Molecular Neurodegeneration

RESEARCH ARTICLE

Open Access

Heparin-enriched plasma proteome is significantly altered in Alzheimer's disease

Qi Guo[1,2,3†], Lingyan Ping[1,2,3†], Eric B. Dammer[1,2,3], Duc M. Duong[1,2,3], Luming Yin[1,2,3], Kaiming Xu[1,2,3], Anantharaman Shantaraman[1,2,3], Edward J. Fox[1,2,3], Todd E Golde[2,4,5], Erik C.B. Johnson[2,3,4], Blaine R. Roberts[1,2,3,4], James J. Lah[2,3,4], Allan I. Levey[2,3,4*] and Nicholas T. Seyfried[1,2,3,4*]

Abstract

Introduction Heparin binding proteins (HBPs) with roles in extracellular matrix assembly are strongly correlated to β-amyloid (Aβ) and tau pathology in Alzheimer's disease (AD) brain and cerebrospinal fluid (CSF). However, it remains challenging to detect these proteins in plasma using standard mass spectrometry-based proteomic approaches.

Methods We employed heparin-affinity chromatography, followed by off-line fractionation and tandem mass tag mass spectrometry (TMT-MS), to enrich HBPs from plasma obtained from AD ($n=62$) and control ($n=47$) samples. These profiles were then correlated to Aβ, tau and phosphorylated tau (pTau) CSF biomarkers and plasma pTau181 from the same individuals, as well as a consensus brain proteome network to assess the overlap with AD brain pathophysiology.

Results Heparin enrichment from plasma was highly reproducible, enriched well-known HBPs like APOE and thrombin, and depleted high-abundant proteins such as albumin. A total of 2865 proteins, spanning 10 orders of magnitude in abundance, were measured across 109 samples. Compared to the consensus AD brain protein co-expression network, we observed that specific plasma proteins exhibited consistent direction of change in both brain and plasma, whereas others displayed divergent changes, highlighting the complex interplay between the two compartments. Elevated proteins in AD plasma, when compared to controls, included members of the matrisome module in brain that accumulate with Aβ deposits, such as SMOC1, SMOC2, SPON1, MDK, OLFML3, FRZB, GPNMB, and the APOE4 proteoform. Additionally, heparin-enriched proteins in plasma demonstrated significant correlations with conventional AD CSF biomarkers, including Aβ, total tau, pTau, and plasma pTau181. A panel of five plasma proteins classified AD from control individuals with an area under the curve (AUC) of 0.85. When combined with plasma pTau181, the panel significantly improved the classification performance of pTau181 alone, increasing the AUC from 0.93 to 0.98. This suggests that the heparin-enriched plasma proteome captures additional variance in cognitive dementia beyond what is explained by pTau181.

†Qi Guo and Lingyan Ping contributed equally to this work.

*Correspondence:
Allan I. Levey
alevey@emory.edu
Nicholas T. Seyfried
nseyfri@emory.edu
Full list of author information is available at the end of the article

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Conclusion These findings support the utility of a heparin-affinity approach coupled with TMT-MS for enriching amyloid-associated proteins, as well as a wide spectrum of plasma biomarkers that reflect pathological changes in the AD brain.

Keywords Alzheimer's disease, Amyloid, Biomarkers, Cerebrospinal fluid, Heparin, Heparan sulfate proteoglycans, Proteomics, Plasma

Background

Alzheimer's disease (AD) is an age driven neurodegenerative disease characterized by the accumulation of two core pathologies, amyloid beta (Aβ) plaques and phosphorylated tau neurofibrillary tangles (NFTs) in the brain, ultimately leading to dementia [1–3]. Notably, both pathologies accumulate over a decade prior to clinical symptoms in a latent preclinical phase of disease, which provides an opportunity for early detection using biomarkers [4]. To this end, progress towards early accurate diagnosis and effective treatments for AD has been mainly focused on the two hallmark pathologies with recent advances in developing assays for Aβ, tau and phosphorylated tau (pTau) species in cerebrospinal fluid (CSF) and plasma [5–7]. Specifically, phosphorylated tau on threonine-181 (pTau181) and on threonine-217 (pTau217) have recently emerged as promising plasma biomarkers for AD, reflecting Aβ deposition as well as subsequent tau aggregation in the brain [5–7]. However, it has been demonstrated that combining multiple proteins in CSF enhances the accuracy and discriminative capability of both pTau and Aβ for dementia [8], which could apply to plasma as well.

Emerging evidence has suggested that Aβ and tau, represent only a fraction of the complex and heterogeneous biology of AD [9, 10]. For example, large-scale bulk RNA-seq and proteomics studies link AD to diverse biological mechanisms beyond Aβ and tau, involving various biochemical pathways and cell types in brain [11–13]. These studies revealed pathophysiological mechanisms such as synapse loss and neuroinflammation linked to immune, vascular, metabolic and extracellular matrix (ECM) dysfunction. Furthermore, integrated analyses across brain and biofluids demonstrated a significant overlap between changes of brain and CSF proteomes in AD, enabling early disease prediction even in the preclinical phase of AD [12, 14–16]. For example, CSF proteomic measurements in autosomal-dominant AD (ADAD) that overlap with brain protein co-expression modules were recently used to define the evolution of AD pathology over a timescale spanning six decades [17]. SMOC1 and SPON1, ECM proteins associated with Aβ plaques, showed elevated levels in CSF almost three decades before symptom onset. Subsequent alterations were observed in synaptic, metabolic, axonal, inflammatory proteins, and finally, reductions in neurosecretory proteins [17]. Similar trends in these biomarkers were observed in late-onset AD (LOAD), where a targeted CSF proteomic panel reflecting diverse brain-based pathophysiology enhanced the ability of Aβ, tau, and pTau in predicting clinical diagnosis, FDG PET, hippocampal volume, and measures of cognitive severity [8]. Notably, in vivo measurements of fibrillary amyloid plaques in the brain of AD patients using the PET ligand florbetapir (AV45) was most strongly associated with SMOC1, further supporting its role as an important surrogate marker of underlying Aβ pathology in brain [8].

In a consensus human brain proteome network, both SMOC1 and SPON1 are hub proteins within module 42 (M42), which was assigned the term 'matrisome', given the collection of ECM-associated proteins and strong enrichment of glycosaminoglycan-binding proteins [11, 18]. This module is most strongly associated with AD neuropathology and cognition within brain network [11], and contained several additional proteins that have previously been identified to be correlated with or directly bind to Aβ [11, 19–21]. This includes amyloid precursor protein (APP), a proteomic surrogate for Aβ deposition in brain, and apolipoprotein E (APOE), the protein product of the AD genetic risk factor *APOE* [22]. Interestingly, levels of the M42 matrisome module were increased in individuals carrying the *APOE ε4* allele, the strongest genetic risk factor for late-onset AD [11]. Furthermore, most M42 proteins are heparan sulfate (HS) or heparin binding proteins (HBPs), including SMOC1, SPON1, MDK, and APOE among others [23–25]. Notably, heparin and HS accelerate the formation of Aβ fibrils [26–28] and matrisome signaling has been associated with *APOE ε4* in mixed cortical cultures [29]. Taken together, these discoveries indicate that HBPs within M42, exhibiting associations with amyloid plaques, are linked to the *APOE ε4* allele, and function as early indicators of AD risk in CSF. Hence, if readily detectable in plasma, members of M42 hold significant potential as biomarkers for amyloid pathology in AD.

Although SMOC1 and SPON1 have been reported to change in AD plasma using antibody- or aptamer-based proteomic technologies [30], members of M42 have been extremely difficult to identify and quantify using mass spectrometry (MS)-based proteomic

approaches in plasma. Much like CSF, human plasma is characterized by a large dynamic range of protein abundance, estimated at 12–13 orders of magnitude [31], in which albumin and other high-abundant proteins can prevent the detection of proteins of interest. However, the concentration of albumin in plasma (~640 μM) is about 200-fold higher than in CSF (~3 μM), which means fractionation methods such as the immunodepletion of albumin and other high-abundant proteins are typically required to enhance the depth [32]. However, our attempts using these approaches have only partially enriched members of M42 in plasma [30]. Therefore, given the shared heparin-binding properties of M42 members, we aimed to capture and quantify M42 matrisome members from plasma using heparin-affinity chromatography followed by MS-based proteomic analysis to enhance the coverage and quantification of these proteins and assess their changes in AD.

Here we describe a heparin-affinity chromatography approach to capture and enrich HBPs from human plasma across normal controls and individuals clinically diagnosed with AD. Collectively we identified over 2800 proteins in the heparin-enriched proteome, spanning 10 orders of magnitude in protein abundance in plasma. We further show that members of M42, including SMOC1 and SPON1, were significantly increased in AD plasma and correlated to CSF levels of amyloid, tau and pTau as well as plasma pTau, suggesting that these proteins are related to AD pathophysiology in both brain and plasma. Finally, we leveraged the consensus brain protein co-expression network and examined the relationship between plasma and brain proteomes. Plasma proteins within certain network modules showed consistent increases or decreases in both the AD brain and plasma, while others displayed a divergent change. In summary, these findings provide strong support for the integration of a heparin enrichment method with MS-based proteomic analysis for identifying a wide spectrum of plasma biomarkers that mirror pathological changes in the AD brain.

Methods

Materials

Primary antibodies used included a mouse monoclonal anti-thrombin antibody (Catalog No. ab17199, Abcam) and a goat polyclonal anti-ApoE antibody (Catalog No. K74180B, Meridian Life Science). Secondary antibodies used were conjugated with either Alexa Fluor 680 (Invitrogen) or IRDye 800 (Rockland) fluorophores for enhanced detection and visualization. Heparin-sepharose (Cytiva, lot#17099801 for Set 1 and lot#17099803 for Set 2) was used to enrich HBPs from plasma samples.

Plasma and CSF samples

All participants providing plasma and CSF samples gave their informed consent following the protocols approved by the Institutional Review Board at Emory University. Comprehensive cognitive assessments, including the Montreal Cognitive Assessment (MoCA), were administered to all patients as part of their evaluation at the Emory Cognitive Neurology Clinic and the Emory Goizueta Alzheimer's Disease Research Center (ADRC). Diagnostic data were sourced from the ADRC (LBC) and the Emory Cognitive Neurology Program (CRIN-NeuCog). Despite the participants being recruited from two studies, the same clinicians classified case–control status in both settings based on clinical criteria, detailed cognitive measures according to National Alzheimer Coordination Center coding guidelines, imaging findings, and CSF biomarker results. Plasma and CSF samples were collected from participants on the same day using standard procedures. These samples were processed and stored in accordance with the 2014 ADC/NIA best practices guidelines. For participants recruited through the Emory Cognitive Neurology Clinic, CSF samples were sent to Athena Diagnostics and assayed for CSF AD biomarkers, including $A\beta_{1\text{-}42}$, tTau, and pTau181, utilizing the INNOTEST assay platform. CSF samples collected from research participants in the ADRC were assayed using the INNO-BIA AlzBio3 Luminex assay. To analyze plasma pTau181 concentrations, EDTA plasma samples were prepared according to manufacturer's instructions from the pTau181 kit v2 (Quanterix Billerica, Massachusetts, USA). Samples were run in a single batch. Plasma was thawed at room temperature for 45 min and then centrifuged at 5000 × g for 10 min. The plasma samples were then diluted four times and measured on the Simoa HDX platform. Mean intra-assay coefficient of variation (CV) were below 10%. In total, a pooled plasma sample and two sets of individual plasma samples were used in the study. Our discovery Set 1 comprised plasma samples from 18 cognitively assessed normal controls and 18 individuals with mild cognitive impairment or AD, whereas replicate Set 2 included plasma samples obtained from 36 cognitively assessed normal controls and 49 individuals with mild cognitive impairment or AD. Out of the 121 samples, 13 overlapped between the two sets, resulting in a total of 108 unique individuals. Further details about demographics, cognitive scores, CSF biomarkers, and pTau181 plasma biomarker levels

Table 1 Characteristics of human subjects used in this study

	Set1, *n*=36			Set2, *n*=85		
	Control, n=18	*AD, n=18*	*p-value**	*Control, n=36*	*AD, n=49*	*p-value**
Age, mean years (SD)	68.6 (9.2)	*65.2 (12.1)*	0.3492	70.4 (8.9)	66.1 (6.6)	0.0125
Sex, *n* (%)			0.7354			0.0884
Male	11 (61.1)	11 (61.1)		11 (30.6)	24 (49.0)	
Female	11 (61.1)	8 (44.4)		25 (69.4)	24 (49.0)	
Race, *n* (%)			0.5023			0.0034
White	11 (61.1)	12 (66.7)		24 (66.7)	45 (91.8)	
Black	7 (38.9)	5 (27.8)		12 (33.3)	4 (8.2)	
American Indian or Alaska Native	0 (0.0)	1 (5.5)		0 (0.0)	0 (0.0)	
Cognition						
MoCA	27.1 (1.6)	13.1 (7.2)	6.28E-09	26.9 (2.0)	16.5 (6.7)	9.19E-14
ApoE Status, *n* (%)			0.029			0.057
APOE 2/2	0 (0.0)	0 (0.0)		0 (0.0)	1 (0.0)	
APOE 2/3	2 (11.1)	0 (0.0)		5 (13.9)	1 (2.0)	
APOE 2/4	1 (5.6)	0 (0.0)		2 (5.5)	1 (2.0)	
APOE 3/3	11 (61.1)	6 (33.3)		18 (50.0)	21 (42.9)	
APOE 3/4	4 (22.2)	4(22.2)		10 (27.8)	14 (28.6)	
APOE 4/4	0 (0.0)	8 (44.5)		1 (2.8)	11 (22.4)	
Biomarkers, mean (SD)#						
CSF $A\beta_{1\text{-}42}$	552.1 (95.0)	223.9 (75.0)	2.86E-13	499.8 (137.7)	313.7 (145.6)	6.02E-08
CSF tTau CSF	41.8 (18.3)	148.5 (70.0)	3.9E-07	54.9 (26.5)	108.6 (43.3)	4.02E-09
pTau181	25.4 (9.0)	58.7 (18.2)	4.99E-08	54.9 (26.5)	61.6 (25.6)	6.72E-10
Plasma pTau181	1.7 (0.7)	4.8 (2.0)	2.94E-06	2.4 (1.0)	3.8 (1.2)	4.48E-07

AD Alzheimer's disease, *SD* Standard Deviation, *pTau181* tau phosphorylated at threonine 181, *tTau* total tau

* Analyses of group differences included Chi-squared test for categorical variables and Student's *t*-test for continuous variables

CSF and plasma biomarker values are in pg/ml

for individuals in both sets of samples are available in Table 1 and Supplemental Tables 4 and 15.

Heparin binding protein enrichment from plasma

The principal experiment of heparin enrichment was conducted in technical triplicates, using 40 µl of pooled neat human plasma per replicate. Initially, 100 µl of heparin-sepharose bead slurry (1:1 w/v) was prepared with 50 µl of beads for each replicate ($n=3$) and binding buffer (50 mM sodium phosphate, pH 7.4). Each slurry was then washed twice with 1 ml of binding buffer. Subsequently, each replicate of neat plasma (40 µl) was diluted with 1 ml of binding buffer to generate diluted plasma (DP) as input and mixed with the heparin-sepharose beads prepared above. The incubation was performed at room temperature for 10 min coupled with rotation. Following the enrichment step, the beads were spun down at 500 × g for 2 min and the supernatant was collected as the heparin-depleted flowthrough (Hp-depleted FT) fraction. The heparin-sepharose beads were then washed twice with 1 ml of binding buffer and resuspended in 1 ml of binding buffer (Hp-enriched fraction). The Hp-enriched fraction was then split into 650 µl for digestion and 350 µl for SDS-PAGE and western blotting. The supernatant was removed before further processing. For Set 1 ($n=36$) and Set 2 ($n=85$) samples, a global pooled standard (GPS) was prepared for each set before enrichment as internal control for enrichment by pooling equal amount of each sample within each set. A nearly identical protocol as described for the pooled plasma was followed for subsequent enrichment of Set 1 and Set 2, using 50 µl of beads (Set 1) and 200 µl of beads (Set 2) correspondingly for each set.

Gel electrophoresis and western blot analysis

Western blotting and Coomassie Blue staining were performed on all three fractions from pooled plasma sample (DP input, $n=3$; Hp-depleted FT, $n=3$; Hp-enriched

fraction, $n=3$) as previously described [33–35]. For Coomassie Blue staining, samples were boiled with 4×Laemmli sample buffer, and an equal volume to 0.2 µl of the neat plasma was loaded from DP inputs and FT fractions onto an SDS-PAGE gel (Invitrogen). To enhance protein visualization, the Hp-enriched fractions were loaded at a fivefold higher amount (equal to 1 µl of the neat plasma) than the input and FT. For the western blotting, an equal volume to 0.25 µl of the neat plasma was loaded from all three fractions. The gels were either stained with Coomassie Blue G250 overnight or semi-dry transferred to nitrocellulose membranes (Invitrogen) using the iBlot2 system (Life Technologies). Subsequently, the membranes were blocked with casein blocking buffer (Sigma B6429) for 30 min at room temperature. They were then probed with two primary antibodies (mouse monoclonal anti-thrombin and goat polyclonal anti-ApoE) at a 1:1000 dilution overnight at 4 °C. On the following day, the membranes were rinsed and incubated with secondary antibodies conjugated to the Alexa Fluor 680 fluorophore (Invitrogen) at a 1:10,000 dilution for one hour at room temperature. After another round of rinsing, the membranes were once again incubated with secondary antibodies conjugated to a second IRDye 800 fluorophore at a 1:10,000 dilution for one hour at room temperature. For Set 1 samples ($n=36$), western blotting was performed on all three fractions from each of the samples as well as GPS as described above. For Set 2 samples ($n=85$), western blotting was performed on all three fractions from only the GPS in triplicates.

Heparin-enriched plasma protein digestion

Sample digestion was carried out on all three fractions for pooled plasma sample (DP input, $n=3$; Hp-depleted FT, $n=3$; Hp-enriched fraction, $n=3$), and only Hp-enriched fraction for Set 1 ($n=36$) and Set 2 ($n=85$) samples. For DP inputs and FT fractions from pooled plasma sample, 25 µl was digested as previously described [36]. For the Hp-enriched fractions, digestion was performed on beads. The supernatant was removed before adding 40 µl of 0.4 M CAA and 8 µl of 0.5 M TCEP with 400 µl of 50 mM ammonium carbonate for reduction and alkylation. Following the same steps as input and FT above [36], 25 mAU of LysC was used for LysC digestion at 37 °C with shaking at 1000 rpm. 10 µg of trypsin was then added the following day for overnight digestion. After trypsin digestion, the digested peptides were acidified and desalted as previously described [36]. Global internal standard (GIS) was prepared for Set 1 and Set 2 separately by pooling 100 µl of the elution from all samples within each set and divided into 900 µl per aliquot. The elution was then dried to completeness via speed vacuum (Labconco).

