medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Multiomics profiling of human plasma and CSF reveals ATN derived networks and highlights causal links in Alzheimer's disease

Liu Shi[a+], Jin Xu[b+], Rebecca Green[c,d,e], Asger Wretlind[f], Jan Homann[g], Noel J. Buckley[a], Betty M. Tijms[h], Stephanie J. B. Vos[i], Christina M. Lill[g,j,k], Mara ten Kate[h], Sebastiaan Engelborghs[l,m], Kristel Sleegers[n,o], Giovanni B. Frisoni[p,q], Anders Wallin[r], Alberto Lleó[s], Julius Pop[t,u], Pablo Martinez-Lage[v], Johannes Streffer[w], Frederik Barkhof[x,y], Henrik Zetterberg[z,aa,bb,cc], Pieter Jelle Visser[h,i], Simon Lovestone[a,dd], Lars Bertram[g,ee], Alejo J. Nevado-Holgado[a], Petroula Proitsi[c§], Cristina Legido-Quigley[b,f*§]

[a] Department of Psychiatry, University of Oxford, UK
[b] Institute of Pharmaceutical Science, King's College London, London, UK
[c] Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[d] UK National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre, South London and Maudsley Trust, London, UK
[e] MRC Unit for Lifelong Health & Ageing at UCL, University College London, London, UK
[f] Steno Diabetes Center Copenhagen, Gentofte, Denmark
[g] Lübeck Interdisciplinary Platform for Genome Analytics, University of Lübeck, Lübeck, Germany
[h] Alzheimer Center, VU University Medical Center, Amsterdam, the Netherlands
[i] Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Alzheimer Centrum Limburg, Maastricht University, Maastricht, the Netherlands
[j] Institute of Epidemiology and Social Medicine, University of Muenster, Muenster, Germany
[K] Ageing Epidemiology Research Unit (AGE), School of Public Health, Imperial College London, London, UK
[l] Reference Center for Biological Markers of Dementia (BIODEM), Institute Born-Bunge, University of Antwerp, Antwerp, Belgium
[m] Department of Neurology, UZ Brussel and Center for Neurociences (C4N), Vrije Universiteit Brussel, Brussels, Belgium
[n] Complex Genetics Group, VIB Center for Molecular Neurology, VIB, Antwerp, Belgium
[o] Institute Born-Bunge, Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium
[p] University of Geneva, Geneva, Switzerland
[q] IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy
[r] Institute of Neuroscience and Physiology, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden
[s] Neurology Department, Hospital Sant Pau, Barcelona, Spain. Centro de Investigación en Red en enfermedades neurodegenerativas (CIBERNED).
[t] University Hospital of Lausanne, Lausanne, Switzerland
[u] Geriatric Psychiatry, Department of Mental Health and Psychiatry, Geneva university Hospitals, Geneva, Switzerland
[v] CITA-Alzheimer Foundation, San Sebastian, Spain
[w] AC Immune SA, Lausanne, Switzerland, formerly Janssen R&D, LLC. Beerse, Belgium at the time of study conduct
[x] Department of Radiology and Nuclear Medicine, Amsterdam UMC, Vrije Universiteit, The Netherland
[y] Queen Square Institute of Neurology and Centre for Medical Image Computing, University College London, UK
[z] Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden
[aa] Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Mölndal, Sweden

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[bb] UK Dementia Research Institute at UCL, London, United Kingdom
[cc] Department of Neurodegenerative Disease, UCL Institute of Neurology, London, United Kingdom
[dd] Janssen Medical (UK), High Wycombe, UK
[ee] Department of Psychology, University of Oslo, Oslo, Norway

+ Contributed equally
§ Senior authors with equal contribution
***correspondence:**
Cristina Legido-Quigley: cristina.legido_quigley@kcl.ac.uk

Abstract (150 words limit)

INTRODUCTION: This study employed an integrative system and causal inference approach to explore molecular signatures in blood and CSF, the Amyloid/Tau/Neurodegeneration [AT(N)] framework, MCI conversion to AD, and genetic risk for AD.

METHODS: Using the EMIF-AD MBD cohort, we measured 696 proteins in cerebrospinal fluid (n=371), 4001 proteins in plasma (n=972), 611 metabolites in plasma (n=696) and genotyped data in whole-blood (7,778,465 autosomal SNPs, n=936). We investigated associations: molecular modules to AT(N), module hubs with AD Polygenic Risk scores and APOE4 genotypes, molecular hubs to MCI conversion and probed for causality with AD using Mendelian Randomization (MR).

RESULTS: AT(N) framework associated key hubs were mostly proteins and few lipids. In MR analyses, Proprotein Convertase Subtilisin/Kexin Type 7 showed weak causal associations with AD, and AD was causally associated with Reticulocalbin 2 and sphingomyelins.

DISCUSSION: This study reveals multi-omics networks associated with AT(N) and MCI conversion and highlights AD causal candidates.

Key words: Alzheimer's disease; multi-omics; AT(N) framework; polygenic risk score; Mendelian randomization; multimodal biomarker

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

1. Introduction

Alzheimer's disease (AD) is characterised by the presence of β-amyloid (Aβ) containing plaques, and neurofibrillary tangles composed of modified tau protein together with the progressive loss of synapses and neurons [1]. The National Institute on Aging and Alzheimer's Association (NIA-AA) have proposed to classify AD based on biomarkers of amyloid pathology (A), tau pathology (T), and neurodegeneration (N) (the ATN framework) [2]. Yet, despite their diagnostic utility, these three markers reflect only a portion of the complex pathophysiology of AD. In prodromal stages, the interplay between AT(N) changes, genetic factors and peripheral molecular changes may affect the rate of disease progression.

Conducting unbiased and high-throughput omics-based research in biological fluids and human brain tissues provides a data-driven approach to identify the many processes involved in AD pathogenesis and to prioritize links to relevant clinical and neuropathological traits. For example, an increasing number of proteomics studies [3-5], including ours [6-8], have identified AD pathophysiological pathways related to immune response and inflammation, oxidative stress, energy metabolism and mitochondrial function. Metabolomics studies have also identified such pathways related to AD [9-11]. A combination of omics, also called multi-omics or deep phenotyping studies, provides an opportunity to explore the molecular interplay with both genotypic and phenotypic variability in AD, bringing in new findings and uncovering novel pathways. Finally, causal inferences approaches allow to scrutinize the causal relationship between molecular markers and AD, highlighting potential interventional targets. Therefore, in this study, we conducted multi-omics analyses with four modalities (cerebrospinal fluid [CSF] proteomics, plasma proteomics, plasma metabolomics and whole blood genetics) from the EMIF-AD MBD study, followed by Mendelian Randomization (MR) analyses (**Figure 1**).

We have four objectives: Firstly, we wanted to test if proteomic and metabolomic molecular signatures were associated with AD endophenotypes including amyloid, CSF total tau (T-tau), CSF

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

phosphorylated tau (P-tau), white matter hyperintensity volume, CSF YKL-40, mini mental state examination (MMSE) score, and mild cognitive impairment (MCI) conversion. Secondly, we wanted to investigate the associations between molecular signatures and molecular hubs (main molecules driving associations) with APOE4 genotypes and AD polygenic risk scores (PRS). Thirdly, we wanted to query our findings in prodromal AD by extracting and integrating hub molecules in MCI individuals that converted to AD by computing a network for MCI converters versus non-converters. Finally, MR analyses interrogated the causal relationship between hub molecules and AD.

2. Methods

2.1. Participants: EMIF-AD Multimodal Biomarker Discovery (MBD) study

The EMIF-AD MBD study is part of the European Medical Information Framework for Alzheimer's disease (EMIF; http://www.emif.eu/emif-ad-2/); a public-private partnership funded through the Innovative Medicines Initiative (IMI). The design of the EMIF-AD MBD study has been described previously [12]. Briefly, 1221 samples from three groups of people (cognitively normal controls [CTL], MCI and AD) were chosen from pre-existing cohorts with the goal of including samples from people with pathology as well as those without. All participating centres have agreed to share data as part of the EMIF-AD MBD study.

