



NEURODEGENERATIVE DISEASE

Proteomics of brain, CSF, and plasma identifies molecular signatures for distinguishing sporadic and genetic Alzheimer's disease

Yun Ju Sung^{1,2,3}, Chengran Yang^{1,2}, Joanne Norton^{1,2}, Matt Johnson^{1,2}, Anne Fagan^{4,5}, Randall J. Bateman^{4,5}, Richard J. Perrin^{4,5,6}, John C. Morris^{4,5,6}, Martin R. Farlow⁷, Jasmeer P. Chhatwal⁸, Peter R. Schofield^{9,10}, Helena Chui¹¹, Fengxian Wang^{1,2}, Brenna Novotny¹, Abdallah Eteleeb¹, Celeste Karch^{1,2}, Suzanne E. Schindler^{4,5}, Herve Rhinn¹², Erik C. B. Johnson¹³, Hamilton Se-Hwee Oh¹⁴, Jarod Evert Rutledge¹⁴, Eric B. Dammer¹³, Nicholas T. Seyfried^{13,15}, Tony Wyss-Coray¹⁴, Oscar Harari¹, Carlos Cruchaga^{1,2,4*}

Copyright © 2023 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

Proteomic studies for Alzheimer's disease (AD) are instrumental in identifying AD pathways but often focus on single tissues and sporadic AD cases. Here, we present a proteomic study analyzing 1305 proteins in brain tissue, cerebrospinal fluid (CSF), and plasma from patients with sporadic AD, *TREM2* risk variant carriers, patients with autosomal dominant AD (ADAD), and healthy individuals. We identified 8 brain, 40 CSF, and 9 plasma proteins that were altered in individuals with sporadic AD, and we replicated these findings in several external datasets. We identified a proteomic signature that differentiated *TREM2* variant carriers from both individuals with sporadic AD and healthy individuals. The proteins associated with sporadic AD were also altered in patients with ADAD, but with a greater effect size. Brain-derived proteins associated with ADAD were also replicated in additional CSF samples. Enrichment analyses highlighted several pathways, including those implicated in AD (calcineurin and Apo E), Parkinson's disease (α -synuclein and LRRK2), and innate immune responses (SHC1, ERK-1, and SPP1). Our findings suggest that combined proteomics across brain tissue, CSF, and plasma can be used to identify markers for sporadic and genetically defined AD.

INTRODUCTION

Alzheimer's disease (AD) is the most common cause of dementia, reducing the quality of life among patients and caregivers (1). AD is characterized by amyloid β (A β)–containing plaques and tau neurofibrillary tangles in the brain, resulting in neuronal loss, neuroinflammation, and memory decline (2). AD is genetically heterogeneous. Around 1 to 3% of cases are autosomal dominant AD (ADAD), carrying pathogenic variants in genes such as amyloid precursor protein (*APP*), presenilin-1 (*PSEN1*), and *PSEN2*, normally with onset before 65 years old (3). Most other AD cases are considered sporadic and manifest after 65 years old (4). We and others have identified several rare coding variants in

triggering receptor expressed on myeloid cells 2 (*TREM2*) that increase the risk to develop AD by almost twofold, making *TREM2* the second strongest genetic risk factor for sporadic AD after apolipoprotein E (*APOE*) (5, 6). Proteomic profiling of genetically defined AD subtypes, including individuals with AD risk variants in *TREM2* and ADAD cases, is important for fully understanding the biology of this heterogeneous disease and for identifying AD subtype-specific molecular markers and therapeutic targets.

Timely diagnosis of AD is critical in clinical practice. Neurofilament light chain (NFL), A β 42, A β 42/40 ratio, and phospho-tau181 (p-tau) are cerebrospinal fluid (CSF) protein biomarkers (7), and there are also several noninvasive plasma biomarkers, including A β 42/40 ratio, p-tau217, p-tau231, and glial fibrillary acidic protein (GFAP) (7–9). Identification of additional marker signatures in CSF and plasma is important; hence, signatures may be more effective disease-modifying targets. Even for clinical trials and therapies that target identified proteins, additional markers that do not depend on the target protein are needed to monitor a treatment outcome. Therefore, it is important to develop prediction models that are independent of A β and tau pathology.

To identify proteomic profiles for sporadic AD and for genetically defined AD subtypes (carriers for *TREM2* risk variants and individuals with ADAD), we measured 1305 proteins in brain, CSF, and plasma from the Knight Alzheimer Disease Research Center (Knight ADRC) and Dominantly Inherited Alzheimer Network (DIAN) cohorts (3, 10). We identified a set of proteins that were differentially altered in sporadic AD and genetically defined AD cases. Many of these proteins were successfully validated through several independent cohorts using multiple orthogonal

¹Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63108, USA. ²NeuroGenomics and Informatics Center, Washington University School of Medicine, St. Louis, MO 63108, USA. ³Division of Biostatistics, Washington University School of Medicine, St. Louis, MO 63108, USA. ⁴Charles F. and Joanne Knight Alzheimer Disease Research Center, Washington University School of Medicine, St. Louis, MO 63108, USA. ⁵Department of Neurology, Washington University School of Medicine, St. Louis, MO 63108, USA. ⁶Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63108, USA. ⁷Department of Neurology, Indiana University School of Medicine, Indianapolis, IN 46202, USA. ⁸Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02115, USA. ⁹Neuroscience Research Australia, Randwick, NSW 2031, Australia. ¹⁰School of Biomedical Sciences, University of New South Wales, Sydney, NSW 2052, Australia. ¹¹Department of Neurology, University of Southern California, Los Angeles, CA 90089, USA. ¹²Leal Therapeutics, 17 Briden St., Office 329, Worcester, MA 01605, USA. ¹³Goizueta Alzheimer's Disease Research Center, Emory University School of Medicine, Atlanta, GA 30329, USA. ¹⁴Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA 94304, USA. ¹⁵Department of Biochemistry, Emory School of Medicine, Atlanta, GA 30329, USA.

*Corresponding author. Email: ccruchaga@wustl.edu

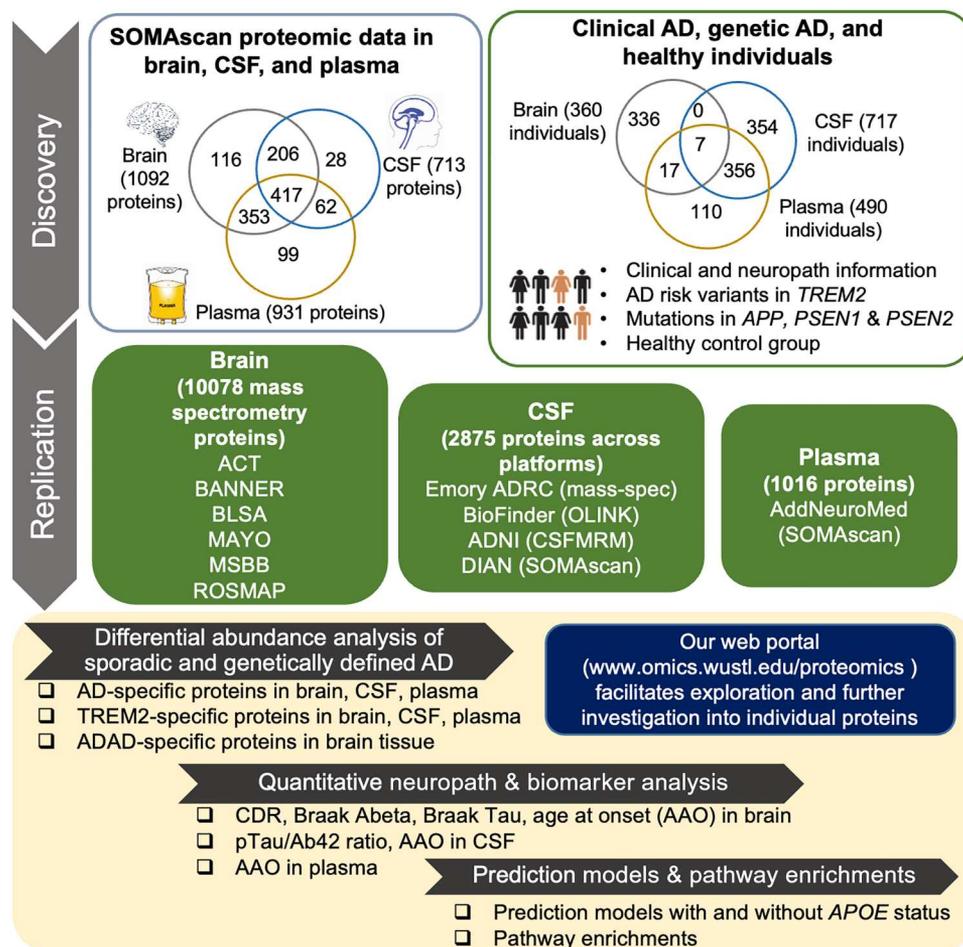


Fig. 1. Study outline. In the discovery stage, protein measures in brain, CSF, and plasma samples were obtained with SOMAScan targeting 1305 proteins from Knight ADRC and DIAN participants with comprehensive clinical information about AD pathology and cognition. This discovery cohort contained patients with sporadic AD (290 brain, 176 CSF, and 105 plasma samples), *TREM2* risk variant carriers (21 brain, 47 CSF, and 131 plasma samples), patients with ADAD (24 brain samples), and healthy controls (25 brain, 494 CSF, and 254 plasma samples). Differential abundance analyses were performed for sporadic AD status, *TREM2* risk variant carrier status, and ADAD status. Several publicly available external proteomic datasets were used to replicate our findings (details in Supplementary Materials and Methods). Last, replicated proteins were used for creating prediction models for brain tissue, CSF, and plasma and pathway enrichment analysis. In addition, we built a web portal (omics.wustl.edu/proteomics) to support interactive visualization and exploration (fig. S2).

platforms. The replicated proteins were used to create brain-, CSF-, and plasma-specific prediction models and to identify the pathways leading to the disease. We also built a web portal (omics.wustl.edu/proteomics) to support interactive visualization and exploration for the scientific community.

RESULTS

Proteomic signatures of sporadic AD

To identify proteomic changes that were associated with sporadic AD across brain, CSF, and plasma, we quantified 1305 proteins using a multiplexed, single-stranded DNA aptamer assay developed by SomaLogic (11). After standard data processing, normalization, and quality control (QC) (Supplementary Materials and Methods), differential abundance analysis was performed on the remaining 1092 proteins in autopsy brain tissue from 290 patients with neuropathologically confirmed AD and 25 cognitively normal individuals with no brain pathology. Differential abundance analysis was also

performed on 713 proteins in CSF samples from 176 patients with AD and 494 healthy controls and on 931 proteins in plasma samples from 105 patients with AD and 254 healthy controls (Fig. 1). In this analysis, genetically defined ADs (*TREM2* variant carriers and ADAD) were excluded. Sample size and patient characteristics are summarized in Table 1. The correlation matrix of these proteomic data is shown in fig. S1. A web portal (omics.wustl.edu/proteomics) was created to facilitate both exploration of our analysis and further investigation into individual proteins across disease status or sex (fig. S2). We performed surrogate variable (SV) analysis (12) to remove batch effects and other unmeasured heterogeneity in the proteomic data. We then performed regression analysis with log-transformed protein abundance as a dependent variable and sporadic AD status as an independent variable while including age, sex, and SVs as covariates. Results were not different when the technical batch was used instead of SVs (figs. S3 and S4). Instead of false discovery rate (FDR), we used a more stringent multiple test correction based on the number of

Table 1. Summary characteristics of participants with proteomic measures in the Knight ADRC and DIAN cohorts. CO, healthy control; AD, sporadic AD cases; *TREM2*, AD risk variant (p.E151K, p.H157Y, p.L211P, p.R136Q, p.R163Q, p.R47H, p.R62H, and p.T96K) carriers in *TREM2*.

Origin	Status	Sample size (N)	% Female	Age (mean ± SD)
Brain	CO	25	61.72	88.24 ± 8.85
	AD	290	33.33	83.98 ± 8.83
	ADAD	24	76.00	55.67 ± 14.58
	<i>TREM2</i>	21	57.14	82.57 ± 7.62
CSF	CO	494	55.26	73.15 ± 6.43
	AD	176	46.02	74.60 ± 7.02
	<i>TREM2</i>	47	44.68	74.00 ± 6.48
Plasma	CO	254	57.48	71.53 ± 7.31
	AD	105	37.14	72.59 ± 7.67
	<i>TREM2</i>	131	64.89	74.98 ± 8.17

independent proteins (see Supplementary Materials and Methods). Several publicly available datasets were downloaded and analyzed to replicate our findings in the discovery cohort (table S1 and Fig. 1).