Label-free mass spectrometry to assess heparin enrichment

Dried peptides from three fractions of pooled plasma sample were reconstituted in peptide loading buffer (0.1% FA, 0.03% TFA, 1% ACN). Using an RSLCnano liquid chromatography (LC) system, approximately 1 µg of peptide was loaded onto an in-house made column (75 µm internal diameter and 50 cm length) packed with 1.9-µm ReproSil-Pur C18-AQ resin (Maisch, Germany) and eluted over a 120-min gradient. Elution was performed at a rate of 300 nl/min with buffer B/buffer (A + B) ratio ranging from 1 to 99% (buffer A, 0.1% FA in water; buffer B, 0.1% FA in 80% ACN). Mass spectrometry was performed with a high-field asymmetric waveform ion mobility spectrometry (FAIMS) Pro-equipped Eclipse Orbitrap mass spectrometer (ThermoFisher) in positive ion mode using data-dependent acquisition (DDA) with 3×1-s top speed cycles and 3 compensation voltages (-40, -60 and -80). Each compensation voltage (CV) top speed cycle consisted of one full MS scan with as many MS/MS events that could fit in the 1-s cycle time. Full MS scans were collected at a resolution of 120 k [350 to 1500 mass/charge ratio (m/z) range, 4×10^{-5} automatic gain control (AGC) target, and 50-ms maximum ion injection time]. All higher-energy collision-induced dissociation (HCD) MS/MS spectra were acquired in the ion trap (1.6 m/z isolation width, 35% collision energy, 1×10^{-4} AGC target, and 35-ms maximum ion time). Dynamic exclusion was set to exclude previously sequenced peaks for 60 s within a 10-ppm (parts per million) isolation window. Only precursor ions with charge states between 2 and 7 were selected for fragmentation.

The raw files of all three fractions from pooled plasma sample ($n=9$) were searched using FragPipe (FP, version 20.0). The FP pipeline for label-free quantification (LFQ) relies on MSFragger [37, 38] (version 3.8) for peptide identification. The peptide search was done against all canonical human proteins downloaded from Uniprot (20,402 total sequencese; accessed 02/11/2019), as well as 51 common contaminants, and all 20,453 reverse sequences (decoys). The prescribed LFQ-MBR workflow in FP was used with parameters specified as follows: precursor mass tolerance was -20 to 20 ppm, fragment mass tolerance of 0.7 Da, mass calibration and parameter optimization were selected, and isotope error was set to 0/1/2. Cleavage type was set to semi-enzymatic. Enzyme specificity was set to strict-trypsin and up to two missing trypsin cleavages were allowed. Peptide length was allowed to range from 7 to 50 and peptide mass from 200 to 5,000 Da. Variable modifications that were allowed in our search included: oxidation on methionine, and N-terminal acetylation on protein. Static modifications included: carbamidomethylation on cysteine.

Peptide-spectrum match (PSM) were validated using Percolator [39]. The false discovery rate (FDR) threshold was set to 1% using Philosopher [40] (version 5.0.0). The peptide and UniprotID-identified protein abundances were quantified using IonQuant [41] (version 1.9.8) for downstream analysis. All raw files, the database, and the FP search parameter settings are provided on https://www.synapse.org/#!Synapse:syn52525880/files/.

To enable protein overlap analysis across the three fractions and conduct gene ontology (GO) analysis for proteins within each fraction, proteins that were detected in at least 2 out of 3 replicates within each fraction were selected. Before performing differential abundance analysis, the data from all nine samples were merged and protein levels were first scaled by dividing each protein intensity by intensity sum of all proteins in each sample followed by multiplying by the maximum protein intensity sum across all nine samples. Instances where the intensity was '0' were treated as 'missing values'. Subsequently, Perseus-style imputation was applied to these missing values for further analysis [42].

Tandem mass tag (TMT) labeling of peptides

TMT labeling of digested Set 1 and Set 2 samples was performed as previously described [36]. The discovery dataset (Set 1) was comprised of 36 individual samples and 3 GIS randomized based on age, sex and diagnosis into three batches and the replication dataset (Set 2) was comprised of 85 individual samples and 5 GIS randomized into five batches. Supplemental Table 4 and 15 provides the sample-to-batch arrangement for each of these samples. These samples were then labeled using TMTpro kits (Thermo Fisher Scientific, A44520, Lot number: VH3111511 for Set 1; UK297033 for Set 2 with XB338618 for channels 134C and 135). First, each peptide digest was resuspended in 75 ml of 100 mM triethylammonium bicarbonate (TEAB) buffer, and 5 mg of TMT reagent was dissolved into 200 ml of anhydrous acetonitrile (ACN). After that, 15 ml of TMT reagent solution was subsequently added to the resuspended peptide digest and incubated for 1 h at room temperature. Following that, the reaction was quenched with 4 ml of 5% hydroxylamine (Thermo Fisher Scientific, 90,115) for 15 min. Then, the peptide solutions were combined according to the batch arrangement. Finally, each TMT batch was desalted with 60 mg HLB columns (Waters) and dried via speed vacuum (Labconco).

Mass spectrometry analysis for TMT labeled samples

High-pH off-line fractionation was performed on TMT labeled dried peptides as previously described [36]. A total of 96 individual equal volume fractions were collected across the gradient per TMT batch and dried via speed vacuum (Labconco). An equal volume of each high-pH peptide fraction was initially resuspended in loading buffer (0.1% FA, 0.03% TFA, and 1% ACN) and was loaded and separated using an EASY-nanoLC system on a self-made 15-cm-long, 150-μM internal diameter (ID) fused silica column packed with 1.9-μm ReproSil-Pur C18-AQ resin from Maisch, Germany. Elution was carried out over a 20-min gradient at a flow rate of 1200 nL/min, with buffer B/buffer (A + B) ratio ranging from 1 to 99%. For Set 1, mass spectrometry was performed on a high-field asymmetric waveform ion mobility spectrometry (FAIMS) Pro-equipped Orbitrap Lumos (Thermo) in positive ion mode. DDA was used with 2 × 1-s top speed cycles and 2 compensation voltages (-45 and -65). Each top speed cycle included one full MS scan with as many MS/MS events as possible within the 1-s cycle time. Full MS scans were acquired at a resolution of 120,000 over a m/z range of 410 to 1600, with an AGC target of 4×10^{-5} and a maximum ion injection time of 50 ms. All HCD MS/MS spectra were obtained at a resolution of 50,000, with a 0.7 m/z isolation width, 35% collision energy, 1×10^{-5} AGC target, and a maximum ion time of 86 ms. Dynamic exclusion was set to exclude previously sequenced peaks for 20 s within a 10-ppm (parts per million) isolation window. Only precursor ions with charge states between 2 and 6 were selected for fragmentation. For Set 2, elution was performed over a 30-min gradient at a flow rate of 1250 nl/min, with buffer B/buffer (A + B) ratio ranging from 1 to 99% (buffer A: 0.1% FA in water; buffer B: 0.1% FA in 80% ACN). Mass spectrometry was conducted on a high-field asymmetric waveform ion mobility spectrometry (FAIMS) Pro-equipped Orbitrap Eclipse (Thermo) in positive ion mode. DDA employed 2 × 1.5-s top speed cycles and 2 compensation voltages (-45 and -65). Each top speed cycle comprised one full MS scan with as many MS/MS events as possible within the 1.5-s cycle time. Full MS scans were acquired at a resolution of 60,000 over an m/z range of 410 to 1600, with an AGC target of 4×10^{-5} and a maximum ion injection time of 50 ms. All HCD MS/MS spectra were collected at a resolution of 30,000, with TurboTMT on, a 0.7 m/z isolation width, 35% collision energy, 250% normalized AGC target, and a maximum ion time of 54 ms. Dynamic exclusion was configured to exclude previously sequenced peaks for 20 s within a 10-ppm isolation window. Precursor ions with charge states between 2 and 6 were selectively chosen for fragmentation.

Database search parameters for TMT labeled samples

FP (version 18.0) was used to search both the discovery (Set 1) and replication (Set 2) datasets as essentially described [37, 43]. First, mzML files were generated from the original MS.raw files (96 raw files/fractions

per batch) of both Set 1 (3 TMT16 batches) and Set 2 (5 TMT18 batches) using the ProteoWizard MSConvert tool (version 3.0) with options including: 'Write index', 'TPP compatibility', and 'Use zlib compression', as well as "peakPicking" filter setting. Then all 96×8 mzML files from both sets were searched together using MSFragger (version 3.5). The human proteome database used comprised of 20,402 sequences (Swiss-Prot, downloaded 2/11/2019) and their corresponding decoys, including common contaminants. We also included the APOE2 and APOE4 coding variant: CLAVYQAGAR (APOE2), and LGADMEDVR (APOE4). Briefly, search settings included: Precursor mass tolerance was -20 to 20 ppm, the fragment mass tolerance was set to 20 ppm, mass calibration and parameter optimization were selected, and the isotope error was set to -1/0/1/2/3. The enzyme specificity was set to strict-trypsin and up to two missed cleavages allowed. Cleavage type was set to semi-enzymatic. Peptide length was allowed in the range from 7 to 35 and peptide mass from 200 to 5,000 Da. Variable modifications that were allowed in our search included: oxidation on methionine, N-terminal acetylation on protein, TMTpro modifications on serine, threonine and histidine as described [44], with a maximum of 3 variable modifications per peptide. Static modifications included: isobaric TMTpro (TMT16) modifications on lysine and the peptide N-termini as well as carbamidomethylationon of cysteine. MSFragger search results were processed using Percolator [39] for PSM validation, followed by Philosopher [40] for protein inference (using ProteinProphet [45]) and FDR filtering. The reports of the quantified peptides and UniprotID-identified proteins with FDR < 1% were generated.

All raw files from the consensus brain dataset [11], including 1080 raw files generated from 45 TMT 10-plexes for the Religious Orders Study and Memory and Aging Project (ROSMAP) BA9 tissues; 624 raw files generated from 26 TMT 11-plexes for ROSMAP BA6/BA37 tissues; 528 raw files generated from 22 TMT 11-plexes for the Banner Sun Health Research Institute (Banner) tissues; and 760 raw files generated from 20 TMT 11-plexes for Mount Sinai tissues, were re-searched using FragPipe (version 20.0) on the same Uniprot database as described above with slight modifications. This mainly included use of the first generation "TMT10" workflow rather than the newer generation TMTpro workflow and the cleavage type was set to enzymatic. Peptide length was allowed to range from 7 to 50. Variable modifications that were allowed in our search included: oxidation on methionine, N-terminal acetylation on protein, TMT10 modifications on serine, with a maximum of 3 variable modifications per peptide. Static modifications included: isobaric TMT10 modifications on lysine and peptide N-termini as well as carbamidomethylation on cysteine. All raw files, the database, the sample to TMT channel information, and the FP search parameter settings are provided on https://www.synapse.org/#!Synapse:syn52525880/files/.

Data normalization and variance correction for TMT labeled samples

The FP outputs of 8 batches from Set 1 and Set 2 or 113 batches from the consensus brain were integrated to get a combined raw abundance file. Protein levels were first scaled by dividing each protein intensity by the sum of all the reporter ion intensities of the TMT channel (each sample) followed by multiplying by the maximum channel-specific protein intensity sum. Proteins with more than 50% of missing values in each analysis were removed from the matrix prior to the further process (no imputation of missing values was performed). A tunable median polish approach (TAMPOR) was used to adjust technical batch variance within each dataset as previously described [46]. The algorithm is fully documented and available as an R function, which can be downloaded from https://github.com/edammer/TAMPOR. Following this, non-parametric bootstrap regression for batch only or for age, sex, race and batch, within each dataset was performed for Set 1 and Set 2. For reprocessed brain dataset, we restricted our analysis to 456 samples of control, asymptomatic AD (AsymAD) and AD from the Banner (control = 26, AsymAD = 57, AD = 77, 22 TMT11 batches) and the ROSMAP (control = 75, AsymAD = 127, AD = 94, 36 TMT10 batches), and regressed for age, sex, postmortem interval (PMI) and batch.

Depletion of high-abundant proteins from plasma

Other than heparin enrichment, depletion of high-abundant proteins was also performed on Set 1 samples ($n = 36$) prior to digestion, using the High Select Top14 Abundant Protein Depletion Resin (Thermo Fisher Scientific, A36372BR). The resin slurry was equilibrated to room temperature, and 500 µl of resin was aliquoted into each well of a filter plate (Nunc, 278011), with the bottom sealed using a silicone sealing mat (AXYGEN, AM-96-PCR-RD). Once the resin settled at the bottom of the spin column, 8 µl of each sample was added, and depletion was performed by gentle rotation for 15 min at room temperature. After incubation, the seal was removed, followed by centrifugation at 1000×g for 2 min. The sample flow-through was collected in a 1 ml deep-well plate, resulting in approximately 350 µl of each immunodepleted sample. Protein digestion, TMT labeling (Thermo Fisher Scientific, A44520, Lot number: YA357799), MS

analysis, database searching, and data normalization were performed in the same way as the Hp-enriched Set 1 described above.

Proteome coverage overlap and gene ontology (GO) enrichment analysis

All proteome overlap was visualized using the venneuler R package (v1.1–3) *venneuler* function. All functional enrichment was determined using the GOparallel function as documented on https://github.com/edammer/GOparallel. The Bader lab monthly updated.GMT formatted ontology gene lists [47] were used for retrieving GO annotation. Z-score and *p*-value from the one-tailed Fisher's exact test (FET) followed by Benjamini-Hochberg (BH) false discovery rate (FDR) correction was used to assess the significance. A cutoff of Z-score > 1.96 (BH FDR corrected $p < 0.05$ and a minimum of five genes per ontology) was used as filter prior to pruning the ontologies.

Protein differential abundance and hierarchical clustering

All differential abundance was presented as volcano plots that were generated with the ggplot2 package in R v4.2.1. Pairwise differentially abundant proteins were identified using Student's *t*-test, followed by Benjamini-Hochberg (BH) false discovery rate (FDR) correction. Supervised clustering analysis on differentially abundant Hp-enriched plasma proteins was performed with the R NMF package [48] in R v4.2.1. A cutoff of BH FDR-corrected $p < 0.0005$ was used to obtain 82 highly significant proteins and clustered with euclidian distance metric, complete linkage method using the *hclust* function is called from the NMF package *aheatmap* function.

Correlation across platforms and replicate datasets

To evaluate the consistency across different platforms, we compared the heparin-enriched TMT-MS-analyzed (Heparin-MS) data with plasma measurements obtained using immunodepletion, SomaScan® aptamer-based technology from SomaLogic (located in Colorado, USA) and proximity extension assay (PEA) technology from Olink® (based in Uppsala, Sweden). Immunodepletion was performed on the same 36 samples in Set 1 as described above. Data from SomaScan and PEA (referred to as "Olink" henceforth) were obtained for 35 (control = 18, AD = 17) out of the 36 individuals in Set 1 previously assessed using the Heparin-MS method. These data are accessible from a prior publication [30] and were subjected to cross-platform analysis using median-centered normalization [46]. Additionally, we performed correlation analyses between the two Heparin-MS sets (Set 1 and Set 2) using the $\log_2$ fold-change (AD vs Control) values. These correlation analyses were carried out and visualized employing the *verboseScatterplot* function from the R WGCNA package, utilizing the Pearson correlation coefficient and Student's *p*-value to determine the statistical significance of these correlations.

Meta-analysis and correlation assessment of protein abundance with AD-related traits

For the (Set 1 + Set 2) *meta*-analysis, 12 non-overlapping cases that did not meet our enforced CSF biomarker criteria (tTau/$A\beta_{1\text{-}42}$ ratio > 0.226 for AD) or MoCA cutoffs (AD ≤ 24, Control ≥ 24) [49] at the time of lumbar puncture were removed, resulting in 109 samples (96 unique cases). *Meta*-analysis of significance on the combined selected samples ($n = 109$) was performed on 2865 proteins using the R survcomp package *combine.test* function to calculate *meta p*-value. Average $\log_2$ fold-change (AD vs Control) between the two sets was used for the x-axis. All proteins along with corresponding *meta p*-value and fold-change were listed in Supplemental Table 21. Variance-corrected protein abundance of all samples was then Z-transformed by subtracting the mean across 109 samples followed by dividing by standard deviation. Correlations and the Student's *p*-value for their significance between Z-scores and immunoassay measures of AD-related traits, including cognition (MoCA score), CSF $A\beta_{1\text{-}42}$, CSF tTau, CSF pTau181, CSF tTau/$A\beta_{1\text{-}42}$ ratio and plasma pTau181, were calculated using the R WGCNA package *corAndPvalue* function. Same correlations performed for SomaScan ($n = 35$ samples) and Olink ($n = 35$ samples) were listed in Supplemental Table 24–25.

ROC analysis of plasma ptau181 and hp-enriched plasma proteins for AD classification

Receiver-operating characteristic (ROC) analysis was performed on plasma pTau181 and 536 significant Hp-enriched plasma proteins (*meta p*-value < 0.0001) with no missing values across 34 control (CSF tTau/$A\beta_{1\text{-}42}$ < 0.226) and 61 AD (CSF tTau/$A\beta_{1\text{-}42}$ > 0.226) individuals (duplicate controls were removed from the analysis). The algorithm was run in R (version 4.3.1) with a generalized linear model binomial fit of selected protein measurements (individual protein or muti-protein panels) to the binary classification (AD vs Control), using the pROC package implementing ROC curve plots, calculations of AUC and DeLong 95% confidence interval [50]. The significance of difference between two correlated AUCs were calculated with DeLong's test within pROC package [51]. Additional ROC curve characteristics including sensitivity, specificity, and accuracy were calculated with the reportROC R package and listed in Supplemental Table 27.

Protein re-assignment to consensus AD brain network modules

We reprocessed 456 of the Banner and ROSMAP brain samples used in our previously published WGCNA consensus brain network from 2022 [11], which underwent re-analysis with Fragpipe (FP) as outlined above. 8956 proteins were identified and the biweight midcorrelation (bicor) of each was calculated to the 44 eigenproteins of the original network, and a module assignment was made for the FP output proteins to the module with the highest positive correlation, if greater than or equal to 0.30, otherwise being assigned as "grey" [52]. As described previously, a module eigenprotein represents the principal component of all proteins within a module. Moreover, we conducted bicor correlations and Student's *p*-value for their significance to assess the association of module eigenproteins with Consortium to Establish a Registry for Alzheimer's Disease (CERAD), Braak staging, and Mini-mental state examination (MMSE), using the *bicorAndPvalue* function from the R WGCNA package.

Cell type enrichment analysis

As outlined previously, cell type enrichment for each module was performed by cross-referencing the corresponding gene symbols of each module with cell-type–specific gene lists derived from previously published RNA-seq data [11, 53]. Significance of cell type enrichment within each module was then determined using a one-tailed FET and corrected for multiple comparisons by the BH FDR method. The algorithm is fully documented and available as an R function, which can be downloaded from https://www.github.com/edammer/CellTypeFET.

Over-representation analysis of differentially abundant plasma proteins in brain network

Hp-enriched plasma proteins that overlap with brain proteome and significantly altered in AD plasma compared to controls (Student's $p < 0.05$) were assessed for over-representation in brain proteome using a one-tailed FET, and those modules with Benjamini-Hochberg (BH) false discovery rate (FDR)-corrected $p < 0.05$ were considered significant. The background for this overrepresentation analysis comprised 8956 UniprotID-identified proteins from consensus brain [11] which was re-searched by FP as described above.