General clinical and demographic information were available for all subjects (including *APOE* ε4 genotype data). Furthermore, each participant had a measure of brain amyloid load, using either CSF Aβ or amyloid positron emission tomography (PET) imaging. CSF T-tau and P-tau analysis data were available for over 90% of the subjects. We used CSF (or where not available, PET) amyloid as "A", CSF P-tau 181 as "T" and CSF T-tau as "N" to define the AT(N) framework. The classification of the status (abnormal/normal) of amyloid, P-tau and T-tau has been described previously [12]. We dichotomized these biomarkers as normal or abnormal and categorized them into four groups: no

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

pathology (A-T-N-, referring as "A-TN-"), amyloid positive but both T and N negative (A+T-N-, referring as "A+TN-"), amyloid positive and T/N positive (including A+T-N+, A+T+N- and A+T+N+, referring as "A+TN+") and Suspected Non-Alzheimer Pathology (SNAP, including A-T-N+, A-T+N- and A-T+N+). In addition, the following AD-related endophenotypes were also measured for the majority of the subjects: (i) CSF YKL-40; (ii) MRI measures of white matter hyperintensities; (iii) clinical assessments including baseline diagnosis, baseline MMSE score and MCI conversion [12].

2.2. Omics analyses

We performed multi-omics analyses for these subjects including CSF proteomics, plasma proteomics and metabolomics as well as genome-wide SNP genotyping analyses (**Figure 1**).

CSF proteomics

We used tandem mass tag (TMT) technique to measure proteins in CSF. More details can be found in [13]. We imputed proteins using K-nearest neighbour (K=10) and removed any missing > 70%, leading to a total of 696 proteins in 371 samples for further analysis.

Plasma proteomics

We used the SOMAscan assay platform (SomaLogic Inc.) to measure proteins in plasma. SOMAscan is an aptamer-based assay allowing for the simultaneous measurement and quantification of large number of proteins. Here we measured 4001 proteins in 972 individuals. The details have been described previously [14].

Plasma metabolomics

We measured plasma metabolites using Metabolon platform (Metabolon Inc.). Metabolites with more than 70% missing were excluded and we imputed the missing metabolites using K-nearest neighbour (K=10), resulting in 611 metabolites in 696 subjects for further analysis. More details can be found in [11].

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Single nucleotide polymorphism (SNP) genotyping

A detailed account of the genotyping procedures and subsequent bioinformatic workflows can be found in [15]. Briefly, a total of 936 DNA samples were sent for genome-wide SNP genotyping using the Infinium Global Screening Array (GSA) with Shared Custom Content (Illumina Inc.). After quality control (QC) and imputation, a total of 7,778,465 autosomal SNPs with minor allele frequency (MAF) ≥0.01 were retained in 898 individuals of European ancestry for downstream analyses and genetic principal components (PCs) were computed [15].

2.3. Statistical analysis

All statistical analyses were completed using R (version 4.1.2). To compare baseline cohort characteristics across three different diagnostic groups (CTL, MCI and AD), we used one-way analysis of variance (ANOVA) and chi-square tests to compare continuous and binary variables, respectively.

Weighted Gene Correlation Network Analysis (WGCNA)

We used the R package WGCNA [16] to construct a weighted and unsigned co-expression network for each individual omics layer. This clustering method is based on calculating correlations between paired variables. The resulting modules or groups of co-expressed analytes were used to calculate module eigenprotein/eigenmetabolite metrics. The eigenprotein/eigenmetabolite-based connectivity (kME) value was used to represent the strength of an analyte's correlation with the module. Analytes with high intramodular kME in the top 90th percentile within a module were considered as hub proteins/metabolites.

The correlations between eigenprotein/eigenmetabolite and AD endophenotypes were calculated using Spearman's correlation, the *p* values were corrected with false discovery rate (FDR) and corrected *p* values are presented in a heat map. Furthermore, we used one way ANOVA test to assess pairwise difference of eigenprotein/eigenmetabolite among different AT(N) framework.

Pathway enrichment analysis

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Protein pathway enrichment analysis was performed using WebGestalt software (http://www.webgestalt.org/). Briefly, proteins within a module were assembled into a "protein list" and all proteins measured were used as "background". This enrichment analysis was performed on the KEGG database. Metabolite enrichment analysis was performed using the hypergeometric test. The original 60 sub-pathways pre-defined by Metabolon based on the KEGG database were employed as reference [17]. We further performed cell type enrichment analysis for CSF proteins using BEST tool (http://best.psych.ac.cn/#).

AD polygenic risk score (PRS) calculation

The genome-wide association study summary statistics from Kunkle et al. [18] (N=63,926; 21,982 AD clinically ascertained cases, 41,944 controls) were used as the reference data. PRS were constructed using PRSice-2 [19], with and without SNPs in the *APOE* region (chr 19, GRCh37 coordinates 44912079 to 45912079) [20]. AD PRS were computed using two p-value thresholds (P_T), previously recommended for PRS including and excluding the *APOE* region: $5x10^{-8}$ (*APOE* region included) and 0.1 (*APOE* region excluded) [21]. SNPs in linkage disequilibrium (r2>0.001 within a 250kb window) were clumped, retaining the SNP with the lowest *p*-value.

Association of AD PRS and AT(N) with modules and hubs

We used linear regression analyses to investigate the association of AD PRS (as predictor) with eigenprotein/eigenmetabolite of AT(N) framework-related modules and hub proteins/metabolites (with kME varying from top 90th percentile to top 98th percentile) in these modules, adjusting for sex, age, and genetic PC1 to PC5 [22] (to control for population stratification). We used logistic regression analyses to explore the association of AT(N) markers (as binary outcome) with hubs, adjusting for sex, age and *APOE* ε4 genotype.

Partial correlation network

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

We used age, sex, *APOE* genotype, AD PRS (P_T = 0.1, *APOE* region excluded) and all hub proteins/metabolites (with kME in the top 90th percentile) as input features for the graphical LASSO algorithm and extended Bayesian information criterion to determine the model complexity for MCI conversion using the R package `huge´ [23]. Data were auto-scaled prior to model-fitting. Partial correlation network of selected metabolites, proteins and genetic variables was computed and visualized with R package 'qgraph'.

Mendelian randomization

We finally investigated whether any of the A/T/N hubs correlating with MCI conversion status were causally linked to AD, by performing bi-directional two-sample Mendelian Randomization (MR) analyses implemented in the "TwoSampleMR" R package [24] and the MendelianRandomization package [25]. A number of sensitivity analyses for both single cis instrument MR and multiple (cis) instruments MR (**Supplementary methods**) were applied to determine the robustness of the MR findings.

3. Results

3.1. Subject demographics

Table 1 shows the demographic information of subjects for each individual omics analysis. Despite the difference in sample size for each omics layer analysis, no significant difference was observed in the distribution of sex across different diagnostic groups. However, the CTL group was younger and had a lower proportion of *APOE* ε4 carriers compared with the MCI and AD groups. Furthermore, the CTL participants had longer education and higher MMSE score. In terms of AD pathology markers, the ratio of abnormality of amyloid, P-tau and T-tau in AD and MCI individuals was, as expected, significantly higher than in controls.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

3.2 Co-expression network analysis of individual omics modalities reveals modules linked to AD endophenotypes

We first performed a clustering analysis of the CSF proteome using WGCNA. We found four modules (M) of co-expressed proteins. We ranked modules based on size from largest (M1 turquoise; n= 526 proteins) to smallest (M4 yellow; n=51 proteins) (**Figure 2A**, Table S1). We further investigated the biological significance of proteins in each module and found that three modules (M1 turquoise, M2 blue and M4 yellow modules) were enriched with various pathways after FDR correction (**Figure 2B**). When checking cell type enrichment, we found that all four modules were enriched with endothelial cells. Furthermore, M1 turquoise module was enriched with oligodendrocytes, neurons and astrocytes. M2 blue and M4 yellow modules were enriched with microglia (**Figure 2A**).

We then assessed the module correlations to AD endophenotypes. We used amyloid-β as "A", CSF P-tau levels as a biomarker of tau ("T"), CSF T-tau as biomarkers of neurodegeneration ("N"), white matter hyperintensity (WMH) volume as a biomarker for vascular disease burden ("V"), CSF YKL-40 as a biomarker of inflammation ("I") and MMSE score as "C" (**Figure 2A**). Overall, two (M1 and M4) and three (M1, M2 and M3) modules were significantly associated with "T" and "N", respectively after FDR correction. Furthermore, three (M1, M2 and M4) modules were associated with "I". None of the modules were correlated with "A", "V", "C" or MCI conversion.

We used the same approach to analyse plasma proteomics and metabolomics data. We obtained nine modules from plasma proteins (**Figure 2C,** previously published [26]). Four modules (M2, M3, M4 and M8) had positive correlations with "A", "T" and "N". One (M3) and four (M1, M3, M8 and M9) modules were associated with "V" and "I", respectively. In comparison, most plasma modules were associated with "C" and MCI conversion. Furthermore, such associations were in concordance with AT(N) markers correlations. For example, M2, M3, M4 and M8 modules were positively associated with "A", "T" and "N" but were negatively correlated with MMSE score. Furthermore, they were increased in MCI converters (n=103) compared with MCI non-converters (n=223) (**Figure**

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

2C). We further investigated the biological significance of proteins in four AT(N) markers-related modules (M2, M3, M4 and M8) and found that three of them were enriched with various pathways, such as cytokine-cytokine receptor interaction and metabolic pathways (**Figure 2D**).

