Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

RESEARCH ARTICLE

Alzheimer's disease CSF biomarkers correlate with early pathology and alterations in neuronal and glial gene expression

Ali S. Ropri[1,2] | Tiffany G. Lam[1,2] | Vrinda Kalia[3] | Heather M. Buchanan[1,2] | Anne Marie W. Bartosch[1,2] | Elliot H. H. Youth[1,2] | Harrison Xiao[1,2] | Sophie K. Ross[1,2] | Anu Jain[4] | Jayanta K. Chakrabarty[4] | Min Suk Kang[2,5] | Deborah Boyett[6] | Eleonora F. Spinazzi[6] | Gail Iodice[7] | Robert A. McGovern[8] | Lawrence S. Honig[2,5] | Lewis M. Brown[4] | Gary W. Miller[3] | Guy M. McKhann[6] | Andrew F. Teich[1,2,5]

[1]Department of Pathology and Cell Biology, Columbia University Irving Medical Center, New York, New York, USA

[2]Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University Irving Medical Center, New York, New York, USA

[3]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York, USA

[4]Quantitative Proteomics and Metabolomics Center, Department of Biological Sciences, Columbia University, New York, New York, USA

[5]Department of Neurology, Columbia University Irving Medical Center, New York, New York, USA

[6]Department of Neurosurgery, Columbia University Irving Medical Center, New York, New York, USA

[7]Ankyra Therapeutics, Cambridge, Massachusetts, USA

[8]Department of Neurosurgery, University of Minnesota, Minneapolis, Minnesota, USA

Correspondence
Guy McKhann, Department of Neurosurgery, Columbia University Irving Medical Center, 710 West 168th Street, Room 411, New York, NY 10032, USA.
Email: gm317@cumc.columbia.edu

Andrew Teich, Department of Pathology and Cell Biology, Columbia University Irving Medical Center, 630 West 168th Street, PH 15-124, New York, NY 10032, USA.
Email: aft25@cumc.columbia.edu

Ali S. Ropria and Tiffany G. Lam share joint first authorship.

Abstract

INTRODUCTION: Normal pressure hydrocephalus (NPH) patients undergoing cortical shunting frequently show early Alzheimer's disease (AD) pathology on cortical biopsy, which is predictive of progression to clinical AD. The objective of this study was to use samples from this cohort to identify cerebrospinal fluid (CSF) biomarkers for AD-related central nervous system (CNS) pathophysiologic changes using tissue and fluids with early pathology, free of *post mortem* artifact.

METHODS: We analyzed Simoa, proteomic, and metabolomic CSF data from 81 patients with previously documented pathologic and transcriptomic changes.

RESULTS: AD pathology on biopsy correlates with CSF β-amyloid-42/40, neurofilament light chain (NfL), and phospho-tau-181(p-tau181)/β-amyloid-42, while several gene expression modules correlate with NfL. Proteomic analysis highlights seven core proteins that correlate with pathology and gene expression changes on biopsy, and metabolomic analysis of CSF identifies disease-relevant groups that correlate with biopsy data.

DISCUSSION: As additional biomarkers are added to AD diagnostic panels, our work provides insight into the CNS pathophysiology these markers are tracking.

Highlights

- AD CSF biomarkers correlate with CNS pathology and transcriptomic changes.
- Seven proteins correlate with CNS pathology and gene expression changes.
- Inflammatory and neuronal gene expression changes correlate with YKL-40 and NPTXR, respectively.
- CSF metabolomic analysis identifies pathways that correlate with biopsy data.
- Fatty acid metabolic pathways correlate with β-amyloid pathology.

1 | BACKGROUND

Chronic hydrocephalus in the elderly may occur for a variety of reasons, although in the absence of a clear etiology most of these cases are classified as idiopathic normal pressure hydrocephalus (NPH).[1,2] Placement of a ventricular cerebrospinal fluid (CSF) shunt can provide symptomatic relief for many elderly patients with NPH.[3–5] At some centers, a cortical brain biopsy is sent to pathology, obtained at the shunt insertion point. Early Alzheimer's disease (AD) pathology has been reported in a percentage of these biopsies, with β-amyloid plaque pathology ranging from 42% to 67%, and tau pathology relatively sparse in NPH cortical biopsies,[6] although some studies have found trace tau pathology at higher levels.[7] Perhaps not surprisingly, AD pathology on biopsy predicts future development of clinical AD, which suggests that at least a subset of the NPH population is in the early stages of AD.[8,9] For these reasons, AD pathology in NPH patients has been studied by several groups as a way to understand early AD pathophysiology.[5,6,8–11] An additional advantage is that biopsy tissue is free of *post mortem* changes, and since these are living subjects, longitudinal follow-up after biopsy is possible.

Our group has recently analyzed RNA-seq data from a cohort of 106 NPH patients with varying degrees of accompanying AD pathology.[10] This analysis identified a limited set of genes that correlated with quantified histologic measurements of β-amyloid and tau pathology, with a significant enrichment of immune response genes. Weighted Correlation Network Analysis (WGCNA) identified 4 out of 58 modules that correlated with AD pathology. Two of the modules were enriched for microglial genes, and we found that these two modules partially captured the reported downregulation of homeostatic genes and upregulation of disease-associated microglial (DAM) genes reported in the mouse AD model literature.[12] We also identified an astrocytic module and a neuronal module that correlated with quantified pathology, suggesting that transcriptomic changes accompanying early AD pathology encompass a multi-cellular response with a prominent immune response component. We subsequently validated that these modules correlate with AD pathology in autopsy brain transcriptomic datasets, although our microglial modules capture the downregulation of homeostatic genes reported in the mouse AD model literature bet-

ter than several autopsy datasets.[10] Our interpretation of these results is that we are capturing an early response to AD pathology in our data, and that this may also be partially captured in the mouse AD model literature. Interestingly, recent snRNA-seq data from human tissue with early pathologic changes have identified a microglial response that is similar to what we reported,[11,13] which further validates the relevance of our data for understanding the early stages of AD.

In summary, our prior work has produced a useful dataset for understanding early transcriptomic changes in AD, and this has motivated us to identify biomarkers that can track these changes. Our study allows us to ask this question in a unique way, as we are able to obtain CSF at the same timepoint as brain tissue, free of *post mortem* artifact. In the present study, we perform proteomics, mass spectrometry-based metabolomics, and Simoa AD biomarker measurements on ventricular CSF that was collected at the same time as the biopsy tissue on 81 NPH patients from our prior study, and analyze these data to identify biomarkers of disease that correlate with biopsy pathology and gene expression changes. We identify a number of previously unreported patterns between brain and CSF, which addresses the challenge of identifying biomarkers for established pathophysiologic central nervous system (CNS) changes that occur in the setting of early AD pathology.

2 | METHODS

This study is a retrospective study that uses tissue and CSF samples not required for clinical diagnosis and associated clinical and demographic data. This study was reviewed and approved by the Columbia University Institutional Review Board (IRB), and all relevant ethical regulations have been followed. Processing and analysis of brain tissue, including RNA-seq and histologic analysis of β-amyloid and tau, has been previously reported,[10] and is briefly summarized below. Biopsies were taken from frontal cortex in two-thirds of the subjects and parietal cortex in one-third of the subjects in our original cohort, with the location for a given patient chosen for cosmetic reasons. This ratio in the subgroup of 81 subjects analyzed in this study is similar. As noted in ref. [10], changes in gene expression that correlate with AD

RESEARCH IN CONTEXT

1. **Systematic review**: The authors reviewed the literature using Pubmed and conference abstracts. While Alzheimer's disease (AD) cerebrospinal fluid (CSF) biomarkers have been extensively studied, no study comprehensively relates CSF biomarkers to central nervous system (CNS) transcriptomic changes and pathology in subjects with early AD pathology using specimens free of *post mortem* artifact.

2. **Interpretation**: AD CNS pathology and associated transcriptomic changes correlate with CSF AD Simoa markers, seven CSF proteins that are supported by prior literature for AD relevance, and several AD-relevant metabolic groups. Notably, we demonstrate specific inflammatory and neuronal gene expression changes that correlate with YKL-40 and NPTXR, respectively.

3. **Future directions**: Future directions include replicating these findings, which is important given the unique nature of this study. In addition, our work will help inform the use of AD biomarkers for tracking pathophysiologic changes in human brain, and also inform animal models that study these relationships (such as the interplay between microglia, astrocytes, and YKL-40).

pathology trend similarly in both regions, and we combined all subjects together[10] for purposes of analysis. The average age of our original 106 subjects is 74.9 years, and the subgroup of 81 subjects analyzed in this study (that have both biopsy and ventricular CSF collected) is 74.5 years (see Table S1 for patient demographics). CSF analyzed in this study was obtained through the shunt catheter into polypropylene tubes, and promptly frozen and stored at −80°C.