Results

Heparin binding protein enrichment enhances plasma proteome coverage

To assess the feasibility and efficacy of a heparin enrichment strategy from plasma, an equal volume (40 µl) of pooled human plasma, diluted into binding buffer, was introduced to heparin-sepharose resin in technical triplicates as described in our workflow (Fig. 1A). Before performing MS analyses, fractions including diluted plasma (DP) inputs, the heparin-depleted flow-through (Hp-depleted FT), and the heparin-enriched (Hp-enriched) fraction underwent gel electrophoresis and Coomassie Blue staining to visualize the proteins. These fractions were also prepared for immunoblotting to detect known HBPs, thrombin and APOE (Fig. 1B). Immunoblotting demonstrated that heparin-affinity chromatography was effective at enriching APOE and thrombin, which exhibited increased levels in the Hp-enriched fractions compared to the inputs and FT. Moreover, by Coomassie Blue staining we observed a reduction in the amount of albumin (ALB, ~ 66 kDa) in the Hp-enriched fractions compared to the inputs and FT. In support of these findings, single-shot label-free MS proteomic analyses revealed a notable disparity in the base peak chromatogram among the inputs, FT, and Hp-enriched fractions (Fig. 1C). In the chromatogram spanning the 120-min time window, the pattern of precursor peptide ion peak intensities in the input and FT fractions remained remarkably consistent, primarily dominated by a few prominent feature peaks, which is consistent with the presence of highly abundant proteins such as albumin. However, the base peak chromatogram for the Hp-enriched fractions exhibited much greater complexity, suggesting the presence of additional peptide species in the sample and a decrease in the dominance of the albumin peptides as compared to those observed in the input and FT fractions. Consistently, following database search, we observed a significant reduction (~ 2.6 fold) in the number of peptide spectral counts, a semi-quantitative measure of protein abundance, for albumin in the Hp-enriched fractions compared to the inputs and FT. Reciprocally, we observed a significant increase in the number of peptide spectral counts for thrombin (~ 3.3 fold) and APOE (~ 3.4 fold) in the Hp-enriched fractions compared to the inputs and FT, confirming our immunoblot findings (Fig. 1D). Collectively, these observations suggest that the heparin-affinity method not only enhanced the concentration of the HBPs, but also concurrently reduced the presence of highly abundant proteins like albumin in the Hp-enriched fraction. To assess the consistency of the Hp-enriched MS proteomic approach, measurements of proteins with no missing values across all three replicates within each fraction were listed in Supplemental Table 1 along with their coefficient of variation (CV). All three fractions show an average CV of ~ 25%, which is consistent with the technical CVs reported for label-free proteomic assays [54, 55]. To evaluate the depth of protein identification across the fractions, we assessed the overlap of proteins that were detected in at least

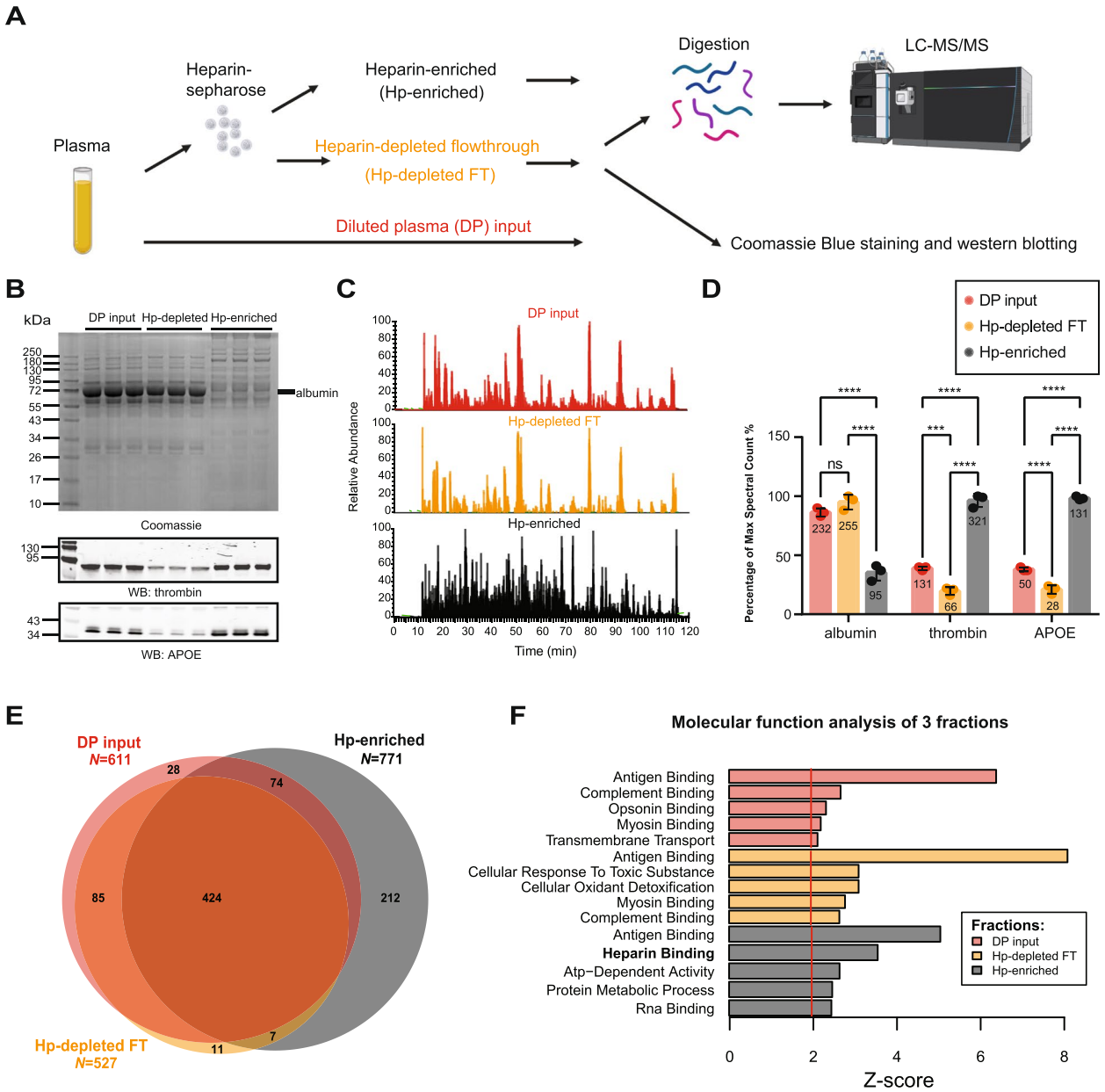

Fig. 1 Heparin enrichment of the plasma proteome. **A** Step-by-step process of heparin enrichment and subsequent mass spectrometry (MS) analysis. Plasma samples were subjected to heparin enrichment, yielding three distinct fractions: DP input ($n=3$), Hp-depleted FT ($n=3$), and Hp-enriched fraction ($n=3$). Each fraction underwent either Coomassie Blue staining, western blotting, or trypsin digestion before subsequent label-free MS analysis utilizing an Orbitrap Eclipse mass spectrometer. **B** Coomassie Blue staining and western blotting was performed in triplicates for all fractions. A reduction of the amount of albumin (~66 kDa) in the Hp-enriched fraction compared to the DP input and FT was illustrated by Coomassie Blue staining. Enrichment was also determined by western blotting for thrombin (~75 kDa) and APOE (~34 kDa) in the Hp-enriched fraction. **C** MS base peak chromatograms of each fraction reveal the variation in sample complexity. The Hp-enriched fraction exhibits a notable increase in sample complexity compared to the DP input and Hp-depleted FT fractions, reflecting successful enrichment of low-abundant proteins in the Hp-enriched fraction. **D** Total number of peptide spectral counts (reported as the maximum percentage) for albumin, thrombin, and APOE in each fraction. Actual spectral counts are labeled within columns. Albumin is significantly depleted from the Hp-enriched fraction compared to the DP input and FT. Conversely, thrombin and APOE are significantly enriched in this fraction. ANOVA with Tukey post-hoc correction was used to determine the *p*-values (* $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$). **E** The number of proteins identified in each fraction: DP input ($N=611$), Hp-depleted FT ($N=527$), and Hp-enriched fraction ($N=771$), and the degree of protein overlaps among them. F) GO terms for proteins detected in each fraction, using 841 total proteins measured in at least 2 out 3 replicates within each fraction as background, highlighting the enrichment of proteins associated with the "Heparin binding" molecular function in the Hp-enriched fraction. Z-score > 1.3 ($p<0.05$) is significant. WB, western blotting; LC–MS, liquid chromatography-mass spectrometry

two out of three replicates of each fraction (Fig. 1E and Supplemental Table 2). Notably, the Hp-enriched fraction identified the greatest number of proteins ($N=771$) compared to the input ($N=611$) and FT ($N=527$). Approximately 29% (221 out of 771) of the proteins in the Hp-enriched fraction were unique to this fraction. Gene ontology (GO) analysis conducted for the proteins identified in each fraction showed significant enrichment (Z-score > 1.3) of the term associated with 'Heparin binding' in the Hp-enriched fraction compared to those in the FT and inputs (Fig. 1F).

To comprehensively evaluate the extent of heparin enrichment, we utilized the protein abundance values based on protein signal intensities from both the DP inputs ($n=3$) and the Hp-enriched fractions ($n=3$), where we applied a significance threshold of $p<0.05$ and a fold-change threshold of > 2 across 821 proteins. In total, 518 proteins exhibited significant changes in their abundance levels across the two fractions (Supplemental Fig. 1A and Supplemental Table 3), with 338 increased and 180 decreased. Among these, we consistently observed the enrichment of key HBPs, such as thrombin (F2), APOE, and APP, which possess a heparin binding domain [56]. Additionally, other neurodegeneration related proteins, such as neurogranin (NRGN) and valosin-containing protein (VCP), also exhibited significant enrichment in the Hp-enriched fractions. As expected, albumin (ALB) was significantly decreased in the Hp-enriched fractions in addition to other highly abundant proteins decreased, including transferrin (TF), A2M, ORM, SERPINA1, and APOA1. Notably, *APOE* has three genetic alleles (*ε 2, ε 3, ε 4*) associated with AD risk: *APOE4* has the highest risk, *APOE2* the lowest risk, and *APOE3* the intermediate risk [22]. *APOE4* and *APOE2* allelic variants can be identified through coding changes that result in variant specific peptides following trypsin digestion [57]. In the pooled human plasma sample used here, all three APOE protein variants were detected and quantified. Interestingly, APOE4 showed the most enrichment in Hp-enriched fractions (15.5-fold, $p=0.0003$), followed by APOE3 (14.7-fold, $p=0.0004$), which was inferred due to a lack of variant specific peptides, and APOE2 (7.4-fold, $p=0.006$) (Supplemental Fig. 1B). This aligns with prior research, highlighting differences in APOE heparin affinity among the variant isoforms (APOE4 > APOE3 > APOE2) [58], suggesting that the genetic susceptibility to AD attributed to APOE may be related to biological differences in binding to HS. Collectively, our findings support successful enrichment of not only APOE, but other HBPs from human plasma.

The heparin-enriched plasma proteome is significantly altered in AD

After successfully demonstrating enrichment of HBPs from human plasma, we aimed to both enhance the depth and determine the differences in the plasma proteome between AD ($n=18$) and control ($n=18$) individuals using discovery Set 1 (Supplemental Table 4). Following heparin-affinity enrichment, we employed tandem mass tag mass spectrometry (TMT-MS) in conjunction with high-pH off-line fractionation to identify a total of 3284 proteins (Fig. 2A and Supplemental Table 5). AD diagnoses were established based on notable cognitive impairment as determined by the Montreal Cognitive Assessment (MoCA), with scores averaging 13.1 ± 7.2 for AD and 27.1 ± 1.6 for controls (Table 1). These diagnoses were further supported by the presence of low Aβ levels and elevated tTau and pTau levels detected in the CSF immunoassays for AD and normal levels of these biomarkers in controls (Table 1). Furthermore, we measured pTau181 in the plasma samples which showed significant differences between control and AD cases (Fig. 2B). Prior to subjecting the samples to TMT-MS, we confirmed the heparin enrichment of APOE and thrombin via

(See figure on next page.)

Fig. 2 Heparin-enriched plasma proteome is significantly altered in AD. **A** Plasma samples were collected from both the control group ($n=18$) and individuals with AD ($n=18$). These samples underwent heparin-sepharose enrichment, followed by trypsin digestion and TMT labeling. Subsequently, high-pH off-line fractionation and LC–MS/MS analysis were conducted using an Orbitrap Lumos mass spectrometer. **B** Measurements of pTau181 across all 36 samples were displayed. Statistical significance was determined by Student's *t*-test ($p=2.94\times10^{-6}$). **C** The volcano plot illustrates the differential abundance of 2,077 proteins between the control and AD groups. The x-axis represents the $\log_2$ fold-change (AD vs CTL), while the y-axis represents the Student's *t*-statistic ($-\log_{10}p$-value) calculated for all proteins in the pairwise group. Proteins significantly increased in AD ($N=579$) are highlighted in red ($p<0.05$), whereas those significantly decreased in AD ($N=661$) are depicted in blue. Grey dots represent proteins with insignificant changes. **D** Top GO terms of the 579 increased (red) or 661 decreased (blue) proteins in AD measured in Set 1 considering the background of 2077 proteins in the plasma proteome. Three GO terms with the highest Z-scores within the domains of biological process, molecular function, and cellular components are presented. **E** A supervised cluster analysis was conducted across the control and AD plasma of discovery samples (Set 1), employing the 82 most significantly altered proteins in the dataset (BH FDR-corrected $p<0.0005$). Other AD-related traits for each sample are also presented by color scale on the top. pTau, phospho-tau; tTau, total tau; CTL, control; LC–MS, liquid chromatography-mass spectrometry

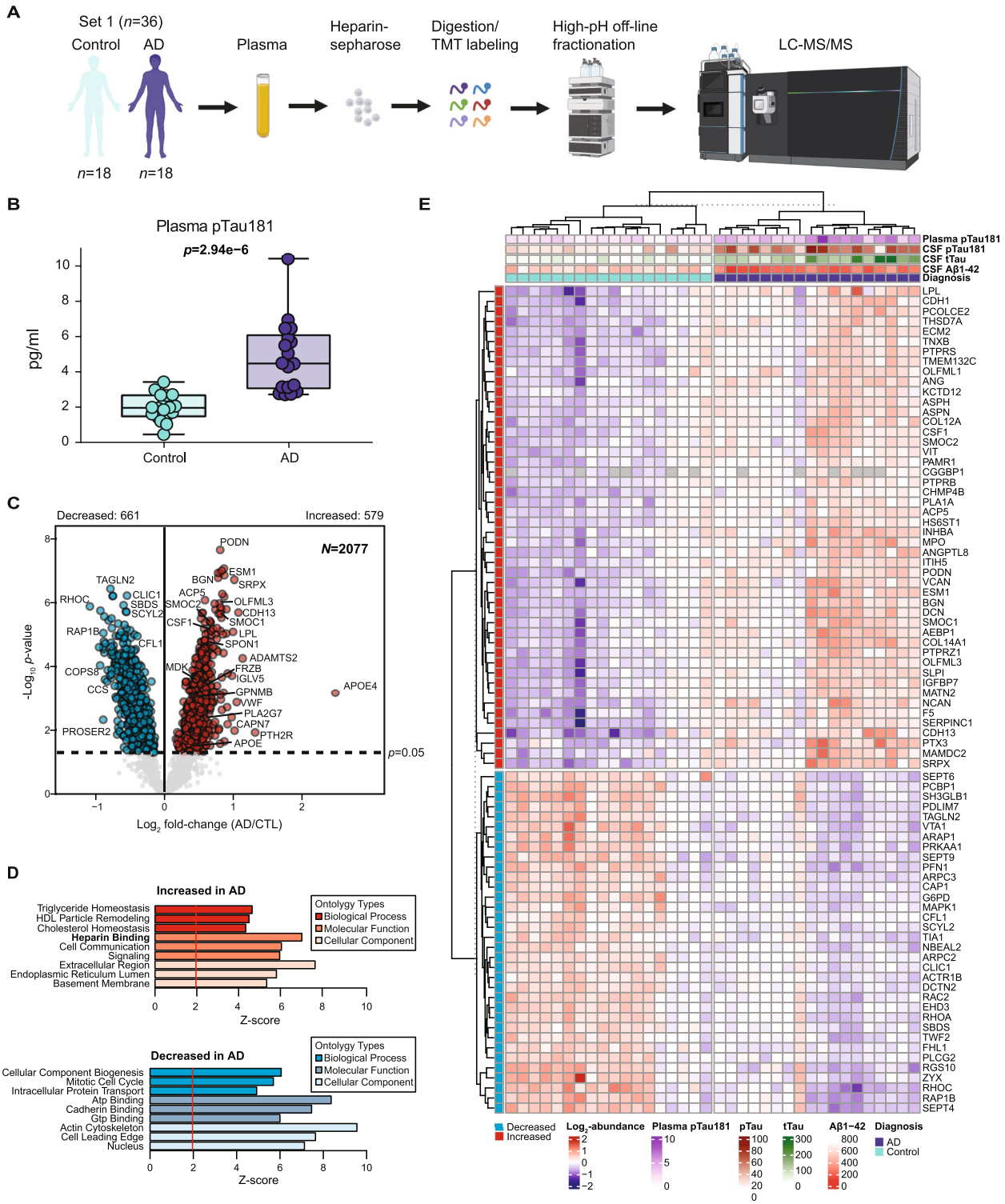

Fig. 2 (See legend on previous page.)

immunoblotting of the input, FT, and Hp-enriched fractions (Supplemental Fig. 2). To enhance data completeness, we considered only those proteins quantified in at least 50% of the samples for subsequent analyses, culminating in the final quantification of 2077 proteins (Supplemental Table 6). To identify differentially abundant

proteins in the AD plasma proteome, we generated a volcano plot for pairwise comparisons of AD versus control individuals (Fig. 2C). Proteins with significant changes in abundance in the AD group were determined using a Student's *t*-test ($p<0.05$). The complete list of differentially abundant proteins is provided in Supplementary Table 7. We identified 579 proteins with significantly increased abundance and 661 proteins with significantly decreased abundance in AD cases. Notably, our findings support an increase of M42 matrisome-associated proteins in AD plasma, which is the most significant module correlated with neuropathology and cognition in AD brain [11]. This includes SMOC1 ($p=2.26\times10^{-6}$), OLFML3 ($p=1.06\times10^{-6}$), APOE ($p=0.028611$), MDK ($p=0.000367$), SPON1 ($p=3.07\times10^{-5}$), GPNMB ($p=0.000935$), and FRZB ($p=0.000555$) [11, 30]. Particularly noteworthy is the identification of SMOC1, MDK, GPNMB, and FRZB, which had not previously been measured in plasma using MS-based technology [30]. Additionally, we discovered other highly significant proteins, including SMOC2 ($p = 4.15 \times 10^{-6}$), APOE4 ($p = 0.000698$), BGN ($p = 1.68 \times 10^{-7}$), ESM1 ($p = 1.40 \times 10^{-7}$), CSF1 ($p = 7.52 \times 10^{-6}$), PLA2G7 ($p = 0.004595$), GCG ($p = 0.004995$), and PODN ($p = 2.25 \times 10^{-8}$). As expected, levels of APOE4 were more increased in the AD group compared to the control group, consistent with the higher frequency of APOE4 carriers in the AD group. GO analysis of the 579 significantly increased proteins in AD indicated that 'Heparin binding' was a major altered pathway (Fig. 2D), which aligns with the strong correlation observed between M42 members and HBPs in AD brain [11]. Conversely, the 661 significantly decreased proteins were strongly linked to 'ATP binding' and 'Mitotic cell cycle' processes among others. To investigate protein differential abundance at the individual sample level, we used proteins with Benjamini-Hochberg (BH) false discovery rate (FDR)-corrected $p<0.0005$ to perform supervised cluster analysis across all 36 samples. As illustrated in Fig. 2E, the expression profiles of 82 highly significant proteins in Hp-enriched plasma correctly distinguished AD from control cases, with only minor exceptions. Among these 82 proteins, 48 were increased in AD, while 34 were decreased.