For plasma metabolomics, we obtained seven modules (**Figure 2E**), among which M4 module was negatively associated with "A", "T" and "N" and M3 module was positively associated with "N". Furthermore, one (M1) and two (M1 and M4) modules were associated with "V" and "I", respectively. Two (M3 and M4) and four (M1, M5, M6 and M7) modules were associated with "C" and MCI conversion respectively. Furthermore, such associations were in concordance with AT(N) markers correlations. We further investigated the biological significance of metabolites in AT(N) markers-related modules (M3 and M4) and found that they were enriched in lipid pathways (**Figure 2F**).

3.3 Correlation of individual omics modules with the AT(N) framework

We dichotomized AT(N) biomarkers as normal or abnormal and categorized individuals into one of four groups: A-T-N- (no pathology), A+TN- (amyloid pathology), A+TN+ (Alzheimer pathology) and A-TN+ (SNAP). We then assessed the expression of each module eigenprotein/eigenmetabolite across different ATN groups. For CSF protein modules, we found that three modules (M1 turquoise, M2 blue and M4 yellow) showed a significant difference across ATN profiles from one-way ANOVA test (**Figure 3A-C**). Four plasma protein modules (M2 blue, M3 brown, M4 yellow and M8 pink, **Figure 3D-G,** adapted from [26]) and three plasma metabolites modules (M4 yellow, M5 green and M3 brown, **Figure 3H-J**) showed a significant difference across ATN profiles.

3.4 Association between AT(N) framework-related modules and AD PRS

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

We firstly selected AT(N) framework-related modules from each individual omics for further analysis. As a result, we selected three CSF protein modules (M1 turquoise, M2 blue and M4 yellow), four plasma protein modules (M2 blue, M3 brown, M4 yellow and M8 pink) and three plasma metabolite modules (M3 brown, M4 yellow and M5 green). We then analysed the correlations between these ten modules as well as between these modules and AD PRS. When analysing associations between modules, we found that the metabolite M5 green module was negatively correlated with three plasma protein modules (M2 blue, M3 brown and M8 pink). Additionally, a negative correlation was observed between metabolite M3 brown module and plasma protein M8 pink module. In contrast, a positive correlation was observed between metabolite M4 yellow module and plasma protein M4 yellow module as well as between metabolite M3 brown module and CSF protein M4 yellow module. In addition, CSF protein M4 yellow module was positively associated with plasma protein M3 brown module (**Figure 3K, Table S1-S3**).

When analysing the associations between these modules and AD PRS, we found that only plasma protein modules were significantly associated with AD PRS. In detail, two plasma protein modules (M2 blue and M4 yellow) were positively associated with PRS (*APOE* region included and excluded) at P_T=0.1. Additionally, M2 blue module was significantly associated with PRS at $5x10^{-8}$ threshold with *APOE* gene region included (**Figure 3K, Table S4**).

3.5 Association of hub proteins/metabolites with AT(N) markers and AD PRS

We selected hub proteins/metabolites within AT(N) framework-related modules and analysed the association between these hub proteins/metabolites (with kME varying from top 90th percentile to top 98th percentile, **Table S5**), as well as the association of these hub proteins/metabolite with AT(N) markers and AD PRS. When checking the associations between hub metabolites and proteins, we found that there was a strong correlation between metabolites and plasma proteins. In detail, five metabolites (four phosphatidylethanolamines (PEs) and one LysoPE) correlated with most hub

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

proteins after controlling for multiple testing. Two metabolites (sphingomyelins [SM] d40:2 and d41:2) in M3 brown module were correlated with proteins in plasma M8 pink module and CSF M4 yellow module. In contrast, relatively week correlations were observed between CSF and plasma proteins. (**Figure 4A, Table S6**).

We also investigated the association of these proteins/metabolites with AD PRS (*APOE* region included and excluded (**Table S7**)). For plasma hub proteins, all 23 proteins in M2 blue module were positively associated with AD PRS both at $P_T=5x10^{-8}$ (*APOE* region included) and $P_T=0.1$ (*APOE* region included and excluded). Similar trend was also observed for most proteins in M4 yellow module, with only six proteins being positively associated with AD PRS at $P_T=5x10^{-8}$ (*APOE* region included), whereas most proteins, except for three, were associated with the $P_T=0.1$ AD PRS (with *APOE* and without *APOE*) (**Figure 4A**). For hub metabolites, three SMs in M3 brown module and three PEs in M5 green module were associated with AD PRS ($P_T=5x10^{-8}$) with and without *APOE* region respectively. However, such associations did not pass FDR correction (**Figure 4A** in light red). No associations were observed between CSF hub proteins and AD PRS.

When investigating the association of hub proteins and metabolites with AT(N) markers, we found that most CSF and plasma hub proteins were positively associated with amyloid, P-tau and T-tau after FDR correction. In contrast, hub metabolites were negatively associated with amyloid, P-tau and T-tau only at nominal level except for sphinganine (**Figure 4A**) (**Table S8**).

3.6 Hub molecules integration in MCI conversion

Having demonstrated the association of hub proteins/metabolites with AT(N) markers and AD PRS, we then sought to find a multimodal signal that might shed insights on MCI conversion. To do this we first used LASSO algorithm and extended Bayesian information criterion to select features from age, sex, AD PRS (P_T = 0.1, APOE region excluded), *APOE* ε4 genotype and all plasma hub metabolites/proteins (with kME in the top 90th percentile, **Table S9**) to predict MCI conversion. As a result, AD PRS, *APOE* ε4 genotype and several metabolites/proteins were selected from LASSO. Of

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

the metabolites/proteins, two SMs, two PEs and one protein (proprotein convertase subtilisin/kexin type 7 [PCSK7]) from the blue module were negatively correlated with MCI conversion while the rest four selected proteins were positively associated with MCI conversion including reticulocalbin 2 (RCN2) from the blue module, and three proteins from the brown module: ephrin receptor tyrosine kinase A2 (EFNA2), Collagen alpha-1(XV) chain (COL15A1) and AP-1 complex subunit gamma-like 2 (AP-1) (**Figure 4B)**. In addition, correlations were also observed between metabolites/proteins and AD PRS and *APOE* ε4 genotype (**Figure 4B**).

3.7 Causal links of hub proteins/metabolites with AD

We finally used a bidirectional two-sample Mendelian randomization to determine whether there was evidence for a causal relationship of MCI conversion related hub proteins/metabolites with Alzheimer's disease. Using Wald ratio estimate, we observed weak associations between PCSK7 and AD as well as between COL15A1 and AD using data from Sun et al. [27]. In sensitivity analyses, the causal relationship between PCSK7 and AD was replicated using summary data from an independent protein GWA study by Suhre et al. [28] (**Table 2**). Further support for causal effects for the association of PCSK7 with AD came from multiple-cis instrument MR ($p<0.001$ for IVW, 95% CI = 0.8 to 0.9, $N_{SNPs} = 4$, **Figure S4**), although this was not the case for COL15A1 (**Table S10**). Multiple-cis instrument MR robust methods (MR-Egger and Weighted-median MR) and sensitivity analyses estimates for PCSK7 were consistent with Wald ratio estimates in direction and magnitude, and showed no horizontal pleiotropy or evidence of heterogeneity, further supporting the validity of the MR assumptions (**Table S10**). In reverse MR analysis, we identified a causal association between Alzheimer's disease, RCN2 and SM (**Table 2, Figure S1-3**). Robust methods and sensitivity analyses provided additional support for such causal effects (**Table S10**).

4. Discussion (words:1161)

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Alzheimer's disease is characterized by non-linear and heterogeneous biological alterations. Multi-level biological networks underlie AD pathophysiology, including but not limited to proteostasis (amyloid-β and tau), synaptic homeostasis, inflammatory and immune responses, lipid and energy metabolism, and oxidative stress [30]. Therefore, a systems-level approach is needed to fully capture AD multifaceted pathophysiology. Here we used unbiased and high throughput multi-omics profiling of AD. We applied correlation network analysis to identify modules linked to a variety of AD endophenotypes including "A", "T", "N", "V", "I" and "C". We found that four modules obtained from CSF proteins were associated with at least one pathology marker of "T" (P-tau), "N" (T-tau) and "I" (YKL-40). Furthermore, the three "I" related modules (M1 turquoise, M2 blue and M4 yellow) were enriched with either microglia or astrocytes, which are key cellular drivers and regulators of neuroinflammation [31], further indicating the consistency between correlation network analysis and cell type enrichment analysis. In addition, of the four modules, three were enriched with various pathways which have been reported being associated with Alzheimer's such as Ras signalling pathway [32], axon guidance [33], cell adhesion molecules (CAMs) [34], and lysosome pathway [35], further demonstrating the relatedness of these proteins with AD.