2.1 | Summary of previously reported data analyzed in this study

As reported in ref. [10], RNA was extracted from biopsy samples using miRNeasy Mini Kit (QIAGEN; Cat No./ID: 217004), and samples with RIN values ≥ 6 were selected for sequencing. RNAs were prepared for sequencing using the Illumina TruSeq mRNA library prep kit, and samples underwent single-end sequencing to 30 M read depth. The quality of all fastq files was confirmed with FastQC v 0.11.8,[14] followed by variance stabilizing transformation (VST),[15] and surrogate variable analysis (SVA)[16] and *removeBatchEffect*[17] were sequentially used to remove confounding variables in our dataset. To generate gene expression modules, we utilized Weighted Gene Co-expression Network Analysis (WGCNA) to identify gene co-expression modules.[18] As highlighted in ref. [10], four gene expression modules correlated

with β-amyloid and tau pathology on biopsy. In ref. [10], we use the color-coded names assigned by WGCNA to label these modules (saddlebrown, orange, darkgrey, and mediumpurple3). In ref. [10], we note that these four modules are enriched for cell-type specific genes, with saddlebrown enriched for microglial disease-associated (DAM) genes, mediumpurple3 enriched for microglial homeostatic genes, orange enriched for astrocytic genes, and darkgrey enriched for neuronal genes. In this paper, we use the cell-type specific labels when describing these modules for easier readability. Note that cell-type specific genes are from snRNA-seq data and are defined as enriched in a specific cell type compared to other cell types, which we have previously used to characterize these modules.[10]

Immunohistochemistry for tau (AT8; Thermo Fisher; Catalog # MN1020), and β-amyloid (6E10; BioLegend; Catalog # 803003) was performed using the Ventana automated slide stainer. β-amyloid plaques were counted per square millimeters; in slides with enough tissue, three fields were averaged together, whereas in slides with less tissue, the largest number of possible fields were counted. For tau quantification, we devised a rating scale to grade the minimal degree of tau pathology seen in NPH biopsies (Figure S1). Grade 0 was given to biopsies with no tau pathology. Grade 1 was given to biopsies that have any tau pathology at all, usually one or more dystrophic neurites, but do not make criteria for Grade 2. Grade 2 was given to biopsies that have at least one tau-positive neuron or neuritic plaque, but do not make criteria for Grade 3. Grade 3 was reserved for biopsies with tau pathology evenly distributed throughout the biopsy.

2.2 | Enzyme-linked immunosorbent assay (ELISA) analysis

Simoa technology (Quanterix, Inc., Billerica, MA) on the SR-X platform was used to measure Aβ–40, Aβ–42, and total-tau with the multiplex Neurology 3-plex A kit #101995, NfL with the NF-light Advantage kt (SR-X) kit #103400, and p-tau181 with the p-tau181 Advantage V2 kit #103714. All assays were performed in duplicate for each sample, using eight calibrators and two positive controls (low and high concentrations) in 96-well plates. CSF was rapidly thawed, gently vortexed, centrifuged, and, diluted as per kit specifications depending on assay, and added to kit beads (100 μL) by pipette in each well. Then in succession, plates are incubated for 15 min at 30°C, shaking at 1000 rpm, magnetic-washing 3× for 5 min total, followed by addition of SBG reagent (100 μ L), another incubation for 10 min at 30°C at 1000 rpm, washing again 5× for 7 min total and then reading on the Simoa SR-X machine. Each plate assays in duplicate 34 samples. This highly sensitive assay system has lower limits of quantitation of about 1, 0.1, and 0.1 pg/mL for Aβ–40, Aβ–42, and tau, 0.1 pg/mL for -p-tau181, and 0.3 pg/mL for NfL; coefficients of variation within duplicates are between 3% and 10%. Simoa measurements (Aβ–42, Aβ–40, Aβ–42/Aβ–40 ratio,[19] total tau, p-tau181, p-tau181/Aβ–42 ratio) were regressed for age and sex with limma::removeBatchEffect (version 3.54.2).

2.3 | Proteomics analysis

Proteins from CSF were studied by quantitative measurement of protein abundance with a mass spectrometry-based proteomic method. Prior to proteome analysis, depletion of high abundance proteins was performed with High-Select Top14 Abundant Protein Depletion Resin columns (Pierce/Thermo Fisher Scientific). CSF proteins were resuspended in 8 M urea, 3 mM dithiothreitol (DTT), 100 mM ammonium bicarbonate in liquid chromatography/mass spectrometry grade water, reduced with DDT, and alkylated with iodoacetamide. For proteolytic digestion, samples were diluted five-fold in 100 mM ammonium bicarbonate and then digested using sequencing grade trypsin (Promega V511) at a protease/protein ratio of 1:50 at 37°C for 16 h as described previously.[20] Samples were then desalted with Nest Group C18 Macrospin columns (Southborough, MA). Peptide concentration was evaluated by NanoDrop spectrophotometry (Thermo Fisher Scientific) at 205 nm and LC/MS inject loading amounts were adjusted (normalized) based on peptide concentration. Peptides were separated with an acetonitrile/formic acid gradient at 300 nL/min on a 75 μm ID × 50 cm Acclaim PepMap C18 2 μm particle size column with an UltiMate 3000 RSLCnano liquid chromatograph. This was coupled to a Q-Exactive HF mass spectrometer (Thermo Scientific). Data were acquired in data dependent acquisition mode (DDA) and proteins were identified by database search using PEAKS Studio (version 10.6; Bioinformatics Solutions Inc.) and a Human UniProt reviewed database with isoforms (UniProt release 2020_04, August 11, 2020). All raw mass spectrometry files produced in this work are publicly available at the MassIVE proteomics repository (https://massive.ucsd.edu).

Protein abundance was measured by label-free quantitation with PEAKS Studio. Peaks software detected 701,261 features across 81 LC/MS/MS runs. Identifications were returned for 1021 proteins with a 1% false discovery rate (FDR) by the Peaks program. Of the 1021 proteins, 14 proteins subjected to partial antibody depletion, 394 proteins were represented by a single peptide and 6 added proteins and lab contaminants were deleted from the analysis. A total of 607 proteins and 528 unique proteins represented by two or more peptides were included in the analysis. We recorded a total of 49,167 quantitative protein determinations with only 174 missing values (0.4% missing values = 99.6% data completeness). Thus the number of missing values was so negligible to render statistical "imputation" unnecessary. Batch correction was achieved using a tunable approach for median polish of ratio (TAMPOR) algorithm[21] for removing technical variance, and protein abundance values were normalized within each batch with no Global Internal Standard (GIS) (1). Effects attributable to age and sex were regressed with limma::removeBatchEffect (version 3.54.2). Log2 normalized protein abundances of 528 CSF proteins from 81 samples were Spearman-correlated to histologic metrics (β-amyloid load, tau load, and glial fibrillary acidic protein (GFAP) staining) and brain transcriptomic eigengenes for WGCNA-inferred modules using Hmisc::rcorr (version 4.4.2).

The major strength of this study is the paired biopsy data, and as such we did not approach our data from the standpoint of attempting to identify novel biomarkers. Instead, our goal was to identify proteins that have been well validated in prior work and examine how these markers correlate with our biopsy data. Note that we are not powered to survive multiple hypothesis correction with our dataset. Power analysis (using pwr.r.test function in r) suggests that with 81 samples, we are powered at 0.8 to detect $r = 0.3$ at $p = 0.05$, which is similar to the nominally significant r values seen in our proteomics data (Table S2). The adjusted p-values using FDR depend on the overall profile of unadjusted p-values, and in that sense are context-specific.[22] Thus, it is difficult to know a priori what unadjusted p-value will correspond to an FDR of 0.05 in a given study, which would be necessary to identify the FDR adjusted p-value we are powered for. However, given that we are near the threshold of being sufficiently powered for discovery with unadjusted p-values, it is a reasonable assumption that multiple hypothesis testing will eliminate many if not all of the significant findings.

Given the above, we started with a recently published analysis of AD CSF that also included data from four additional published validation datasets (five cohorts total[21,23,24]). We focused on proteins in our own data that significantly increased or decreased in AD in at least one of the cohorts from this study with a corrected p-value (FDR) of 0.05, and at least nominally (unadjusted p-value less than 0.05) trended in a similar direction in at least one other cohort. This resulted in 45 proteins. We then tested whether any of these 45 proteins were nominally significantly correlated with either the histologic measurements of β-amyloid or tau on biopsy or any of our four gene expression modules. For module correlations, we only considered CSF proteins that were significantly correlated with a module eigengene if the same gene also correlated significantly and in a similar direction with the module eigengene in the brain RNA-seq data, as this would suggest that the CSF protein correlation is reflective of underlying CNS biology. Using these thresholds, we focused on seven proteins from the proteomics data for further analysis. Note that all seven CSF proteins that passed our criteria are correlating with our biopsy data in a similar direction to the reported correlations in the AD literature (i.e., if a CSF protein is reported to increase in AD, in our data, it positively correlates with either AD pathology or a module that increases in AD, and vice versa). We did not apply an explicit threshold that this be the case in our filtering, and this supports the idea that our filtering has focused on CSF biomarkers of disease relevant for AD. We also tested whether proteins that are nominally insignificant (unadjusted p-value > 0.05) in all of the above five AD cohorts (115 proteins total) are nominally significant with any of the NPH biopsy variables; see the Results section for discussion of this analysis. Similar analysis among the above five AD cohorts to identify proteins nominally significant in one cohort and nominally insignificant in the other four cohorts revealed the following (cohorts labeled using names assigned in ref. [24]): four significant proteins in the CSF discovery cohort, two proteins in replication cohort 1, one protein in replication cohort 2, two proteins in replication cohort 3, and one protein in replication cohort 4.