To assess any biases that heparin enrichment may have caused, and to compare the differences in proteome coverage to traditional immunodepletion methods [59, 60], we benchmarked the Hp-enriched proteome against the immunodepleted proteome from the same 36 samples in Set 1 (Supplemental Fig. 3A). A total of 1,129 proteins were identified (< 50% missing values) after immunodepletion across control and AD samples, nearly half as many identified by heparin enrichment. Among these, 432 proteins were increased in AD and 383 were decreased (Supplemental Fig. 3B and Supplemental Table 8). We found that 804 out of the 1,129 proteins (71%) in the immunodepletion dataset overlapped with the Hp-enriched dataset, with 325 proteins not detected by heparin enrichment (Supplemental Fig. 3C and Supplemental Table 9). However, the Hp-enriched proteome identified nearly four fold as many additional proteins ($N=1,273$) that were not detected by the immunodepletion approach, suggesting that the heparin enrichment method captures a broader spectrum of the plasma proteome compared to the traditional immunodepletion method. These include many members of the M42 matrisome module, such as SMOC1, SMOC2, OLFML3, SPON1, FRZB, and MDK (Supplemental Fig. 3D). Additionally, significant protein levels (AD vs. Control) measured following either heparin enrichment or immunodepletion exhibited a strong correlation regarding the directional changes observed in AD plasma (Supplemental Fig. 3E). There was only approximately 1% discordance in the overlapping proteins with a p-value < 0.05, and when restricted to a p-value < 0.01, all proteins showed the same direction of change in AD. This indicates that the heparin enrichment does not significantly bias the direction of change in the AD plasma proteome compared to immunodepletion approaches. In summary, these results reveal hundreds of differentially abundant plasma proteins that are altered in AD and are not captured by immunodepletion approaches.

Directional changes in the AD plasma proteome across different proteomic platforms

To assess the consistency of the direction of change in AD plasma proteome measured by our heparin enrichment approach coupled with TMT-MS (Heparin-MS), we utilized protein measurements, previously obtained through either the SomaScan® aptamer-based technology or the proximity extension assay (PEA) technology from Olink®, from 35 overlapping samples in our discovery set (Set 1) of control ($n=18$) and AD ($n=17$) plasma samples (Fig. 3). The protein measures were obtained from these two platforms following cross-platform median-centered normalization as previously described [30] and are provided in Supplemental Table 10 and 11. As anticipated, the aptamer-based SomaScan yielded the largest set of protein measurements from the plasma samples ($N=7284$), followed by our Heparin-MS method ($N=2077$) and Olink ($N=979$). We used gene symbols from each output to calculate the number of measurements and their overlap across the various platforms (Fig. 3A and Supplemental Table 12). Each platform exhibited its own unique group of non-overlapping identified proteins (gene products), which we subsequently subjected to

GO analyses to assess their associations with biological pathways. Notably, the unique gene products identified in the Heparin-MS method exhibited significant association with pathways related to 'Alzheimer's disease and miRNA effects' and 'Parkin-ubiquitin proteasomal system pathway' (Fig. 3B), which further underscores the value of heparin enrichment in capturing biology related to AD and related neurodegenerative diseases. To gauge the consistency in the direction of change in AD plasma proteome compared to controls across independent platforms, we conducted correlation analyses of the $\log_2$ fold-change (AD vs Control) for overlapping proteins between Heparin-MS and Olink ($N=279$ gene products, $cor=0.73$, $p=1.1e^{-47}$) or SomaScan ($N=1183$ gene products, $cor=0.62$, $p=1.4e^{-126}$). Both comparisons revealed strong positive correlations between the two platforms (Fig. 3C). However, consistent with our previous results [30], we observed a notable bias towards decreased protein levels in AD plasma on the SomaScan and Olink platform even following median-centered normalization (Supplemental Fig. 4A-B and Supplemental Tables 13–14). This bias was more pronounced in SomaScan, even when restricted to only significant proteins (p-value < 0.05) (Supplemental Fig. 4C-D). While 95% of the overlapping significant proteins ($N=55/58$) in Olink and the Heparin-MS exhibited the same direction of change in AD, 18% of proteins ($N=78/445$) were discordant in the direction of change between the Heparin-MS and SomaScan (i.e., significantly decreased in AD by SomaScan, yet significantly increased by Heparin-MS). Given that the significant protein levels (p-value < 0.05) measured following either heparin enrichment or immunodepletion exhibited strong correlation regarding the directional changes observed in AD plasma (Supplemental Fig. 3E), this discrepancy may be attributed to the methodology. For example, the SomaScan aptamer-based platform requires multiple separate dilutions to accommodate the wide dynamic range of plasma protein concentrations, which may impact the affinity-based aptamer binding for low-abundant protein measurements [61]. Nevertheless, the findings indicate that the observed direction of change of Hp-enriched plasma proteins measured by TMT-MS in the AD group largely aligns with those proteins measured by Olink and SomaScan. Notably, increases of several M42 members in AD were consistent across at least two platforms, including SMOC1, OLFML3, GPNMB, HTRA1, and APOE, whereas SPON1, PTN, APP and FRZB showed the same direction of change in AD across all three platforms. Additional plasma proteins such as ESM1, PLA2G7, BGN, CSF1, GCG, VWF, and LPL were also consistently increased in AD across at least two platforms. Boxplots for SPON1, WAS, ESM1 and PLA2G7 that showed concordant changes in all three platforms are provided (Fig. 3D). Thus, with some exceptions, these data support an overall positive correlation among the three complementary proteomic platforms for shared protein measurements in AD plasma.

Reproducibility and depth of the heparin enrichment approach coupled with TMT-MS in plasma

To demonstrate the reproducibility of the Heparin-MS approach in plasma, we analyzed a separate set of samples (Set 2) involving 49 controls and 36 AD individuals (Supplemental Table 15), which included 13 overlapping controls with the discovery Set 1. These samples underwent heparin enrichment processing similar to the procedures applied to Set 1, albeit using a different volume and lot of heparin-sepharose beads (Fig. 4A). TMT-MS was employed for the analysis of the Hp-enriched fraction, carried out across five batches, resulting in the identification of 2618 proteins measured in 50% or more of the samples. (Supplemental Table 16). Collectively, the two datasets identified a total of 3284 unique proteins, but only 2866 in 50% or more of the samples across two

(See figure on next page.)

Fig. 3 Heparin-enriched plasma proteome is largely consistent and complementary to other independent proteomic platforms. **A** The number and overlap of proteins (i.e., unique gene products) quantified in plasma across three platforms, including the TMT-MS approach in Hp-enriched samples (Heparin-MS, control = 18, AD = 18), the PEA-based assay (Olink, control = 18, AD = 17), and the aptamer-based method (SomaScan, control = 18, AD = 17). The inclusion criterion required at least 18 measurements across all samples. **B** Wiki-pathway analysis highlighting specific pathways from uniquely identified gene products in each of the three platforms. The Heparin-MS method exhibited significant association with pathways related to 'Alzheimer's disease and miRNA effects' and 'Parkin ubiquitin proteasomal system pathway', underscoring the neurodegenerative disease specificity. A total of 5503 platform-unique gene symbols were used as background for GO analysis. **C** Pearson correlation between $\log_2$ fold-change (AD vs CTL) of common gene products measured by the Heparin-MS and Olink (left, $N=279$ gene products, $cor=0.73$, $p=1.1e^{-47}$) as well as the Heparin-MS and the SomaScan (right, $N=1183$ gene products, $cor=0.62$, $p=1.4e^{-126}$). The significance of Pearson correlation was determined by Student's p-value. Several M42 members and associated proteins showed concordant directions in both comparisons, including SPON1, APP, PTN, FRZB, ESM1, PLA2G7, VWF, GCG, and LPL. **D** Boxplots display the consistent and statistically significant changes in AD observed for SPON1, WAS, ESM1, and PLA2G7 across all three platforms. Significance was determined by Student's t-test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). CTL, control; *cor*, Pearson correlation coefficient

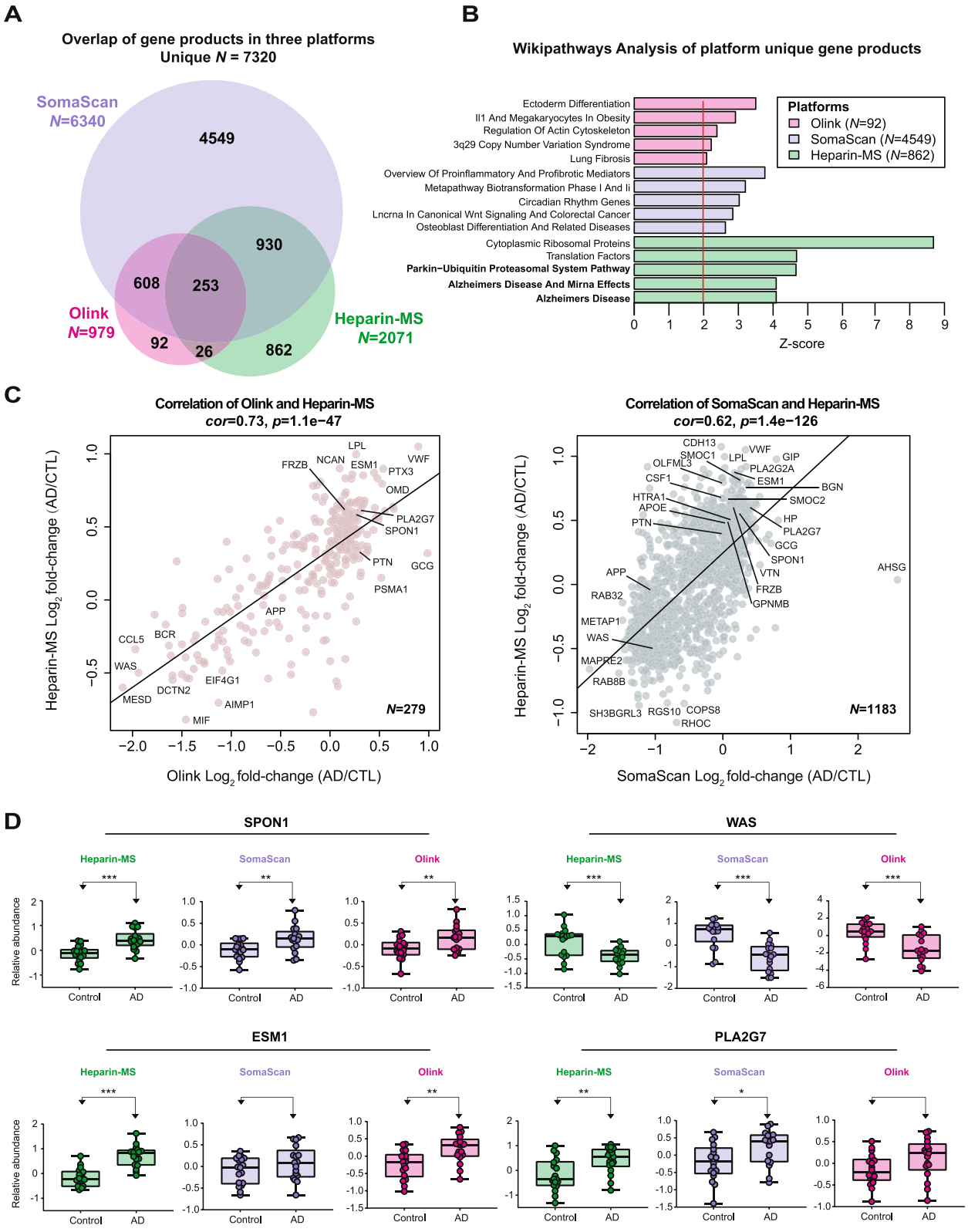

Fig. 3 (See legend on previous page.)

sets were used for further analysis (Fig. 4B). This number is more than twofold greater than the coverage generated with immunodepletion method followed by TMT-MS (Supplemental Fig. 3C), supporting the depth and sensitivity of the heparin enrichment approach when coupled to TMT-MS. Impressively, this encompassed proteins across 10 orders of magnitude in concentration within the plasma, even extending to those with the lowest concentration (1.47 pg/ml), as evidenced by inclusion of proteins such as LAG3 and RNF213 as well as many M42 matrisome members based on The Human Protein Atlas Database (Fig. 4C and Supplemental Table 17). In total, 1829 out of 2866 (64%) proteins overlapped between the two independent TMT-MS datasets (Fig. 4D and Supplemental Table 18). Differential expression analyses were performed on all 2618 proteins within Set 2, resulting in 826 proteins being increased and 795 proteins decreased in AD (Supplemental Table 19). Notably, significant protein changes between AD and control samples across the two sets were highly reproducible (Fig. 4E). Among the 732 overlapping proteins with a BH FDR-corrected p-value < 0.05 in both datasets, only 10 (1.4%) exhibited discordant changes in AD across the two sets ($cor = 0.93$, $p < 1e^{-200}$). However, when increasing the significance threshold to BH FDR-corrected p-value < 0.01, we observed no discordant differences among the 390 overlapping proteins across the two sets. Among these highly significant proteins, we confirmed several targets outside of M42 matrisome members that were consistently validated across different proteomic platforms (Fig. 3C), including ESM1, PLA2G7, BGN, CSF1, and GCG, further reinforcing their changes in AD plasma. Overall, our heparin enrichment approach coupled to TMT-MS consistently and reliably captured a diverse spectrum of plasma proteins, spanning an impressive dynamic range of magnitude.

Association of heparin-enriched plasma proteins with AD biomarkers and cognitive measures

CSF biomarkers, including $A\beta_{1\text{-}42}$, tTau, and pTau181, have played a pivotal role in identifying individuals who are either at risk or already manifesting underlying AD pathology, collectively referred to as AT + individuals [49]. Furthermore, recent advancements have brought plasma pTau species, specifically pTau181, pTau231 and pTau217, into focus due to their associations with both underlying amyloid and tau pathology [62]. Here, we aimed to evaluate proteins across the two sets of Hp-enriched plasma proteome that are associated with cognition and AT + status by correlating their abundances to MoCA scores, CSF AD biomarkers ($A\beta_{1\text{-}42}$, tTau, tTau/$A\beta_{1\text{-}42}$, pTau181) and plasma pTau181 collected from the same patients. To enhance the specificity of our analysis, we excluded cases from the initial 121 samples (including 13 overlapping controls as described above) that did not meet strict CSF biomarker criteria (tTau/$A\beta_{1\text{-}42}$ ratio) or MoCA cutoffs [49] as described in methods (Supplemental Fig. 5A). As a result of this selection process, 12 non-overlapping samples were excluded from Set 2, resulting in a total of 109 samples (Set 1 = 36 samples, Set 2 = 73 samples) for further analysis (Supplemental Table 20). Boxplots for each measurement across all unique cases ($n = 96$) were generated to show their separation based on disease status (Supplemental Fig. 5B). Subsequently, we employed a *meta*-analysis of Set 1 and Set 2 to generate a composite *meta* p-value that considered the significance of association and effect size (AD vs Control) for 2865 proteins, all of which exhibited less than 50% missing values within each set. This *meta*-analysis revealed a substantial increase in differentially abundant proteins in the Hp-enriched AD plasma, with 945 proteins showing an increase and 923 proteins showing a decrease in AD, as compared to analyzing either dataset alone (Fig. 5A and Supplemental Table 21). Importantly, this approach

(See figure on next page.)

Fig. 4 Heparin enrichment enhances the depth of plasma proteome and is reproducible. **A** The second replication dataset (Set 2) was comprised of control ($n = 36$) and AD individuals ($n = 49$), which underwent similar heparin enrichment and TMT-MS analysis as the discovery dataset (Set 1). **B** The combined datasets yielded a total of 3284 unique proteins, and 2866 of them being identified and quantified in 50% or more of the samples (< 50% missing) across two sets. Set 1 and Set 2 identified 2077 and 2618 proteins respectively, with 50% missing. **C** The 2866 proteins were ranked by their $\log_2$ estimated concentration (pg/L) in plasma. Protein concentration information was obtained from The Human Protein Atlas Database. Notably, this set of proteins covered an impressive range of concentrations, spanning 10 orders of magnitude. Even proteins with the lowest concentration, such as LAG3 and RNF213, were included, along with numerous members of the M42 matrisome that are highlighted. **D** The number and overlap of proteins quantified in Set 1 ($N = 2077$) and Set 2 ($N = 2618$) with less than 50% missing values. Among the 2866 total unique proteins identified, approximately 64% (1829 out of 2866) were overlapping between the datasets. **E** A scatter plot illustrates the Pearson correlation between $\log_2$ fold-change (AD vs CTL) of significantly altered proteins in both Set 1 and Set 2. There're 732 proteins overlapping in the two sets and being significantly changed in AD with a BH FDR-corrected p-value < 0.05 in both datasets. Only 10 out of 732 proteins exhibited discordant changes in the two sets, demonstrating a high degree of concordance ($cor = 0.93$, $p < 1e^{-200}$). Furthermore, all 390 overlapping proteins selected with a BH FDR-corrected p-value < 0.01 displayed consistent changes in both datasets when comparing AD and control samples, with a remarkable correlation of 0.96 ($p < 1e^{-200}$). The significance of Pearson correlation was determined by Student's t-test. CTL, control; LC–MS, liquid chromatography-mass spectrometry; *cor*, Pearson correlation coefficient

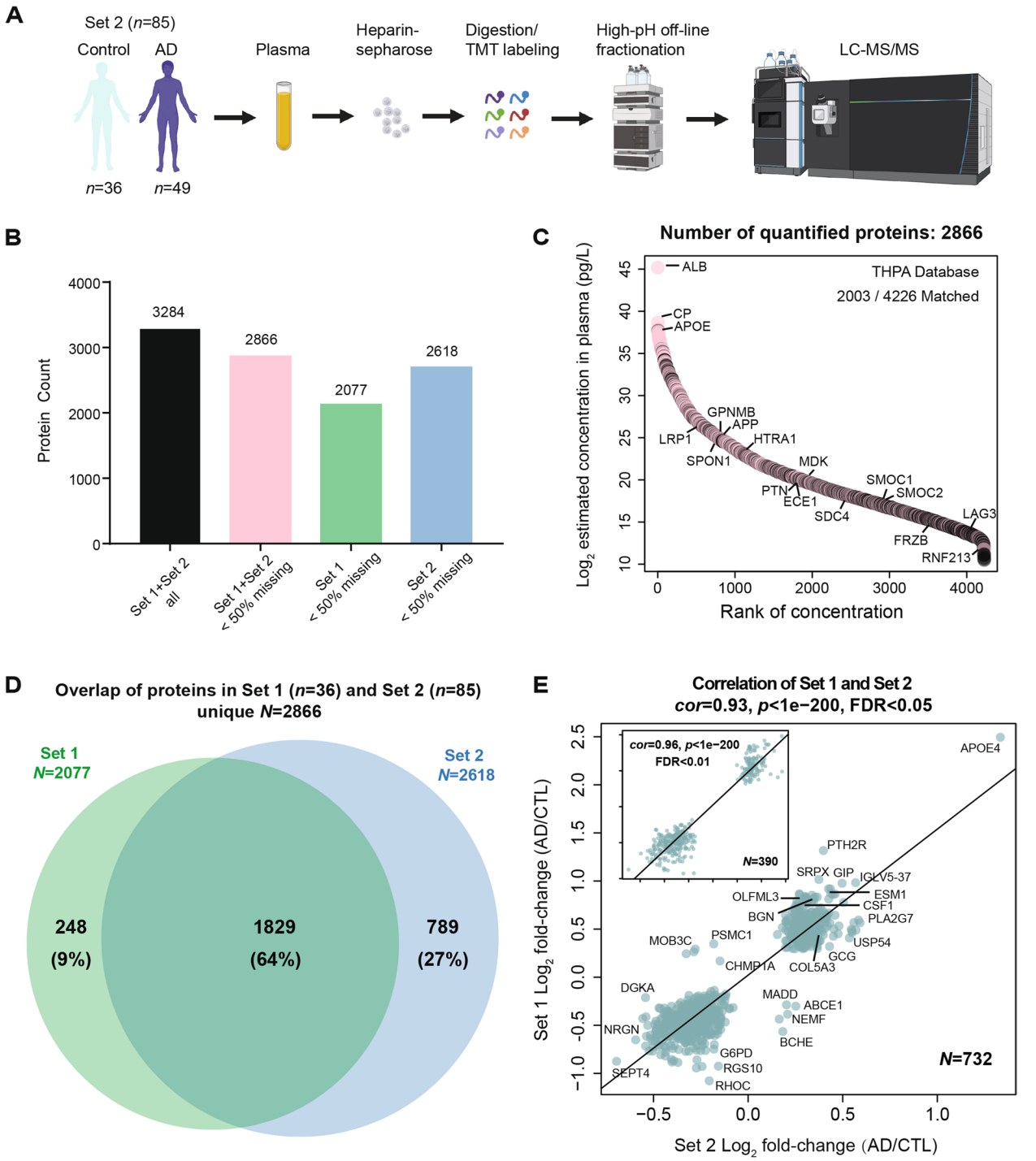

Fig. 4 (See legend on previous page.)