Brain proteomic profiles for sporadic AD

In the brain, 12 proteins showed significant association with AD status (Fig. 2A; see table S2 for *P* values). All these proteins were also significantly associated with other AD-related traits, including age at onset and AD neuropathological characteristics [such as Braak scores and clinical dementia rating (CDR) (13) at death] (figs. S5 and S6; see table S3 for *P* values). To examine consistency across brain, CSF, and plasma, we checked whether these proteins were also associated with AD risk or onset in CSF and plasma. Of the 12 brain proteins associated with AD status, only 6 were found in both CSF and plasma samples. Of these, five proteins [SPARC-related modular calcium-binding protein 1 (SMOC1), hepatocyte growth factor (HGF), follistatin-like 1 (FSTL1), ubiquitin-conjugating enzyme E2 (UBC9), and neuroepithelial cell transforming 1 (NET1)] were associated with AD status or age at onset in both CSF and plasma data ($P < 0.05$; table S2), which represents a 333-fold enrichment ($P = 5.8 \times 10^{-13}$) to what would be expected by chance.

To replicate our findings in the discovery cohort, we downloaded the mass spectrometry data from the Adult Changes in Thought (ACT), Banner Sun Health Research Institute (BANNER), Baltimore Longitudinal Study of Aging (BLSA), Mayo Clinic (MAYO), Mount Sinai Brain Bank (MSBB), and the Religious Orders Study and the Memory and Aging Project (ROSMAP). We integrated these mass spectrometry datasets including 10,078 proteins from 415 patients with AD and 194 controls (referred as MassSpec Joint) and subsequently performed differential abundance analysis for AD status (table S4). Of the nine proteins that were present in these datasets, eight replicated [Midkine, SMOC1, chromogranin-A (CgA), HGF, neurexin-1- β (NRX1B), UBC9, NET1, and serum amyloid P-component (SAP)] with the consistent direction at $P < 0.05$ (Fig. 2D). This represents a 35-fold enrichment to what would be expected by chance ($P = 1.3 \times 10^{-12}$). In addition, to confirm that our results were not false positives due to the joint analysis that

included six studies, we checked the published results of each individual study [Johnson *et al.* (14), Higginbotham *et al.* (15), and Wingo *et al.* (16)]. Individual study analysis also provided 25- to 34-fold enrichments (table S1). Overall protein changes with AD status in our discovery data and the merged replication data, MassSpec Joint, were similar ($P < 3.6 \times 10^{-3}$; fig. S7A). Here, we identified 12 proteins that were altered in brain tissue from patients with sporadic AD, of which 5 were validated in CSF and plasma and 8 were replicated in the external datasets that had been generated with orthogonal proteomic platforms.

CSF proteomic profiles for sporadic AD

In CSF, 117 proteins were significantly associated with clinical AD status (Fig. 2B; see table S5 for *P* values). Of these 117 proteins, 78 were also found in brain and plasma, and 27 proteins [including extracellular signal-regulated kinase 1 (ERK-1) and leucine-rich repeat kinase 2 (LRRK2)] were validated across brain tissue, CSF, and plasma (138-fold enrichment, $P = 3.3 \times 10^{-50}$). For external replication, we obtained and analyzed Alzheimer's Disease Neuroimaging Initiative (ADNI) multiple reaction monitoring (MRM) proteomic data. We also used results based on BioFinder OLINK data from Whelan *et al.* (17) and Emory-ADRC mass spectrometry data from Higginbotham *et al.* (15). Of the 117 CSF proteins identified in the discovery cohort, 90 were present in external datasets (tables S5 and S6). In these external datasets, 40 proteins [including 14-3-3, calcineurin, SMOC1, GFAP, secreted phosphoprotein 1 (SPP1), and peroxiredoxin-1 (PRDX1)] were replicated at $P < 0.05$ and in the consistent direction (14- to 34-fold enrichments, $P \leq 4.4 \times 10^{-5}$; Fig. 2D). Separately in the ADNI data (320 CSF samples), 8 proteins were available among the 117 identified in discovery, and 7 were replicated. In Higginbotham *et al.* (15), there were 88 proteins among our identified proteins, and 34 were replicated. We speculate that a small sample size of this study ($N = 40$) is a primary reason for a limited power in replicating the discovery findings. The correlation of altered protein changes between discovery and replication data was strong ($r = 0.43$ to 0.82 , $P < 3.4 \times 10^{-7}$; fig. S7B). We therefore expect that more proteins would replicate in larger studies.

Plasma proteomic profiles for sporadic AD

In plasma, 26 proteins were associated with sporadic AD status after the multiple testing correction (Fig. 2C and table S7). Similar to previous analyses, we leveraged data from brain tissue, CSF, and plasma to replicate these findings. Of the 26 plasma proteins associated with AD status, 16 were found in brain and CSF, and 7 proteins [including ERK-1, cell adhesion-associated, oncogene regulated (CDON), and SHC-adaptor protein 1 (SHC1)] were replicated (175-fold enrichment, $P = 6.8 \times 10^{-15}$). For external replication, we downloaded and analyzed the AddNeuroMed SOMAscan 1.1K proteomic data that were processed and deposited by Sattler *et al.* (18). Of 26 associated proteins, we were able to test 19 in these data (table S8). Nine proteins (including CAMK2D and HMG-1) were replicated (Fig. 2D). This represented a 18.9-fold enrichment ($P = 2.8 \times 10^{-10}$) to what would be expected by chance.

Here, we identified 8 proteins in brain, 40 proteins in CSF, and 9 proteins in plasma that were altered by AD status through a traditional discovery and replication strategy by leveraging several external proteomic datasets generated from orthogonal platforms. They included Apo E2 and SMOC1 and other previously identified proteins (table S9) related to amyloid/tau pathology. Several proteins identified in this study showed a very weak correlation with CSF

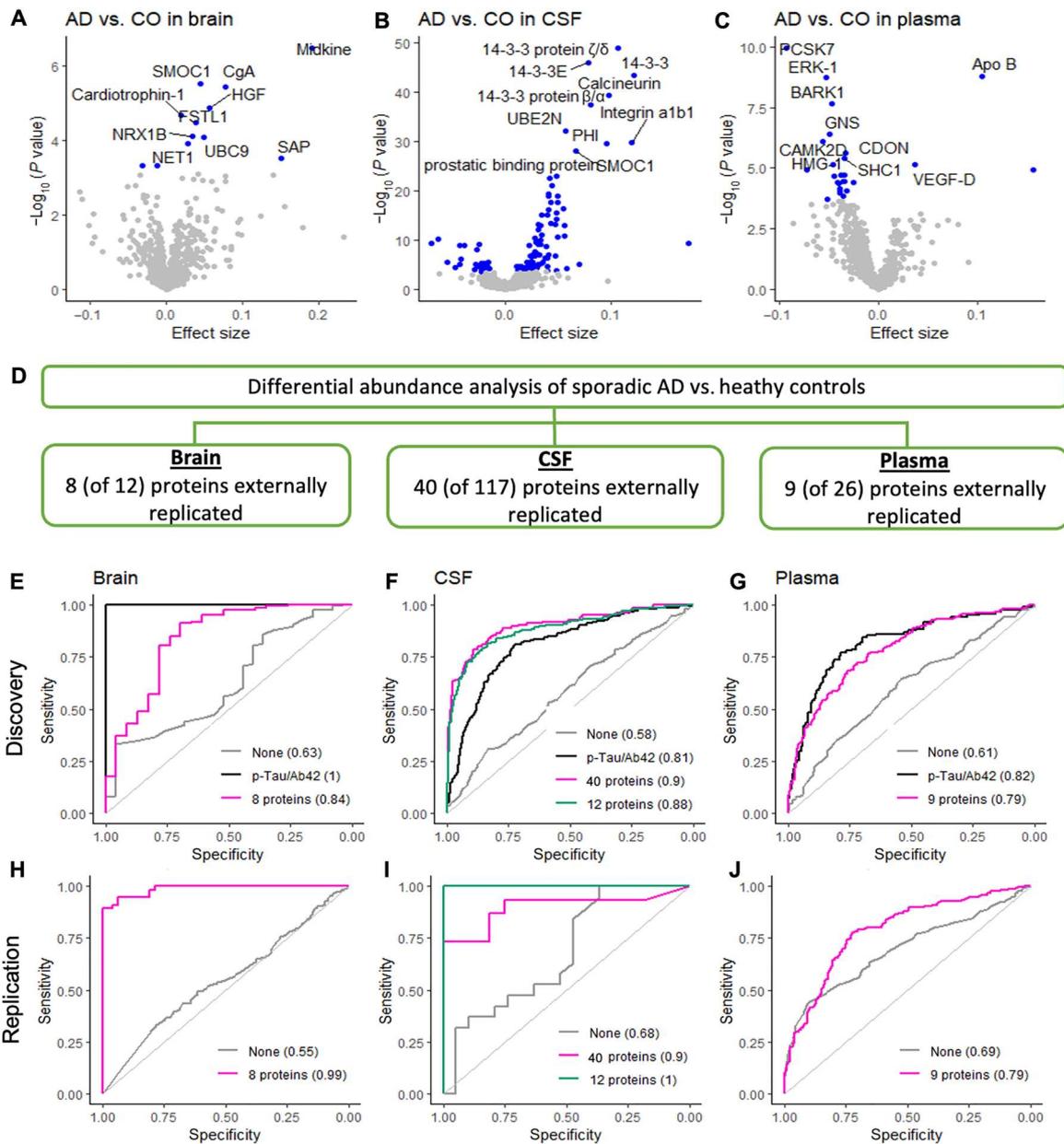


Fig. 2. Proteomic profiling of sporadic AD and external replication. (A to C) Volcano plots displaying the log₂ fold change (x axis) against statistical significance (y axis) for all proteins tested for differential abundance between control (CO) and AD cases of the brain (A), CSF (B), and plasma (C) discovery proteome. The blue points show the proteins significant at the multiple testing–corrected threshold. Although the top 10 proteins are labeled here, the volcano plots in the web portal (omics.wustl.edu/proteomics) support interactive exploration for all proteins. (D) Summary showing the number of identified and externally replicated brain, CSF, and plasma proteins. (E to J) Receiver operating characteristic (ROC) curves show the performance of externally replicated proteins (8 in brain, 40 in CSF, and 9 in plasma) for the discrimination of patients with AD and healthy controls (CO) for both discovery [(E) in brain, (F) in CSF, and (G) in plasma] and replication data [(H) in brain, (I) in CSF, and (J) in plasma]. Sex and age were included as covariates for all models. “None” corresponds to the model including age and sex only, without any proteins or other biomarkers. In CSF, a model with a subset of 12 proteins was also examined (in green curve). Replication data were MassSpec Joint combining all mass spectrometry–based cohorts in brain, Emory-ADRC mass spectrometry data in CSF, and AddNeuroMed in plasma.

Aβ42 (median correlation value for Aβ42 was 0.06 in brain, 0.002 in CSF, and 0.04 in plasma). Whereas CSF proteins were modestly correlated with CSF p-Tau181 (median correlation = 0.35), proteins in brain and plasma were very weakly correlated (median correlation = −0.03 and 0.006, respectively). Correlation plots for all these externally replicated proteins are in fig. S8.

Because in some scenarios it may not be possible to use independent datasets for replication, we tested whether protein identification across brain tissue, CSF, and plasma can serve as an alternative option for replication. An enrichment test showed that the proteins in brain tissue, CSF, and plasma were more likely to be replicated in external independent datasets than proteins identified in samples of

a single origin (15- to 40-fold enrichments, $P \leq 3.63 \times 10^{-3}$; table S10). This suggests that multitissue and fluid proteomic data can be used as a viable replication strategy to support the findings.

Proteomic profiles based on AD biomarkers

Most AD biomarker studies in CSF and plasma published so far examined clinical AD status and not biomarker-based status (18). To allow for an easy comparison with these earlier studies, we also performed our analysis using clinical status as AD classifier. However, several studies indicated that up to 30% of cognitively normal elderly individuals could be presymptomatic for AD (19) and that other neurodegenerative diseases can clinically masquerade as AD dementia (20). It was shown that the CSF p-tau/A β 42 ratio is a gold-standard biomarker not only for AD status but also for predicting AD progression from normal to dementia within 5 years (10). To examine whether our findings based on clinical status were robust, we subsequently performed differential analyses with both AT(N) classification and CSF p-tau/A β 42 ratio. We had access to CSF p-tau/A β 42 measures for 689 (of 717) CSF samples and 393 (of 490) plasma samples in the discovery study. Following the biomarker-based AT(N) classification (21), we obtained AT classification of amyloid/tau positivity in these CSF and plasma samples (Supplementary Materials and Methods). Of the 117 proteins associated with clinical AD status, 97 were significant for CSF p-tau/A β 42, and 102 were significant for the AT(N) classification (see table S11 for P values). There was a strong correlation between protein changes by clinical AD status and those by biomarker-based status ($r = 0.86$ and 0.88 , respectively; $P < 1.0 \times 10^{-16}$; fig. S9). Similar results were found for plasma (fig. S9 and table S12). This high correlation indicates that the results found using clinical AD status can be interchangeable with those using biomarker status in this discovery study.