From plasma metabolomics, we found that the M3 brown module was associated with "N" (T-tau) and "C" (cognition) and enriched with sphingolipid and ceramide metabolism. These findings align with literature report as the lipids within this module have been reported being associated with cognitive progression [36] and hippocampal atrophy [37]. In addition, M4 yellow module was associated with five AD pathology markers ("A", "T", "N", "I" and "C") and enriched with three pathways including gamma-glutamyl amino acid, plasmalogen, and polyamine metabolism. The findings are also consistent with previous reports showing that these pathways were associated with AD pathogenesis [38] and inflammatory cascade [39].

Two modules (M2 blue and M4 yellow) from plasma proteomics were associated with AD PRS (both with and without *APOE* gene) at 0.1 level. Of the two modules, M2 blue module was associated with

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

PRS at $5x10^{-8}$ thresholds only when PRS included SNPs in the *APOE* region, indicating that such association may be driven by *APOE*. Hub proteins in M2 blue module were correlated with PRS at $5x10^{-8}$ thresholds only with SNPs in the *APOE* region included, further indicating that associations may be driven by *APOE*. For plasma metabolomics, three sphingomyelins (SMs) from M3 brown module were associated PRS ($P_T = 5x10^{-8}$) nominally only when the *APOE* region was included, also indicating *APOE* gene dependence. This is in line with literature findings that nominal association between SMs and PRS was reported [40].

We identified several closely correlated networks for metabolites, proteins, genetic factors, and MCI conversion. Interestingly *APOE* and MCI conversion status were correlated to PCSK7 and sphingomyelins SMs have been previously associated with cognitive progression in AD [41-43]. Furthermore, the integration of AD PRS showed that phosphatidylethanolamines [44, 45] and EFNA2 [46] were associated to both (MCI converter and AD PRS), with potential as early targets.

We finally investigated the causal relationship between A/T/N hubs associated with MCI conversion status and AD. Our MR analyses highlighted a potential weak causal relationship between PCSK7 and AD which was robust in both single and multiple cis instruments MR analyses and was replicated using an independent pQTL dataset. We also found a causal relationship in the opposite direction, whereby AD status is potentially causally linked to RCN2 that has been proposed as a therapeutic target for atherosclerosis [47]. Finally, although we didn't have GWA summary data for the SM and PE hubs examined in this study, our MR analyses showed that AD was causally linked to SM levels, as previously shown when NMR data were used [48].

These findings are of great translational potential, particularly PCSK7 for which studies in Alzheimer's disease are lacking. This convertase protein is very interesting as it is found in the BACE1 locus region which encompasses several genes (PCSK7, RNF214, BACE1, CEP164) making it a plausible protein activator of downstream amyloid deposition [49].

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The causal associations from AD to RCN2 and sphingomyelins are also intriguing as both highlight a possible vascular component caused by AD genetic liability, bringing new directionality between vascular disease and dementia. These molecules and their potential causal links to AD suggest novel avenues of research and intervention.

There are limitations for our study. First, the population in this study is of European ancestry and mainly included participants who had high ratio of amyloid pathology and *APOE* ε4 carriers. Therefore, they are not necessarily representative of the broader community. Validation in independent cohorts and particularly in other ethnic groups and community-based populations are needed to see if the results are generalizable.

Despite this, our study is the largest study we are aware of to report multi-omics relating to AD endophenotypes, particularly to the AT(N) framework. Our findings offer new insights into changes in individual proteins/metabolites linked to AD endophenotypes, the AT(N) framework and AD PRS. The nominated causal proteins/metabolites may be tractable targets for mechanistic studies of AD pathology. Furthermore, they may represent promising drug targets in the early stages of AD.

Conflict of Interest

SL is named as an inventor on biomarker intellectual property protected by Proteome Sciences and Kings College London unrelated to the current study and within the past five years has advised for Optum labs, Merck, SomaLogic and been the recipient of funding from AstraZeneca and other companies via the IMI funding scheme. HZ has served at scientific advisory boards and/or as a consultant for Abbvie, Alector, ALZPath, Annexon, Apellis, Artery Therapeutics, AZTherapies, CogRx, Denali, Eisai, Nervgen, Novo Nordisk, Pinteon Therapeutics, Red Abbey Labs, reMYND, Passage Bio, Roche, Samumed, Siemens Healthineers, Triplet Therapeutics, and Wave, has given lectures in symposia sponsored by Cellectricon, Fujirebio, Alzecure, Biogen, and Roche, and is a co-founder of

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Brain Biomarker Solutions in Gothenburg AB (BBS), which is a part of the GU Ventures Incubator Program (outside submitted work). AL has served at scientific advisory boards of Fujirebio Europe, Eli Lilly, Novartis, Nutricia and Otsuka and is the inventor of a patent on synaptic markers in CSF (all unrelated to this study). JP has served at scientific advisory boards of Fujirebio Europe, Eli Lilly and Nestlé Institute of Health Sciences, all unrelated to this study. SE has received unrestricted research grants from Janssen Pharmaceutica and ADx Neurosciences and has served at scientific advisory boards of Biogen, Eisai, icometrix, Novartis, Nutricia / Danone, Pfizer, Roche, all unrelated to this study. FB is a steering committee or iDMC member for Biogen, Merck, Roche, EISAI and Prothena. Consultant for Roche, Biogen, Merck, IXICO, Jansen, Combinostics. Research agreements with Merck, Biogen, GE Healthcare, Roche. Co-founder and shareholder of Queen Square Analytics LTD, all unrelated to this study.

Acknowledgements

This research was conducted as part of the EMIF-AD MBD project which has received support from the Innovative Medicines Initiative Joint Undertaking under EMIF grant agreement no 115372, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in-kind contribution. The DESCRIPA study was funded by the European Commission within the 5th framework program (QLRT-2001-2455). The EDAR study was funded by the European Commission within the 5th framework program (contract # 37670). The Leuven cohort was funded by the Stichting voor Alzheimer Onderzoek (grant numbers #11020, #13007 and #15005). RV is a senior clinical investigator of the Flemish Research Foundation (FWO). The San Sebastian GAP study is partially funded by the Department of Health of the Basque Government (allocation 17.0.1.08.12.0000.2.454.01.41142.001.H). We acknowledge the contribution of the personnel of the Genomic Service Facility at the VIB-U Antwerp Center for Molecular Neurology. The research at VIB-CMN is funded in part by the University of Antwerp Research Fund. HZ is a Wallenberg Scholar supported by grants from the Swedish Research Council

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

(#2018-02532), the European Research Council (#681712), Swedish State Support for Clinical Research (#ALFGBG-720931), the Alzheimer Drug Discovery Foundation (ADDF), USA (#201809-2016862), and the UK Dementia Research Institute at UCL. FB is supported by the NIHR biomedical research centre at UCLH. LS is funded by the Virtual Brain Cloud from European commission (grant no. H2020-SC1-DTH-2018-1). R.G. was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. This paper represents independent research part-funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. JX and CLQ thank Lundbeck Fonden for the support (grant no. R344-2020-989).

Ethics Statement

Written informed consent was obtained from all participants before inclusion in the study. The medical ethics committee at each site approved the study.

Data Availability

The datasets generated and analysed during the current study are available from the EMIF-AD Catalogue via submitted research proposals which have to be approved by the data-owners from each parent cohort.

References

1. Knopman, D.S., et al., *Alzheimer disease.* Nat Rev Dis Primers, 2021. **7**(1): p. 33.
2. Jack, C.R., Jr., et al., *NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease.* Alzheimers Dement, 2018. **14**(4): p. 535-562.
3. Johnson, E.C.B., 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): p. 769-780.
4. Higginbotham, L., et al., *Integrated proteomics reveals brain-based cerebrospinal fluid biomarkers in asymptomatic and symptomatic Alzheimer's disease.* Sci Adv, 2020. **6**(43).