In an effort to further reinforce our findings, we identified commercially available ELISAs that have been used extensively in the literature for two proteins of interest (YKL-40 and NPTXR).[25–30] We performed ELISA measurements for these two proteins on CSF aliquots from all 81

subjects for ELISA validation, according to the manufacturers' instructions. Levels of NPTXR (#ELH-NPTXR; Ray Biotech, GA, USA), and YKL-40 (# DC3L10; R&D systems, MN, USA), were assessed using a 1:10 and 1:100 dilution, respectively.

2.4 | Metabolomics analysis

The details on the acquisition of the untargeted metabolomics has been previously described.[31,32] Briefly, the metabolites were extracted from plasma using acetonitrile and the extracts were injected in triplicate on two chromatographic columns: a hydrophilic interaction column (HILIC) under positive ionization (HILIC+) and a C18 column under negative ionization (C18-), to obtain three technical replicates per sample per column. Data were collected in full scan mode for molecules within 85-1250 kDa on a Thermo Orbitrap HFX Q-Exactive mass spectrometer. The untargeted metabolomic data were processed through a computational pipeline that leverages open source feature detection and peak alignment software, *apLCMS*[33] and *xMSanalyzer*.[34] Correction for batch effects was performed using ComBat, which uses an empirical Bayesian framework to adjust for known batches in which the samples were run.[35] Metabolic features detected in at least 70% of the samples were retained, leaving 3638 features from the HILIC+ column and 4532 features from the C18- column for further analysis. Zero-intensity values were considered below the detection limit of the instrument and were imputed with half the minimum intensity observed for each metabolic feature. The intensity of each metabolic feature was log10 transformed, quantile normalized, and auto-scaled for normalization and standardization. Normalized feature values for HILIC+ and C18- columns were regressed for sex and age and Spearman correlated with each variable of interest (histological β-amyloid, tau, and GFAP, and the four modules). Correlation values of metabolic features with each variable of interest from each column were combined and fed into mummichog (MetaboAnalyst version 5.0)[36] to highlight significantly enriched metabolomic pathways (Gamma FDR adjusted $p < 0.05$). For the mummichog algorithm parameters, only pathways containing at least five metabolites were considered.

2.5 | Statistics

All statistical analyses were performed in R (version 4.3.0), except those conducted with ELISA data which were performed using Graph-Pad Prism software (version 9.4.1). All correlations were performed using Spearman's rank correlation coefficient, and *p*-values are two-tailed (null hypothesis $r = 0$). Correlations for metabolomics and proteomics were assessed using the rcorr function as implemented in the Hmisc package (version 4.4.2) in R. FDR correction was used to adjust *p*-values for multiple comparisons where indicated. Hub genes were identified using the intramodularConnectivity function as implemented in WGCNA (version 1.72.1).[18] The TAMPOR algorithm used for batch correction for the proteomics data can be downloaded from https://github.com/edammer/TAMPOR. Regressions for the pro-

teomics and simoa data were performed using the function removeBatchEffect (version 3.54.2) as implemented in the package limma in R. Regressions for the metabolomics data were performed using the function ComBat as implemented in the package sva in R. Pathway enrichment analysis for the metabolomics data was conducted by the Mummichog algorithm in MetaboAnalyst (version 5.0). Of the 81 CSF samples, all analyses were performed successfully on the full set of 81 with the following exceptions; 80 have CSF Aβ−40 values, 78 have CSF Aβ−42 values, 78 have CSF Aβ42/40 values, 80 have CSF ptau 181 values, 77 have CSF ptau 181/Aβ42 values, 80 have CSF tau values, and 80 have immunohistochemical GFAP values.

2.6 | Immunohistochemistry

Immunohistochemistry for GFAP was performed on sections of formalin fixed, paraffin embedded tissue using the rabbit monoclonal EP672Y antibody run on the Ventana Ultra platform. Brightfield microscopy was used to capture images of the GFAP stained slides at 4× magnification. Images were processed with CellProfiler v4.2.5.[37] Tissue region within each image was distinguished from slide background through global minimum cross entropy thresholding, excluding objects below 50 pixels and any object that did not overlap with a manually drawn outline of the tissue region. Regions of histological artefact to exclude from tissue regions such as folded sections were identified through the global robust background thresholding for pixels over two standard deviations above the mean after discarding 85% of the bottom intensity pixels. Average GFAP intensity was then measured over the entire thresholded tissue area for each case. We were able to perform GFAP staining on 80 of the 81 specimens.

3 | RESULTS

3.1 | Study overview and review of cohort biopsy data

This study involves comparisons across a range of datatypes, summarized in Figure 1A. Patients undergoing shunt surgery for NPH had biopsy tissue removed at the shunt insertion point, and in a subset of these patients ventricular CSF was banked at the same time (see the Methods section). Comparison of RNA-seq analysis and quantified β-amyloid and tau pathology from biopsy tissue has previously been reported on 106 patients.[10] This prior analysis identified four key gene expression modules that correlate with AD pathology. These modules consist of two microglial modules, one of which is enriched for DAM genes and positively correlates with AD pathology on biopsy, and one of which is enriched for microglial homeostatic genes and negatively correlates with AD pathology on biopsy. We also identified an astrocytic module that positively correlates with AD pathology on biopsy, and a neuronal module that negatively correlates with AD pathology on biopsy (see ontology analysis for these modules reproduced from ref. [10] in Table S3, and Table S4 for reproduced module genes). In

FIGURE 1 Study overview and review of cohort biopsy data. (A) Schematic for the NPH study in this paper (see text for details). (B) Our four modules correlate with quantified β-amyloid and tau pathology on the 81 biopsies with CSF similarly to the correlations reported in ref. [10]. For this study, we also added quantified GFAP staining, and correlations with the four modules are shown (* = FDR adjusted p-value < 0.05, see Table S5 for numbers used in this figure). (C) Schematic for our filtering of proteins for proteomic analysis. We selected proteins that passed an FDR of 0.05 in at least one previously published study and trended in the same direction (i.e., up or down in AD) with an unadjusted *p*-value of 0.05 in at least one other study, and which also correlated with one of our pathology variables or gene expression modules with an unadjusted *p*-value of 0.05. (See the Methods section for all details of our filtering steps, cohorts labeled using names assigned in ref. [24]). AD, Alzheimer's disease; NPH, normal pressure hydrocephalus.
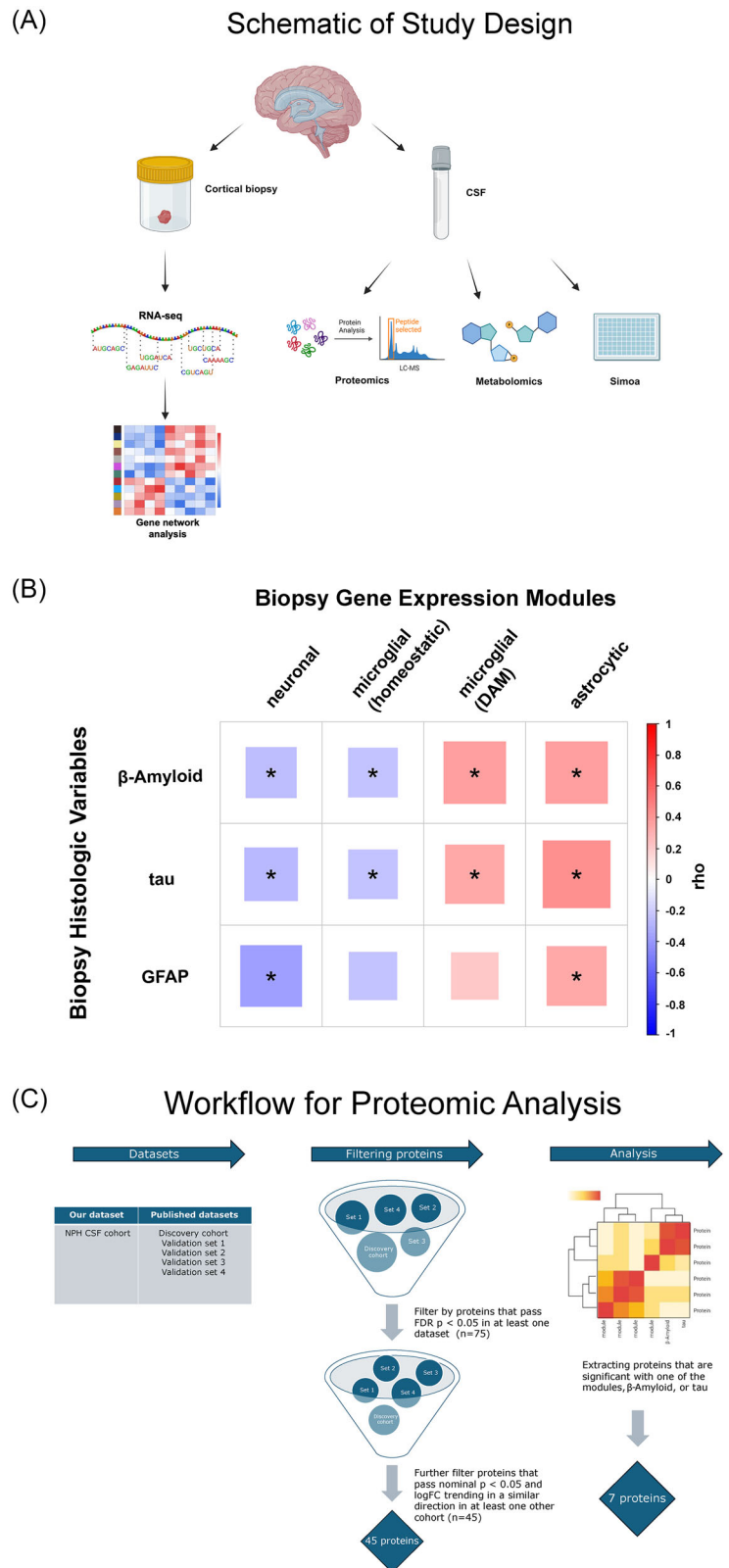