also ensured the retention of proteins that were identified in only one dataset (e.g., SMOC1 in Set 1), thereby preserving their availability for subsequent analyses. To assess the relationship between Hp-enriched plasma proteins and AD biomarkers and cognitive measures, we conducted correlation analyses between Z-transformed protein abundance (Supplemental Table 22) and MoCA scores, as well as immunoassay values of CSF ($A\beta_{1\text{-}42}$, tTau, pTau181, tTau/$A\beta_{1\text{-}42}$) and plasma pTau181 (Supplemental Table 23). Correlation between SomaScan or Olink protein abundance and MoCA scores, as well as immunoassay values of CSF ($A\beta_{1\text{-}42}$, tTau, pTau181,

tTau/A$\beta_{1\text{-}42}$) and plasma pTau181 were also provided in Supplemental Table 24–25. In Fig. 5B, we highlighted 77 Hp-enriched plasma proteins that exhibit significant correlations with at least three of the AD biomarker measurements or MoCA scores. Individual correlation scatterplots and linear fit lines are presented for ESM1, BGN, PLA2G7, and CSF1, all of which exhibit significant correlations with AD biomarkers and cognitive measures (Fig. 5C). Moreover, these proteins have consistently demonstrated increased abundance in AD plasma in both Set 1 and Set 2, as well as other platforms as described above. To ensure that the observed changes in these proteins in AD were not influenced by age, sex, or race, we performed regression analyses for these traits within Set 1 and Set 2, followed by the same *meta*-analysis described earlier (Supplemental Fig. 6 and Supplemental Table 26). Abundance differences of significant proteins ($N = 1765$, *meta* $p < 0.05$) in AD versus the control group, with or without regression, were highly correlated ($cor = 0.99$, $p < 1e^{-200}$). Only 9 proteins (0.5%) showed discordant changes in AD, and all matrisome proteins remained significant. This indicates that the vast majority of Hp-enriched proteins differing in AD plasma in this study are not affected by age, sex, or race. Therefore, the correlation between these Hp-enriched plasma proteins to established AD biomarkers and cognitive measures suggests the potential utility of these proteins for disease classification or staging in the context of AD progression.

Evaluation of heparin-enriched plasma proteins for classifying AD

To evaluate the ability of Hp-enriched plasma proteins to classify CSF biomarker confirmed AD cases from control individuals, we performed a receiver-operating characteristic (ROC) curve analysis. This analysis used plasma pTau181 measurements and only included plasma proteins with no missing values and those that were highly significant (*meta* p-value < 0.0001, $N = 536$ proteins) across 34 control and 61 AD individuals (Supplemental Table 27). Our results show that the top ten performing proteins in plasma individually achieved an area under the curve (AUC) value ranging from 0.81 to 0.82. In comparison, plasma pTau181 demonstrated a higher AUC of 0.93, consistent with previous findings [63] (Fig. 6A). When we combined the top five plasma proteins (CHST15, BGN, THSD7A, SLPI, and RGS10) into a panel, the AUC modestly increased to 0.85 (Fig. 6B). Although the panel did not outperform plasma pTau181 alone, combining it with pTau181 significantly improved classification performance, increasing the AUC from 0.93 (pTau181 alone) to 0.98, with a p-value of 0.00645. This indicates that the Hp-enriched plasma proteome captures some of the variance in cognitive dementia not solely explained by pTau181 levels.

Overlap between the heparin-enriched plasma and human brain proteomes

Previously integrated analysis of the human brain and CSF proteomes has revealed a substantial overlap of approximately 70%, strongly supporting a hypothesis that CSF serves as a valuable window into the brain [14]. However, less is known for the overlap between plasma and brain proteomes. While studies have explored the overlap between brain and plasma proteomes using TMT-MS proteomic datasets following the immunodepletion of highly abundant proteins [30], we sought to leverage the depth of our Heparin-MS proteomic datasets to assess the overlap with human postmortem brain proteome. To ensure the consistency of protein overlap analysis, we employed the same Uniprot database and utilized the FragPipe search algorithm we used for Heparin-MS plasma analysis to re-analyze 456 dorsolateral prefrontal cortex (DLPFC) tissues from control, asymptomatic AD (AsymAD) and AD brains from the ROSMAP and the Banner cohorts (Supplemental Fig. 7A and Supplemental Table 28) [11]. The classification of

(See figure on next page.)

Fig. 5 Association of heparin binding proteins in plasma with cognition (MoCA) and conventional AD biomarkers: CSF A$\beta_{1\text{-}42}$, tTau and pTau181 and plasma pTau181. **A** A *meta*-analysis of significant differences between control and AD on 2865 Hp-enriched plasma proteins that were measured in 50% or more samples within filtered Set 1 (control = 18, AD = 18) and Set 2 (control = 29, AD = 44). The x-axis represents the mean $\log_2$ fold-change (AD vs CTL), indicating an average abundance difference between Set 1 and Set 2. The y-axis shows the Student's *t*-statistic ($-\log_{10}$ *meta* p-value) for all proteins in each pairwise group. Significantly increased proteins in AD are marked in red (*meta* $p < 0.05$), while proteins with significantly decreased levels in AD are denoted in blue. Grey dots represent proteins with insignificant changes. **B** The heatmap highlights 77 Hp-enriched plasma proteins that have strong correlations to AD biomarkers. The color scale represents the degree of Pearson correlation (positive in red and negative in blue) between Z-transformed plasma protein abundances and immunoassay measures of various AD-related traits, including cognition (MoCA score), CSF A$\beta_{1\text{-}42}$, CSF tTau, CSF pTau181, CSF ratio of tTau/A$\beta_{1\text{-}42}$, and plasma pTau181. Significance levels determined by Student's *t*-test are denoted by overlain asterisks; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. **C** Individual scatterplots illustrate the correlations with CSF A$\beta_{1\text{-}42}$, CSF pTau181, and plasma pTau181 of four specific Hp-enriched plasma proteins: ESM1, BGN, CSF1, and PLA2G7. *Cor* and *p*-values for each correlation are provided above each plot. The colors are differentiated by sets, with red representing Set 1 and blue denoting Set 2. pTau, phospho-tau; tTau, total tau; CTL, control; *cor*, Pearson correlation coefficient

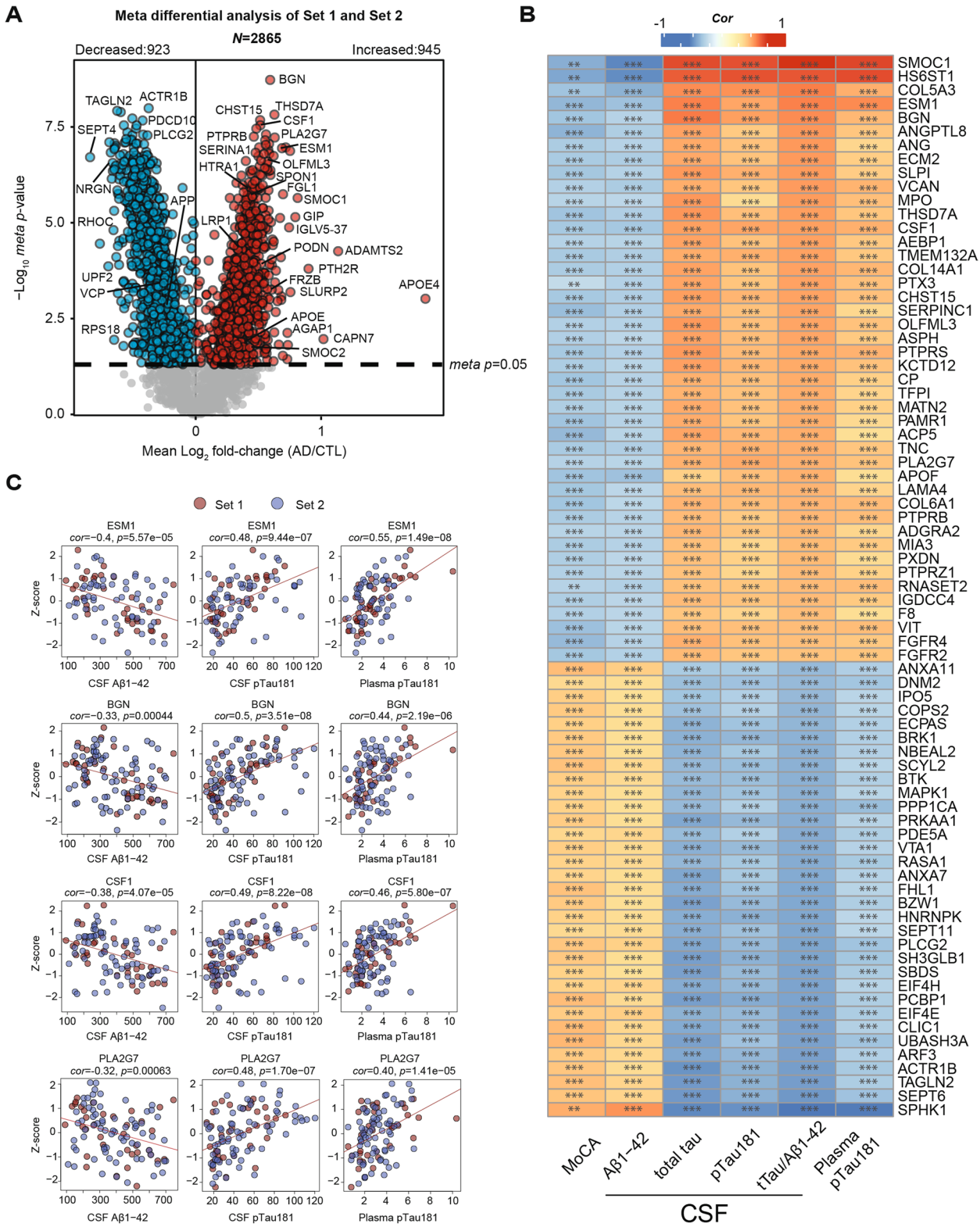

Fig. 5 (See legend on previous page.)

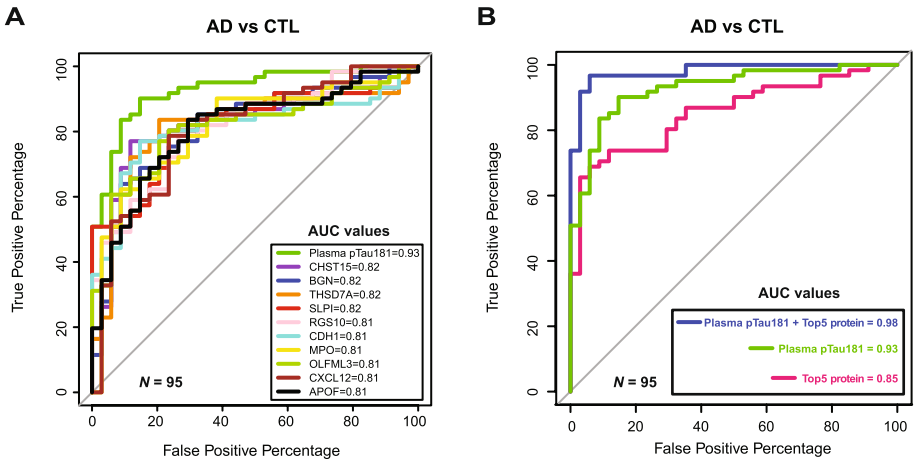

Fig. 6 Evaluation of heparin-enriched plasma proteins for classifying AD. ROC curves for CSF biomarker confirmed AD ($n=61$) versus biomarker-negative control individuals ($n=34$) were generated to determine the top-ranked diagnostic proteins among the 536 highly significant plasma proteins (*meta* $p<0.0001$) with no missing values across all unique 95 individuals. **A** Individual ROC curves and AUCs for the top ten ranked plasma proteins, as well as plasma pTau181 alone. **B** ROC curves and AUCs of i) the combined top five performing proteins as a panel (pink), ii) plasma pTau181 alone (green), and iii) plasma pTau181 plus the protein panel (blue). ROC curve statistics for highly significant proteins, including AUC, *p*-value, 95% DeLong confidence interval, accuracy, specificity, and sensitivity for AD vs CTL, are provided in Supplemental Table 27. CTL, control; AUC, area under the curve

cases incorporated semi-quantitative histopathological assessments of Aβ and tau neurofibrillary tangle deposition, along with an evaluation of cognitive function in proximity to the time of death, as detailed in our previous work [11]. Specifically, AsymAD cases were characterized by a neuropathological burden of Aβ plaques and tau tangles that closely resembled AD cases. However, these individuals did not exhibit substantial cognitive impairment near the time of death, indicating an early preclinical stage of AD. By applying a data normalization approach similar to the one used for the Hp-enriched plasma samples, we successfully identified 8,956 proteins detected in 50% or more of the samples from brain tissues corresponding to 8,904 unique gene symbols (Supplemental Table 29). When compared to a previous brain proteome dataset generated using Proteome Discoverer (PD) and an older 2015 Uniprot protein database [11], it was found that 7,476 of these unique gene symbols overlapped between the two outputs, while 1,428 were exclusive to the FragPipe, representing an approximately 18% increase in the number of identifications (Supplemental Fig. 7B and Supplemental Table 30). This included proteins such as SMOC2, which is highly homologous paralog of SMOC1 [64] and also increased in AD brain and plasma (Fig. 5A). Furthermore, to ensure the consistency of the direction of change in the AD brain proteome compared to controls, we correlated the differentially abundant proteins across these two sets of results (Supplemental Table 31–32), revealing a correlation coefficient of 0.9 between AD and control groups and a correlation coefficient of 0.84 between AsymAD and control groups (Supplemental Fig. 7C). This robust

(See figure on next page.)

Fig. 7 Mapping differential abundant heparin-enriched plasma proteins in AD within a human consensus brain network. **A** The I-graph displays the updated M42 membership following FP database search, which includes 35 total proteins. Members with increased abundance in Hp-enriched AD plasma are highlighted in red, while those with decreased abundance are indicated in blue. **B** The pie chart shows the number of proteins that overlap between the Hp-enriched plasma dataset ($N=2865$) and consensus human brain datasets ($N=8956$ proteins). 2211 out of 2865 (77%) proteins identified in plasma are also identified in the brain. The percentage coverage of proteins in each module is also presented. **C** 44 modules of previously generated consensus human brain co-expression network [11] visualized in the order of module relatedness following protein re-assignment as described in method. **D** Overlap of Hp-enriched plasma proteins (y-axis) with increased abundance in AD depicted in red or decreased abundance in AD shown in blue within brain network modules. The intensity of color shading indicates the degree of overlap. Statistical significance is indicated in the heatmap regions using stars (* $p<0.05$, ** $p<0.01$, *** $p<0.001$). The *p*-values derived from FET were BH FDR-corrected. **E** The heatmap demonstrates the bicor correlations of each module with CERAD, Braak, and MMSE cognitive scores (* $p<0.05$, ** $p<0.01$, *** $p<0.001$). As mentioned previously, M42 'Matrisome' exhibits the strongest correlation with AD pathology, and several synaptic modules (M1, M4, M5, and M22) display an overall decrease in AD brain. **F** Using FET, the cell type nature of each module was assessed by module protein overlap with brain cell-type-specific markers of astrocytes, microglia, neurons, oligodendrocytes and endothelia. The strength of the color shading indicates the degree of cell type enrichment with asterisks denoting statistical significance (* $p<0.05$, ** $p<0.01$, *** $p<0.001$). The FET-derived *p*-values were BH FDR-corrected. CTL, control

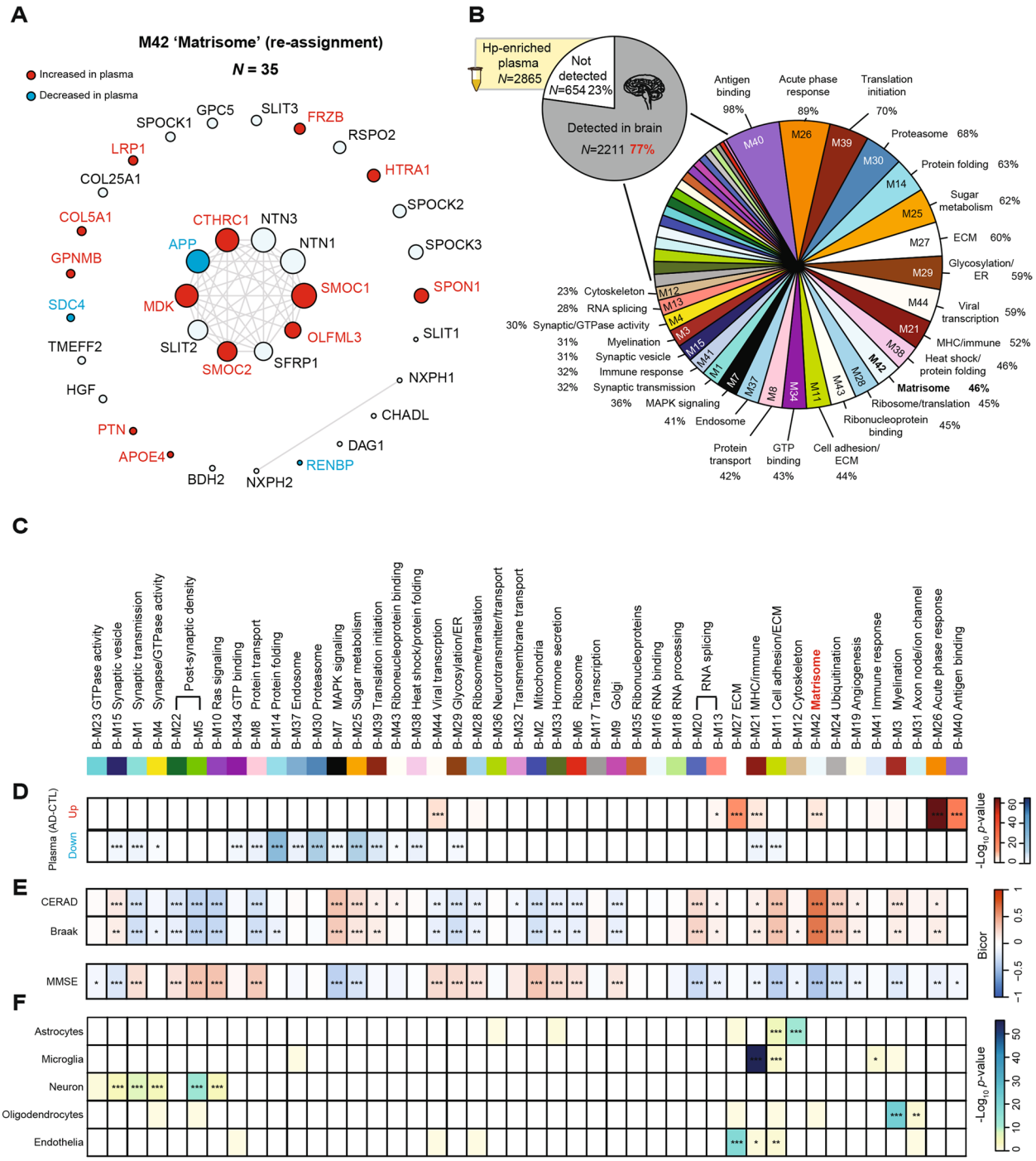

Fig. 7 (See legend on previous page.)

correlation provides strong support for the reliability of the FragPipe quantitative outputs. Subsequently, we systematically mapped all the measured brain proteins by FragPipe to one of the 44 pre-existing network eigenproteins from a consensus brain network [11]. This mapping was achieved by recalculating the kME (bicor correlation to module eigenprotein) for each protein and assigning it to the module with the highest correlation (Supplemental Table 33). A revised M42 was generated (Fig. 7A), comprising 35 members, marking a 10% increase compared to the number observed in the previous TMT-MS brain dataset using PD [11]. This new M42 included additions

such as SMOC2, COL5A1, SLIT3, HGF, NXPH2, RSPO2, and CHADL, among which SMOC2 and COL5A1 exhibited significant increases in AD plasma levels. Moreover, the incorporation of APOE-specific protein isoforms into the database led to the assignment of APOE4 to M42, rather than APOE (inferred as APOE3), following the FragPipe search of the brain proteome. Although APOE exhibited the strongest association with M42 (kME = 0.287) compared to all other modules in the network, it did not meet the 0.3 kME cutoff for inclusion into a network module. This finding further bolsters the genetic connection between the *APOE ε4* genotype and increased levels of M42 members in AD [11, 65]. When comparing the proteins identified in plasma with those in the brain proteome, we discovered that a substantial portion, approximately 77% (2211 out of 2685) of the plasma-identified proteins were also quantified in the brain (Fig. 7B and Supplemental Table 34). Surprisingly, this was a higher percentage than we have reported for the CSF-brain overlaps (~70%) [14]. To assess the overlap between the brain network and the plasma proteome, we mapped Hp-enriched plasma proteins to one of the 44 brain co-expression modules in Fig. 7B. Among these modules, those associated with M40 'Antigen binding' (98% of members) and M26 'Acute phase response' (89% of members) exhibited the highest degree of overlap with the plasma proteome. Notably, M42 'Matrisome', which exhibited the most robust correlation with AD pathology in the brain [11, 30], demonstrated a 46% overlap (16 out of 35 members) with the plasma proteome. This represents a significant improvement compared to the immunodepletion approach, which only detected 5 out of 35 members (APOE4, APP, DAG1, GPNMB, and LRP1), reflecting a mere 14% of the module (Supplemental Tables 8 and 33). Collectively, the significant overlap between brain and plasma proteomes highlights the potential of the Hp-enriched plasma proteome in capturing signatures related to AD brain pathophysiology.