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

5. Butterfield, D.A. and D. Boyd-Kimball, *Oxidative Stress, Amyloid-β Peptide, and Altered Key Molecular Pathways in the Pathogenesis and Progression of Alzheimer's Disease.* J Alzheimers Dis, 2018. **62**(3): p. 1345-1367.
6. Ashton, N.J., et al., *Blood protein predictors of brain amyloid for enrichment in clinical trials?* Alzheimers Dement (Amst), 2015. **1**(1): p. 48-60.
7. Hye, A., et al., *Plasma proteins predict conversion to dementia from prodromal disease.* Alzheimers Dement, 2014. **10**(6): p. 799-807.e2.
8. Shi, L., et al., *Discovery and validation of plasma proteomic biomarkers relating to brain amyloid burden by SOMAscan assay.* Alzheimers Dement, 2019. **15**(11): p. 1478-1488.
9. Wilkins, J.M. and E. Trushina, *Application of Metabolomics in Alzheimer's Disease.* Front Neurol, 2017. **8**: p. 719.
10. Tönnies, E. and E. Trushina, *Oxidative Stress, Synaptic Dysfunction, and Alzheimer's Disease.* J Alzheimers Dis, 2017. **57**(4): p. 1105-1121.
11. Kim, M., et al., *Primary fatty amides in plasma associated with brain amyloid burden, hippocampal volume, and memory in the European Medical Information Framework for Alzheimer's Disease biomarker discovery cohort.* Alzheimers Dement, 2019. **15**(6): p. 817-827.
12. Bos, I., et al., *The EMIF-AD Multimodal Biomarker Discovery study: design, methods and cohort characteristics.* Alzheimers Res Ther, 2018. **10**(1): p. 64.
13. Tijms, B.M., et al., *Pathophysiological subtypes of Alzheimer's disease based on cerebrospinal fluid proteomics.* Brain, 2020. **143**(12): p. 3776-3792.
14. Shi, L., et al., *Discovery and validation of plasma proteomic biomarkers relating to brain amyloid burden by SOMAscan assay.* Alzheimer's & Dementia, 2019. **15**(11): p. 1478-1488.
15. Hong, S., et al., *Genome-wide association study of Alzheimer's disease CSF biomarkers in the EMIF-AD Multimodal Biomarker Discovery dataset.* Transl Psychiatry, 2020. **10**(1): p. 403.
16. Langfelder, P. and S. Horvath, *WGCNA: an R package for weighted correlation network analysis.* BMC Bioinformatics, 2008. **9**: p. 559.
17. Xu, J., et al., *Sex-Specific Metabolic Pathways Were Associated with Alzheimer's Disease (AD) Endophenotypes in the European Medical Information Framework for AD Multimodal Biomarker Discovery Cohort.* Biomedicines, 2021. **9**(11).
18. Kunkle, B.W., et al., *Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing.* Nat Genet, 2019. **51**(3): p. 414-430.
19. Euesden, J., C.M. Lewis, and P.F. O'Reilly, *PRSice: Polygenic Risk Score software.* Bioinformatics, 2015. **31**(9): p. 1466-8.
20. Zettergren, A., et al., *Association between polygenic risk score of Alzheimer's disease and plasma phosphorylated tau in individuals from the Alzheimer's Disease Neuroimaging Initiative.* Alzheimers Res Ther, 2021. **13**(1): p. 17.
21. Leonenko, G., et al., *Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores.* Nat Commun, 2021. **12**(1): p. 4506.
22. Hong, S., et al., *TMEM106B and CPOX are genetic determinants of cerebrospinal fluid Alzheimer's disease biomarker levels.* Alzheimers Dement, 2021. **17**(10): p. 1628-1640.
23. Tofte, N., et al., *Metabolomic Assessment Reveals Alteration in Polyols and Branched Chain Amino Acids Associated With Present and Future Renal Impairment in a Discovery Cohort of 637 Persons With Type 1 Diabetes.* Front Endocrinol (Lausanne), 2019. **10**: p. 818.
24. Hemani, G., et al., *The MR-Base platform supports systematic causal inference across the human phenome.* Elife, 2018. **7**.
25. Yavorska, O.O. and S. Burgess, *MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data.* Int J Epidemiol, 2017. **46**(6): p. 1734-1739.
26. Shi, L., et al., *Replication study of plasma proteins relating to Alzheimer's pathology.* Alzheimers Dement, 2021. **17**(9): p. 1452-1464.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

27. Sun, B.B., et al., *Genomic atlas of the human plasma proteome.* Nature, 2018. **558**(7708): p. 73-79.
28. Suhre, K., et al., *Connecting genetic risk to disease end points through the human blood plasma proteome.* Nat Commun, 2017. **8**: p. 14357.
29. Borges, M.C., et al., *Circulating Fatty Acids and Risk of Coronary Heart Disease and Stroke: Individual Participant Data Meta-Analysis in Up to 16 126 Participants.* J Am Heart Assoc, 2020. **9**(5): p. e013131.
30. Hampel, H., et al., *Omics sciences for systems biology in Alzheimer's disease: State-of-the-art of the evidence.* Ageing Res Rev, 2021. **69**: p. 101346.
31. Hampel, H., et al., *A Path Toward Precision Medicine for Neuroinflammatory Mechanisms in Alzheimer's Disease.* Front Immunol, 2020. **11**: p. 456.
32. Kirouac, L., et al., *Activation of Ras-ERK Signaling and GSK-3 by Amyloid Precursor Protein and Amyloid Beta Facilitates Neurodegeneration in Alzheimer's Disease.* eNeuro, 2017. **4**(2).
33. Zhang, L., et al., *Roles and Mechanisms of Axon-Guidance Molecules in Alzheimer's Disease.* Mol Neurobiol, 2021. **58**(7): p. 3290-3307.
34. Nagappan-Chettiar, S., E.M. Johnson-Venkatesh, and H. Umemori, *Activity-dependent proteolytic cleavage of cell adhesion molecules regulates excitatory synaptic development and function.* Neurosci Res, 2017. **116**: p. 60-69.
35. Zhang, Y., et al., *The role of ubiquitin proteasomal system and autophagy-lysosome pathway in Alzheimer's disease.* Rev Neurosci, 2017. **28**(8): p. 861-868.
36. Kalinichenko, L.S., et al., *Sphingolipid control of cognitive functions in health and disease.* Prog Lipid Res, 2022. **86**: p. 101162.
37. Kim, M., et al., *Association between Plasma Ceramides and Phosphatidylcholines and Hippocampal Brain Volume in Late Onset Alzheimer's Disease.* J Alzheimers Dis, 2017. **60**(3): p. 809-817.
38. Mahajan, U.V., et al., *Dysregulation of multiple metabolic networks related to brain transmethylation and polyamine pathways in Alzheimer disease: A targeted metabolomic and transcriptomic study.* PLoS Med, 2020. **17**(1): p. e1003012.
39. Bozelli, J.C., Jr., S. Azher, and R.M. Epand, *Plasmalogens and Chronic Inflammatory Diseases.* Front Physiol, 2021. **12**: p. 730829.
40. Liu, Y., et al., *Plasma lipidome is dysregulated in Alzheimer's disease and is associated with disease risk genes.* Transl Psychiatry, 2021. **11**(1): p. 344.
41. Mielke, M.M., et al., *Plasma sphingomyelins are associated with cognitive progression in Alzheimer's disease.* J Alzheimers Dis, 2011. **27**(2): p. 259-69.
42. He, X., et al., *Deregulation of sphingolipid metabolism in Alzheimer's disease.* Neurobiol Aging, 2010. **31**(3): p. 398-408.
43. Morrow, A.R., et al., *CSF sphingomyelin metabolites in Alzheimer's disease, neurodegeneration, and neuroinflammation.* Alzheimer's & Dementia, 2021. **17**: p. e052290.
44. Llano, D.A. and V. Devanarayan, *Serum Phosphatidylethanolamine and Lysophosphatidylethanolamine Levels Differentiate Alzheimer's Disease from Controls and Predict Progression from Mild Cognitive Impairment.* J Alzheimers Dis, 2021. **80**(1): p. 311-319.
45. Calzada, E., O. Onguka, and S.M. Claypool, *Phosphatidylethanolamine Metabolism in Health and Disease.* Int Rev Cell Mol Biol, 2016. **321**: p. 29-88.
46. Ma, X., et al., *Metabolic Reprogramming of Microglia Enhances Proinflammatory Cytokine Release through EphA2/p38 MAPK Pathway in Alzheimer's Disease.* J Alzheimers Dis, 2022.
47. Li, J., et al., *Reticulocalbin 2 as a Potential Biomarker and Therapeutic Target for Atherosclerosis.* Cells, 2022. **11**(7).
48. Compton, H., et al., *Effects of genetic liability to Alzheimer's disease on circulating metabolites across the life course.* medRxiv, 2022: p. 2022.03.24.22272867.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

49. Damotte, V., et al., *Plasma amyloid β levels are driven by genetic variants near APOE, BACE1, APP, PSEN2: A genome-wide association study in over 12,000 non-demented participants.* Alzheimers Dement, 2021. **17**(10): p. 1663-1674.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

575 **Table 1. Demographics of participants included in multi-omics analysis by diagnosis.** One-way
576 analysis of variance (ANOVA) and chi-square tests were used to compare continuous and binary
577 variables, respectively. Percentage of cases is shown in brackets for male sex, *APOE* ε4 carriers and
578 the abnormality of amyloid, P-tau and T-tau.

