121986 (2023).
33. Stanga, S. et al. Unfolded p53 in the pathogenesis of Alzheimer's disease: is HIPK2 the link? *Aging* **2**, 545–554 (2010).
34. Qiu, N. -z et al. Phosphodiesterase 8 (PDE8): distribution and cellular expression and association with Alzheimer's disease. *Neurochem. Res.* **49**, 1993–2004 (2024).
35. Wang, Y. et al. KDM6B cooperates with Tau and regulates synaptic plasticity and cognition via inducing VGLUT1/2. *Mol. Psychiatry* **27**, 5213–5226 (2022).
36. Yuan, Z. et al. Testing and controlling for horizontal pleiotropy with probabilistic Mendelian randomization in transcriptome-wide association studies. *Nat. Commun.* **11**, 3861 (2020).
37. Real, R. & Vargas, J. M. The probabilistic basis of Jaccard's index of similarity. *Syst. Biol.* **45**, 380–385 (1996).
38. Yuzaki, M. The C1q complement family of synaptic organizers: not just complementary. *Curr. Opin. Neurobiol.* **45**, 9–15 (2017).
39. Xu, X. et al. Characterization of brain resilience in Alzheimer's disease using polygenic risk scores and further improvement by integrating mitochondria-associated loci. *J. Adv. Res.* **56**, 113–124 (2024).
40. Huang, Z., Ang, W., Huang, H. & Wang, Y. VD/VDR-mediated ATG16L1 activation reduces Alzheimer's disease-like pathology and cognitive decline. *Mol. Cell. Toxicol.* **21**, 151–162 (2025).
41. Li, X. et al. Phosphorylated eukaryotic translation factor 4E is elevated in Alzheimer brain. *NeuroReport* **15**, 2237–2240 (2004).
42. Astillero-Lopez, V. et al. Proteomic analysis identifies HSP90AA1, PTK2B, and ANXA2 in the human entorhinal cortex in Alzheimer's disease: potential role in synaptic homeostasis and Aβ pathology through microglial and astroglial cells. *Brain Pathol.* **34**, e13235 (2024).
43. Molina, C. R. et al. Elucidating the role of mitochondrial Alzheimer's disease risk gene LACTB in myeloid cells. *Alzheimer's. Dement.* **20**, e086622 (2024).
44. Jordan, K. L., Koss, D. J., Outeiro, T. F. & Giorgini, F. Therapeutic targeting of Rab GTPases: relevance for Alzheimer's disease. *Biomedicines* **10**, 1141 (2022).
45. Htet, M. et al. HEXIM1 is correlated with Alzheimer's disease pathology and regulates immediate early gene dynamics in neurons. Preprint at https://www.biorxiv.org/content/biorxiv/early/2024/09/28/2024.09.27.615234.full.pdf (2024).
46. Tao, Y. et al. The predicted key molecules, functions, and pathways that bridge mild cognitive impairment (MCI) and Alzheimer's disease (AD). *Front. Neurol.* **11**, 233 (2020).
47. Vieira, N., Rito, T., Correia-Neves, M. & Sousa, N. Sorting out sorting nexins functions in the nervous system in health and disease. *Mol. Neurobiol.* **58**, 4070–4106 (2021).
48. Tsai, A. P. et al. INPP5D expression is associated with risk for Alzheimer's disease and induced by plaque-associated microglia. *Neurobiol. Dis.* **153**, 105303 (2021).
49. Peretti, D. E. et al. Association of glial fibrillary acid protein, Alzheimer's disease pathology and cognitive decline. *Brain* **147**, 4094–4104 (2024).
50. Dharshini, S. A. P., Taguchi, Y. H. & Gromiha, M. M. Investigating the energy crisis in Alzheimer disease using transcriptome study. *Sci. Rep.* **9**, 18509 (2019).
51. Szklarczyk, D. et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
52. Pastrello, C., Kotlyar, M., Abovsky, M., Lu, R. & Jurisica, I. PathDIP 5: improving coverage and making enrichment analysis more biologically meaningful. *Nucleic Acids Res.* **52**, D663–D671 (2024).
53. Liu, X. et al. Clusterin transduces Alzheimer-risk signals to amyloidogenesis. *Signal Transduct. Target. Ther.* **7**, 325 (2022).
54. Lambert, J.-C. et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).