The heparin-enriched AD plasma proteome contextualized through integration with a consensus AD brain proteome network

To gain deeper insights into the biological connection between differentially abundant plasma proteins and the brain network in AD, we conducted a comprehensive integrated analysis of the two distinct proteomes. Specifically, we examined which brain co-expression modules (Fig. 7C) exhibited strong overlap with differentially abundant Hp-enriched plasma proteins in AD using a Fisher's exact test (FET). As depicted in Fig. 7D, Hp-enriched plasma proteins that increased in AD (red) are significantly overlapped with seven brain modules, while 16 modules exhibited overlaps with proteins significantly decreased in AD plasma (blue). Interestingly M21, which is linked to microglia and inflammation and modestly increased in symptomatic AD brain, contains both increased and decreased plasma proteins. As described previously [11], many of these brain modules displayed significant correlations with AD pathology (i.e., CERAD for neuritic plaque burden and Braak staging for the progression of tau neurofibrillary tangles), cognitive status prior to death (MMSE), and brain cell types as illustrated in Fig. 7E-F. This allowed us to prioritize modules that exhibited significant correlations with AD clinicopathological traits and specific brain cell types and that were also enriched with differentially abundant plasma proteins in AD. We categorized these modules into groups based on their expression trends in disease (Fig. 8). For example, certain modules displayed concordant changes in AD plasma and brain, even in some instances in the asymptomatic phase of disease in the brain. This included modules such as the 'Matrisome' (M42), whereas modules associated with 'Acute-phase response' (M26) and 'RNA-splicing' (M13) were increased in both plasma and brain only in the symptomatic phase of disease. We also observed significant overlap between plasma proteins that are significantly decreased in AD and modules related to synaptic proteins or ER/glycosylation (M1, M8, and M29) that exhibit a gradual reduction in levels within the brain across control, AsymAD, and AD cases. Interestingly, within the modules, the trends were not consistent for all plasma proteins. For example, in M42, while most of the matrisome members were significantly increased in AD brain and plasma, SDC4 and APP exhibited significant decreases in AD plasma, which was in opposition to their changes in the brain. Another class of modules demonstrated divergent abundance trends in AD brain and plasma (M11, M7, M25, M39, M15, and M44). For instance, proteins specifically in M11 'Cell adhesion/ECM', M7 'MAPK signaling' and M25 'Sugar metabolism', such as PTBP1, FHL1 and DCTN2, exhibited increased levels in the AD brain, yet decreased levels in the Hp-enriched AD plasma. Finally, modules that showed significant overlap with the Hp-enriched plasma proteome, yet exhibited moderate to low correlations with AD clinicopathological traits are detailed in Supplemental Fig. 8. These included M40 'Antigen binding', which is highly increased only in the AD plasma, potentially reflecting a peripheral-derived inflammation and immune response to AD pathology in the brain. On the contrary, M37 'Endosome' and M30 'Proteasome' were significantly decreased in AD plasma while not changing in the brain, suggesting a potential failure of protein clearance in AD plasma. In summary, these analyses facilitated the categorization of groups

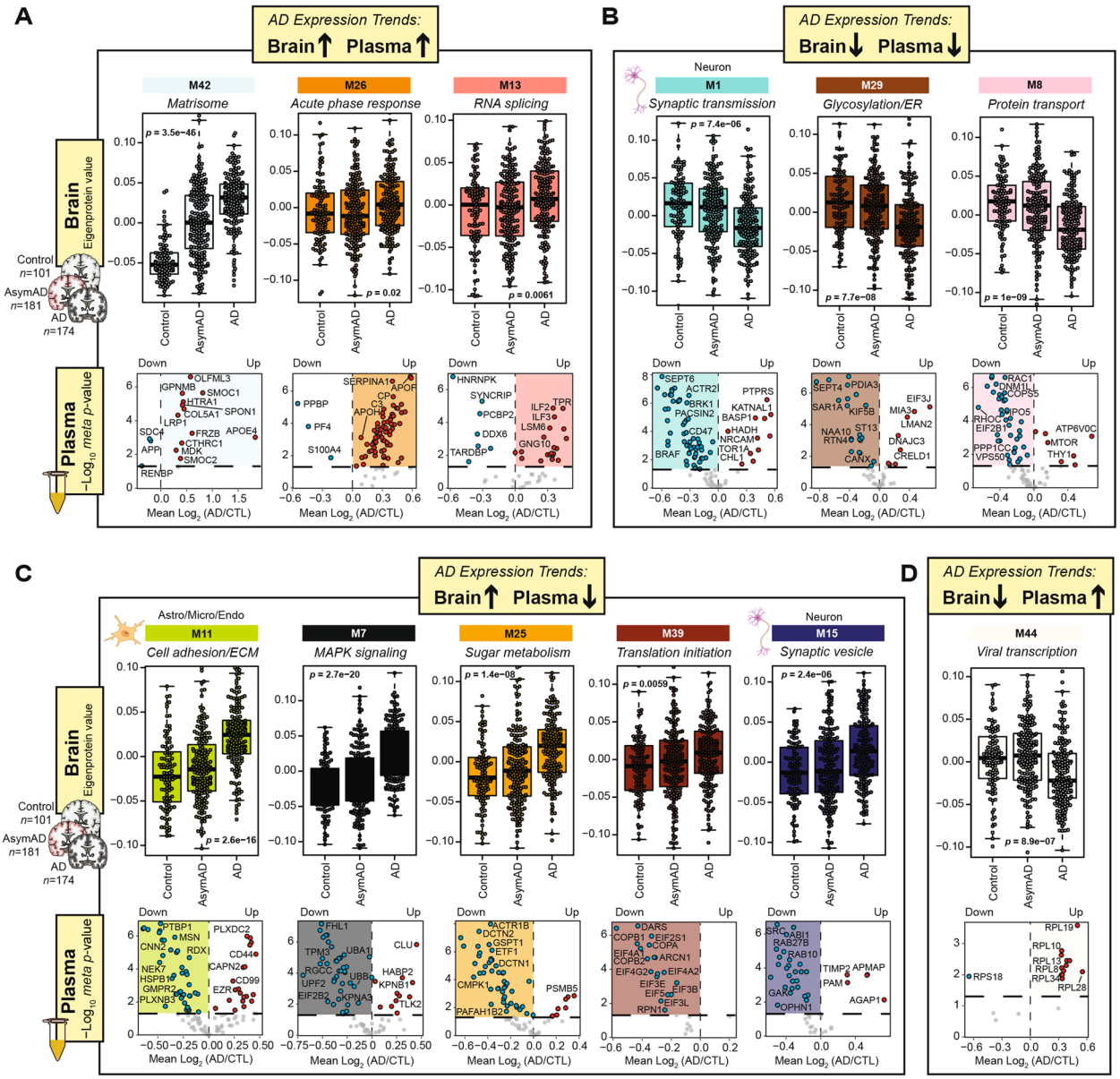

Fig. 8 Overlap between human brain network modules and differentially abundant heparin-enriched plasma proteins in AD. Protein expression trends of the 12 out of 44 (27%) consensus brain network modules [11] that exhibited significant correlations with AD clinicopathological traits and were enriched with differentially abundant plasma proteins in AD. Brain module abundance is represented by eigenprotein values of the consensus brain network (control = 101, AsymAD = 181, AD = 174), while volcano plots illustrate the differential abundance (log_2 fold-change AD vs CTL) of module proteins overlapped with the Hp-enriched plasma proteome. The statistical significance of changes in module eigenprotein abundance across the three groups of the consensus brain cohort was assessed using ANOVA with Tukey post-hoc correction. Modules with $p < 0.05$ were considered significant. **A** The M42 'Matrisome', M26 'Acute-phase response', and M13 'RNA-splicing' displayed increased protein levels in both the AD brain and plasma. **B** The M1 'Synaptic transmission', M29 'Glycosylation/ER' and M8 'Protein transport' showed decreased protein levels in both AD brain and plasma. **C**-**D** The remaining six modules, including M11 'Cell adhesion/ECM', M7 'MAPK signaling', M25 'Sugar metabolism', M39 'Translation initiation', M15 'Synaptic vesicle', and M44 'Viral transcription', exhibited divergent abundance changes in AD brain and plasma. CTL, control; AsymAD, asymptomatic AD; Astro, astrocytes; Micro, microglia; Endo, endothelia

of Hp-enriched plasma proteins based on their associations with AD brain pathophysiology. Additionally, they shed light on the directionality of specific plasma proteins within the brain network modules. Some of these proteins exhibited consistent increases or decreases in both the brain and plasma, while others displayed divergent changes.

Discussion

Here we show that heparin enrichment of plasma followed by MS-based proteomic analyses (Heparin-MS) is a reproducible methodology to enrich a subset of plasma proteins that distinguish AD from controls. More than 2800 plasma proteins spanning over 10 orders of magnitude in protein concentration in plasma can be reliably assessed using this methodology. Contextualizing these data by integrating them with previous studies and our extensive human AD brain and CSF data, we show that Hp-enriched AD plasma proteome provides novel insight into the pathobiology of AD by reflecting protein changes in AD brain.

Our primary rational for exploring heparin enrichment of plasma was to evaluate whether this enrichment would enable us to detect and quantify members of an AD brain protein module in plasma, referred to as M42, the matrisome [11]. Our previous proteomic data had identified M42 as a highly conserved protein module in the human AD brain and the CRND8 mouse model of amyloid deposition [11, 19]. Notably, selected M42 proteins show some of the highest fold increases in the human AD and CRND8 mouse brain proteome, including APP/Ab and APOE, as well as many proteins known to bind heparin [19]. We have recently shown many members of this module bind amyloid and co-accumulate with Aβ in plaques, cerebrovascular amyloid, and/or dystrophic neurites [19]. As in the AD brain proteome, we find that many M42 members in Hp-enriched plasma proteome can also distinguish AD from controls. Thus, our Hp-enriched plasma data indicates that M42 proteins are not only intimately linked to AD amyloid pathology, but also appear to be robust AD plasma protein biomarkers. More unexpectedly, our current data suggests that many proteins identified initially in the human AD brain proteome are also robustly detectable in the Hp-enriched AD plasma proteome. Targeted analyses of AD plasma such as Aβ42, Aβ40, various pTau species, GFAP and neurofilament light chain (NEFL) have provided highly informative blood-based biomarkers that clearly reflect underlying pathological processes in the brain [66]. However, broader integration of AD plasma proteome with the human AD brain proteome has not yielded many novel insights into AD pathobiology. As discussed in detail below, the evaluation of our Hp-enriched AD plasma proteome through the lens of our AD brain proteome suggests that protein changes in multiple brain modules are detectable in the AD plasma proteome.

These findings have many implications relevant to our understanding of AD as a brain only versus a whole-body disorder. Currently it is unclear whether the changes observed in the Hp-enriched AD plasma are simply readouts of protein level changes in the AD brain reflected in the plasma or evidence that AD pathophysiology directly impacts peripheral organ function in a way that drives biomarkers changes, which can be observed in plasma. Given the complex relationship between changes in the AD brain and Hp-enriched plasma proteome that is not always conserved in directionality within selected protein modules, previous evidence that plasma biomarkers may reflect changes in the brain proteome, and precedence for robust bidirectional crosstalk between the brain and the periphery, our own bias is that the plasma proteomic changes represent an admixture of central and peripheral processes.

For some brain protein modules, such as M42, we observed overall consistency among brain, CSF and plasma in terms of direction of change [11, 20, 21]. For example, not only are SMOC1 and SPON1 increased in CSF upwards of 30 years in advance of AD [17], these proteins are also increased in the preclinical or asymptomatic stage of disease in brain, suggesting each as high value biomarkers if detected in plasma [11]. SPON1 was one of the most consistent M42 proteins measured across proteomic platforms as it was increased in AD by aptamer-based (SomaScan) and antibody-based (Olink) measurements. In total, 16 matrisome proteins in M42 were differentially abundant in AD plasma following Heparin-MS, where SMOC1, SPON1, OLFML3, GPNMB, and HTRA1 are among the most significant ones elevated. Other newly identified members of M42 in plasma include SMOC2 and HGF, which have also been shown to be elevated in AD CSF [67]. However, the biological basis for the consistency of changes in these compartments is not yet understood. Future studies aiming to measure matrisome proteins in plasma from cohorts like Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Dominantly Inherited Alzheimer Network (DIAN) will be essential for assessing their prognostic values in predicting disease outcomes.

In our previous consensus analysis of the human brain, we did not incorporate the APOE-specific isoforms (APOE2 and APOE4) into the human protein database [11]. Since these isoforms differ due to substitutions of cysteine residues with arginine residues, they can be easily distinguished in the human proteome after trypsin cleavage that releases novel peptides, which can then be accurately mapped and quantified using mass spectrometry [57]. Notably, by integrating genomics and proteomics data from the same individuals, we previously demonstrated that the individuals carrying an *APOE ε4* allele exhibited higher M42 levels in brain, and this regulation was not solely driven through the levels of the APOE protein itself [11]. In this study, the inclusion of

the APOE4-specific protein isoform rather than APOE in M42 serves to further strengthen the genetic association between *APOE ε4* genotype and M42 levels. This also indicates that the APOE4 protein isoform may have a stronger predisposition for interaction with heparin and HBPs within M42 in the brain than other APOE isoforms [58]. In the future, the implementation of integrated genomics and proteomics pipelines will be needed for assessing whether M42 proteins in plasma are under genetic regulation by *APOE ε4*. This will provide valuable insights into whether this relationship is consistent across both the central nervous system and the peripheral system.

Heparin and HS accelerate the formation of Aβ fibrils [26–28] and loss of this heparin-APOE binding interaction has been suggested as a possible mechanism for the resistance to AD of the *APOE* Christchurch loss-of-function mutation recently described in a *PSEN1* ADAD mutation carrier [58]. However, more recently, a rare *RELN* variant has been proposed to delay the age of onset of siblings with ADAD, who do not carry the Christchurch *APOE* variant [68]. Like APOE, Reelin (RELN) is also an HBP, but in contrast to the *APOE* Christchurch variant, the *RELN* variant is associated with heightened interactions with heparin and a consequent reduction in tau phosphorylation via the Dab1 signaling pathway [68]. It is worth noting that in our study, we observed increased levels of RELN in the Hp-enriched fraction compared to input. Hence, there appears to be an intricate relationship between heparin interactions, APOE, and AD risk. The exact mechanisms by which HBPs, including APOE and members of M42, influence amyloid deposition and potential clearance still requires further investigation.

Beyond their interactions with APOE and M42 members, HSPGs alone have been implicated in AD progression and pathogenesis [69]. Specifically, HSPGs are believed to play a crucial role in mediating the internalization and propagation of specific proteopathic seeds of tau [70]. Namely, the HSPG glypican-4 (GPC4) has been identified as a contributor to APOE4-induced tau hyperphosphorylation [71]. It is also noteworthy that 6-O sulfation on HSPGs is presumed to regulate the cell-to-cell propagation of tau [72]. Interestingly, glypican-5 (GPC5), which is structurally homologous to glypican-4, is a core member of M42. While it is yet to be established whether glypican-5 plays a role in the regulation of tau internalization, the co-expression between APOE4 and GPC5 in the brain suggests the possibility of such involvement. Evidence supporting the role of HS in the etiology of AD is also emerging in genome-wide association studies (GWAS). For example, GWAS *meta*-analysis identified the heparan sulfate-glucosamine 3-sulfotransferase 1 gene (*HS3ST1*) as a risk locus associated with late-onset AD [73]. Furthermore, a recent study reported a seven-fold increase in total brain HS in AD compared to controls and other tauopathies [74]. These findings collectively suggest that dysfunction in HS and HBPs in brain, CSF and plasma may play a central role in the etiology and the clinicopathological presentation of AD.

We also identified a significant number of novel plasma proteins that exhibited increased levels in AD. Among these proteins were the proteoglycan biglycan (BGN), which is typically found in the extracellular matrix of blood vessels [75], and Endothelial Cell–Specific Molecule 1 (ESM1), also known as endocan. Both proteins play roles in regulating endothelial cell function and angiogenesis, and they have been implicated in processes related to inflammation and vascular disease [75, 76]. Furthermore, we observed an elevation in Colony-Stimulating Factor 1 (CSF1) in the plasma of individuals with AD. CSF1 primarily functions in the regulation of the immune system [77]. Elevated levels of CSF1 in plasma have been associated with various diseases and conditions, including inflammation, cancer, and certain autoimmune disorders [78, 79]. Taken together, these findings suggest that widespread systemic inflammation may also manifest in the plasma of individuals with AD. However, whether this phenomenon is specific to AD or extends to other subtypes of dementia remains a subject that requires further investigation.