Prediction models based on brain, CSF, and plasma proteins

A prediction model (or predictive model) using protein markers is critical for early diagnosis and monitoring disease progression. We created prediction models based on the 8, 40, and 9 proteins that were detected in brain, CSF, and plasma, respectively, and that were replicated in external datasets while including sex and age as covariates. The prediction model based on the eight brain proteins provided high accuracy in distinguishing AD cases from cognitively normal individuals: an area under curve (AUC) of 0.84 in discovery and 0.99 in the independent MassSpec Joint data (Fig. 2, E and H). The prediction model based on the 40 proteins detected in CSF provided an AUC of 0.89 in discovery and 0.90 in the Emory-ADRC mass spectrometry replication study (Fig. 2, F and I). Because this prediction model included too many proteins to translate into clinical practice, we performed a stepwise model selection and identified a panel of 12 proteins (table S13). These 12 proteins provided accuracy almost as high as all 40 proteins, leading to an AUC of 0.88 in discovery and 0.99 in replication data. This was significantly higher than the AUC based on the well-known CSF p-tau/A β 42 ratio (AUC = 0.81; $P = 2.4 \times 10^{-6}$). Using the same approach for plasma, the prediction model based on the nine proteins led to an AUC of 0.79 in both discovery and AddNeuroMed replication data (Fig. 2, G and J). This was not statistically different from the AUC with the CSF p-tau/A β 42 ratio (AUC = 0.82; $P > 0.05$). The prediction model based on each protein was similar between the discovery and replication data (fig. S10).

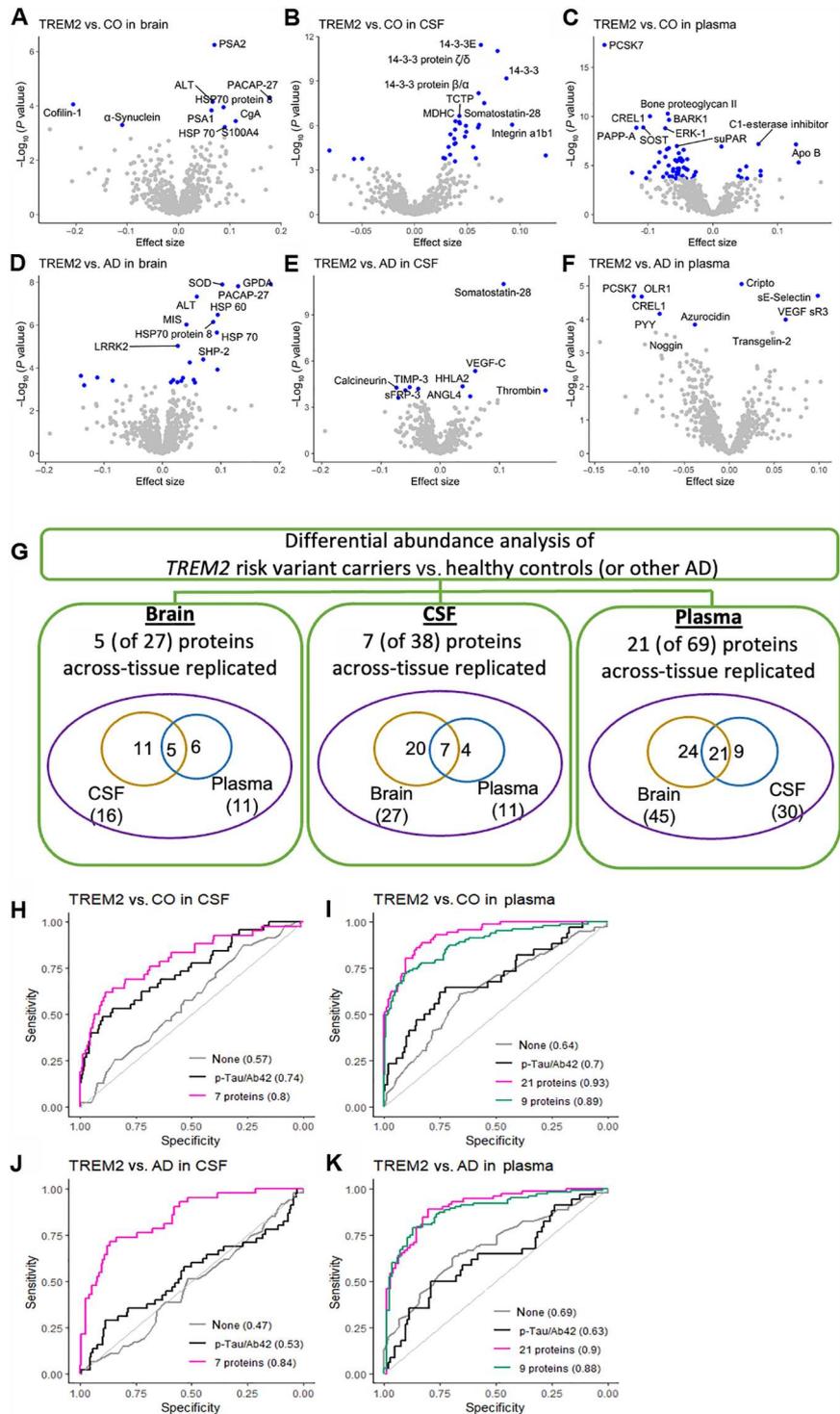
Proteomic signatures of *TREM2* risk variant carriers based on brain, CSF, and plasma proteins

Several rare coding variants in *TREM2* that increase risk of AD by almost twofold have been identified (4). We therefore aimed to identify proteomic signatures across brain tissue, CSF, and plasma of individuals carrying AD risk variants in *TREM2*. Given the low frequency of occurrence of different *TREM2* variants, all of them were combined into one group of *TREM2* risk variant carriers. We generated brain, CSF, and plasma proteomic data for 21, 47, and 131 *TREM2* variant carriers, respectively (Table 1), and compared their protein abundances with those of cognitively normal individuals and patients with sporadic AD who did not carry any *TREM2* variant (Fig. 3, A to F).

In the brain, nine proteins (including α -synuclein) were altered in *TREM2* risk variant carriers compared with cognitively normal individuals at the multiple testing-corrected threshold (Fig. 3A and table S14). In addition, 23 proteins (including LRRK2) were altered in *TREM2* risk variant carriers when compared with other sporadic AD cases (Fig. 3D and table S15). Five proteins [alanine aminotransferase (ALT), heat shock protein 70 (HSP70), HSP70 protein 8, pituitary adenylate cyclase-activating polypeptide 27 (PACAP-27), and proteasome subunit α type-1 (PSA1)] were commonly found, indicating their distinct protein abundance across all three groups (*TREM2* variant carriers versus healthy controls versus other sporadic AD cases). The external data included only four *TREM2* variant carriers in MAYO, seven in MSBB, and eight in ROSMAP, which did not provide any statistical power to replicate our findings. As we demonstrated, our study design is a viable alternative approach to identify proteins that would replicate in external datasets, and we leveraged our data to identify those proteins that replicate across brain tissue, CSF, and plasma. Of these 27 *TREM2*-associated proteins found in the brain (combining 9 and 23 proteins minus 5 commonly found), 11 were replicated only in CSF, 6 were replicated only in plasma, and 5 [ALT, α -synuclein, muellerian-inhibiting factor (MIS), LRRK2, and platelet-activating factor acetylhydrolase (PAFAH) β subunit] were replicated in both fluids (Fig. 3G). This represents a 74-fold enrichment ($P = 7.5 \times 10^{-9}$) to what would be expected by chance.

In CSF, we identified 31 proteins altered in *TREM2* variant carriers compared with healthy controls (Fig. 3B and table S16) and 10 proteins altered in *TREM2* variant carriers compared with other sporadic AD cases (Fig. 3E and table S17). Three proteins (nucleoside diphosphate kinase A, somatostatin-28, and thrombin) were common. Of these 38 proteins (31 and 10 minus 3 commonly found), 20 were replicated in the brain, 4 were replicated in the plasma, and 7 [14-3-3E, 14-3-3 protein ζ/δ , somatostatin-28, SMOC1, Ubiquitin+1, quinone oxidoreductase-like 1 (QORL1), and calcineurin] were replicated in brain tissue and plasma (Fig. 3G). This represents a 73-fold enrichment ($P = 7.19 \times 10^{-12}$) to what would be expected by chance. In plasma, we identified a total of 69 proteins: 65 proteins altered in *TREM2* variant carriers compared with healthy controls (Fig. 3C and table S18); 7 proteins altered in *TREM2* variant carriers compared with other sporadic AD (Fig. 3F and table S19); and 3 overlaps [cysteine rich with epidermal growth factor-like domains 1 (CREL1), Cripto, and proprotein convertase subtilisin/kexin type 7 (PCSK7)]. Among these 69 proteins, 24 proteins were present in the brain, 9 proteins were replicated in the CSF, and 21 [including bone proteoglycan II, pappalysin 1 (PAPP-A), ERK-1, and vascular cell adhesion molecule 1

Fig. 3. Proteomic profiling of *TREM2* variant carrier status and replication across brain tissue, CSF, and plasma. (A to C) Volcano plots displaying the \log_2 fold change (x axis) against statistical significance (y axis) for all proteins tested for differential abundance between *TREM2* variant carriers and healthy controls (CO) [in brain (A), in CSF (B), and in plasma (C)]. **(D to F)** Volcano plots displaying the \log_2 fold change (x axis) against statistical significance (y axis) for all proteins tested for differential abundance in *TREM2* variant carriers and sporadic AD cases without any *TREM2* variants in brain (D), in CSF (E), and in plasma (F). Although the top 10 proteins are labeled here, the volcano plots in the web portal (omics.wustl.edu/teomics) support interactive exploration for all proteins. **(G)** Summary showing identified and across-tissue replicated proteins in three tissues. Multiple proteins showed differential abundance in *TREM2* variant carriers (compared with controls or other sporadic AD cases), several of which are replicated across tissue. **(H to K)** Tissue-specific ROC curves show the performance of prediction models based on proteins replicated across tissues [(H) and (J) in CSF and (I) and (K) in plasma], including sex and age as covariates. “None” corresponds to the model including age and sex only, without any proteins or other biomarkers.



(VCAM-1)] were replicated in plasma (Fig. 3G), representing a 122-fold enrichment ($P = 5.47 \times 10^{-38}$) to what would be expected by chance.

We also created prediction models that could distinguish *TREM2* variant carriers from noncarriers in both sporadic AD cases and controls for CSF and plasma. In CSF, the prediction model based on the seven proteins replicated across brain tissue,

CSF, and plasma provided an AUC of 0.79 for distinguishing *TREM2* variant carriers from cognitively normal individuals (Fig. 3H). The same model showed an AUC of 0.84 for distinguishing *TREM2* variant carriers from the other sporadic AD cases (Fig. 3J). Although the CSF p-tau/Aβ42 ratio is a very good biomarker to distinguish AD cases from controls, no previous studies examined how the CSF p-tau/Aβ42 ratio provides prediction

for *TREM2* variant carriers. In this study, CSF p-tau/A β 42 showed an AUC of 0.74 for *TREM2* variant carriers versus cognitively normal individuals and 0.53 for *TREM2* variant carriers versus other ADs. Both AUC values from our *TREM2*-associated prediction model with seven proteins were significantly higher than the model based on the CSF p-tau/A β 42 ratio ($P < 1.6 \times 10^{-5}$; Fig. 3H).

In plasma, our prediction model based on the 21 proteins replicated across brain tissue, CSF, and plasma provided an AUC of 0.93 in distinguishing *TREM2* variant carriers from healthy individuals (Fig. 3I). Predictive performance was significantly higher than the model based on the CSF p-tau/A β 42 ratio (AUC = 0.69; $P = 1.1 \times 10^{-3}$). Similarly, in distinguishing *TREM2* risk variant carriers from other AD cases, the prediction model based on the same 21 proteins provided an AUC of 0.90 (Fig. 3K), which is significantly higher ($P = 1.5 \times 10^{-4}$) than the AUC with the CSF p-tau/A β 42 ratio (AUC = 0.63). Because the number of proteins is large, we performed a stepwise model selection and found a subset of nine proteins (table S13) that provided AUCs of 0.89 and 0.88 in distinguishing *TREM2* variant carriers from cognitively normal individuals and from other sporadic AD cases, respectively (Fig. 3, I and K). The prediction models including age, sex, and *APOE* ϵ 4 status as covariates provided similar performance (fig. S11).