55. Harcha, P. A. et al. Mast cell and astrocyte hemichannels and their role in Alzheimer's disease, ALS, and harmful stress conditions. *Int. J. Mol. Sci.* **22**, 1924 (2021).
56. Ma, J., Yu, J.-T. & Tan, L. MS4A cluster in Alzheimer's disease. *Mol. Neurobiol.* **51**, 1240–1248 (2015).
57. Rosner, D. et al. The Alzheimer's disease risk genes MS4A4A and MS4A6A cooperate to negatively regulate TREM2 and microglia states. *Neuron* **114**, 836–849.e11 (2026).
58. Tong, Q. & Chen, L. Associations of Alzheimer's disease neuropathologic changes with clinical presentations of Parkinson's disease. *J. Alzheimer's. Dis.* **81**, 201–207 (2021).
59. Li, X. et al. Systematic analysis and biomarker study for Alzheimer's disease. *Sci. Rep.* **8**, 17394 (2018).
60. Henriques, A. G. et al. Altered protein phosphorylation as a resource for potential AD biomarkers. *Sci. Rep.* **6**, 30319 (2016).
61. Dai, L. & Shen, Y. Insights into T-cell dysfunction in Alzheimer's disease. *Aging Cell* **20**, e13511 (2021).
62. Seto, M. et al. Multi-omic characterization of brain changes in the vascular endothelial growth factor family during aging and Alzheimer's disease. *Neurobiol. Aging* **126**, 25–33 (2023).
63. Porro, C., Cianciulli, A. & Panaro, M. A. The regulatory role of IL-10 in neurodegenerative diseases. *Biomolecules* **10**, 1017 (2020).
64. Deaton, C. A. & Johnson, G. V. W. Presenilin 1 regulates membrane homeostatic pathways that are dysregulated in Alzheimer's disease. *J. Alzheimer's. Dis.* **77**, 961–977 (2020).
65. Jun, G. et al. A novel Alzheimer disease locus located near the gene encoding tau protein. *Mol. Psychiatry* **21**, 108–117 (2016).
66. Schneeberger, S. et al. Interleukin-12 signaling drives Alzheimer's disease pathology through disrupting neuronal and oligodendrocyte homeostasis. *Nat. Aging* **5**, 622–641 (2025).
67. Haglund, A. et al. Cell state-dependent allelic effects and contextual Mendelian randomization analysis for human brain phenotypes. *Nat. Genet.* **57**, 358–368 (2025).
68. Bennett, D. A. et al. Religious orders study and rush memory and aging project. *J. Alzheimers Dis.* **64**, S161–S189 (2018).
69. Wingo, A. P. et al. Integrating human brain proteomes with genome-wide association data implicates new proteins in Alzheimer's disease pathogenesis. *Nat. Genet.* **53**, 143–146 (2021).
70. Jansen, I. E. et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019).
71. Wightman, D. P. et al. A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat. Genet.* **53**, 1276–1282 (2021).
72. Maccioni, R. et al. Signal peptide peptidase-like 2b modulates the amyloidogenic pathway and exhibits an Aβ-dependent expression in Alzheimer's disease. *Prog. Neurobiol.* **235**, 102585 (2024).
73. Shang, H., Liu, T., Gong, W. & Zhao, Y. Exploring the bidirectional relationships between alzheimer's disease and cerebral small vessel disease: Insights from Mendelian randomization. *J. Stroke Cerebrovasc. Dis.* **34**, 108259 (2025).
74. Culbert, A. A. et al. MAPK-activated protein kinase 2 deficiency in microglia inhibits pro-inflammatory mediator release and resultant neurotoxicity: relevance to neuroinflammation in a transgenic mouse model of alzheimer disease*. *J. Biol. Chem.* **281**, 23658–23667 (2006).
75. Stegle, O., Parts, L., Piipari, M., Winn, J. & Durbin, R. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat. Protoc.* **7**, 500–507 (2012).
76. Barbeira, A. N. et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, 1825 (2018).
77. Dai, Q. et al. OTTERS: a powerful TWAS framework leveraging summary-level reference data. *Nat. Commun.* **14**, 1271 (2023).
78. Sudlow, C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
79. Szklarczyk, D. et al. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **49**, D605–D612 (2021).
80. Szklarczyk, D. et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **51**, D638–D646 (2023).
81. Thomas, P. D. et al. PANTHER: making genome-scale phylogenetics accessible to all. *Protein Sci.* **31**, 8–22 (2022).
82. Wishart, D. S. et al. PathBank: a comprehensive pathway database for model organisms. *Nucleic Acids Res.* **48**, D470–D478 (2020).
83. Gillespie, M. et al. The Reactome Pathway Knowledgebase 2022. *Nucleic Acids Res.* **50**, D687–D692 (2022).
84. Cain, A. et al. Multicellular communities are perturbed in the aging human brain and Alzheimer's disease. *Nat. Neurosci.* **26**, 1267–1280 (2023).
85. Wingo, A. P. et al. Shared proteomic effects of cerebral atherosclerosis and Alzheimer's disease on the human brain. *Nat. Neurosci.* **23**, 696–700 (2020).
86. De Jager, P. L. et al. A multi-omic atlas of the human frontal cortex for aging and Alzheimer's disease research. *Sci. Data* **5**, 180142 (2018).

Acknowledgements

This work is supported by the National Institutes of Health (NIH) (R35GM138313 and R01AG089703 for Q.L. and J.Y). ROSMAP is supported by P30AG10161, P30AG72975, R01AG17917, R01AG015819, U01 AG072572, and U01 AG046152.

Author contributions

Q.L. analyzed all data and drafted the manuscript; R.L.P. and S.Tang analyzed the data; S.Tasaki, D.A.B., N.T.S., P.L. De Jager, V.M., and A.S.B. generated the ROS/MAP omics data and participated in data interpretation; J.Y. conceived and supervised the study. All authors edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-026-10030-4.

Correspondence and requests for materials should be addressed to Jingjing Yang.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Qiao Fan & Rosie Bunton-Stasyshyn. A peer review file is available.