Characteristics	Sample size	CTL	MCI	AD	*P* value
CSF proteomics					
n	371	123	154	94	NA
Age mean (SD), y	371	64.4 (7.8)	69.0 (7.4)	68.1 (8.1)	<0.001
Male sex N (%)	371	66 (54)	77 (50)	49 (52)	0.83
APOE ε4+ N (%)	371	45 (37)	78 (51)	59 (63)	<0.001
MMSE (SD)	370	28.7 (1.3)	26.5 (2.7)	22.1 (3.8)	<0.001
Education mean (SD), y	371	12.4 (3.5)	10.8 (3.6)	10.1 (3.8)	<0.001
Amyloid + N (%)	371	41 (33)	77 (50)	81 (86)	<0.001
P-tau + N (%)	367	29 (24)	87 (58)	69 (73)	<0.001
T-tau + N (%)	365	28 (23)	82 (55)	74 (80)	<0.001
Plasma proteomics					
n	972	372	409	191	NA
Age mean (SD), y	972	64.6 (8.0)	69.9 (8.0)	70.5 (8.8)	<0.001
Male sex N (%)	972	209 (56)	216 (53)	103 (54)	0.64
APOE ε4+ N (%)	972	139 (37)	195 (48)	116 (61)	<0.001
MMSE (SD)	967	28.8 (1.2)	26.2 (2.6)	21.4 (4.7)	<0.001
Education mean (SD), y	972	12.8 (3.7)	11.0 (3.7)	10.3 (3.9)	<0.001
Amyloid + N (%)	972	112 (30)	254 (62)	168 (88)	<0.001
P-tau + N (%)	876	53 (19)	215 (53)	128 (67)	<0.001
T-tau + N (%)	880	54 (19)	235 (58)	152 (80)	<0.001
Plasma metabolomics					
n	696	284	275	137	NA
Age mean (SD), y	696	65.0 (7.9)	70.0 (8.1)	70.1 (8.5)	<0.001
Male sex N (%)	696	155 (55)	141 (51)	81 (59)	0.60
APOE ε4+ N (%)	696	111 (39)	153 (56)	84 (61)	<0.001
MMSE (SD)	691	28.8 (1.1)	25.7 (2.8)	21.5 (4.8)	<0.001
Education mean (SD), y	696	12.8 (3.8)	11.1 (3.4)	10.4 (3.7)	<0.001
Amyloid + N (%)	696	114 (40)	197 (72)	122 (89)	<0.001
P-tau + N (%)	641	44 (19)	161 (59)	93 (68)	<0.001
T-tau + N (%)	641	45 (19)	177 (65)	107 (79)	<0.001

579 CTL, cognitively normal controls; MCI, mild cognitive impairment; AD, Alzheimer's disease; CSF,
580 cerebrospinal fluid; SD, standard deviation; MMSE, mini mental state examination; +, abnormality;
581 P-tau, phosphorylated tau; T-tau, total tau.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 2. Examination of the causal relationship between hub proteins/metabolites and Alzheimer's using bidirectional Mendelian randomization

Protein	Estimate*					
	Forward (hub proteins/metabolites →Alzheimer's disease)			Backward (Alzheimer's disease→ hub proteins/metabolites)		
	No. of SNPs	Slope (95% CI)	P value	No. of SNPs	Slope (95% CI)	P value
PCSK7	1 [27]	0.88 (0.79 - 0.99)	**0.029**	20	0.06 (-0.14 - 0.02)	0.135
PCSK7	1 [28]	0.96 (0.93 - 0.99)	**0.027**	NA	NA	NA
RCN2	1 [27]	1.04 (0.87 - 1.24)	0.648	20	0.10 (0.03 - 0.17)	**0.004**
EFNA2	1 [27]	1.04 (0.81 - 1.34)	0.741	20	-0.02 (-0.09 - 0.06)	0.679
AP-1	1 [27]	0.84 (0.67 - 1.04)	0.114	20	-0.01 (-0.09 - 0.07)	0.804
COL15A1	1 [27]	1.15 (1.01 - 1.30)	**0.032**	20	0.02 (-0.05 - 0.09)	0.544
SM	58 [29]	1.28 (0.87 - 1.89)	0.212**	21	0.14 (0.03 - 0.24) ***	**0.012**

*The Wald ratio estimate and inverse variance weighting (IVW) estimate were used for MR analyses with a single and multiple SNPs, respectively.