Proteomic signatures of ADAD

To identify proteins associated with ADAD status, we generated proteomic data from the parietal cortices of 24 individuals carrying pathogenic ADAD variants (19 individuals with *PSEN1*, 1 with *PSEN2*, and 4 with *APP* variants) recruited from the DIAN and the Knight ADRC studies. We identified 109 proteins altered in ADAD mutation carriers when compared with cognitively normal individuals with no brain pathology at the multiple testing-corrected threshold (Fig. 4A). Because ADAD cases were much younger than the cognitively normal control group, age was not included in the model to avoid colinearity. However, this could lead to false positives because some of the significant proteins could be associated with age rather than ADAD status. To address this, we identified 98 proteins associated with age in the control group at nominal significance ($P < 0.05$; Fig. 4B and table S20). Of these age-associated proteins, 17 (of 109) were associated with ADAD status and excluded from any downstream analyses.

To validate whether the remaining 92 proteins were also associated with ADAD status in CSF, we analyzed CSF proteins from 289 ADAD mutation carriers and 184 noncarriers from the DIAN study. Of the 92 proteins identified in brain, 89 passed QC in CSF proteomic data, and 14 were altered by ADAD status in CSF in the consistent direction (Fig. 4C and table S21), representing a 6.3-fold enrichment ($P = 4.88 \times 10^{-8}$) to what would be expected by chance. We next leveraged these 14 proteins to create potential prediction models for distinguishing ADAD mutation carriers from noncarriers. A prediction model using these 14 proteins in the brain provided an AUC of 1, fully separating ADAD cases from the healthy controls (Fig. 4D). In CSF data, the same 14 proteins provided higher predictive performance than the model with sex alone (AUC = 0.86 versus 0.52, $P < 2.2 \times 10^{-16}$; Fig. 4E).

As presented earlier, we identified 12 proteins associated with sporadic AD status in brain tissue (Fig. 2A and table S2). We also sought to determine whether the proteins altered by sporadic AD status showed similar alteration in individuals with ADAD. We found that most of the proteins associated with sporadic AD

brains displayed even stronger alteration in ADAD individuals compared with healthy individuals (table S22). The proteins associated with sporadic AD status showed 36% higher protein changes in ADAD brain samples on average ($P = 1.0 \times 10^{-4}$; Fig. 4F). For example, SMOC1 showed a significant association not only for sporadic AD status (protein change = 0.04; $P = 3.1 \times 10^{-6}$) but also for ADAD at a larger magnitude (protein change = 0.13; $P = 2.3 \times 10^{-6}$; Fig. 4G).

Similarly, other proteins, including SAP, HGF, CgA, and NET1, were more dysregulated in ADAD compared with AD (Fig. 4, H to K). As presented earlier, SMOC1 was also associated in sporadic AD status in both CSF ($P = 8.4 \times 10^{-29}$) and plasma ($P = 0.002$), suggesting that it could be used to create a new prediction model for AD.

Validation of SomaLogic data

We recently demonstrated that SomaLogic data correlated with classic enzyme-linked immunosorbent assay (ELISA) for several well-known CSF biomarkers (22). We found high correlation (over 0.91) for NFL, neurogranin, and visinin-like 1 (VILIP-1). Prediction for AD status based on the CSF SOMAscan proteins was also comparable to that based on the ELISA markers.

To further validate our SOMAscan data, we examined the 150 proteins we identified across different platforms in samples from the Emory, ADNI, and DIAN cohorts (tables S23 and S24). In the Emory cohort, the overall correlation for the overlapping proteins was 0.75 for CSF and 0.74 for plasma when comparing SomaLogic versus mass spectrometry measures and 0.74 for both CSF and plasma when comparing SomaLogic versus Olink (table S23). Many proteins showed higher correlation, including SMOC1, α -synuclein, and 14-3-3E (figs. S12 to S15). In ADNI, the overall correlation for the eight proteins that overlap was 0.72 (fig. S16). Of these proteins, GFAP, aldolase A (ALDOA), PRDX1, PRDX6, and polyubiquitin K48 (UBB; fig. S16), all with very high correlation, were replicated in CSF analysis using ADNI mass spectrometry data (table S17). APOB was identified in our plasma analyses. For the DIAN cohort, we found that the overall correlation for the overlapping 14 proteins were 0.79 (table S23). These proteins included SMOC1 (identified in brain and CSF), osteopontin (SPP1), tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein ζ (YWHAZ), and ALDOA (identified in CSF) (figs. S17 and S18).

Functional pathways

Last, we examined functional pathways for the proteins identified in our study using Enrichr (23). As expected, the identified proteins were enriched for the AD pathway (FDR = 1.9×10^{-2} ; Fig. 5A and table S25), including Apo E (*APOE*), calcineurin [protein phosphatase 3 regulatory subunit B, α (*PPP3R1*) and protein phosphatase 3, catalytic subunit (*PPP3CA*)], and ERK-1 [mitogen-activated protein kinase 3 (*MAPK3*)] (Fig. 5B and fig. S19). This finding supports our study design to identify and validate proteins for AD. In addition to the AD pathway, the identified proteins altered in sporadic AD and *TREM2* variant carriers were enriched for the Parkinson's disease (PD) pathway (FDR = 2.1×10^{-12} for sporadic AD; FDR = 1.4×10^{-4} for *TREM2* variant carriers; table S25). Among the proteins in the PD pathway, α -synuclein (*SNCA*) was altered in both sporadic AD ($P = 9.3 \times 10^{-8}$) and *TREM2* carriers ($P = 5.0 \times 10^{-4}$) compared with controls. Similarly, LRRK2 was altered

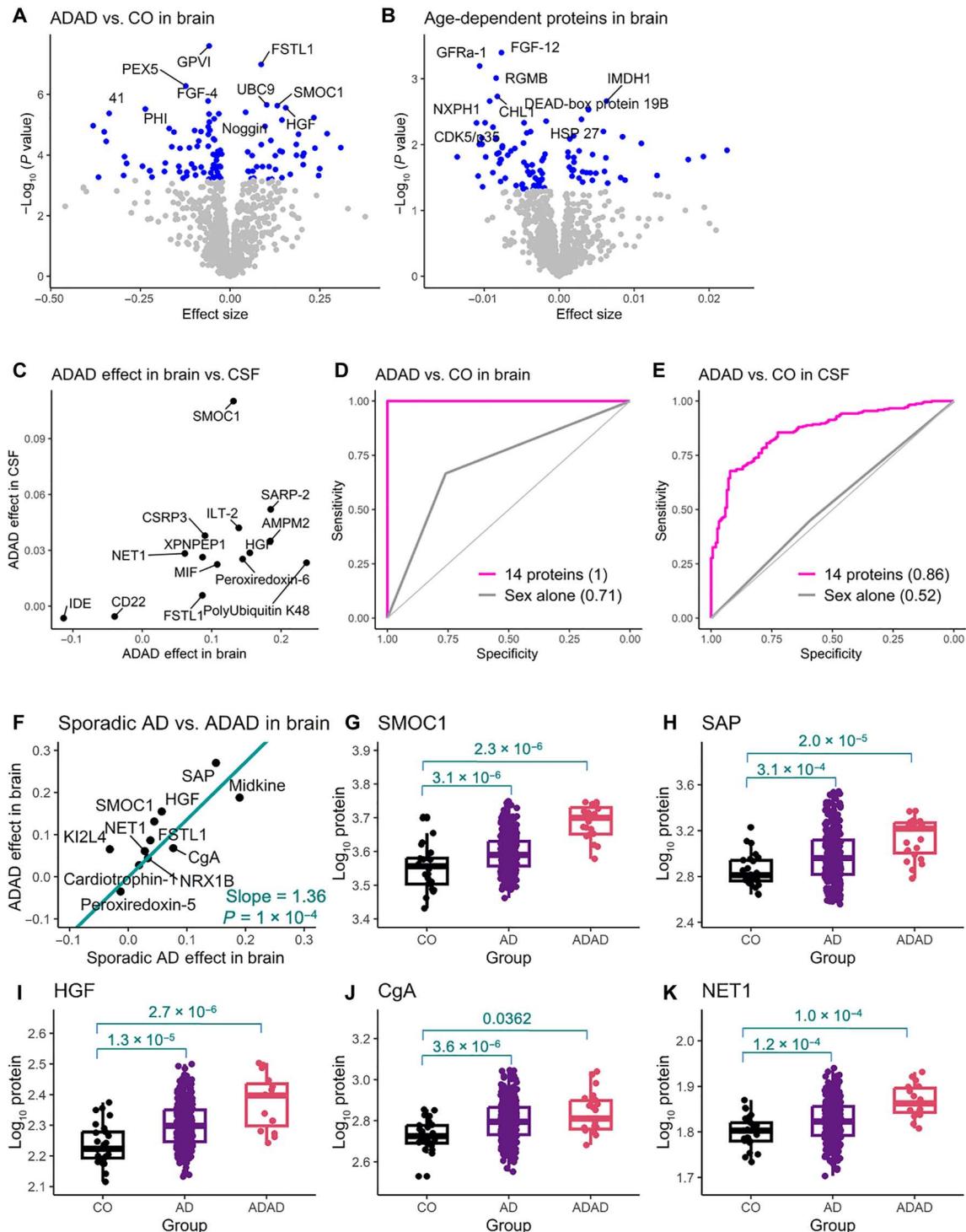


Fig. 4. Proteomic profiling of ADAD status. (A) The volcano plots display the \log_2 fold change of protein abundance in brain (x axis) against statistical significance (y axis) between individuals with ADAD and healthy controls (CO). (B) Volcano plots displaying the \log_2 fold change of protein abundance in brain (x axis) against statistical significance (y axis) dependent on age. (C to E) Scatterplot of the 14 proteins replicated in CSF (C) and their prediction models in brain (D) and CSF (E). (F) Scatterplot of the 12 proteins associated with sporadic AD status. The y axis shows the effect of ADAD status on log-transformed protein abundance; the x axis shows the effect of sporadic AD on log-transformed protein abundance. (G to K) Box plots for the select five proteins (SMOC1, SAP, HGF, CgA, and NET1).

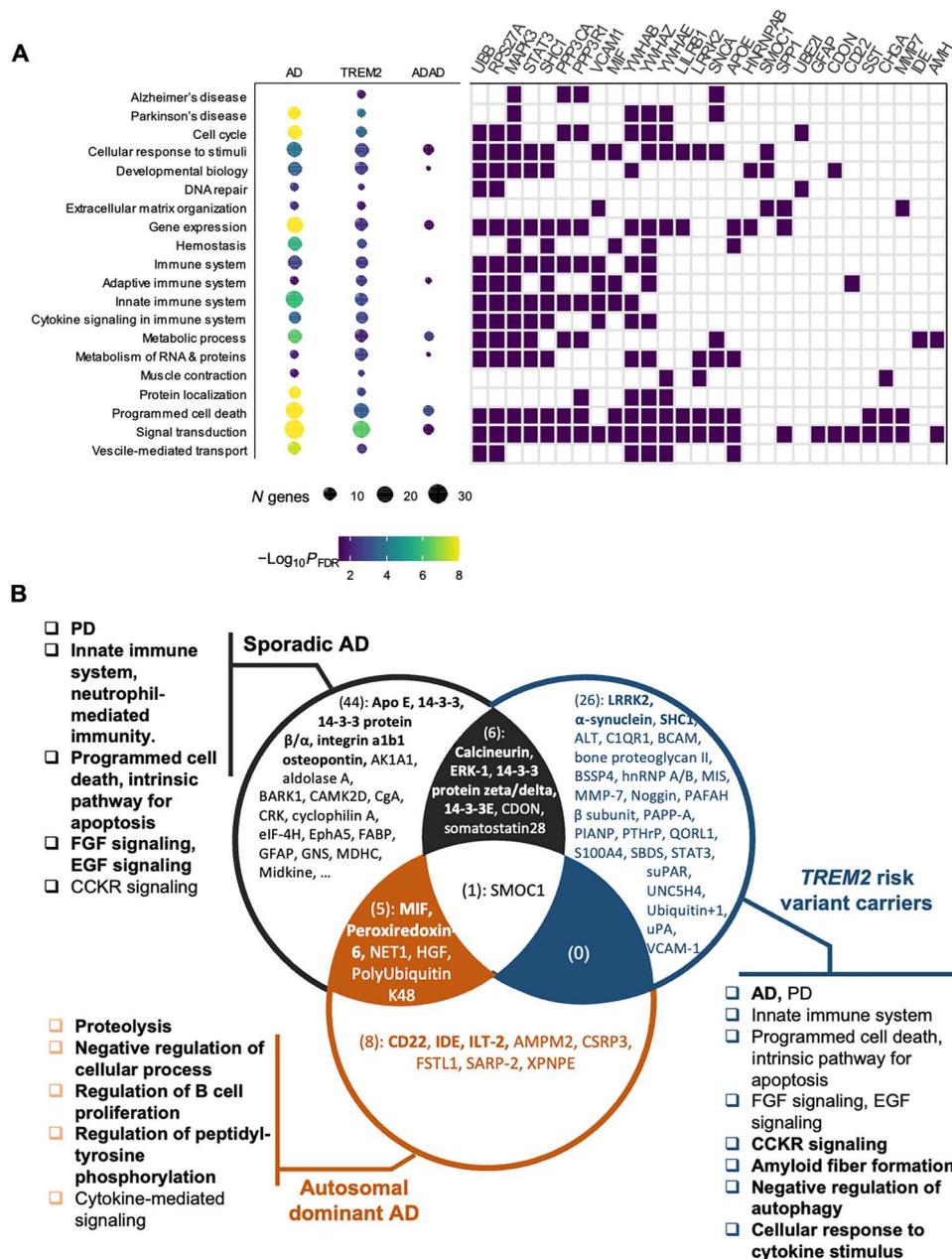


Fig. 5. Pathway enrichment for sporadic and genetically defined AD. (A) (Left) Dot chart illustrating several pathways shared across patients with sporadic AD, *TREM2* carriers, and patients with ADAD. The size of the dot corresponds to the number of identified genes. The color of the dot corresponds to the FDR-corrected significance (see table S25). (Right) The tile plot shows differentially expressed genes that belong to a specific pathway. A full list of 81 genes is shown in fig. S12. (B) The Venn diagram shows the overlap of identified proteins across the three groups of AD: 56 externally replicated proteins for sporadic AD, 33 across-tissue replicated proteins for *TREM2*, and 14 proteins replicated in DIAN CSF data for ADAD. Proteins and pathways described in (A) are highlighted. The most enriched pathways across three groups are in boldface.