In this study, we also leveraged the consensus brain protein co-expression network to explore the relationship between the Hp-enriched plasma and brain proteomes in AD. Within some brain network modules, certain plasma proteins consistently exhibited increases or decreases in AD that mirrored changes in the brain, while others displayed divergent alterations as previously described [30]. Among those consistently increased, beyond M42 members, were proteins mapping to M26 in the brain, associated with the 'Acute phase response,' which suggests that proteins related to complement activation potentially associated with immune function are enriched in AD plasma. Notably, proteins of interest in M26 included SERPINA3, which was recently identified through a large-scale analysis of the plasma proteome using Mendelian randomization as potentially causal in AD pathogenesis [80]. Additionally, we observed proteins in plasma that were mapped to neuronal modules related to synaptic biology in the brain, which displayed consistent decreases in AD. Whether this change in plasma reflects synapse loss in the brain will require further investigation. Nevertheless, it is intriguing that a signature of synaptic loss typically associated with cognitive decline in AD brain appears in the Hp-enriched

AD plasma. There was also some discordance in the direction of change between AD plasma and brain proteomes. For instance, proteins specifically in M7 'MAPK signaling' and M25 'Sugar metabolism,' exhibited increased levels in the AD brain, yet decreased levels in the Hp-enriched AD plasma. This trend contrasts with the direction of change observed in CSF [30], where glycolytic signature is increased in AD even in the preclinical phase [14, 30, 50]. The precise mechanisms underlying this discordance remain unclear. However, it is plausible that blood–brain barrier dysfunction might contribute to these differences [81], where proteins increased in plasma are decreased in CSF, and vice versa.

Although we successfully captured heparin-binding proteins (HBPs) in plasma, future studies employing high salt concentrations in the washing steps could further enhance the specificity of enrichment for HBPs from human plasma [82]. However, an unintended consequence of the current method was the clearance of highly abundant proteins such as albumin. This reduction of albumin from the Hp-enriched fraction resulted in the comprehensive coverage of nearly 3,000 plasma proteins following TMT-MS coupled with high-pH off-line fractionation. Thus, in contrast to immunodepletion methods [83], antibody-free affinity enrichment-based approaches utilizing nanoparticle, cationic/anionic or hydrophobic/hydrophilic-based strategies appear to substantially enhance plasma proteome coverage through MS-based technologies [84]. Collectively, this progress marks a step toward overcoming one of the major limitations of MS-based plasma proteomics, which is the vast dynamic range of protein abundances. Utilization of more advanced mass spectrometers, such as the Orbitrap Astral, which quantifies five times more peptides per unit time than other state-of-the-art Orbitrap mass spectrometers [85], is expected to significantly enhance the depth of proteome coverage in plasma when employing these affinity enrichment strategies. Further studies that couple the Orbitrap Astral MS platform with data-independent acquisition mass spectrometry (DIA-MS) will likely enhance scalability for larger sample throughput. This is of particular significance due to the complementary coverage of the Heparin-MS plasma proteome in contrast to the SomaScan and Olink platforms, which will further enhance the depth of the plasma proteome when measurements are integrated across platforms [30].

While this study provided a comprehensive proteomic analysis of Hp-enriched plasma from human subjects, several limitations should be acknowledged. Notably, the study participants predominantly consisted of non-Latino white individuals. Recent reports have highlighted disparities in AD prevalence, with Black and Hispanic populations showing a higher likelihood of developing AD compared to older white Americans [86–88]. Additionally, it has been observed that cognitively impaired African American individuals have lower levels of CSF tTau and pTau compared to Caucasians [36]. An important ongoing initiative of the Accelerating Medicines partnership for AD (AMP-AD) [89] is the inclusion of African American and Hispanic individuals in plasma biomarker studies. Research efforts employing heparin enrichment techniques should aim to encompass a more diverse participant population to better capture the complexities of AD across different racial and ethnic groups. Future studies that investigate the interplay between age, sex, and race within the Hp-enriched plasma proteome will yield valuable insights. Moreover, our plasma proteomics study exclusively focused on control and symptomatic individuals who were AD biomarker positive based on their tau/amyloid ratio in the CSF. Future plasma proteomic studies aimed at exploring the pre-symptomatic stages of AD before cognitive impairment manifests will be needed to identify plasma proteins that undergo early changes in the disease course. Nevertheless, this study offers a global view into the Hp-enriched plasma proteome, reinforcing a hypothesis that increased matrisome proteins are shared between the brain and blood in AD.

Conclusion

In summary, these findings provide support for the integration of a heparin enrichment method with MS-based proteomics for identifying a wide spectrum of plasma protein signatures that reflect pathological changes in the AD brain.

Abbreviations

Hp	Heparin
HS	Heparan Sulfate
HSPG	Heparan Sulfate Proteoglycan
ALB	Albumin
AD	Alzheimer's disease
Aβ	Amyloid beta
APP	Amyloid precursor protein
ADAD	Autosomal dominant AD
CSF	Cerebrospinal fluid
CV	Coefficient of variation
DP	Diluted plasma
FDR	False discovery rate
GIS	Global internal standard
GO	Gene ontology
HBPs	Heparin binding proteins
Hp-depleted FT	Heparin-depleted flowthrough
Hp-enriched	Heparin-enriched
LOAD	Late-onset AD
M42	Module 42
MoCA	Montreal Cognitive Assessment
NRGN	Neurogranin
TMT-MS	Tandem mass tag mass spectrometry
TF	Transferrin
VCP	Valosin-containing protein

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13024-024-00757-1.

Additional file 1: Supplemental Figure 1. Comparing heparin enrichment across APOE isoforms.A)Thevolcano plot shows differentially enriched proteins in the Hp-enriched fractions ($n = 3$) of pooled plasma sample compared to the DP inputs ($n = 3$). Red circles represent significant Hp-enriched proteins in plasma and examples of AD-related HBPs are highlighted, in addition to APOE isoforms. Blue symbols represent proteins significantly depleted from the Hp-enriched fractions. The significance cutoff is $p < 0.05$ (ANOVA with Tukey post-hoc correction) and fold-change > 2.B)The bar plot shows average abundance changes of APOE2, APOE3 and APOE4 isoforms across three replicates within each fraction of pooled plasma (DP input = 3, Hp-depleted FT = 3, Hp-enriched = 3). The y-axis is calculated by average $\log_2$ fold-change of protein intensity from three replicates over their geomean across all 9 samples. The significance of difference was determined by ANOVA with Tukey post-hoc correction and denoted with stars (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Supplemental Figure 2. Western blotting of thrombin and APOE in Set 1 ($n = 36$). A) Western blotting was conducted on DP input, Hp-depleted FT, and Hp-enriched fraction obtained from Set 1 samples, which included 18 control and 18 AD cases. Thrombin (~75 kDa) and APOE (~34 kDa) were the target HBPs of interest. On each gel, 2 control samples, 2 AD samples, and 1 GPS sample (see method) were loaded for each fraction. The results show that both thrombin and APOE were depleted from the DP input and Hp-depleted FT and enriched in the Hp-enriched fraction. WB: western blotting. Supplemental Figure 3. Comparison of immunodepleted and heparin-enriched plasma proteomes. A) The Set 1 (control = 18, AD = 18) underwent top 14 removal and the same TMT-MS analysis as performed on the Hp-enriched Set 1. B) The volcano plot illustrates the differential abundance of 1,129 proteins between the control and AD groups. The x-axis represents the $\log_2$ fold-change (AD vs CTL), while the y-axis represents the Student's *t*-statistic ($-\log_{10}$ *p*-value) calculated for all proteins in the pairwise comparison. Proteins significantly increased in AD ($N = 432$) are highlighted in red ($p < 0.05$), whereas those significantly decreased in AD ($N = 383$) are depicted in blue. Grey dots represent proteins with insignificant changes. C) The number and overlap of proteins quantified in the top 14 ($N = 1,129$) and Hp-enriched ($N = 2,077$) datasets with less than 50% missing values are shown. There are 1,273 proteins only detected by heparin enrichment, while the top 14 method identifies only 325 specific proteins. D) The volcano plot illustrates the differential abundance of 1,273 proteins (AD vs Control) uniquely identified in the Hp-enriched dataset. The x-axis represents the $\log_2$ fold-change (AD vs CTL), while the y-axis represents the Student's *t*-statistic ($-\log_{10}$ *p*-value) calculated for all proteins in the pairwise comparison. Proteins significantly increased in AD ($N = 339$) are highlighted in red ($p < 0.05$), whereas those significantly decreased in AD ($N = 409$) are depicted in blue. Grey dots represent proteins with insignificant changes. E) A scatter plot shows the Pearson correlation between $\log_2$ fold-change (AD vs CTL) of significantly altered proteins in both the top 14 and Hp-enriched datasets. There are 380 proteins overlapping between the two sets, showing significant changes in AD with a BH FDR-corrected *p*-value< 0.05 in both datasets. Only 5 out of 380 proteins exhibited discordant changes in AD vs Control, demonstrating a high degree of concordance ($cor = 0.93$, $p < 2.2e^{-166}$). All 215 overlapping proteins selected with a BH FDR-corrected *p*-value < 0.01 displayed consistent changes in both datasets when comparing AD and control samples, with a remarkable correlation of 0.96 ($p < 1e^{-119}$). The significance of the Pearson correlation was determined by Student's *t*-test. CTL, control; *cor*, Pearson correlation coefficient. Supplemental Figure 4. Differential protein abundance of SomaScan and Olink and their correlations with Heparin-MS.A-B) The volcano plot illustrates the differential abundance of proteins measured by SomaScan ($N = 7,284$) and Olink ($N = 979$) between the control and AD groups. The x-axis represents the $\log_2$ fold-change (AD vs CTL), while the y-axis represents the Student's *t*-statistic ($-\log_{10}$ *p*-value) calculated for all proteins in each pairwise comparison. Proteins significantly increased in AD are highlighted in red ($p < 0.05$), whereas those significantly decreased in AD are depicted in blue. Grey dots represent proteins with insignificant changes. For SomaScan, 308 proteins are increased in AD while 3,917 are decreased. For Olink, 108 proteins are increased in AD while 284 are decreased. Protein levels in AD are generally lower than in controls in SomaScan, resulting in a lower $\log_2$ fold-change (AD vs CTL) compared to the Heparin-MS data. C-D) Pearson correlation between $\log_2$ fold-change (AD vs CTL) of significant common gene products (FDR < 0.05) measured by the Heparin-MS and SomaScan ($N = 445$ gene products, $cor = 0.7$, $p = 9.1e^{-67}$), as well as the Heparin-MS and Olink ($N = 58$ gene products, $cor = 0.88$, $p = 9.4e^{-20}$). The significance of the Pearson correlation was determined by Student's *t*-test. CTL, control; *cor*, Pearson correlation coefficient. Supplemental Figure 5. Workflow of the data analysis pipeline across two Hp-enriched plasma datasets. A) Set 1 and Set 2 were jointly analyzed using FP, followed by independent batch-specific variance correction procedures, which included TAMPOR and batch-regression. Subsequently, 12 cases from Set 2 that did not meet the AT+ threshold criteria were excluded (see methods). A total of 109 samples and 2865 total proteins were selected for further analysis, with 13 overlapping control samples between the two datasets. B) Measurements of various AD-related traits, including cognition (MoCA score), CSF $A\beta_{1-42}$, CSF tTau, CSF pTau181, CSF ratio of tTau/$A\beta_{1-42}$, and plasma pTau181, were shown for the selected unique cases ($n = 96$). Significance levels determined by Student's *t*-test are denoted by overlain asterisks;* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. Supplemental Figure 6. Comparison of Set 1 and Set 2 with or without regression for age, sex, and race. A-D) Variance partition analysis is visualized by violin plots, using experimental factors to evaluate the percentage of explained variance in samples. The y-axis represents the percentage of explained variance, while the x-axis shows factors contributing to variance, including age, sex, race, group, batch, and residuals. A) Set 1 post-TAMPOR and before regression. B) Set 1 after regression. C) Set 2 post-TAMPOR and before regression. D) Set 2 after regression. Notably, variance due to age, sex, race, and batch was significantly reduced after correction, underscoring the efficacy of the correction procedure in removing trait-related variability from the proteomic data. E) A scatter plot illustrates the correlation between $\log_2$ fold-changes for AD vs CTL ($cor = 0.99$, $p < 1e^{-200}$) on significant proteins ($N = 1765$, *meta* $p < 0.05$) in the *meta*-analysis (Set 1 + Set 2), with only batch regression or after regression for age, sex, race, and batch. CTL, control; *cor*, Pearson correlation coefficient. Supplemental Figure 7. Comparing TMT-MS proteomic measurements of human brain generated by FragPipe (FP) and Proteome Discoverer (PD). A) 456 raw files collected from the ROSMAP and Banner cohorts as previously described (11) underwent a database search using FP (see methods), resulting in the identification of 8956 UniprotID-identified proteins, each with measurements available in 50% or more across 456 individual cases (control = 101, AsymAD = 181, AD = 174). These proteins were subsequently assigned to one of the 44 pre-existing consensus network modules (11) by re-calculating the kME (bicor correlation to module eigenprotein) for each protein and assigning it to the module that exhibited the highest correlation (52). B) The number and overlap of unique gene products identified in the FP and the PD outputs, with 7476 overlapping between the two datasets. FP provided an additional 1428 unique gene products compared to PD, resulting in an 18% increase in proteome coverage. C) Scatter plots illustrate the correlation between $\log_2$ fold-change for AD vs CTL (left, $cor = 0.9$, $p < 1e^{-200}$) and AsymAD vs CTL (right, $cor = 0.84$, $p < 1e^{-200}$), using common gene products ($N = 7467$) found in both FP and PD search results. CTL, control; *cor*, Pearson correlation of coefficient. Supplemental Figure 8. Overlap between additional human brain network modules and differentially abundant Hp-enriched plasma proteins in AD. Protein expression trends are examined for the 10 modules that exhibit significant overlap with differentially abundant Hp-enriched plasma proteome but demonstrate moderate to low correlation with AD clinicopathological traits in the brain. Brain module abundance is quantified by eigenprotein values derived from the consensus brain dataset (11) (control = 101, AsymAD = 181, AD = 174), while volcano plots illustrate the differential abundance ($\log_2$ AD vs CTL) of module proteins overlapped with the Hp-enriched plasma proteome. The statistical significance of changes in module eigenprotein abundance

across the three groups in the consensus brain cohort was assessed using ANOVA with Tukey post-hoc correction. Modules with $p < 0.05$ were considered significant. Among these modules, M4 'Synaptic transmission', M34 'GTP binding', M14 'Protein folding', M37 'Endosome', M30'Proteasome', M43 'Ribonucleoprotein binding' and M38 'Heat shock/protein folding' consist of proteins with decreased abundance in AD plasma, while M40'Antigen binding' exclusively contains increased plasma proteins in AD. M27'ECM' and M21 'MHC/immune' exhibit a balanced representation of both increased and decreased plasma proteins in AD. CTL, control; AsymAD, asymptomatic AD.

Additional file 2: Table S1 : Protein abundance and CVs of 3 fractions of pooled plasma. Table S2: Protein overlap across 3 fractions of pooled plasma. Table S3: Pooled plasma ANOVA. Table S4: Set 1 sample information ($n = 36$). Table S5: (Set 1 + Set 2) raw protein abundance ($n = 121$). Table S6: Set 1 variance-corrected protein abundance. Table S7: Set 1 Student's *t*-test (heparin enrichment). Table S8: Set 1 Student's *t*-test (top 14 removal). Table S9: Protein overlap between heparin enrichment and top 14 removal. Table S10: SomaScan median-centered data Table S11: Olink median-centered data. Table S12: Protein overlap across Heparin-MS, SomaScan and Olink. Table S13: SomaScan Student's *t*-test. Table S14: Olink Student's *t*-test. Table S15: Set 2 sample information ($n = 85$). Table S16: Set 2 variance-corrected protein abundance. Table S17: Plasma protein concentration. Table S18: Protein overlap between Set 1 and Set 2. Table S19: Set 2 Student's *t*-test. Table S20: (Set 1 + Set 2) sample information without outliers ($n = 109$). Table S21: (Set 1 + Set 2) *meta*-analysis of significance. Table S22: (Set 1 + Set 2) Z-transformed protein abundance. Table S23: Correlations between Z-score and AD biomarkers. Table S24: Correlations between SomaScan and AD biomarkers. Table S25: Correlations between Olink and AD biomarkers. Table S26: (Set 1 + Set 2) *meta*-analysis of significance after regression for age, sex, race and batch. Table S27: ROC curve statistics for highly significant plasma proteins and plasma pTau181 ($n = 95$, duplicated samples and missing values removed) .Table S28: (ROSMAP + BANNER) brain sample information ($n = 456$). Table S29: (ROSMAP + BANNER) variance-corrected protein abundance. Table S30: Protein overlap between FragPipe and Proteome Discoverer search of (ROSMAP + BANNER) brain dataset. Table S31: (ROSMAP + BANNER) ANOVA from FragPipe search. Table S32: (ROSMAP + BANNER) ANOVA from Proteome Discoverer search. Table S33: (ROSMAP + BANNER) module re-assignment. Table S34: Protein overlap between Hp-enriched plasma and brain datasets.

Acknowledgements
We thank the Seyfried lab and members of the Emory Center for Neurodegenerative Disease (CND) for helpful comments and discussions.

Authors' contributions
Conceptualization: DMD, EJF. and NTS Methodology: QG, LP, YL, KX, DMD, BRR, JJL and NTS Investigation: QG, LP, EBD, JJL, AIL, and NTS Formal Analysis: QG, EBD, AS Writing – Original Draft: Q.G., L.P., D.M.D., and N.T.S.; Writing – Review & Editing: QG, LP, EJF, DMD, EBD, ECBJ, JJL, AIL, and NTS Funding Acquisition: JJL, AIL and NTS Resources: JJL, AIL Supervision: AIL, and NTS.

Funding
This study was supported by the following National Institutes of Health (NIH) funding mechanisms: U01AG061357 (A.I.L and N.T.S), RF1AG062181 (N.T.S.) and P30AG066511 (A.I.L) and the Foundations for the National Institute of Health AMP-AD 2.0 grant.

Availability data and materials
Raw mass spectrometry data and pre- and post-processed plasma protein abundance data and case traits related to this manuscript are available at https://www.synapse.org/HeparinPlasma. The results published here are in whole or in part based on data obtained from the AMP-AD Knowledge Portal (https://adknowledgeportal.synapse.org). The AMP-AD Knowledge Portal is a platform for accessing data, analyses and tools generated by the AMP-AD Target Discovery Program and other programs supported by the National Institute on Aging to enable open-science practices and accelerate translational learning. The data, analyses and tools are shared early in the research cycle without a publication embargo on secondary use. Data are available for general research use according to the following requirements for data access and data attribution (https://adknowledgeportal.synapse.org/#/DataAccess/Instructions). All data are available in the main text or the supplementary materials.

Declarations

Ethics approval and consent to participate
Local institutional review boards approved the procedures for this study and written informed consent was obtained.

Consent for publication
Not applicable.

Competing interests
A.I.L, N.T.S. and D.M.D. are co-founders of Emtherapro Inc.

Author details
[1]Department of Biochemistry, School of Medicine, Emory School of Medicine, 505J Whitehead Biomedical Research Building, 615 Michael St, Atlanta, GA 30322, USA. [2]Center for Neurodegenerative Disease Center, Emory University School of Medicine, Atlanta, GA 30322, USA. [3]Goizueta Alzheimer's Disease Research Center, Emory University School of Medicine, Atlanta, GA 30322, USA. [4]Department of Neurology, Emory University School of Medicine, Atlanta, GA 30322, USA. [5]Department of Pharmacology and Chemical Biology, School of Medicine, Emory University, Atlanta, GA 30322, USA.