** Cochran's Q p<0.001

***there was significant heterogeneity even after removing APOE (Cochran's Q p<0.05).

medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

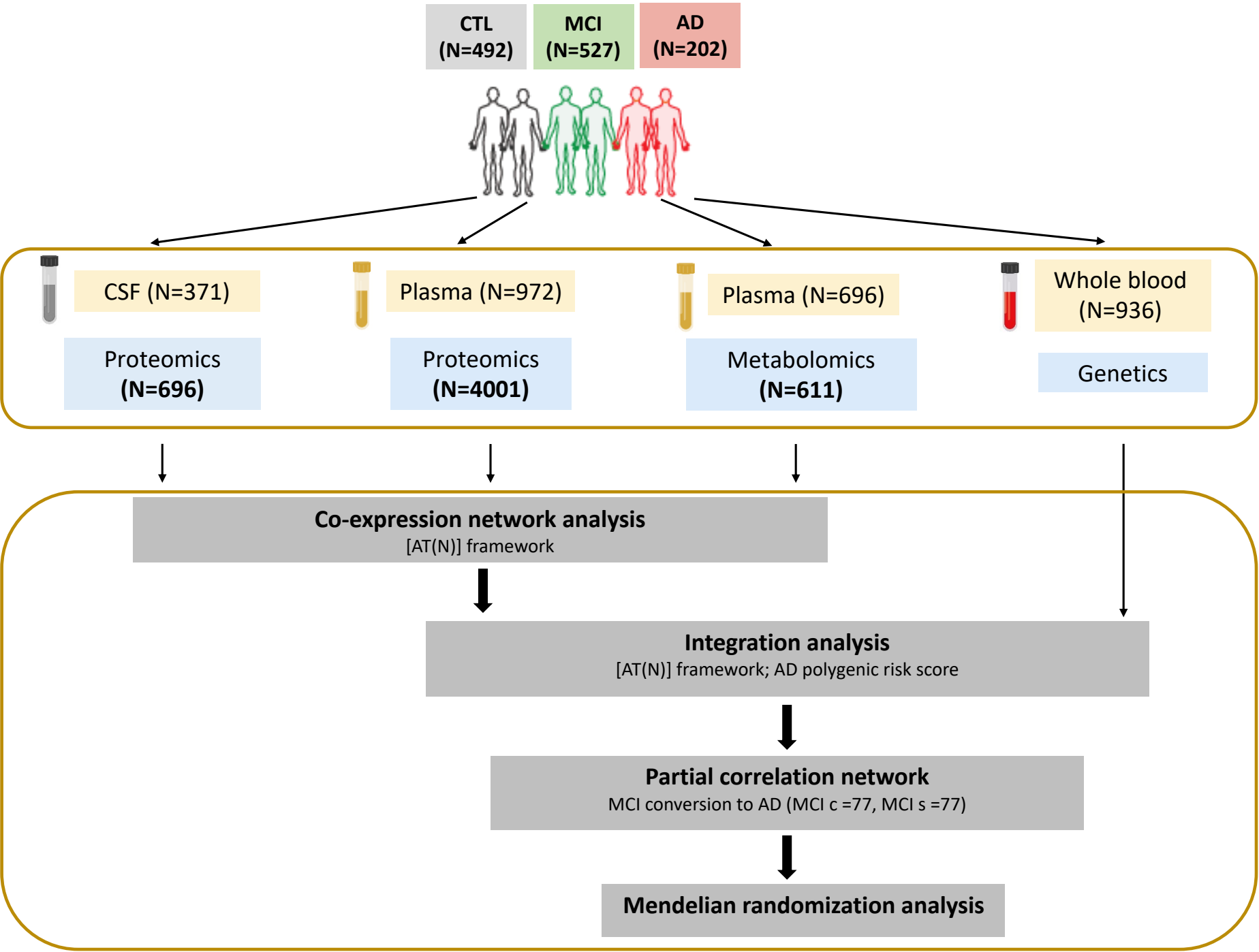
Figure Legends

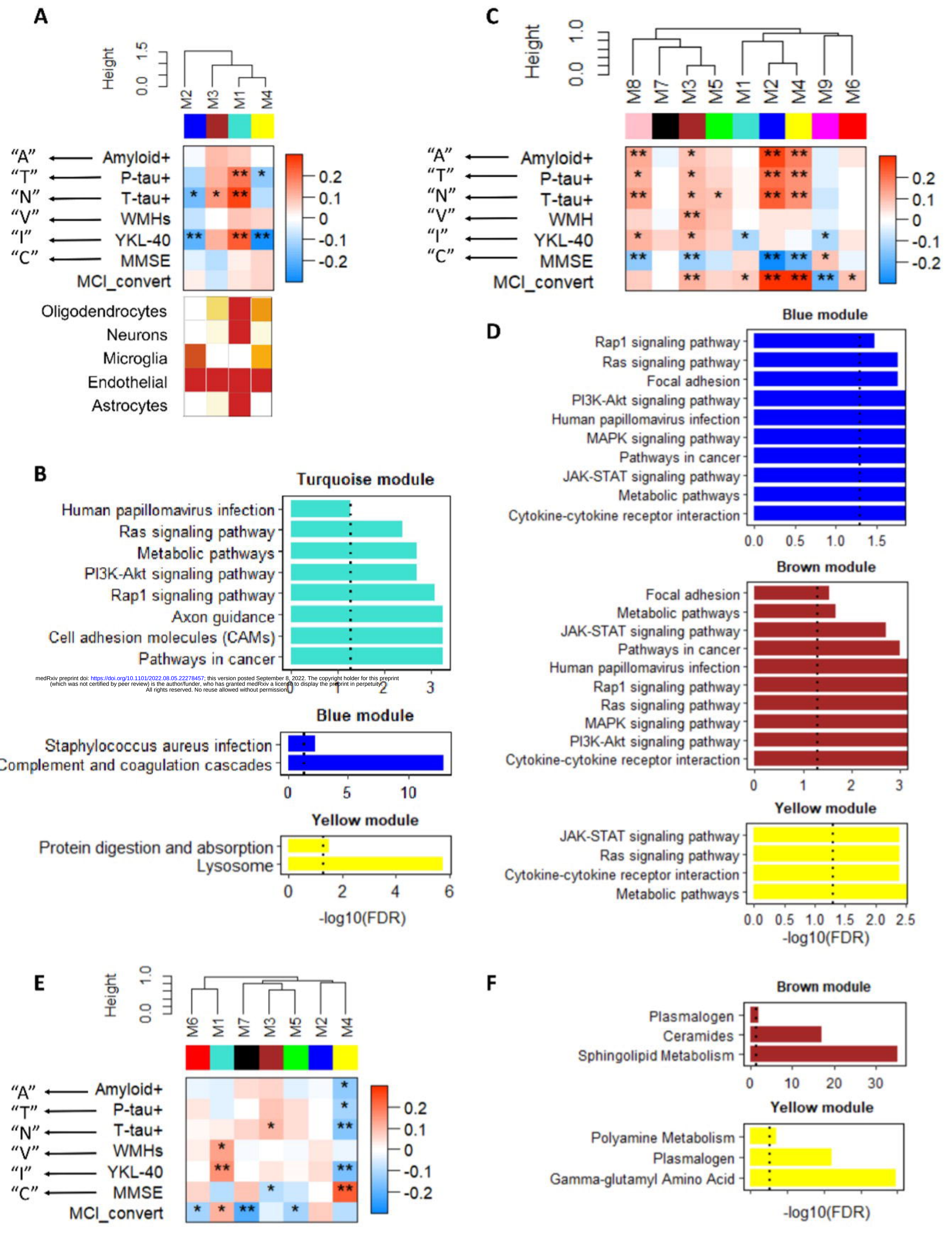
Figure 1. Flowchart of study design. CTL, cognitively normal controls; MCI, mild cognitive impairment; AD, Alzheimer's disease; Aβ, β-amyloid; CSF, cerebrospinal fluid; A, amyloid pathology; T, tau pathology; N, neurodegeneration; PRS, polygenic risk score; MR, mendelian randomization.

Figure 2. Individual omics modules correlating to AD endophenotypes. (A) Weighted gene correlation network analysis (WGCNA) of the CSF proteomics and cell type enrichment analysis of modules; (B) Enriched KEGG pathways of three modules in CSF proteins; (C) WGCNA of plasma proteomics; (D) Enriched KEGG pathways of three modules in plasma proteins; (E) WGCNA of plasma metabolomics; (F) Enriched KEGG pathways of two modules in plasma metabolites. * and ** denote significant correlations $p < 0.05$ and $p < 0.001$ after false discovery rate (FDR) correction respectively; CSF, cerebrospinal fluid; "A", amyloid; "T", tau; "N", neurodegeneration; "V", vascular; "I", inflammation; "C", cognition; +, abnormality; P-tau, phosphorylated tau; T-tau, total tau; WMH, white matter hyperintensity; MMSE, mini mental state examination; MCI, mild cognitive impairment.

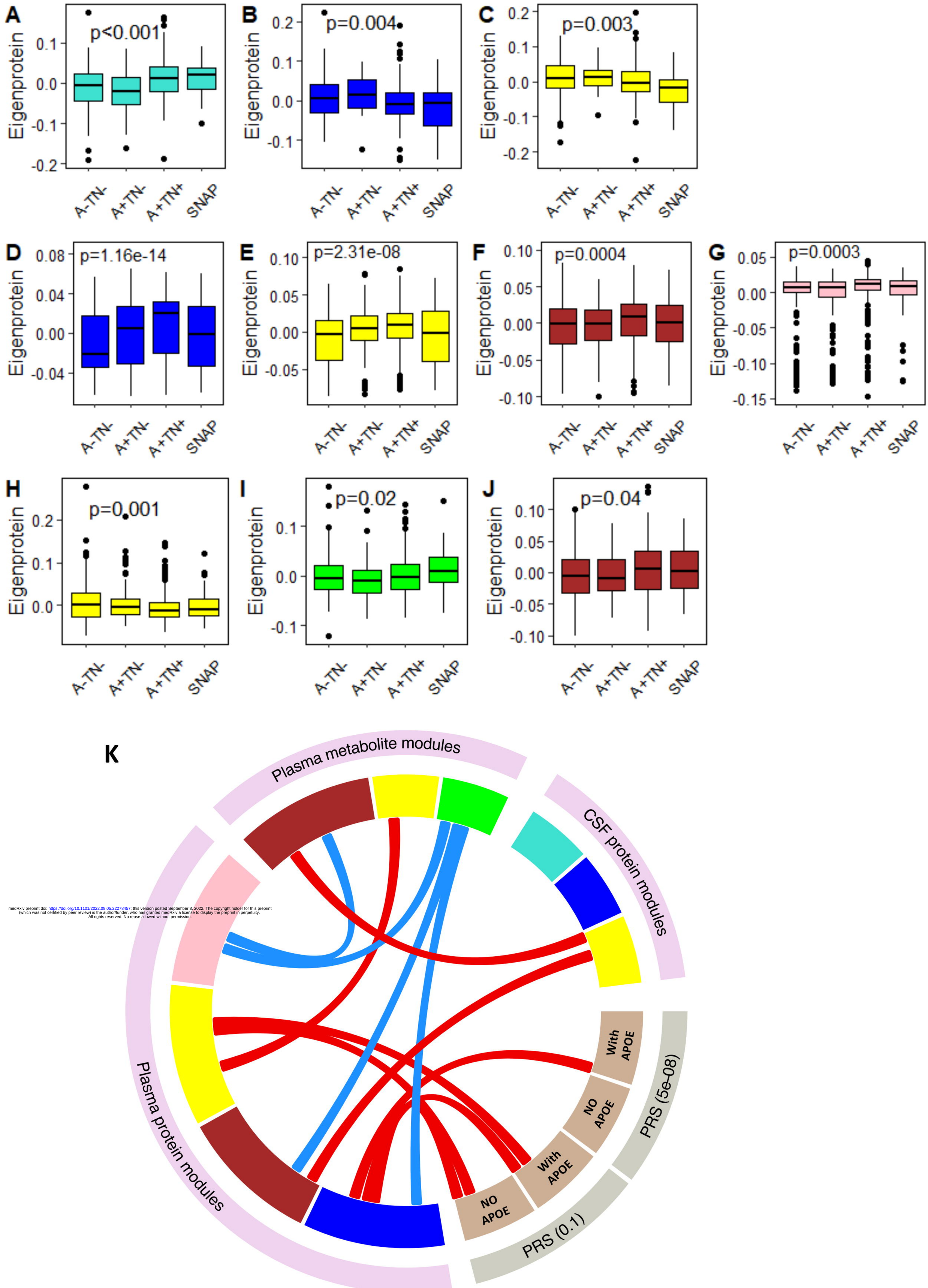
Figure 3. Protein and metabolite modules correlate to AT(N) profile and AD PRS. The relationship of the AT(N) framework with (A-C) three CSF protein modules, (D-G) four plasma protein modules, (H-J) three plasma metabolite modules, (K) Relation of AT(N) framework-related modules with AD PRS (with and without *APOE* region) at two thresholds (PT=$5x10^{-8}$ & 0.1); red and blue links denoted positive and negative correlations, respectively. CSF, cerebrospinal fluid; SNAP, Suspected Non-Alzheimer Pathology.