in sporadic AD ($P = 8.3 \times 10^{-9}$) and *TREM2* carriers ($P = 9.3 \times 10^{-6}$). We also found that identified proteins in sporadic AD and *TREM2* carriers were enriched for the innate immune system pathway ($FDR = 2.0 \times 10^{-3}$ for sporadic AD; $FDR = 8.4 \times 10^{-3}$ for *TREM2* carriers). In particular, proteins in sporadic AD were enriched for the neutrophil-mediated immunity ($FDR = 1.6 \times 10^{-6}$). The proteins in this pathway include osteopontin (*SPP1*) and integrin $\alpha 1\beta 1$ (*ITGB1*), among others.

The most strongly enriched pathways for sporadic AD were the programmed cell death ($FDR = 8.6 \times 10^{-9}$) and intrinsic pathway for apoptosis ($FDR = 7.0 \times 10^{-11}$), which included several 14-3-3 proteins (*YWHAE* and *YWHAZ*). This suggests that some of the cell death associated with AD pathogenesis is regulated by apoptosis and not due to necrosis, entosis, ferroptosis, or lysosomal-dependent cell death. Signaling pathways, including the fibroblast growth factor signaling pathway ($FDR = 9.0 \times 10^{-9}$) and epidermal

growth factor (EGF) receptor signaling pathway (FDR = 8.7×10^{-10}), were also among the top pathways for sporadic AD.

The most strongly enriched pathway by *TREM2*-specific proteins was cholecystokinin (CCK) receptor signaling (FDR = 5.4×10^{-7}). This pathway includes *ITGB1*, *PPP3CA*, *YWHA B*, and *MAPK3*. CCK is a satiety hormone that is highly expressed in the brain, including hippocampus. It has been shown that CSF CCK was related to memory scores, higher CSF tau, and p-tau values in the ADNI cohort (24). Other notable pathways related to *TREM2*-specific proteins were amyloid fiber formation (FDR = 2.5×10^{-3}), negative regulation of autophagy (FDR = 5.3×10^{-3}), and cellular response to cytokine stimulus (FDR = 5.7×10^{-3}).

Pathways enriched by ADAD-specific proteins also pointed to biological processes involved in this AD subtype. Notable pathways were proteolysis (FDR = 3.2×10^{-3}), negative regulation of cellular process (FDR = 4.0×10^{-2}), regulation of B cell proliferation (FDR = 2.7×10^{-2}), and regulation of peptidyl-tyrosine phosphorylation (FDR = 3.2×10^{-3}). They include insulin-degrading enzyme (IDE) and macrophage migration inhibitory factor (MIF), a proinflammatory cytokine involved in the innate immune response. IDE is involved in the cellular breakdown of insulin and reported to be involved in the degradation and clearance of naturally secreted A β protein by neurons and microglia (25). These proteins were also enriched for the cytokine-mediated signaling pathway (FDR = 4.1×10^{-2}), indicating that inflammation also plays a role in ADAD.

DISCUSSION

We performed a proteomic characterization of sporadic and genetically defined AD subtypes. In particular, we obtained proteomic measures from Knight ADRC and DIAN cohorts and identified proteomic profiles for sporadic AD, *TREM2* variant carriers, and ADAD cases in brain tissues, CSF, and plasma. These proteomic profiles, replicated in independent datasets and across brain tissue and fluids, were used to create specific prediction models and to identify functional pathways. Our CSF-specific model for sporadic AD provided higher predictive performance than that with CSF p-tau/A β 42 values (AUC = 0.88 to 0.90 versus 0.81; Fig. 2). The predictive performance using plasma proteins was similar to that of CSF p-tau/A β 42. These models were replicated in independent datasets with different protein quantification, indicating that our model is robust, reproducible, and reliable.

In addition, we demonstrated that (i) SOMAscan measurements were in good agreements with mass spectrometry and OLINK measurements, with an overall correlation coefficient of around 0.74 (table S23 and fig. S18); (ii) many proteins identified here showed strong correlation, including SMOC1, α -synuclein (*SNCA*), 14-3-3E (*YWHA E*), GFAP, and ALDOA (correlation >0.81; figs. S12 to S18); and (iii) the correlation of these proteins between SomaLogic and mass spectrometry or Olink was consistent regardless of cohorts and tissue or fluid. Similar findings have been reported recently by Dammer *et al.* (26). That these proteins replicated in multiple mass spectrometry datasets and show a high correlation across platforms not only strengthens our findings but also indicates that the SomaLogic measurements are robust.

We created the prediction models for clinical AD status instead of biomarker-based status because the latter was unavailable in most of the replication cohorts. The biomarker-based AT(N) status provides classification based on core AD pathophysiological features:

the A β pathway (A), aggregated tau pathophysiology (T), and neuronal injury and neurodegeneration (N) (21). The AT(N) classification identifies individuals in the presymptomatic stage of AD. Because not all AT(N)-positive individuals develop cognitive impairment, because some may carry protective factors (27), our prediction models are optimized for symptomatic AD. AT(N)-based models can identify individuals with AD pathology, not those with clinical symptoms. We excluded any samples showing symptoms beyond AD. Because our controls may include presymptomatic individuals, we also performed analyses using CSF p-tau/A β 42 and AT(N) as an outcome. The top proteins from these analyses [with clinical status or CSF p-tau/A β 42 or AT(N)] were the same, and the effect size for all proteins was highly correlated in both CSF and plasma (tables S11 and S12 and fig. S8), indicating that in the Knight ADRC cohort using clinical status or biomarker-defined status would lead to the same results.

Another advantage of using clinical status is the ability to compare our model with CSF p-tau/A β 42-based models (tables S11 and S12 and fig. S8). We demonstrated that our model provided higher performance than p-tau/A β 42 in CSF and comparable performance in plasma. Nakamura *et al.* (28) presented blood-based biomarkers APP₆₆₉₋₇₁₁/A β ₄₂ and A β ₄₀/A β ₄₂ for predicting A β positivity (AUC: 0.88 to 0.96). However, because several U.S. Food and Drug Administration-approved therapeutic targets are developed, non-A β -based biomarkers are needed to monitor target engagement and treatment outcome. Zhang *et al.* (29) identified 11 plasma proteins for predicting A β positivity (AUC = 0.73), 2 proteins for p-tau positivity (AUC = 0.67), and 7 proteins for AT classification (AUC = 0.77) in one cohort. Our model based on the nine plasma proteins provided higher predictive performance (AUC = 0.79) in both discovery and replication cohorts. Palmqvist *et al.* (30) created prediction models with plasma p-tau₂₁₇, NFL, *APOE* genotype, magnetic resonance imaging (MRI), and cognitive tests, providing an AUC of 0.92 in the discovery dataset and 0.86 in the replication dataset. The AUC for p-tau₂₁₇ alone was 0.72 in the discovery dataset and 0.78 in the replication dataset for predicting clinical AD. The prediction power of plasma p-tau₂₁₇ was similar to those using CSF tau and A β 42/A β 40 (7–9). Our plasma models including only basic demographic information (age and sex) along with identified proteins provided comparable AUC (0.82 to 0.79) as those using blood p-tau₂₁₇ (30). Similar to Palmqvist *et al.* (30) and others (7), we considered independent replication cohorts to validate the model. Better prediction models could be constructed by combining our proteomic profiles with MRI and cognitive tests, as demonstrated in previous studies.

In this study, we identified Apo E (*APOE*), calcineurin (*PPP3R1* and *PPP3CA*), and ERK-1 (*MAPK3*) associated with sporadic AD. *APOE* is the strongest and most common genetic risk factor for AD (31), and individuals with the *APOE* ϵ 4 allele have lower CSF A β 42 values (31) and lower A β 42 clearance in brain (32). Genetic variants in *PPP3R1* have been associated with higher CSF p-tau values and earlier age at onset (33). *MAPK3* was also reported to be involved in AD pathology (34), likely by affecting tau phosphorylation. Calcineurin and ERK-1 were recently reported as part of the causal AD pathway by pQTL and Mendelian randomization analyses (35). In addition, we identified α -synuclein and LRRK2 associated with AD. On autopsy, around 30% of AD cases, including ADAD, present with Lewy bodies, which are deposits of α -synuclein (36). With the discovery of pathogenic mutations in *LRRK2* causing

autosomal dominant PD (37), LRRK2 has emerged as a promising target for PD (38). Those reports, together with our analyses, support the notion that there are pathological events and proteins shared across neurodegenerative diseases, specifically in this case for AD and PD. These findings could also help to develop markers to identify AD cases with PD pathology without waiting for autopsy.

The most strongly enriched pathways for sporadic AD were programmed cell death (FDR = 8.6×10^{-9}), the intrinsic pathway for apoptosis (FDR = 7.0×10^{-11}), the EGF receptor signaling pathway (FDR = 8.7×10^{-10}), and the neutrophil-mediated immunity (FDR = 1.6×10^{-6}), among others. The neutrophil-mediated immunity includes osteopontin (*SPP1*) and integrin $\alpha 1b1$ (*ITGB1*). *SPP1* was recently implicated in microglia activation and AD (39). *ITGB1* is a microglia gene and shown to be differentially expressed in the hippocampus and peripheral blood mononuclear cells (PBMCs) in AD (40), important in microglia activation (41), and part of the causal AD pathway (42). Consistent with recent findings that meningeal lymphatics and endothelial-specific proteins affect microglia and AD risk (43), we found several endothelial-specific proteins [ERK-1, SHC1, and basal cell adhesion molecule (BCAM)]. This may be useful for fully understanding how changes in brain endothelial cells and in the blood-brain barrier contribute to the disease. The EGF receptor signaling pathway regulates growth, survival, proliferation, and differentiation in mammalian cells. Dysregulation of this pathway is associated with white matter injury (44) and downstream to A β oligomers (45) and prevents *APOE4* and A β -induced cognitive and cerebrovascular deficits (46), suggesting that developing strategies targeting this pathway may be particularly efficacious once A β pathology is present.

We identified proteins that are altered in *TREM2* risk variant carriers and ADAD cases. *TREM2* is a microglia gene, and activation of microglia results in their production of proinflammatory cytokines such as interleukin-1 (IL-1), IL-6, and tumor necrosis factor- α (TNF- α). Multiple proinflammatory [C-C motif chemokine ligand 21 (CCL21), colony stimulating factor 1 receptor (CSF1R), and IL-24] and anti-inflammatory [IL-1 receptor antagonist (IL1RN) and annexin A1 (ANXA1)] proteins are part of the cellular response to cytokine stimulus, extending the number of proteins that are part of the pathological pathways in *TREM2* carriers. Once individuals with specific genetic profiles are identified, it is possible to create customized prediction models, an instrumental step toward individualized, specific disease risk evaluation and treatment. Integrating proteomic data with other omics (including genomics, transcriptomics, and metabolomics) can help to better understand the complex etiology of AD, further improving the discovery of AD biomarkers and their use in clinical management, as reviewed by Aerqin *et al.* (47). In addition, there may be different downstream, and potentially causal, pathways leading to disease in these individuals. For example, in *TREM2* variant carriers, we found a higher enrichment of the negative regulation of the autophagy pathway (FDR = 5.3×10^{-3}), which is also involved in cell death and seems to play a major role in AD pathogenesis in *TREM2* variant carriers based on our own analyses. Other unique pathways were CCK receptor signaling (FDR = 5.4×10^{-7}) for *TREM2* risk variant carriers and, for ADAD cases, proteolysis (FDR = 3.2×10^{-3}) and regulation of peptidyl-tyrosine phosphorylation (FDR = 3.2×10^{-3}). These findings highlight the role of these biological

processes and their relationship to *TREM2* and ADAD downstream pathological events, which may represent new therapeutic targets.