(A) Schematic of Study Design

(B) Biopsy Gene Expression Modules

(C) Workflow for Proteomic Analysis

Figure 1B, we show that all four modules originally identified in ref. [10], significantly correlate with previously quantified β-amyloid and tau pathology in the subset of 81 subjects with banked CSF used in this study, similarly to the original group of 106 (see Table S5 for values used in this figure; note that in ref. [10] we use the WGCNA assigned color names for these modules – here we use names that reflect the predominant cell type genes in these modules for simplicity – see the Methods section for details on prior color code assignments from ref. [10]). In this study, we have added quantified GFAP staining as a measure of astrogliosis in order to assess how this metric might also relate to our

CSF analysis. Here, we note that GFAP staining correlates not only with the astrocytic module, but also with the neuronal module, suggesting that neuronal function may decline in tandem with astrogliosis.

Eighty-one of the 106 subjects from the previous study had banked CSF that was available for the present study. We analyzed this CSF using three methods: (1) Simoa measurement of CSF for β-amyloid-42 and −40, total tau, phospho-tau-181 (p-tau181), and neurofilament light chain (NfL); (2) proteomic analysis of CSF; and (3) metabolomic analysis of CSF. Note that we are not powered to perform multiple testing correction in our proteomic mass spectrometry data (see the Methods section), and the major strength of this study is the paired biopsy data. With this in mind, we did not approach our proteomic data from the standpoint of attempting to identify novel biomarkers, and instead used our data to study proteins that have been well validated in prior work and examine how these proteins correlate with our biopsy data. Specifically, we selected proteins that passed an FDR of 0.05 in at least one study and trended in the same direction (i.e., up or down in AD) with an unadjusted p-value of 0.05 in at least one other study, and which also correlated with one of our pathology variables or gene expression modules with an unadjusted p-value of 0.05. (Figure 1C; see the Methods section for all details of our filtering steps). We focused on seven proteins from the proteomics data that passed these thresholds for further analysis, which are highlighted below. Our approach parallels other groups that have reported unadjuted p-values and used alternate rationales for validation (see the Discussion section) and we also validate two key proteins with ELISAs. Nevertheless, this lack of statistical power is a weakness of our study, which we further discuss in detail in the Discussion section.

3.2 | Histologic measurements of AD pathology correlate with CSF biomarkers

We first determined how histologic measurements of β-amyloid, tau, and GFAP on biopsy relate to CSF data (Figure 2A). Quantified β-amyloid plaques and tau pathology on biopsy negatively correlate with Simoa measurements of the β-amyloid-42/40 ratio, and tau pathology also correlates significantly with NfL. Both β-amyloid plaques and tau pathology trend positively with p-tau181 but are not significant, and p-tau181/β-amyloid-42 positively correlates with both. GFAP negatively correlates with both β-amyloid-40 and β-amyloid-42, and is not significantly associated with the β-amyloid-42/40 ratio, suggesting that this metric is less disease-specific than histologic measurements of β-amyloid and tau pathology for AD. Of note, GFAP staining on biopsy does correlate with quantified β-amyloid plaque pathology on biopsy (Spearman's $r = 0.22$, uncorrected p-value = 0.047) and tau pathology on biopsy (Spearman's $r = 0.32$, uncorrected p-value = 0.0044), suggesting that astrogliosis in these biopsies is at least partially reflective of AD pathology, particularly tau pathology. Along the same lines, the concordance between quantified β-amyloid plaques and tau pathology and AD CSF Simoa markers supports the view that the CSF of these patients partially reflects the ongoing AD-related disease captured on biopsy. Also encouragingly, this suggests that analysis of a very small

piece of cortical tissue can give information that is partially predictive of CSF analysis. This should not immediately be assumed, as the CSF presumably captures changes found throughout the neuraxis, and may not correspond to the highly local analysis seen in a small piece of tissue. The logical reciprocal argument is that the findings in our biopsies are in fact reflecting similar biology seen in other areas of cortex.

As noted above, we focused on seven key proteins from the proteomics data that are previously validated in the AD CSF literature. Tau pathology on biopsy is the primary histologic measurement that correlates with selected proteins, with three proteins showing nominally significant trends, and GFAP shows a nominal inverse relationship with VGF (Figure 2B). Finally, we analyzed our CSF with mass spectroscopy-based metabolomics to identify any biological pathways that may correspond to early AD pathology in brain tissue using mummichog, an analysis pipeline that identifies metabolic pathways enriched in metabolome correlations with variables of interest (see the Methods section). Pathways from several metabolic processes that have previously been reported in AD are also predicted to be altered in our CSF (Figure 2C; see Table S6 for all pathways identified and the Discussion section for commentary).

3.3 | Biopsy gene expression modules correlate with CSF biomarkers

We next determined how the four gene expression modules identified in our prior work relate to CSF data. All four modules have trends with β-amyloid-42/40, p-tau181, and p-tau181/β-amyloid-42 in directions consistent with the modules' relationship with AD pathology, with modules that positively correlate with pathology having trends similar to β-amyloid plaque and tau pathology, and modules that decline with increasing pathology showing opposing trends (Figure 3A). The neuronal and microglial homeostatic module both significantly correlate with β-amyloid-42, and trend positively with β-amyloid-40, with the neuronal module achieving significance at a weaker correlation (see Table S7 for all values). Correlations with NfL pass significance with three of the modules, with the microglial (DAM) module and astrocytic modules positively correlating with NfL, and the neuronal module negatively correlating with NfL. Proteomics analysis shows a range of relationships between the four gene expression modules and our seven proteins of interest (Figure 3B). Three neuronal proteins (NPTXR, SCG2, and VGF) all positively correlate with the neuronal module, and the microglial homeostatic module (which negatively correlates with AD pathology) also positively correlates with two of the neuronal markers. YKL-40 correlates with the astrocytic module and the microglial (DAM) module, suggesting that a multicellular response accompanies YKL-40 upregulation in CSF.