Figure 4. (A) Relation of hub proteins/metabolites with AT(N) markers and PRS; hub proteins/metabolites from three CSF protein modules (M1 turquoise, M2 blue and M4 yellow), four plasma protein modules (M2 blue, M3 brown, M4 yellow and M8 pink), and three plasma metabolite modules (M4 yellow, M5 green and M3 brown); red, blue, light red and light blue squares denoted positive association at FDR level (*pFDR*<0.05), negative association at FDR level (*pFDR*<0.05), positive association at nominal level (*p*<0.05, *pFDR*>0.05), and negative association at nominal level (*p*<0.05, *pFDR*>0.05), respectively. **(B) Partial correlation network selected for hub metabolites/proteins, genetic factors, and MCI conversion.** PRS, polygenic risk score; MCI, mild cognitive impairment.

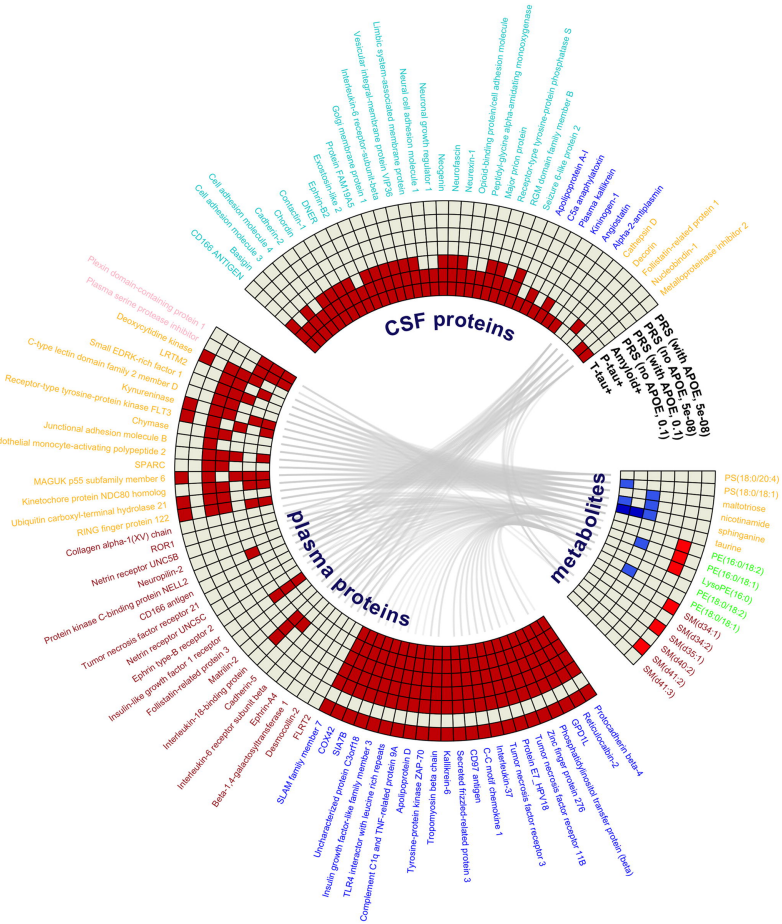

CTL (N=492)
MCI (N=527)
AD (N=202)
CSF (N=371)
Plasma (N=972)
Plasma (N=696)
Whole blood (N=936)
Proteomics (N=696)
Proteomics (N=4001)
Metabolomics (N=611)
Genetics
Co-expression network analysis
[AT(N)] framework
Integration analysis
[AT(N)] framework; AD polygenic risk score
Partial correlation network
MCI conversion to AD (MCI c =77, MCI s =77)
Mendelian randomization analysis



medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



medRxiv preprint doi: https://doi.org/10.1101/2022.08.05.22278457; this version posted September 8, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.


A
CSF proteins
plasma proteins
metabolites
PRS (with APOE, 5e-08)
PRS (no APOE, 5e-08)
PRS (with APOE, 0.1)
PRS (no APOE, 0.1)
Amyloid+
P-tau+
T-tau+
PS(18:0/20:4)
PS(18:0/18:1)
maltotriose
nicotinamide
sphinganine
taurine
PE(16:0/18:2)
PE(16:0/18:1)
LysoPE(16:0)
PE(18:0/18:2)
PE(18:0/18:1)
SM(d34:1)
SM(d34:2)
SM(d35:1)
SM(d40:2)
SM(d41:2)
SM(d41:3)
Protocadherin beta-4
Reticulocalbin-2
GPD1L
Phosphatidylinositol transfer protein (beta)
Zinc finger protein 276
Tumor necrosis factor receptor 11B
Protein E7_HPV18
Tumor necrosis factor receptor 3
Interleukin-37
C-C motif chemokine 1
CD97 antigen
Secreted frizzled-related protein 3
Kallikrein-6
Tropomyosin beta chain
Tyrosine-protein kinase ZAP-70
Apolipoprotein D
Complement C1q and TNF-related protein 9A
TLR4 interactor with leucine rich repeats
Insulin growth factor-like family member 3
Uncharacterized protein C3orf18
SIAT8
COX42
SLAM family member 7
FLRT2
Desmocollin-2
Beta-1,4-galactosyltransferase 1
Ephrin-A4
Cadherin-5
Interleukin-6 receptor subunit beta
Interleukin-18-binding protein
Matrilin-2
Follistatin-related protein 3
Insulin-like growth factor 1 receptor
Ephrin type-B receptor 2
Netrin receptor UNC5C
Tumor necrosis factor receptor 21
Protein kinase C-binding protein NELL2
CD166 antigen
Neuropilin-2
Netrin receptor UNC5B
ROR1
Collagen alpha-1(XV) chain
RING finger protein 122
Ubiquitin carboxyl-terminal hydrolase 21
Kinetochore protein NDC80 homolog
MAGUK p55 subfamily member 6
SPARC
Endothelial monocyte-activating polypeptide 2
Junctional adhesion molecule B
Chymase
Receptor-type tyrosine-protein kinase FLT3
Kynureninase
C-type lectin domain family 2 member D
Small EDRK-rich factor 1
LRTM2
Deoxycytidine kinase
Plasma serine protease inhibitor
Plexin domain-containing protein 1
CD166 ANTIGEN
Basigin
Cell adhesion molecule 3
Cell adhesion molecule 4
Cadherin-2
Chordin
Contactin-1
DNER
Ephrin-B2
Exostosin-like 2
Protein FAM19A5
Golgi membrane protein 1
Interleukin-6 receptor-subunit-beta
Vesicular integral-membrane protein VIP36
Limbic system-associated membrane protein
Neural cell adhesion molecule 1
Neuronal growth regulator 1
Neogenin
Neurofascin
Neurexin-1
Opioid-binding protein/cell adhesion molecule
Peptidyl-glycine alpha-amidating monooxygenase
Major prion protein
Receptor-type tyrosine-protein phosphatase S
RGM domain family member B
Seizure 6-like protein 2
Apolipoprotein A-I
C5a anaphylatoxin
Plasma kallikrein
Kininogen-1
Angiostatin
Alpha-2-antiplasmin
Cathepsin D
Decorin
Follistatin-related protein 1
Nucleobindin-1
Metalloproteinase inhibitor 2
B
PRS
APOE4+
AP−1
MCI Convert
RCN2
SM (d40:2)
PCSK7
SM (d34:1)
EFNA2
PE (18:0/18:2)
PE (16:0/18:2)
COL15A1