The strengths of this study are relatively large sample sizes, replication of proteins associated with sporadic AD cases in multiple external datasets generated using orthogonal proteomic methods, and uniquely identifying proteins associated with the genetically defined AD cases. However, there are several limitations. First, several AD biomarkers in CSF and plasma, including A β 42/A β 40 ratio, p-tau217, and p-tau231, which provide good performance, were not available in our samples. Therefore, we were not able to compare our prediction models with the models based on these biomarkers. Second, to replicate our findings, we sought multiple external datasets (Supplementary Materials and Methods). Because the proteomic data available in these studies were generated using a different platform (OLINK and mass spectrometry), not all proteins identified during discovery were assayed. In addition, several replication datasets had smaller sample sizes than our discovery dataset, providing limited power. The replicated proteins therefore likely represent only a subset of proteins associated with AD status. Third, genetically defined AD cases (*TREM2* variant carriers and individuals with ADAD) are extremely rare. There were not enough carriers (four *TREM2* variant carriers in MAYO, seven in MSBB, eight in ROSMAP, and no ADAD cases) in public datasets, and we were unable to pursue replication of the proteins dysregulated in these genetically defined AD cases. Fourth, we performed very stringent QC (Supplementary Materials and Methods). Because of this, not all proteins passed QC across brain tissue, CSF, and plasma. This limited replications and should be accounted for in future studies.

In summary, we identified previously unidentified proteins and pathways implicated in sporadic and genetically defined AD. We also demonstrated that, by leveraging proteins associated with AD status commonly across brain tissue, CSF, and plasma, one could validate and confirm potential new markers for AD. Although additional validation of some of our findings will be needed, these results highlight the need to combine brain tissue, CSF, and plasma proteomics to fully understand the biology of AD and to create prediction models for individuals with AD with specific genetic profiles.

MATERIALS AND METHODS

Study design

The goal of this study was to identify brain-, CSF-, and plasma-specific proteomic profiles for sporadic AD and genetically defined AD cases (Fig. 1). This study included the brain ($n = 360$), CSF ($n = 717$), and plasma ($n = 490$) data from the Knight ADRC (10) and the DIAN (3) cohorts. The recruited individuals were evaluated by Clinical Core personnel of the Knight ADRC (see Supplementary Materials and Methods). Brain samples were obtained from 290 autopsy-confirmed AD cases, 21 *TREM2* risk variant carriers, and 25 cognitively normal individuals with no brain pathology (Table 1). CSF samples were from 176 individuals with a clinical diagnosis of AD, 47 *TREM2* risk variant carriers, and 494 cognitively normal individuals. Plasma samples were from 105 individuals with a clinical diagnosis of AD, 131 *TREM2* risk variant carriers, and 254 cognitively normal individuals. In CSF and plasma data, AD cases had a diagnosis of dementia of the Alzheimer's type (DAT) using criteria equivalent to the National Institute of Neurological and

Communication Disorders and Stroke–Alzheimer’s Disease and Related Disorders Association for probable AD (48). In addition, we obtained the brain samples from 24 individuals carrying ADAD mutations, of whom 18 were from the DIAN cohort. Among these individuals with ADAD, 19, 1, and 4 carried pathogenic mutations in *PSEN1*, *PSEN2*, and *APP*, respectively. The Institutional Review Board of Washington University School of Medicine in St. Louis approved the study, and research was performed in accordance with the approved protocols.

Proteomic data

For omics characterization in brain tissue, CSF, and plasma, we quantified 1305 proteins using a multiplexed, single-stranded DNA aptamer assay developed by SomaLogic (11). The assay covers a dynamic range of 10^8 and measures all three major categories: secreted, membrane, and intracellular proteins. The proteins cover a wide range of molecular functions and include proteins known to be relevant to human disease. Aliquots of gray matter homogenate (150 μ l) of tissue were provided to the Genome Technology Access Center at Washington University in St. Louis for protein measurement. As previously described (11), modified single-stranded DNA aptamers are used to bind specific protein targets, which are then quantified by a DNA microarray. Protein concentrations are quantified as relative fluorescent units (RFU) of intensity in this DNA microarray. We performed standard data processing, normalization, and extensive QC (see Supplementary Materials and Methods).

Validation of SomaLogic data

To further validate our SOMAscan data, we examined the consistencies of protein abundance of the 150 proteins we identified across different platforms in samples from the Emory, ADNI, and DIAN cohorts. In the Emory cohort, 35 samples had CSF and plasma proteomic data measured in SOMAscan, Olink, and mass spectrometry (2). The ADNI and DIAN cohorts had 110 and 457 samples, respectively, which were measured in both SOMAscan and mass spectrometry in CSF (table S23).

Pathway enrichments

Functional enrichment analysis was performed with Enrichr (23). Within each AD subtype, we merged all proteins identified and validated across brain tissue, CSF, and plasma. The genes that target these merged proteins were used as an input for enrichment analysis. This corresponded to 62 genes (targeting 56 externally replicated proteins) for sporadic AD, 37 genes (targeting 33 proteins replicated across brain tissue, CSF, and plasma) for *TREM2* variant carriers, and 14 genes (targeting 14 proteins replicated in independent DIAN CSF data) for ADAD. Among multiple gene set libraries, KEGG (Kyoto Encyclopedia of Genes and Genomes), Reactome, Panther pathways, and GO biological process were considered. The significance of functional enrichment was reported as the *P* value of Fisher’s exact test, followed by Benjamini-Hochberg adjustment for FDR in testing multiple hypotheses. We considered results with $FDR < 0.05$ as significant and included them while creating the dot chart and tile plots to graphically display our findings.

Statistical analysis

To remove batch effects in our proteomic data (17 batches in brain, 50 batches in CSF, and 27 batches in plasma data) and correct for

other unmeasured heterogeneity, we performed SV analysis using the *sva* (12) R package. We used the *sva* function with AD status (variable of interest) and age at measurement (variable to adjust). We obtained the number of resulting SVs as 10, 32, and 14 in brain, CSF, and plasma, respectively, using the *num.sv* function with the default permutation approach.

To obtain proteomic signatures of sporadic AD status and *TREM2* risk variant carrier status, we performed differential abundance analysis using the following linear regression model.

$$\text{Log}_{10}(\text{protein abundance}) = \text{Status} + \text{Age} + \text{Sex} + \text{SVs}$$

The protein abundance was \log_{10} -transformed to follow the normal distribution. Status corresponds to 1 for sporadic AD cases and 0 for cognitively normal individuals in the analysis of sporadic AD cases. In *TREM2* analysis, status corresponds to 1 for *TREM2* variant carriers and 0 for cognitively normal individuals. We also ran an analysis with *TREM2* variant carriers (as 1) and individuals who were diagnosed with AD dementia but did not carry any *TREM2* or autosomal dominant variant (as 0). Age corresponds to the age at death (in brain tissue) or the age at measurement (in CSF and plasma). SVs corresponds to the resulting number of SVs (10 in brain tissue, 32 in CSF, and 14 in plasma).

For ADAD status, we performed analysis using

$$\text{Log}_{10}(\text{protein abundance}) = \text{ADAD status} + \text{Sex} + \text{SVs}$$

where ADAD status corresponds to 1 for ADAD mutation carriers and 0 for cognitively normal individuals with no brain pathology. Age was excluded because it was confounded with ADAD status. ADAD individuals were much younger than the control group (56 versus 88 years old on average). Because of confounding between age and ADAD status, we performed analysis using

$$\text{Log}_{10}(\text{protein abundance}) = \text{Sex} + \text{Sex} + \text{SVs}$$

within control group only. Any proteins showing changes with age in the control group (with nominal $P < 0.05$) were excluded from our ADAD findings. To check whether SVs actually removed batch and covariate effects, we also tested whether the residuals (after SV adjustment) still had batch and covariate effects. We found that all batch effects had been removed after SV adjustment (fig. S4).

To check the sensitivity of our analysis based on SVs, we also performed an analysis using

$$\text{Log}_{10}(\text{protein abundance}) = \text{Status} + \text{Age} + \text{Sex} + \text{PlateID}$$

where PlateID corresponds to the technical batch effects. We found that the effect estimates between the two analyses were very similar in the analysis for sporadic AD status (correlation = 0.61 in brain, 0.90 in CSF, and 0.91 in plasma; fig. S3A). Although they showed less similarity in *TREM2* analysis, the results at the top proteins were more consistent.

To compare brain AD status and neuropathological characteristics, we performed analyses using Braak neurofibrillary tangle scores and CDR at death in brain data while including the same covariates (fig. S6). To compare clinical and biomarker-based AD status in CSF and plasma, we performed analysis using CSF pTau/A β 42 ratio and the corresponding AT classification in both CSF and plasma data (fig. S8). On the basis of CSF A β 42, A β deposition (A+ versus A–) was determined by dichotomization of a

mixture of two Gaussian distributions with the “mclust” package (V 6.0.0) in R. Tau deposition (T+ versus T−) was similarly determined on the basis of CSF pTau. For the analysis with AT classification, we considered individuals with A+T+ (both deposition) or A−T− (no deposition) by excluding A+T− or A−T+.

We examined the consistency between effect sizes of AD status and AD neuropathology in brain (or biomarkers in CSF and plasma) through the scatterplots. We performed correlation tests using `cor.test` in R to test association between effect sizes with Pearson’s product moment correlation coefficient and two-sided alternative hypothesis.

For age at onset for all three tissues, we performed survival analysis while considering age, sex, and SVs as covariates (fig. S5 and table S3). We created a survival object using the R function `Surv` and performed a Cox proportional hazards regression model using the `coxph` function.

To determine a multiple testing correction threshold for the proteomic data in each tissue, we performed an analysis of principal components (PCs). We found that 75, 169, and 230 PCs in brain, CSF, and plasma data, respectively, explain 95% of the variance of proteomic data after QC. We chose a multiple testing–corrected threshold as 0.05 divided by this number of PCs. The thresholds corresponded to 0.67×10^{-4} in brain, 2.96×10^{-4} in CSF, and 2.21×10^{-4} in plasma. When we applied our multiple testing correction and FDR, we found that the use of our thresholds usually provided fewer significant results and is therefore more conservative than the use of FDR (table S26).

Supplementary Materials

This PDF file includes:

Materials and Methods
Figs. S1 to S19

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S26
MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