Received: 5 February 2024 Accepted: 24 September 2024
Published online: 08 October 2024

References

1. Masters CL, Simms G, Weinman NA, Multhaup G, McDonald BL, Beyreuther K. Amyloid plaque core protein in Alzheimer disease and down syndrome. Proc Natl Acad Sci U S A. 1985;82(12):4245–9.
2. Long JM, Holtzman DM. Alzheimer disease: an update on pathobiology and treatment strategies. Cell. 2019;179(2):312–39.
3. Grundke-Iqbal I, Iqbal K, Tung YC, Quinlan M, Wisniewski HM, Binder LI. Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. Proc Natl Acad Sci U S A. 1986;83(13):4913–7.
4. Dubois B, Hampel H, Feldman HH, Scheltens P, Aisen P, Andrieu S, et al. Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria. Alzheimers Dement. 2016;12(3):292–323.
5. Hampel H, Hu Y, Cummings J, Mattke S, Iwatsubo T, Nakamura A, et al. Blood-based biomarkers for Alzheimer's disease: current state and future use in a transformed global healthcare landscape. Neuron. 2023;111(18):2781–99.
6. Gonzalez-Ortiz F, Kac PR, Brum WS, Zetterberg H, Blennow K, Karikari TK. Plasma phospho-tau in Alzheimer's disease: towards diagnostic and therapeutic trial applications. Mol Neurodegener. 2023;18(1):18.
7. Ashton NJ, Janelidze S, Mattsson-Carlgren N, Binette AP, Strandberg O, Brum WS, et al. Differential roles of Abeta42/40, p-tau231 and p-tau217 for Alzheimer's trial selection and disease monitoring. Nat Med. 2022;28(12):2555–62.
8. Haque R, Watson CM, Liu J, Carter EK, Duong DM, Lah JJ, et al. A protein panel in cerebrospinal fluid for diagnostic and predictive assessment of Alzheimer's disease. Sci Transl Med. 2023;15(712):eadg4122.
9. De Strooper B, Karran E. The cellular phase of Alzheimer's disease. Cell. 2016;164(4):603–15.
10. Rayaprolu S, Higginbotham L, Bagchi P, Watson CM, Zhang T, Levey AI, et al. Systems-based proteomics to resolve the biology of Alzheimer's disease beyond amyloid and tau. Neuropsychopharmacology. 2021;46(1):98–115.
11. Johnson ECB, Carter EK, Dammer EB, Duong DM, Gerasimov ES, Liu Y, et al. Large-scale deep multi-layer analysis of Alzheimer's disease brain reveals strong proteomic disease-related changes not observed at the RNA level. Nat Neurosci. 2022;25(2):213–25.
12. Johnson ECB, Dammer EB, Duong DM, Ping LY, Zhou MT, Yin LM, 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. Nat Med. 2020;26(5):769-+.
13. Zhang B, Gaiteri C, Bodea LG, Wang Z, McElwee J, Podtelezhnikov AA, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. Cell. 2013;153(3):707–20.
14. Higginbotham L, Ping L, Dammer EB, Duong DM, Zhou M, Gearing M, et al. Integrated proteomics reveals brain-based cerebrospinal fluid biomarkers in asymptomatic and symptomatic Alzheimer's disease. Sci adv. 2020;6(43):eaaz9360.
15. Sung YJ, Yang C, Norton J, Johnson M, Fagan A, Bateman RJ, et al. Proteomics of brain, CSF, and plasma identifies molecular signatures for distinguishing sporadic and genetic Alzheimer's disease. Sci Transl Med. 2023;15(703):eabq5923.
16. van der Ende EL, In 't Veld S, Hanskamp I, van der Lee S, Dijkstra JIR, Hok AHYS, et al. CSF proteomics in autosomal dominant Alzheimer's disease highlights parallels with sporadic disease. Brain. 2023;146(11):4495–07.
17. Johnson ECB, Bian S, Haque RU, Carter EK, Watson CM, Gordon BA, et al. Cerebrospinal fluid proteomics define the natural history of autosomal dominant Alzheimer's disease. Nat Med. 2023;29(8):1979–88.
18. Shao X, Gomez CD, Kapoor N, Considine JM, Grams C, Gao YT, et al. MatrisomeDB 2.0: 2023 updates to the ECM-protein knowledge database. Nucleic Acids Res. 2023;51(D1):D1519–30.
19. Levites Y, Dammer EB, Ran Y, Tsering W, Duong D, Abreha M, et al. Abeta amyloid scaffolds the accumulation of matrisome and additional proteins in Alzheimer's disease. BioRxiv. 2023;2023:11.29.568318.
20. Bai B, Wang XS, Li YX, Chen PC, Yu KW, Dey KK, et al. Deep Multilayer brain proteomics identifies molecular networks in Alzheimer's disease progression (vol 105, pg 975, 2020). Neuron. 2020;106(4):700.
21. Drummond E, Kavanagh T, Pires G, Marta-Ariza M, Kanshin E, Nayak S, et al. The amyloid plaque proteome in early onset Alzheimer's disease and Down syndrome. Acta Neuropathol Commun. 2022;10(1):53.
22. Kim J, Basak JM, Holtzman DM. The role of apolipoprotein E in Alzheimer's disease. Neuron. 2009;63(3):287–303.
23. Muramatsu T. Midkine, a heparin-binding cytokine with multiple roles in development, repair and diseases. Proc Jpn Acad Ser B Phys Biol Sci. 2010;86(4):410–25.
24. Saito H, Dhanasekaran P, Nguyen D, Baldwin F, Weisgraber KH, Wehrli S, et al. Characterization of the heparin binding sites in human apolipoprotein E. J Biol Chem. 2003;278(17):14782–7.
25. Klemencic M, Novinec M, Maier S, Hartmann U, Lenarcic B. The heparin-binding activity of secreted modular calcium-binding protein 1 (SMOC-1) modulates its cell adhesion properties. PLoS ONE. 2013;8(2): e56839.
26. McLaurin J, Franklin T, Zhang X, Deng J, Fraser PE. Interactions of Alzheimer amyloid-beta peptides with glycosaminoglycans effects on fibril nucleation and growth. Eur J Biochem. 1999;266(3):1101–10.
27. Iannuzzi C, Irace G, Sirangelo I. The effect of glycosaminoglycans (GAGs) on amyloid aggregation and toxicity. Molecules. 2015;20(2):2510–28.
28. van Horssen J, Wesseling P, van den Heuvel LP, de Waal RM, Verbeek MM. Heparan sulphate proteoglycans in Alzheimer's disease and amyloid-related disorders. Lancet Neurol. 2003;2(8):482–92.
29. Tcw J, Qian L, Pipalia NH, Chao MJ, Liang SA, Shi Y, et al. Cholesterol and matrisome pathways dysregulated in astrocytes and microglia. Cell. 2022;185(13):2213–33 e25.
30. Dammer EB, Ping L, Duong DM, Modeste ES, Seyfried NT, Lah JJ, et al. Multi-platform proteomic analysis of Alzheimer's disease cerebrospinal fluid and plasma reveals network biomarkers associated with proteostasis and the matrisome. Alzheimers Res Ther. 2022;14(1):174.
31. Anderson NL, Anderson NG. The human plasma proteome - history, character, and diagnostic prospects. Mol Cell Proteomics. 2002;1(11):845–67.
32. Raoufinia R, Mota A, Keyhanvar N, Safari F, Shamekhi S, Abdolalizadeh J. Overview of albumin and its purification methods. Adv Pharm Bull. 2016;6(4):495–507.
33. Dammer EB, Fallini C, Gozal YM, Duong DM, Rossoll W, Xu P, et al. Coaggregation of RNA-binding proteins in a model of TDP-43 proteinopathy with selective RGG motif methylation and a role for RRM1 ubiquitination. PLoS ONE. 2012;7(6): e38658.
34. Diner I, Hales CM, Bishof I, Rabenold L, Duong DM, Yi H, et al. Aggregation properties of the small nuclear ribonucleoprotein U1–70K in Alzheimer disease. J Biol Chem. 2014;289(51):35296–313.
35. Seyfried NT, Gozal YM, Donovan LE, Herskowitz JH, Dammer EB, Xia Q, et al. Quantitative analysis of the detergent-insoluble brain proteome in frontotemporal lobar degeneration using SILAC internal standards. J Proteome Res. 2012;11(5):2721–38.
36. Modeste ES, Ping L, Watson CM, Duong DM, Dammer EB, Johnson ECB, et al. Quantitative proteomics of cerebrospinal fluid from African Americans and Caucasians reveals shared and divergent changes in Alzheimer's disease. Mol Neurodegener. 2023;18(1):48.
37. Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. Nat Methods. 2017;14(5):513–20.
38. Yu F, Teo GC, Kong AT, Haynes SE, Avtonomov DM, Geiszler DJ, et al. Identification of modified peptides using localization-aware open search. Nat Commun. 2020;11(1):4065.
39. Kall L, Canterbury JD, Weston J, Noble WS, MacCoss MJ. Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nat Methods. 2007;4(11):923–5.
40. da Veiga LF, Haynes SE, Avtonomov DM, Chang H-Y, Shanmugam AK, Mellacheruvu D, et al. Philosopher: a versatile toolkit for shotgun proteomics data analysis. Nat Methods. 2020;17(9):869–70.
41. Yu F, Haynes SE, Nesvizhskii AI. IonQuant enables accurate and sensitive label-free quantification with FDR-controlled match-between-runs. Mol Cell Proteomics. 2021;20: 100077.
42. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods. 2016;13(9):731–40.
43. He T, Liu Y, Zhou Y, Li L, Wang H, Chen S, et al. Comparative evaluation of proteome discoverer and FragPipe for the TMT-Based Proteome Quantification. J Proteome Res. 2022;21(12):3007–15.
44. Zecha J, Satpathy S, Kanashova T, Avanessian SC, Kane MH, Clauser KR, et al. TMT labeling for the masses: a robust and cost-efficient in-solution labeling approach. Mol Cell Proteomics. 2019;18(7):1468–78.
45. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem. 2003;75(17):4646–58.
46. Dammer EB, Seyfried NT, Johnson ECB. Batch correction and harmonization of -Omics datasets with a tunable median polish of ratio. Front Syst Biol. 2023;3:1092341.
47. Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat Protoc. 2019;14(2):482–517.
48. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. BMC Bioinformatics. 2010;11: 367.
49. Jack CR Jr, Bennett DA, Blennow K, Carrillo MC, Feldman HH, Frisoni GB, et al. A/T/N: an unbiased descriptive classification scheme for Alzheimer disease biomarkers. Neurology. 2016;87(5):539–47.
50. Watson CM, Dammer EB, Ping L, Duong DM, Modeste E, Carter EK, et al. Quantitative mass spectrometry analysis of cerebrospinal fluid protein biomarkers in Alzheimer's disease. Sci Data. 2023;10(1):261.
51. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837–45.
52. Johnson ECB, Dammer EB, Duong DM, Yin L, Thambisetty M, Troncoso JC, et al. Deep proteomic network analysis of Alzheimer's disease brain reveals alterations in RNA binding proteins and RNA splicing associated with disease. Mol Neurodegener. 2018;13(1):52.
53. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keeffe S, et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J Neurosci. 2014;34(36):11929–47.
54. Wang M, You J, Bemis KG, Tegeler TJ, Brown DP. Label-free mass spectrometry-based protein quantification technologies in proteomic analysis. Brief Funct Genomic Proteomic. 2008;7(5):329–39.
55. Higgs RE, Knierman MD, Gelfanova V, Butler JP, Hale JE. Comprehensive label-free method for the relative quantification of proteins from biological samples. J Proteome Res. 2005;4(4):1442–50.
56. Small DH, Nurcombe V, Reed G, Clarris H, Moir R, Beyreuther K, et al. A heparin-binding domain in the amyloid protein precursor of Alzheimer's disease is involved in the regulation of neurite outgrowth. J Neurosci. 1994;14(4):2117–27.
57. Simon R, Girod M, Fonbonne C, Salvador A, Clement Y, Lanteri P, et al. Total ApoE and ApoE4 isoform assays in an Alzheimer's disease

case-control study by targeted mass spectrometry (n=669): a pilot assay for methionine-containing proteotypic peptides. Mol Cell Proteomics. 2012;11(11):1389–403.
58. Arboleda-Velasquez JF, Lopera F, O'Hare M, Delgado-Tirado S, Marino C, Chmielewska N, et al. Resistance to autosomal dominant Alzheimer's disease in an APOE3 Christchurch homozygote: a case report. Nat Med. 2019;25(11):1680–3.
59. Bader JM, Albrecht V, Mann M. MS-based proteomics of body fluids: the end of the beginning. Mol Cell Proteomics. 2023;22(7): 100577.
60. Cao X, Sandberg A, Araujo JE, Cvetkovski F, Berglund E, Eriksson LE, et al. Evaluation of spin columns for human plasma depletion to facilitate MS-based proteomics analysis of plasma. J Proteome Res. 2021;20(9):4610–20.
61. Candia J, Daya GN, Tanaka T, Ferrucci L, Walker KA. Assessment of variability in the plasma 7k SomaScan proteomics assay. Sci Rep. 2022;12(1):17147.
62. Suarez-Calvet M, Karikari TK, Ashton NJ, Lantero Rodriguez J, Mila-Aloma M, Gispert JD, et al. Novel tau biomarkers phosphorylated at T181, T217 or T231 rise in the initial stages of the preclinical Alzheimer's continuum when only subtle changes in Abeta pathology are detected. EMBO Mol Med. 2020;12(12): e12921.
63. Bayoumy S, Verberk IMW, den Dulk B, Hussainali Z, Zwan M, van der Flier WM, et al. Clinical and analytical comparison of six Simoa assays for plasma P-tau isoforms P-tau181, P-tau217, and P-tau231. Alzheimers Res Ther. 2021;13(1):198.
64. Vannahme C, Gösling S, Paulsson M, Maurer P, Hartmann U. Characterization of SMOC-2, a modular extracellular calcium-binding protein. Biochem J. 2003;373(Pt 3):805–14.
65. Higginbotham L, Carter EK, Dammer EB, Haque RU, Johnson ECB, Duong DM, et al. Unbiased classification of the elderly human brain proteome resolves distinct clinical and pathophysiological subtypes of cognitive impairment. Neurobiol Dis. 2023;186: 106286.
66. Hansson O, Blennow K, Zetterberg H, Dage J. Blood biomarkers for Alzheimer's disease in clinical practice and trials. Nature Aging. 2023;3(5):506–19.
67. Whelan CD, Mattsson N, Nagle MW, Vijayaraghavan S, Hyde C, Janelidze S, et al. Multiplex proteomics identifies novel CSF and plasma biomarkers of early Alzheimer's disease. Acta Neuropathol Commun. 2019;7(1):169.
68. Lopera F, Marino C, Chandrahas AS, O'Hare M, Villalba-Moreno ND, Aguillon D, et al. Resilience to autosomal dominant Alzheimer's disease in a Reelin-COLBOS heterozygous man. Nat Med. 2023;29(5):1243–52.
69. Snow AD, Cummings JA, Lake T. The Unifying hypothesis of Alzheimer's disease: heparan sulfate proteoglycans/glycosaminoglycans are key as first hypothesized over 30 years ago. Front Aging Neurosci. 2021;13: 710683.
70. Holmes BB, DeVos SL, Kfoury N, Li M, Jacks R, Yanamandra K, et al. Heparan sulfate proteoglycans mediate internalization and propagation of specific proteopathic seeds. Proc Natl Acad Sci U S A. 2013;110(33):E3138–47.
71. Saroja SR, Gorbachev K, Julia T, Goate AM, Pereira AC. Astrocyte-secreted glypican-4 drives APOE4-dependent tau hyperphosphorylation. Proc Natl Acad Sci U S A. 2022;119(34): e2108870119.
72. Rauch JN, Chen JJ, Sorum AW, Miller GM, Sharf T, See SK, et al. Tau internalization is regulated by 6-O Sulfation on Heparan Sulfate Proteoglycans (HSPGs). Sci Rep. 2018;8(1):6382.
73. Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet. 2019;51(3):404–13.
74. Wang Z, Patel VN, Song X, Xu Y, Kaminski AM, Doan VU, et al. Increased 3-O-sulfated heparan sulfate in Alzheimer's disease brain is associated with genetic risk gene HS3ST1. Sci Adv. 2023;9(21):eadf6232.
75. Shimizu-Hirota R, Sasamura H, Kuroda M, Kobayashi E, Hayashi M, Saruta T. Extracellular matrix glycoprotein biglycan enhances vascular smooth muscle cell proliferation and migration. Circ Res. 2004;94(8):1067–74.
76. Sarrazin S, Adam E, Lyon M, Depontieu F, Motte V, Landolfi C, et al. Endocan or endothelial cell specific molecule-1 (ESM-1): a potential novel endothelial cell marker and a new target for cancer therapy. Biochim Biophys Acta. 2006;1765(1):25–37.
77. Stanley ER, Chitu V. CSF-1 receptor signaling in myeloid cells. Cold Spring Harb Perspect Biol. 2014;6(6):a021857.
78. West RB, Rubin BP, Miller MA, Subramanian S, Kaygusuz G, Montgomery K, et al. A landscape effect in tenosynovial giant-cell tumor from activation of CSF1 expression by a translocation in a minority of tumor cells. Proc Natl Acad Sci U S A. 2006;103(3):690–5.
79. Lin W, Xu D, Austin CD, Caplazi P, Senger K, Sun Y, et al. Function of CSF1 and IL34 in macrophage homeostasis, inflammation, and cancer. Front Immunol. 2019;10: 2019.
80. Walker KA, Chen J, Shi L, Yang Y, Fornage M, Zhou L, et al. Proteomics analysis of plasma from middle-aged adults identifies protein markers of dementia risk in later life. Sci Transl Med. 2023;15(705):eadf5681.
81. Bowman GL, Dayon L, Kirkland R, Wojcik J, Peyratout G, Severin IC, et al. Blood-brain barrier breakdown, neuroinflammation, and cognitive decline in older adults. Alzheimers Dement. 2018;14(12):1640–50.
82. Sandoval DR, Gomez Toledo A, Painter CD, Tota EM, Sheikh MO, West AMV, et al. Proteomics-based screening of the endothelial heparan sulfate interactome reveals that C-type lectin 14a (CLEC14A) is a heparin-binding protein. J Biol Chem. 2020;295(9):2804–21.
83. Gillette MA, Carr SA. Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. Nat Methods. 2013;10(1):28–34.
84. Lee WC, Lee KH. Applications of affinity chromatography in proteomics. Anal Biochem. 2004;324(1):1–10.
85. Heil LR, Damoc E, Arrey TN, Pashkova A, Denisov E, Petzoldt J, et al. Evaluating the performance of the astral mass analyzer for quantitative proteomics using data-independent acquisition. J Proteome Res. 2023;22(10):3290–300.
86. Potter GG, Plassman BL, Burke JR, Kabeto MU, Langa KM, Llewellyn DJ, et al. Cognitive performance and informant reports in the diagnosis of cognitive impairment and dementia in African Americans and whites. Alzheimers Dement. 2009;5(6):445–53.
87. Rajan KB, Weuve J, Barnes LL, Wilson RS, Evans DA. Prevalence and incidence of clinically diagnosed Alzheimer's disease dementia from 1994 to 2012 in a population study. Alzheimers Dement. 2019;15(1):1–7.
88. Gurland BJ, Wilder DE, Lantigua R, Stern Y, Chen J, Killeffer EH, et al. Rates of dementia in three ethnoracial groups. Int J Geriatr Psychiatry. 1999;14(6):481–93.
89. Hodes RJ, Buckholtz N. Accelerating medicines partnership: Alzheimer's disease (AMP-AD) knowledge portal aids Alzheimer's drug discovery through open data sharing. Expert Opin Ther Targets. 2016;20(4):389–91.

Publisher's Note

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