Although we are guiding our proteomic analysis by proteins that are significant in more than one published dataset, one could also ask the converse question (i.e., Are proteins that are consistently not associated with AD in the literature also not significant in our data?). To address this, we returned to the five published proteomic datasets that informed our selection criteria, and identified CSF proteins that
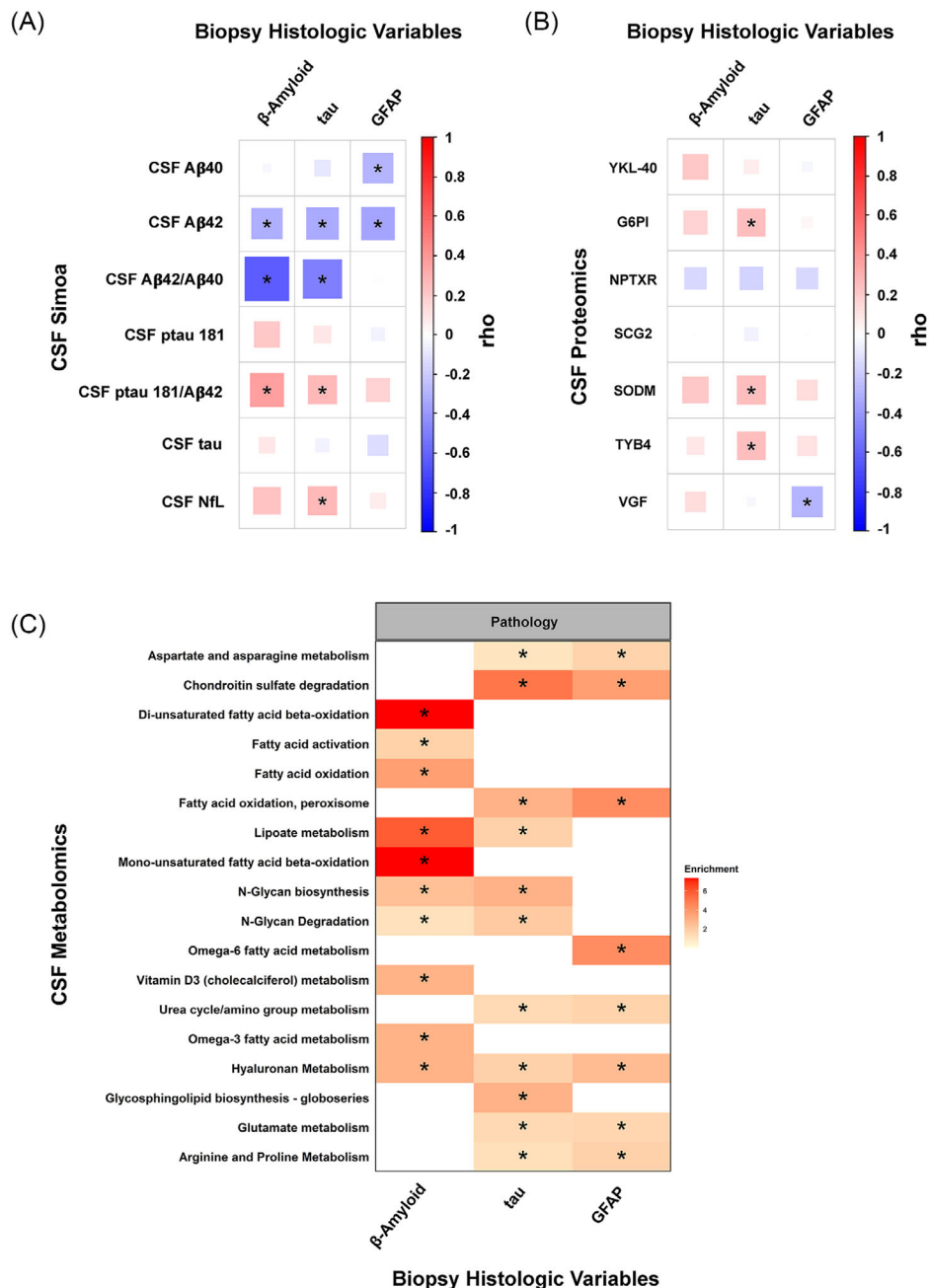
FIGURE 2 Histologic measurements of AD pathology correlate with CSF biomarkers. (A) Correlations of histologic data with CSF Simoa measurements of AD biomarkers. All correlations shown in this figure are Spearman's rank correlation coefficient. The n for each Simoa analysis is variable due to some sample failure. In summary, 80 samples have CSF Aβ40 values, 78 samples have CSF Aβ42 values, 78 samples have CSF Aβ42/40 values, 80 have CSF ptau 181 values, 77 have CSF ptau 181/Aβ42 values, 80 have CSF tau values, and all 81 have CSF NfL values. GFAP staining was also only achieved on 80 samples. All other analyses here and in the rest of the study are completed on all 81 samples. (B) Spearman's correlations of the seven core proteins highlighted in this study with quantified β-amyloid, tau, and GFAP on biopsy. (C) Biological pathways highlighted by mummichog analysis of metabolite correlations with histologic variables (see the Methods section); *FDR adjusted p-value < 0.05 in panels A and C, *p-value < 0.05 in panel B. See text for details, and Tables S2, S6, and S7 for numbers used in this figure. AD, Alzheimer's disease; CFS, Cerebrospinal fluid.

were not nominally significantly associated with AD (unadjusted p-value > 0.05) in all five datasets and that were also present in the NPH CSF proteome (115 proteins total). Of these 115 proteins, we found that at a nominal, unadjusted p-value threshold of 0.05, one protein significantly correlates with quantified tau pathology, one pro-

tein significantly correlates with quantified β-amyloid pathology, six proteins significantly correlate with GFAP staining, no proteins significantly correlate with the astrocyte module, three proteins significantly correlate with the microglial (DAM) module, 24 proteins significantly correlate with the microglial homeostatic module, and 18 proteins
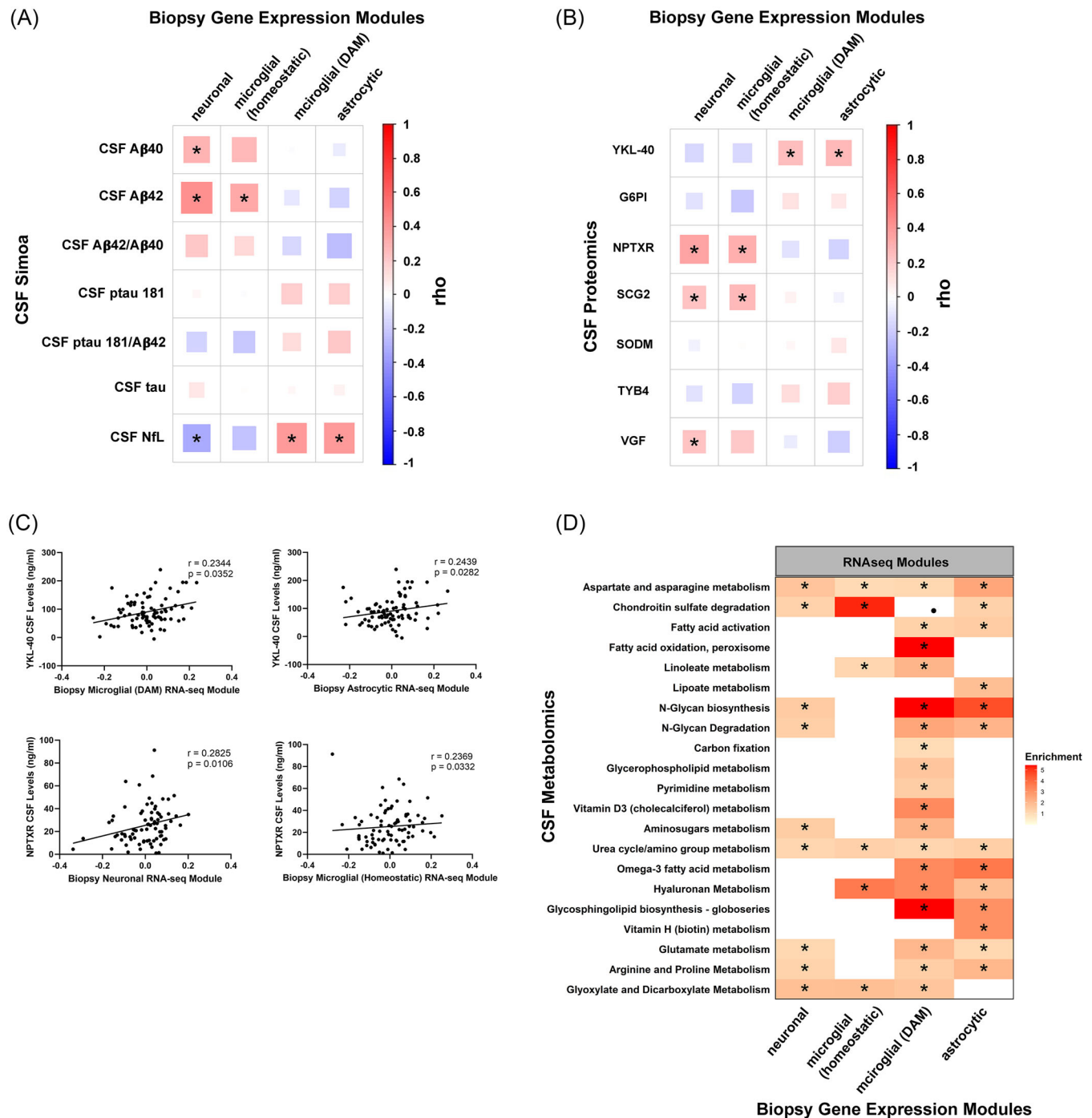
FIGURE 3 Biopsy gene expression modules correlate with CSF biomarkers. (A) Correlations of gene expression modules from Figure 1B with CSF Simoa measurements of AD biomarkers. All correlations shown in this figure are Spearman's rank correlation coefficient. (B) Spearman's correlations of the seven core proteins highlighted in this study with gene expression modules. (C) Spearman's correlation across 81 CSF samples of YKL-40 ELISA values versus microglial (DAM) and astrocytic module eigengenes and NPTXR ELISA values versus neuronal and microglial homeostatic module eigengenes. r- and p-values indicated. (D) Biological pathways highlighted by mummichog analysis of metabolite correlations with gene expression modules (see the Methods section); *FDR adjusted p-value < 0.05 in panels A and D , *p-value < 0.05 in panel B. See text for details, and Tables S2, S6, S7, and S9 for numbers used in this figure. AD, Alzheimer's disease; CFS, Cerebrospinal fluid; DAM, disease-associated microglia.

significantly correlate with the neuronal module (see Table S8). For comparison, we then did an analysis among the five published datasets themselves to see how many nominally significant proteins (unadjusted *p*-value < 0.05) in each dataset are not nominally significant in the other four datasets. All of the published datasets had low numbers of nominally significant proteins that were not significant in the other four, ranging from one to four proteins across the different datasets (see the Methods section). In our own data, the neuronal module and the microglial homeostatic module are the two biopsy signals that stand out as having a relatively high number of significant proteins (18 and 24, respectively) that are not significant in the AD CSF literature we reviewed. Note that these are the two modules that decline in AD, whereas the other two modules and all three histologic measurements (*β*-amyloid, tau, and GFAP) are low at baseline and increase in AD. One could therefore speculate that one reason why a large number of non-AD related CSF proteins correlate with the neuronal and microglial homeostatic modules is because these modules represent homeostatic functions of neurons and microglia respectively, and may be responding to a variety of signals at baseline in addition to AD-associated pathophysiology. Although we cannot definitively say why these modules associate with a large number of non-AD associated proteins, these CSF associated proteins could be a useful starting point to explore the normal CNS biology represented by these modules in a future study.