- 2020 Alzheimer’s disease facts and figures. *Alzheimers Dement.* **16**, 391–460 (2020).
- G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, E. M. Stadlan, Clinical diagnosis of Alzheimer’s disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer’s Disease. *Neurology* **34**, 939–944 (1984).
- R. J. Bateman, C. Xiong, T. L. Benzinger, A. M. Fagan, A. Goate, N. C. Fox, D. S. Marcus, N. J. Cairns, X. Xie, T. M. Blazey, D. M. Holtzman, A. Santacruz, V. Buckles, A. Oliver, K. Moulder, P. S. Aisen, B. Ghetti, W. E. Klunk, E. McDade, R. N. Martins, C. L. Masters, R. Mayeux, J. M. Ringman, M. N. Rossor, P. R. Schofield, R. A. Sperling, S. Salloway, J. C. Morris; Dominantly Inherited Alzheimer Network, Clinical and biomarker changes in dominantly inherited Alzheimer’s disease. *N. Engl. J. Med.* **367**, 795–804 (2012).
- J. D. Ulrich, D. M. Holtzman, TREM2 function in Alzheimer’s disease and neurodegeneration. *ACS Chem. Neurosci.* **7**, 420–427 (2016).
- R. Guerreiro, A. Wojtas, J. Bras, M. Carrasquillo, E. Rogava, E. Majounie, C. Cruchaga, C. Sassi, J. S. Kauwe, S. Younkin, L. Hazrati, J. Collinge, J. Pockock, T. Lashley, J. Williams, J. C. Lambert, P. Amouyel, A. Goate, R. Rademakers, K. Morgan, J. Powell, P. St George-Hyslop, A. Singleton, J. Hardy; Alzheimer Genetic Analysis Group, TREM2 variants in Alzheimer’s disease. *N. Engl. J. Med.* **368**, 117–127 (2013).
- B. A. Benitez, B. Cooper, P. Pastor, S. C. Jin, E. Lorenzo, S. Cervantes, C. Cruchaga, TREM2 is associated with the risk of Alzheimer’s disease in Spanish population. *Neurobiol. Aging* **34**, 1711.e15–1711.e17 (2013).
- B. Olsson, R. Lautner, U. Andreasson, A. Ohrfelt, E. Portelius, M. Bjerke, M. Holtt, C. Rosen, C. Olsson, G. Strobel, E. Wu, K. Dakin, M. Petzold, K. Blennow, H. Zetterberg, CSF and blood biomarkers for the diagnosis of Alzheimer’s disease: A systematic review and meta-analysis. *Lancet Neurol.* **15**, 673–684 (2016).
- S. E. Schindler, J. G. Bollinger, V. Ovod, K. G. Mawuenyega, Y. Li, B. A. Gordon, D. M. Holtzman, J. C. Morris, T. L. S. Benzinger, C. Xiong, A. M. Fagan, R. J. Bateman, High-precision plasma β -amyloid 42/40 predicts current and future brain amyloidosis. *Neurology* **93**, e1647–e1659 (2019).
- Y. Li, S. E. Schindler, J. G. Bollinger, V. Ovod, K. G. Mawuenyega, M. W. Weiner, S. M. Leslie, C. L. Masters, C. J. Fowler, J. Q. Trojanowski, M. Korecka, R. N. Martins, S. Janelidze, O. Hansson, R. J. Bateman, Validation of plasma amyloid- β 42/40 for detecting Alzheimer disease amyloid plaques. *Neurology* **98**, e688–e699 (2022).
- A. M. Fagan, C. M. Roe, C. Xiong, M. A. Mintun, J. C. Morris, D. M. Holtzman, Cerebrospinal fluid tau/beta-amyloid(42) ratio as a prediction of cognitive decline in nondemented older adults. *Arch. Neurol.* **64**, 343–349 (2007).
- L. Gold, D. Ayers, J. Bertino, C. Bock, A. Bock, E. N. Brody, J. Carter, A. B. Dalby, B. E. Eaton, T. Fitzwater, D. Flather, A. Forbes, T. Foreman, C. Fowler, B. Gawande, M. Goss, M. Gunn, S. Gupta, D. Halladay, J. Heil, J. Heilig, B. Hicke, G. Husar, N. Janjic, T. Jarvis, S. Jennings, E. Katilius, T. R. Keeney, N. Kim, T. H. Koch, S. Kraemer, L. Kroiss, N. Le, D. Levine, W. Lindsey, B. Lollo, W. Mayfield, M. Mehan, R. Mehler, S. K. Nelson, M. Nelson, D. Nieuwlandt, M. Nikrad, U. Ochsner, R. M. Ostroff, M. Otis, T. Parker, S. Pietrasiewicz, D. I. Resnicow, J. Rohloff, G. Sanders, S. Sattin, D. Schneider, B. Singer, M. Stanton, A. Sterkel, A. Stewart, S. Stratford, J. D. Vaught, M. Vrkljan, J. J. Walker, M. Watrobka, S. Waugh, A. Weiss, S. K. Wilcox, A. Wolfson, S. K. Wolk, C. Zhang, D. Zichi, Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS ONE* **5**, e15004 (2010).
- J. T. Leek, W. E. Johnson, H. S. Parker, A. E. Jaffe, J. D. Storey, The `sva` package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883 (2012).
- J. C. Morris, The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology* **43**, 2412–2414 (1993).
- E. C. B. Johnson, E. B. Dammer, D. M. Duong, L. Ping, M. Zhou, L. Yin, L. A. Higginbotham, A. Guajardo, B. White, J. C. Troncoso, M. Thambisetty, T. J. Montine, E. B. Lee, J. Q. Trojanowski, T. G. Beach, E. M. Reiman, V. Haroutunian, M. Wang, E. Schadt, B. Zhang, D. W. Dickson, N. Ertekin-Taner, T. E. Golde, V. A. Petyuk, P. L. De Jager, D. A. Bennett, T. S. Wingo, S. Rangaraju, I. Hajjar, J. M. Shulman, J. J. Lah, A. I. Levey, N. T. Seyfried, Large-scale proteomic analysis of Alzheimer’s disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. *Nat. Med.* **26**, 769–780 (2020).
- L. Higginbotham, L. Ping, E. B. Dammer, D. M. Duong, M. Zhou, M. Gearing, C. Hurst, J. D. Glass, S. A. Factor, E. C. B. Johnson, I. Hajjar, J. J. Lah, A. I. Levey, N. T. Seyfried, Integrated proteomics reveals brain-based cerebrospinal fluid biomarkers in asymptomatic and symptomatic Alzheimer’s disease. *Sci. Adv.* **6**, eaaz9360 (2020).
- A. P. Wingo, W. Fan, D. M. Duong, E. S. Gerasimov, E. B. Dammer, Y. Liu, N. V. Harerimana, B. White, M. Thambisetty, J. C. Troncoso, N. Kim, J. A. Schneider, I. M. Hajjar, J. J. Lah, D. A. Bennett, N. T. Seyfried, A. I. Levey, T. S. Wingo, Shared proteomic effects of cerebral atherosclerosis and Alzheimer’s disease on the human brain. *Nat. Neurosci.* **23**, 696–700 (2020).
- C. D. Whelan, N. Mattsson, M. W. Nagle, S. Vijayaraghavan, C. Hyde, S. Janelidze, E. Stomrud, J. Lee, L. Fitz, T. A. Samad, G. Ramaswamy, R. A. Margolin, A. Malarstig, O. Hansson, Multiplex proteomics identifies novel CSF and plasma biomarkers of early Alzheimer’s disease. *Acta Neuropathol. Commun.* **7**, 169 (2019).
- M. Sattler, S. J. Kiddle, S. Newhouse, P. Proitsis, S. Nelson, S. Williams, C. Johnston, R. Killick, A. Simmons, E. Westman, A. Hodges, H. Soininen, I. Kloszewska, P. Mecocci, M. Tsolaki, B. Vellas, S. Lovestone; AddNeuroMed Consortium, R. J. B. Dobson, Alzheimer’s disease biomarker discovery using SOMAScan multiplexed protein technology. *Alzheimers Dement.* **10**, 724–734 (2014).
- A. G. Vlassenko, L. McCue, M. S. Jasele, Y. Su, B. A. Gordon, C. Xiong, D. M. Holtzman, T. L. Benzinger, J. C. Morris, A. M. Fagan, Imaging and cerebrospinal fluid biomarkers in early preclinical Alzheimer disease. *Ann. Neurol.* **80**, 379–387 (2016).
- T. Guo, L. M. Shaw, J. Q. Trojanowski, W. J. Jagust, S. M. Landau; Alzheimer’s Disease Neuroimaging Initiative, Association of CSF A β , amyloid PET, and cognition in cognitively unimpaired elderly adults. *Neurology* **95**, e2075–e2085 (2020).
- C. R. Jack Jr., D. A. Bennett, K. Blennow, M. C. Carrillo, H. H. Feldman, G. B. Frisoni, H. Hampel, W. J. Jagust, K. A. Johnson, D. S. Knopman, R. C. Petersen, P. Scheltens, R. A. Sperling, B. Dubois, A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology* **87**, 539–547 (2016).

22. J. Timsina, D. Gomez-Fonseca, L. Wang, A. Do, D. Western, I. Alvarez, M. Aguilar, P. Pastor, R. L. Henson, E. Herries, C. Xiong, S. E. Schindler, A. M. Fagan, R. J. Bateman, M. Farlow, J. C. Morris, R. Perrin, K. Moulder, J. Hassenstab, J. Chhatwal, H. Mori; Alzheimer's Disease Neuroimaging Initiative; Dominantly Inherited Alzheimer Network Consortia, Y. J. Sung, C. Cruchaga, Comparative analysis of Alzheimer's disease cerebrospinal fluid biomarkers measurement by multiplex SOMAscan platform and immunoassay-based approach. *J. Alzheimer's Dis.* **89**, 193–207 (2022).
23. E. Y. Chen, C. M. Tan, Y. Kou, Q. Duan, Z. Wang, G. V. Meirelles, N. R. Clark, A. Ma'ayan, Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
24. A. Plagman, S. Hoscheidt, K. E. McLimans, B. Klinedinst, C. Pappas, V. Anantharam, A. Kanthasamy, A. A. Willette; Alzheimer's Disease Neuroimaging Initiative, Cholecystokinin and Alzheimer's disease: A biomarker of metabolic function, neural integrity, and cognitive performance. *Neurobiol. Aging* **76**, 201–207 (2019).
25. H. Fu, B. Liu, L. Li, C. A. Lemere, Microglia do not take up soluble amyloid-beta peptides, but partially degrade them by secreting insulin-degrading enzyme. *Neuroscience* **443**, 30–43 (2020).
26. E. B. Dammer, L. Ping, D. M. Duong, E. S. Modeste, N. T. Seyfried, J. J. Lah, A. I. Levey, E. C. B. Johnson, Multi-platform proteomic analysis of Alzheimer's disease cerebrospinal fluid and plasma reveals network biomarkers associated with proteostasis and the matrix. *Alzheimer's Res. Ther.* **14**, 174 (2022).
27. P. Scheltens, B. De Strooper, M. Kivipelto, H. Holstege, G. Chetelat, C. E. Teunissen, J. Cummings, W. M. van der Flier, Alzheimer's disease. *Lancet* **397**, 1577–1590 (2021).
28. A. Nakamura, N. Kaneko, V. L. Villemagne, T. Kato, J. Doecke, V. Dore, C. Fowler, Q. X. Li, R. Martins, C. Rowe, T. Tomita, K. Matsuzaki, K. Ishii, Y. Arahata, S. Iwamoto, K. Ito, K. Tanaka, C. L. Masters, K. Yanagisawa, High performance plasma amyloid- β biomarkers for Alzheimer's disease. *Nature* **554**, 249–254 (2018).
29. Y. Zhang, U. Ghose, N. J. Buckley, S. Engelborghs, G. B. Frisoni, A. Wallin, A. Lleo, J. Popp, P. Martinez-Lage, C. Legido-Quigley, F. Barkhof, H. Zetterberg, P. J. Visser, L. Bertram, S. Lovestone, A. J. Nevado-Holgado, L. Shi, Predicting AT(N) pathologies in Alzheimer's disease from blood-based proteomic data using neural networks. *Front Aging Neurosci.* **14**, 1040001 (2022).
30. S. Palmqvist, P. Tideman, N. Cullen, H. Zetterberg, K. Blennow; Alzheimer's Disease Neuroimaging Initiative, J. L. Dage, E. Stomrud, S. Janelidze, N. Mattsson-Carlgen, O. Hansson, Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. *Nat. Med.* **27**, 1034–1042 (2021).
31. C. Cruchaga, J. S. Kauwe, P. Nowotny, K. Bales, E. H. Pickering, K. Mayo, S. Bertelsen, A. Hinrichs; Alzheimer's Disease Neuroimaging Initiative, A. M. Fagan, D. M. Holtzman, J. C. Morris, A. M. Goate, Cerebrospinal fluid APOE levels: An endophenotype for genetic studies for Alzheimer's disease. *Hum. Mol. Genet.* **21**, 4558–4571 (2012).
32. J. M. Castellano, J. Kim, F. R. Stewart, H. Jiang, R. B. DeMattos, B. W. Patterson, A. M. Fagan, J. C. Morris, K. G. Mawuenyega, C. Cruchaga, A. M. Goate, K. R. Bales, S. M. Paul, R. J. Bateman, D. M. Holtzman, Human apoE isoforms differentially regulate brain amyloid-beta peptide clearance. *Sci. Transl. Med.* **3**, 89ra57 (2011).
33. C. Cruchaga, J. S. Kauwe, K. Mayo, N. Spiegel, S. Bertelsen, P. Nowotny, A. R. Shah, R. Abraham, P. Hollingworth, D. Harold, M. M. Owen, J. Williams, S. Lovestone, E. R. Peskind, G. Li, J. B. Leverenz, D. Galasko; Alzheimer's Disease Neuroimaging Initiative, J. C. Morris, A. M. Fagan, D. M. Holtzman, A. M. Goate, SNPs associated with cerebrospinal fluid phospho-tau levels influence rate of decline in Alzheimer's disease. *PLoS Genet.* **6**, e1001101 (2010).
34. W. Q. Zhao, L. Ravindranath, A. S. Mohamed, O. Zohar, G. H. Chen, C. G. Lyketsos, R. Etcheberrigaray, D. L. Alkon, MAP kinase signaling cascade dysfunction specific to Alzheimer's disease in fibroblasts. *Neurobiol. Dis.* **11**, 166–183 (2002).
35. C. Yang, F. H. G. Farias, L. Ibanez, A. Suhy, B. Sadler, M. V. Fernandez, F. Wang, J. L. Bradley, B. Eiffert, J. A. Bahena, J. P. Budde, Z. Li, U. Dube, Y. J. Sung, K. A. Mihindukulasuriya, J. C. Morris, A. M. Fagan, R. J. Perrin, B. A. Benitez, H. Rhinn, O. Harari, C. Cruchaga, Genomic atlas of the proteome from brain, CSF and plasma prioritizes proteins implicated in neurological disorders. *Nat. Neurosci.* **24**, 1302–1312 (2021).
36. J. M. Ringman, S. Monsell, D. W. Ng, Y. Zhou, A. Nguyen, G. Coppola, V. Van Berlo, M. F. Mendez, S. Tung, S. Weintraub, M. M. Mesulam, E. H. Bigio, D. R. Gitelman, A. O. Fisher-Hubbard, R. L. Albin, H. V. Vinters, Neuropathology of autosomal dominant Alzheimer disease in the national Alzheimer coordinating center database. *J. Neuropathol. Exp. Neurol.* **75**, 284–290 (2016).
37. A. Zimprich, S. Biskup, P. Leitner, P. Lichtner, M. Farrer, S. Lincoln, J. Kachergus, M. Hulihan, R. J. Uitti, D. B. Calne, A. J. Stoessl, R. F. Pfeiffer, N. Patenge, I. C. Carbajal, P. Vieregge, F. Asmus, B. Muller-Myhsok, D. W. Dickson, T. Meitinger, T. M. Strom, Z. K. Wszolek, T. Gasser, Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* **44**, 601–607 (2004).
38. E. Tolosa, M. Vila, C. Klein, O. Rascol, LRRK2 in Parkinson disease: Challenges of clinical trials. *Nat. Rev. Neurol.* **16**, 97–107 (2020).
39. N. Thrupp, C. Sala Frigerio, L. Wolfs, N. G. Skene, N. Fattorelli, S. Poovathingal, Y. Fournier, P. M. Matthews, T. Theys, R. Mancuso, B. de Strooper, M. Fiers, Single-nucleus RNA-seq is not suitable for detection of microglial activation genes in humans. *Cell Rep.* **32**, 108189 (2020).
40. G. Ma, M. Liu, K. Du, X. Zhong, S. Gong, L. Jiao, M. Wei, Differential expression of mRNAs in the brain tissues of patients with Alzheimer's disease based on GEO expression profile and its clinical significance. *Biomed. Res. Int.* **2019**, 8179145 (2019).
41. X. D. Pan, Y. G. Zhu, N. Lin, J. Zhang, Q. Y. Ye, H. P. Huang, X. C. Chen, Microglial phagocytosis induced by fibrillar β -amyloid is attenuated by oligomeric β -amyloid: Implications for Alzheimer's disease. *Mol. Neurodegener.* **6**, 45 (2011).
42. Z. Li, Z. Xiong, L. C. Manor, H. Cao, T. Li, Integrative computational evaluation of genetic markers for Alzheimer's disease. *Saudi J. Biol. Sci.* **25**, 996–1002 (2018).
43. S. Da Mesquita, Z. Papadopoulos, T. Dykstra, L. Brase, F. G. Farias, M. Wall, H. Jiang, C. D. Kodira, K. A. de Lima, J. Herz, A. Louveau, D. H. Goldman, A. F. Salvador, S. Onengut-Gumuscu, E. Farber, N. Dabhi, T. Kennedy, M. G. Milam, W. Baker, I. Smirnov, S. S. Rich; Dominantly Inherited Alzheimer Network, B. A. Benitez, C. M. Karch, R. J. Perrin, M. Farlow, J. P. Chhatwal, D. M. Holtzman, C. Cruchaga, O. Harari, J. Kipnis, Meningeal lymphatics affect microglia responses and anti-A β immunotherapy. *Nature* **593**, 255–260 (2021).
44. J. Scafidi, T. H. Hammond, S. Scafidi, J. Ritter, B. Jablonska, M. Roncal, K. Szigeti-Buck, D. Coman, Y. Huang, R. J. McCarter Jr., F. Hyder, T. L. Horvath, V. Gallo, Intranasal epidermal growth factor treatment rescues neonatal brain injury. *Nature* **506**, 230–234 (2014).
45. L. Wang, H. C. Chiang, W. Wu, B. Liang, Z. Xie, X. Yao, W. Ma, S. Du, Y. Zhong, Epidermal growth factor receptor is a preferred target for treating amyloid- β -induced memory loss. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16743–16748 (2012).
46. R. Thomas, P. Zuchowska, A. W. Morris, F. M. Marottoli, S. Sunny, R. Deaton, P. H. Gann, L. M. Tai, Epidermal growth factor prevents APOE4 and amyloid-beta-induced cognitive and cerebrovascular deficits in female mice. *Acta Neuropathol. Commun.* **4**, 111 (2016).
47. Q. Aeqin, Z. T. Wang, K. M. Wu, X. Y. He, Q. Dong, J. T. Yu, Omics-based biomarkers discovery for Alzheimer's disease. *Cell. Mol. Life Sci.* **79**, 585 (2022).
48. L. Berg, D. W. McKeel Jr., J. P. Miller, M. Storandt, E. H. Rubin, J. C. Morris, J. Baty, M. Coats, J. Norton, A. M. Goate, J. L. Price, M. Gearing, S. S. Mirra, A. M. Saunders, Clinicopathologic studies in cognitively healthy aging and Alzheimer's disease: Relation of histologic markers to dementia severity, age, sex, and apolipoprotein E genotype. *Arch. Neurol.* **55**, 326–335 (1998).

Acknowledgments: We thank all the participants and their families, as well as the many involved institutions and their staff. **Funding:** This work was supported by grants from the NIH [RF1AG074007 (to Y.J.S.), R01AG044546 (to C.C.), P01AG003991 (to C.C. and J.C.M.), RF1AG053303 (to C.C.), RF1AG058501 (to C.C.), and U01AG058922 (to C.C.)] and the Chan Zuckerberg Initiative (CZI). The recruitment and clinical characterization of research participants at Washington University were supported by NIH P30AG066444 (to J.C.M.), P01AG03991 (to J.C.M.), and P01AG026276 (to J.C.M.). This work was supported by access to equipment made possible by the Hope Center for Neurological Disorders, the Neurogenomics and Informatics Center (NGI: <https://neurogenomics.wustl.edu/>), and the Departments of Neurology and Psychiatry at Washington University School of Medicine. The funders of the study had no role in the collection, analysis, or interpretation of data; in the writing of the report; or in the decision to submit the paper for publication. DIAN: Data collection and sharing for this project was supported by the Dominantly Inherited Alzheimer Network (DIAN, U19AG032438) funded by the National Institute on Aging (NIA), the Alzheimer's Association (SG-20-690363-DIAN), the German Center for Neurodegenerative Diseases (DZNE), Raul Carrea Institute for Neurological Research (FLENI), and partial support by the Research and Development Grants for Dementia from Japan Agency for Medical Research and Development, AMED, the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), Spanish Institute of Health Carlos III (ISCIII), Canadian Institutes of Health Research (CIHR), Canadian Consortium of Neurodegeneration and Aging, Brain Canada Foundation, and Fonds de Recherche du Québec Santé. This manuscript has been reviewed by DIAN Study investigators for scientific content and consistency of data interpretation with previous DIAN Study publications. We acknowledge the altruism of the participants and their families and contributions of the DIAN research and support staff at each of the participating sites for their contributions to this study. ADNI: Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). Hence, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf. Data collection and sharing for this project were funded by ADNI (NIH grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging and the National Institute of Biomedical Imaging and Bioengineering and through generous contributions from the following: AbbVie, Alzheimer's Association, Alzheimer's Drug Discovery Foundation, Araclon Biotech, BioClinica Inc., Biogen, Bristol-Myers Squibb Company, CereSpir Inc., Cogstate, Eisai Inc., Elan Pharmaceuticals Inc., Eli Lilly and Company, EuroImmun,

F. Hoffmann–La Roche Ltd. and its affiliated company Genentech Inc., Fujirebio, GE Healthcare, IXICO Ltd., Janssen Alzheimer Immunotherapy Research & Development LLC., Johnson & Johnson Pharmaceutical Research & Development LLC., Lumosity, Lundbeck, Merck & Co. Inc., Meso Scale Diagnostics LLC., NeuroRx Research, Neurotrack Technologies, Novartis Pharmaceuticals Corporation, Pfizer Inc., Piramal Imaging, Servier, Takeda Pharmaceutical Company, and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. AMP-AD: The results published here are in whole or in part based on data obtained from the AD Knowledge Portal (<https://adknowledgeportal.org>). Study data were provided by the Baltimore Longitudinal Study of Aging (BLSA), the Banner Sun Health Research Institute (Banner), the Mount Sinai School of Medicine Brain Bank (MSSB), the Adult Changes in Thought Study (ACT), the Mayo Clinic Brain Bank, and the Religious Orders Study (ROS) and the Memory and Aging Project (MAP). **Author contributions:** Y.J.S. and C.C. designed the study, processed and analyzed the data, and wrote the manuscript. C.Y., F.W., M.J., J.N., A.F., R.J.B., J.C.M., M.R.F., S.E.S., and O.H. acquired Knight ADRC samples and data. J.P.C., P.R.S., H.C., and C.K. acquired DIAN samples and data. E.C.B.J., E.B.D., and N.T.S. acquired and analyzed Emory samples and data. H.S.-H.O., J.E.R., T.W.-C., and O.H. acquired and analyzed Stanford samples and data. Y.J.S., B.N., A.E., and O.H. developed the web portal. All authors read and approved the manuscript. **Competing interests:** C.C. has received research support from GSK and Eisai. C.C. is a member of the advisory board of Vivid Genomics and Circular Genomics and owns stocks. A.F. has received research funding from the National Institute on Aging of the NIH, Biogen, Centene,

Fujirebio, and Roche Diagnostics. She is a member of the scientific advisory boards for Roche Diagnostics and Genentech and also consults for Diadem, DiamiR, and Siemens Healthcare Diagnostics Inc. N.T.S. is a cofounder of Emtherapro Corp. H.R. is an employee of Leal Therapeutics, and participation in this work was funded by Alector. M.R.F. has received funding from AbbVie, Accera, ADCS Posiphen, Eisai, Eli Lilly, Genentech, Novartis, Suven Life Sciences, and vTv Therapeutics. M.R.F. consults for Artery Therapeutics (ATI), Athira, Avanir, AZTherapies, Biogen MA, Cognition Therapeutics, Cyclo Therapeutics, Gemvax, Green Valley, Ionis, Lexeo, Longeveron, McClena, Merck, Nervive, Neurotrope Bioscience, Oligomerix, Otsuka Pharmaceutical, Pinteon, Proclara, Prothema, SToP-AD, and Vaxinity. **Data and materials availability:** All data associated with this study are present in the paper or the Supplementary Materials. Codes to reproduce the main figures and results were deposited in Zenodo under the DOI: 10.5281/zenodo.8028621. Proteomic data from the Knight ADRC participants are available at the NIAGADS and can be accessed at <https://niagads.org/Knight> ADRC-collection under accession no. NG00102: <https://niagads.org/datasets/ng00102>. The summary results using these data are also available to the scientific community through a public web browser: omics.wustl.edu/proteomics. Data generated from the DIAN cohort can be obtained by applying through the DIAN website at <https://dian.wustl.edu/our-research/for-investigators/diantu-investigator-resources/dian-tu-biospecimen-request-form/>.

Submitted 22 April 2022

Resubmitted 17 October 2022

Accepted 13 June 2023

Published 5 July 2023

10.1126/scitranslmed.abq5923

Science Translational Medicine

Proteomics of brain, CSF, and plasma identifies molecular signatures for distinguishing sporadic and genetic Alzheimer's disease

Yun Ju Sung, Chengran Yang, Joanne Norton, Matt Johnson, Anne Fagan, Randall J. Bateman, Richard J. Perrin, John C. Morris, Martin R. Farlow, Jasmeer P. Chhatwal, Peter R. Schofield, Helena Chui, Fengxian Wang, Brenna Novotny, Abdallah Eteleeb, Celeste Karch, Suzanne E. Schindler, Herve Rhinn, Erik C. B. Johnson, Hamilton Se-Hwee Oh, Jarod Evert Rutledge, Eric B. Dammer, Nicholas T. Seyfried, Tony Wyss-Coray, Oscar Harari, and Carlos Cruchaga

Sci. Transl. Med. **15** (703), eabq5923. DOI: 10.1126/scitranslmed.abq5923

View the article online

<https://www.science.org/doi/10.1126/scitranslmed.abq5923>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science Translational Medicine (ISSN 1946-6242) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Translational Medicine* is a registered trademark of AAAS.

Copyright © 2023 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works