In an effort to further reinforce our findings, we identified commercially available ELISAs that have been used extensively in the literature for two proteins of interest (YKL-40 and NPTXR). We performed ELISA measurements for these two proteins on CSF aliquots from all 81 subjects in order to test the reliability of our proteomics data using a method orthogonal to mass spectrometry (in this case, sandwich ELISA). In Figure 3C, we show that ELISA measurements for these two proteins demonstrate a similar correlational profile to the proteomics data (additional correlations not shown are not significant—see Table S9). Intriguingly, YKL-40 does not appear to correlate with AD pathology on biopsy. Although the lack of a significant correlation between CSF YKL-40 and AD pathology in biopsy tissue may be a power issue, at minimum it appears that in our cohort, CSF YKL-40 correlates more strongly with gene expression changes than AD pathology or traditional measures of astrogliosis (i.e., GFAP staining). Finally, metabolomic analysis (Figure 3D) identified enrichment for many of the same pathways that are enriched in the metabolite correlations with AD histologic measurements (Figure 2C).

Our previous publication highlighted the immune response-associated transcriptomic signature in these biopsy samples,[10] and so we were particularly interested in the relationship seen in this study between CSF YKL-40 and the microglial (DAM) and astrocytic modules. In Figure 4, we show hub genes for both of these modules, with microglia and astrocytic genes highlighted in the microglial (DAM) and astrocytic modules respectively (see the Methods section for hub gene and cell type specific gene definitions and Table S10 for lists of hub genes and connectivity values). Mean gene expression vectors composed of only cell-type specific genes from these modules correlate as well with CSF YKL-40 as the module eigengenes themselves. This suggests that the cell-type specific changes reported by these modules

are correlating with CSF YKL-40, which raises the possibility that CSF YKL-40 may be a useful marker for disease-relevant interactions between these cell types (see the Discussion section).

4 | DISCUSSION

The goal of this study is to address the major challenge of identifying biomarkers for established pathophysiologic CNS changes that occur in the setting of early AD pathology. Of particular interest is our finding that CSF YKL-40 protein correlates primarily with two biopsy RNA-seq modules enriched for astrocytic and microglial genes, and not with quantified biopsy AD pathology. YKL-40 (otherwise known as CHI3L1) is a well-established marker of inflammation, and is widely studied in AD.[38,39] The interaction between microglia and astrocytes in AD is also an area of active research, and recent work suggests that this interplay may contribute to neurodegeneration.[40,41] YKL-40 is secreted by astrocytes, although its secretion is thought to be in part modulated through activated microglia,[42,43] which further supports our data showing that specific gene expression changes relating to these two cell types correlate with CSF YKL-40. Of note, our astrocytic and microglial (DAM) modules are also highly correlated ($r = 0.47$, unadjusted *p*-value = 5.07×10^{-07}). One could speculate that we may be partially capturing a disease-relevant interplay between these two cell types in our data, and YKL-40 may be a useful marker to track this process. Future work could use these findings to examine this hypothesis in AD model systems. It should also be noted that our work identifies trends in several neuronal markers of neurodegeneration. Three neuronal proteins that have been shown to decrease in AD (NPTXR, SCG2, and VGF),[38] all positively correlate with the neuronal module, which is also consistent with this module declining in tandem with AD pathology on biopsy. Interestingly, GFAP staining also (inversely) correlates with our neuronal module as well as VGF (a well-documented biomarker that declines in AD[44]), and the microglial homeostatic module declines in parallel with some of the neuronal CSF proteins. This also points to aspects of the microglial/astrocytic response that may be most proximal to early neuronal dysfunction.

Metabolomics analysis identified several processes that have previously been linked to AD that are also predicted to be altered in our CSF. Interestingly, fatty acid oxidation is predicted to be altered in tandem with *β*-amyloid pathology. Fatty acid oxidation has previously been linked to AD though several lines of investigation (reviewed in ref. [45]). Fatty acid oxidation is relatively limited in neurons, in contrast to astrocytes, which metabolize fatty acids transported from neurons in apolipoprotein E (ApoE) -positive lipid particles as a protective mechanism in the setting of lipid peroxidation.[46,47] Lipoic acid is an antioxidant and cofactor for several metabolic enzymes, and similarly to fatty acid oxidation, is linked to the generation of acetyl-CoA,[48,49] suggesting that there may be alterations in oxidative metabolism across multiple pathways as *β*-amyloid accumulates in the brain.

Additional findings in our metabolomics data are alterations in aspartate and asparagine metabolism; marked alterations in aspartate metabolism have recently been found in AD brain tissue using
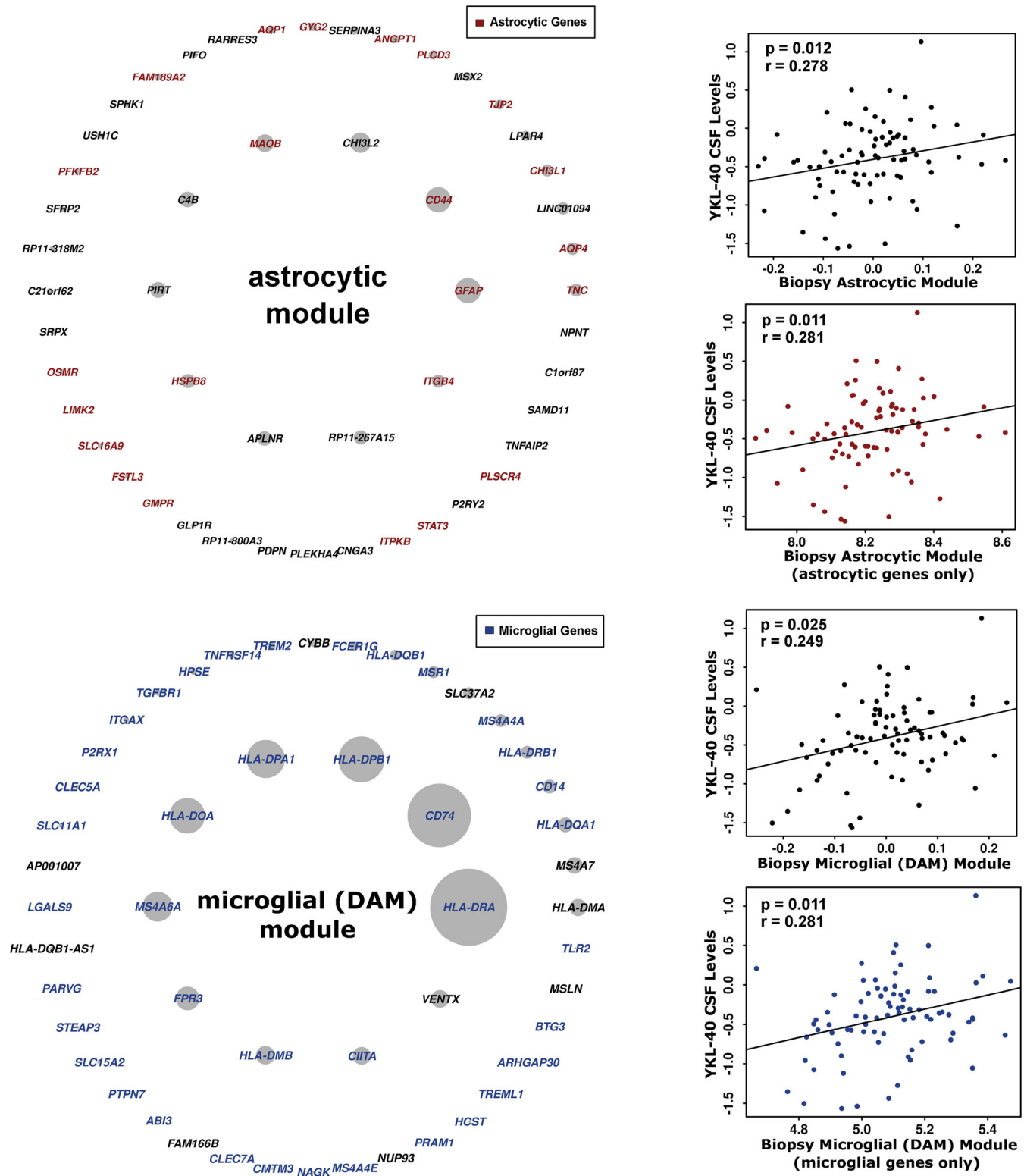
FIGURE 4 CSF YKL-40 correlates with microglial and astrocytic genes. Shown are the hub genes for the astrocytic and microglial (DAM) modules, with astrocytic genes highlighted for the astrocytic module and microglial genes highlighted for the microglial (DAM) module. Both modules correlate with CSF YKL-40. The mean gene expression vector of the astrocytic genes from the astrocytic module and microglial genes from the microglial (DAM) module also correlate with YKL-40, supporting a role for these genes in the relationship between brain pathophysiology and CSF YKL-40. See, Table S10 for hub gene analysis. DAM, disease-associated microglia.

metabolomics analysis.[50] N-glycan degradation is also significantly affected. Glycosylation abnormalities are common in AD, and a variety of abnormalities have been described.[51,52] Tau undergoes glycosylation in AD brain tissue and not in control brain tissue, and this has been shown to be important for maintenance of paired helical filament structure.[53] In summary, several metabolic pathways that are relevant to AD are predicted to be altered in our CSF analysis, and suggest future avenues for investigating how these changes may correlate with AD pathophysiology.

Although this study offers a unique opportunity to correlate analysis of CSF and brain tissue taken at the same timepoint from living patients, there are important limitations. First of all, unlike some large-scale studies[21,24] we are not powered to perform multiple hypothesis testing on our proteomics data. The variability of the CSF proteome is well documented,[54] and it has been noted that the uncritical application of multiple hypothesis testing for some datasets will result in a failure to detect any true positives even when many exist.[55] Along these lines, other groups have reported unadjusted p-values and used alternate rationales for validation,[38,55–59] including significant protein-protein interactions,[56] coherent ontology groupings,[59] and utility with predictive algorithms.[57,58] Here, we have relied on prior validation in AD CSF in at least two other studies, as well as ELISA validation for two of our markers. We consider the two proteins validated by ELISA (NPTXR and YKL-40) to be the proteins where we are able to make the strongest argument concerning relationships with brain transcriptomic data. The other findings from our proteomics data are trends we consider reportable, but requiring additional validation in a future study. To our knowledge, this is the first study to directly link CSF biomarkers and AD-related changes in CNS gene expression in the same patients. While this certainly lends novelty to our findings, it also limits us in options to validate our results. As NPH cohorts are increasingly studied by AD researchers, there will hopefully be more reports where cross-comparision with the data presented here is possible.

In addition, all of the patients in this study have the comorbitidy of hydrocephalus. Although it is not easy to disentangle what effect this might have on the data presented here, it should be noted that AD is usually accompanied by co-morbid neurologic disease, and that "pure AD" actually constitutes a minority of AD cases.[60–62] Thus, pure AD is actually less common among patients with dementia than mixed pathology, and there is no a priori reason to expect hydrocephalus to uniquely affect our analysis more than other common confounders found in various autopsy and clinical cohorts. Although the interaction of NPH and AD is an area of ongoing research,[5,8,9] the fact that we find correlations between AD pathology on biopsy and AD biomarkers in CSF taken at the same timepoint is itself internal validation that aspects of AD pathophysiology can be studied using our approach, even in the presence of potential confounders.

Finally, it should be noted that, in the United States, patients presenting with NPH are disproportionately white and middle/high income which likely reflects disparities in healthcare delivery.[63] In the retrospective study reported here, we were not able to recover race/ethnicity for all subjects, although 52 out of 60 patients with available racial information were white/caucasian, consistent with national trends for patients treated for NPH. The results reported here should therefore also be examined in the future in a multi-ethnic population, in order to determine how applicable they are to other demographic/ethnic groups.

In summary, we show for the first time how CNS transcriptomic changes (and accompanying early AD pathology) are related to CSF biomarkers. As new disease-modifying therapies are developed targeting specific physiologic aspects of AD (such as synaptic dysfunction or the immune response), biomarkers that track these changes will be crucial. The data presented here offer both biomarkers that can be used for these purposes as well directions for future work.

CONFLICT OF INTERESTS STATEMENT

L.S.H. reports grants from NIH, New York State Dept of Health, Lewy Body Disease Association, CurePSP, Abbvie, Acumen, Alector, Biogen, Bristol-Myer Squibb, Cognition, EIP, Eisai, Genentech/Roche, Janssen/Johnson and Johnson, Transposon Therapeutics, UCB, and Vaccinex, as well as consulting fees from Biogen, Corium, Eisai, Genentech/Roche, and New Amsterdam, Payment or honoraria from Eisai Pharmaceuticals, Medscape, and Biogen, Payment for expert testimony from Monsanto and legal firms, support for attending meetings and/or travel from Eisai Pharmaceuticals, participation on a data safety monitoring or advisory board from Prevail Therapeutics/Lilly, Cortexyme, and Eisai, and a leadership role in the Alzheimer's Association. G.W.M. reports grant funding from NIH, CancerUK, Department of Defense (USAMRAA), Alley Corp, and SPARK-NS. A.F.T. reports grant funding from NIH and Regeneron, stock ownership in Biogen and Ionis, paid committee work for DOD and NIH, and unpaid committee work for the Alzheimer's Association. R.A.M. reports grants from NIH, Minnesota Partnership for Biotechnology and Genomics, Minnesota Robotics Institute, and MnDRIVE Data Science Initiative. G.M.M. reports grants from NIH, consulting with Koh Young Inc and NeuroOne Technologies, participation on the Medronic SLATE trial Publication Committee, and leadership/committee roles in the Neurosurgical Society of America, AANS, and ASSFN. L.M.B. reports support from NIH. The other authors have nothing to report. Author disclosures are available in the supporting information.

CONSENT STATEMENT

This study is a retrospective study that uses tissue and CSF samples not required for clinical diagnosis and associated clinical and demographic data. Patients undergoing surgery signed an umbrella consent form, providing consent to use the remaining biopsies and body fluids for research. This study was reviewed and approved by the Columbia University Institutional Review Board (IRB), and all relevant ethical regulations have been followed.

ORCID

Andrew F. Teich https://orcid.org/0000-0002-1916-8490

REFERENCES

1. Oliveira LM, Nitrini R, Roman GC. Normal-pressure hydrocephalus: A critical review. *Dement Neuropsychol*. 2019;13:133-143.
2. Williams MA, Malm J. Diagnosis and treatment of idiopathic normal pressure hydrocephalus. *Continuum*. 2016;22:579-599.
3. Borgesen SE, Gjerris F. The predictive value of conductance to outflow of CSF in normal pressure hydrocephalus. *Brain*. 1982;105:65-86.
4. Symon L, Dorsch NW. Use of long-term intracranial pressure measurement to assess hydrocephalic patients prior to shunt surgery. *J Neurosurg*. 1975;42:258-273.
5. McGovern RA, Nelp TB, Kelly KM, et al. Predicting cognitive improvement in normal pressure hydrocephalus patients using preoperative neuropsychological testing and cerebrospinal fluid biomarkers. *Neurosurgery*. 2019;85:E662-E669.
6. Hamilton R, Patel S, Lee EB, et al. Lack of shunt response in suspected idiopathic normal pressure hydrocephalus with Alzheimer disease pathology. *Ann Neurol*. 2010;68:535-540.
7. Libard S, Alafuzoff I. Alzheimer's disease neuropathological change and loss of matrix/neuropil in patients with idiopathic normal pressure hydrocephalus, a model of Alzheimer's disease. *Acta Neuropathol Commun*. 2019;7:98.
8. Leinonen V, Koivisto AM, Savolainen S, et al. Amyloid and tau proteins in cortical brain biopsy and Alzheimer's disease. *Ann Neurol*. 2010;68:446-453.
9. Luikku AJ, Hall A, Nerg O, et al. Predicting development of Alzheimer's disease in patients with shunted idiopathic normal pressure hydrocephalus. *J Alzheimers Dis*. 2019;71:1233-1243.
10. Huang W, Bartosch AM, Xiao H, et al. An immune response characterizes early Alzheimer's disease pathology and subjective cognitive impairment in hydrocephalus biopsies. *Nat Commun*. 2021;12:5659.
11. Gazestani V, Kamath T, Nadaf NM, et al. Early Alzheimer's disease pathology in human cortex involves transient cell states. *Cell*. 2023;186:4438-4453.e23.
12. Keren-Shaul H, Spinrad A, Weiner A, et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell*. 2017;169:1276-1290.e17.
13. Gerrits E, Brouwer N, Kooistra SM, et al. Distinct amyloid-beta and tau-associated microglia profiles in Alzheimer's disease. *Acta Neuropathol*. 2021;141:681-696.
14. Andrews S, FastQC: a quality control tool for high throughput sequence data. 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc
15. Durbin BP, Hardin JS, Hawkins DM, Rocke DM. A variance-stabilizing transformation for gene-expression microarray data. *Bioinformatics*. 2002;18(Suppl 1):S105-S110.
16. Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet*. 2007;3:1724-1735.
17. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
18. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
19. Hansson O, Lehmann S, Otto M, Zetterberg H, Lewczuk P. Advantages and disadvantages of the use of the CSF Amyloid beta (Abeta) 42/40 ratio in the diagnosis of Alzheimer's Disease. *Alzheimers Res Ther*. 2019;11:34.
20. Szeto B, Valentini C, Aksit A, et al. Impact of systemic versus intratympanic dexamethasone administration on the perilymph proteome. *J Proteome Res*. 2021;20:4001-4009.
21. Johnson ECB, Dammer EB, Duong DM, et al. Large-scale proteomic analysis of Alzheimer's disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. *Nature medicine*. 2020;26:769-780.
22. Benjamini Y, Hochbert Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B*. 1995;57:289-300.
23. Dayon L, Nunez Galindo A, Wojcik J, et al. Alzheimer disease pathology and the cerebrospinal fluid proteome. *Alzheimers Res Ther*. 2018;10:66.
24. Higginbotham L, Ping L, Dammer EB, et al. Integrated proteomics reveals brain-based cerebrospinal fluid biomarkers in asymptomatic and symptomatic Alzheimer's disease. *Sci Adv*. 2020;6:eaaz9360.
25. Babic Leko M, Mihelcic M, Jurasovic J, et al. Heavy metals and essential metals are associated with cerebrospinal fluid biomarkers of Alzheimer's disease. *Int J Mol Sci*. 2022;24:467.
26. Dulewicz M, Kulczynska-Przybik A, Borawska R, Slowik A, Mroczko B. Evaluation of synaptic and axonal dysfunction biomarkers in Alzheimer's disease and mild cognitive impairment based on CSF and bioinformatic analysis. *Int J Mol Sci*. 2022;23:10867.
27. Kusnierova P, Zeman D, Hradilek P, Zapletalova O, Stejskal D. Determination of chitinase 3-like 1 in cerebrospinal fluid in multiple sclerosis and other neurological diseases. *PLoS One*. 2020;15:e0233519.
28. Lim B, Sando SB, Grontvedt GR, Brathen G, Diamandis EP. Cerebrospinal fluid neuronal pentraxin receptor as a biomarker of long-term progression of Alzheimer's disease: a 24-month follow-up study. *Neurobiol Aging*. 2020;93:e1-97.e7.
29. Teitsdottir UD, Jonsdottir MK, Lund SH, Darreh-Shori T, Snaedal J, Petersen PH. Association of glial and neuronal degeneration markers with Alzheimer's disease cerebrospinal fluid profile and cognitive functions. *Alzheimers Res Ther*. 2020;12:92.
30. Dulewicz M, Kulczynska-Przybik A, Slowik A, Borawska R, Mroczko B. Neurogranin and neuronal pentraxin receptor as synaptic dysfunction biomarkers in Alzheimer's disease. *J Clin Med*. 2021;10:4575.
31. Liu KH, Walker DI, Uppal K, et al. High-resolution metabolomics assessment of military personnel: evaluating analytical strategies for chemical detection. *J Occup Environ Med*. 2016;58:S53-S61.
32. Vardarajan B, Kalia V, Manly J, et al. Differences in plasma metabolites related to Alzheimer's disease, APOE epsilon4 status, and ethnicity. *Alzheimers Dement*. 2020;6:e12025.
33. Yu T, Park Y, Johnson JM, Jones DP. apLCMS–adaptive processing of high-resolution LC/MS data. *Bioinformatics*. 2009;25:1930-1936.
34. Uppal K, Soltow QA, Strobel FH, et al. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. *BMC Bioinformatics*. 2013;14:15.
35. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012;28:882-883.
36. Li S, Park Y, Duraisingham S, et al. Predicting network activity from high throughput metabolomics. *PLoS Comput Biol*. 2013;9:e1003123.
37. Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A. CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics*. 2021;22:433.
38. Pedrero-Prieto CM, Garcia-Carpintero S, Frontinan-Rubio J, et al. A comprehensive systematic review of CSF proteins and peptides that define Alzheimer's disease. *Clin Proteomics*. 2020;17:21.
39. Whelan CD, Mattsson N, Nagle MW, et al. Multiplex proteomics identifies novel CSF and plasma biomarkers of early Alzheimer's disease. *Acta neuropathologica communications*. 2019;7:169.
40. Liddelow SA, Barres BA. Reactive astrocytes: production, function, and therapeutic potential. *Immunity*. 2017;46:957-967.
41. Liddelow SA, Guttenplan KA, Clarke LE, et al. Neurotoxic reactive astrocytes are induced by activated microglia. *Nature*. 2017;541:481-487.
42. Bonneh-Barkay D, Bissel SJ, Kofler J, Starkey A, Wang G, Wiley CA. Astrocyte and macrophage regulation of YKL-40 expression and cellular response in neuroinflammation. *Brain Pathol*. 2012;22:530-546.

43. Connolly K, Lehoux M, O'Rourke R, et al. Potential role of chitinase-3-like protein 1 (CHI3L1/YKL-40) in neurodegeneration and Alzheimer's disease. *Alzheimers Dement*. 2023;19:9-24.

44. Quinn JP, Kandigian SE, Trombetta BA, Arnold SE, Carlyle BC. VGF as a biomarker and therapeutic target in neurodegenerative and psychiatric diseases. *Brain Commun*. 2021;3:fcab261.

45. Yin F. Lipid metabolism and Alzheimer's disease: clinical evidence, mechanistic link and therapeutic promise. *FEBS J*. 2023;290:1420-1453.

46. Ioannou MS, Jackson J, Sheu SH, et al. Neuron-astrocyte metabolic coupling protects against activity-induced fatty acid toxicity. *Cell*. 2019;177:1522-1535.e14.

47. Baxter PS, Hardingham GE. Adaptive regulation of the brain's antioxidant defences by neurons and astrocytes. *Free Radic Biol Med*. 2016;100:147-152.

48. Kohlmeier M. *Nutrient Metabolism*. Academic Press; 2003.

49. Solmonson A, DeBerardinis RJ. Lipoic acid metabolism and mitochondrial redox regulation. *J Biol Chem*. 2018;293:7522-7530.

50. Paglia G, Stocchero M, Cacciatore S, et al. Unbiased metabolomic investigation of Alzheimer's disease brain points to dysregulation of mitochondrial aspartate metabolism. *J Proteome Res*. 2016;15:608-618.

51. Haukedal H, Freude KK. Implications of glycosylation in Alzheimer's disease. *Front Neurosci*. 2020;14:625348.

52. Kizuka Y, Kitazume S, Taniguchi N. N-glycan and Alzheimer's disease. *Biochim Biophys Acta Gen Subj*. 2017;1861:2447-2454.

53. Wang JZ, Grundke-Iqbal I, Iqbal K. Glycosylation of microtubule-associated protein tau: an abnormal posttranslational modification in Alzheimer's disease. *Nature medicine*. 1996;2:871-875.

54. Schilde LM, Kosters S, Steinbach S, et al. Protein variability in cerebrospinal fluid and its possible implications for neurological protein biomarker research. *PLoS One*. 2018;13:e0206478.

55. Pascovici D, Handler DC, Wu JX, Haynes PA. Multiple testing corrections in quantitative proteomics: a useful but blunt tool. *Proteomics*. 2016;16:2448-2453.

56. Loupy KM, Lee T, Zambrano CA, et al. Alzheimer's disease: protective effects of mycobacterium vaccae, a soil-derived mycobacterium with anti-inflammatory and anti-tubercular properties, on the proteomic profiles of plasma and cerebrospinal fluid in rats. *J Alzheimers Dis*. 2020;78:965-987.

57. Khan MJ, Desaire H, Lopez OL, Kamboh MI, Robinson RAS. Why inclusion matters for Alzheimer's disease biomarker discovery in plasma. *J Alzheimers Dis*. 2021;79:1327-1344.

58. Kim Y, Kim J, Son M, et al. Plasma protein biomarker model for screening Alzheimer disease using multiple reaction monitoring-mass spectrometry. *Sci Rep*. 2022;12:1282.

59. King CD, Robinson RAS. Evaluating combined precursor isotopic labeling and isobaric tagging performance on orbitraps to study the peripheral proteome of Alzheimer's disease. *Anal Chem*. 2020;92:2911-2916.

60. Schneider JA, Arvanitakis Z, Bang W, Bennett DA. Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology*. 2007;69:2197-2204.

61. James BD, Wilson RS, Boyle PA, TDP-43 stage, mixed pathologies, and clinical Alzheimer's-type dementia. *Brain*. 2016;139:2983-2993.

62. Barnes LL, Leurgans S, Aggarwal NT, et al. Mixed pathology is more likely in black than white decedents with Alzheimer dementia. *Neurology*. 2015;85:528-534.

63. Ghaffari-Rafi A, Mehdizadeh R, Ghaffari-Rafi S, Leon-Rojas J. Inpatient diagnoses of idiopathic normal pressure hydrocephalus in the United States: demographic and socioeconomic disparities. *J Neurol Sci*. 2020;418:117152.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.