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

3

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. Vosⁱ, 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,t*§}
^a Department of Psychiatry, University of Oxford, UK

- ^b Institute of Pharmaceutical Science, King's College London, London, UK
- ^cInstitute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
- ^d UK National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre, South
- 14 London and Maudsley Trust, London, UK
- ^e MRC Unit for Lifelong Health & Ageing at UCL, University College London, London, UK
- 16 ^f Steno Diabetes Center Copenhagen, Gentofte, Denmark
- 17 ^gLübeck Interdisciplinary Platform for Genome Analytics, University of Lübeck, Lübeck, Germany
- ^h Alzheimer Center, VU University Medical Center, Amsterdam, the Netherlands
- 19 ⁱ Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience,
- 20 Alzheimer Centrum Limburg, Maastricht University, Maastricht, the Netherlands
- 21 ^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
- 24 Reference Center for Biological Markers of Dementia (BIODEM), Institute Born-Bunge, University of
- 25 Antwerp, Antwerp, Belgium
- ^m Department of Neurology, UZ Brussel and Center for Neurociences (C4N), Vrije Universiteit Brussel,
- 27 Brussels, Belgium
- 28 Complex Genetics Group, VIB Center for Molecular Neurology, VIB, Antwerp, Belgium
- 29 ° Institute Born-Bunge, Department of Biomedical Sciences, University of Antwerp, Antwerp,
- 30 Belgium
- 31 ^pUniversity of Geneva, Geneva, Switzerland
- 32 ^qIRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy
- ³³ ^r Institute of Neuroscience and Physiology, Sahlgrenska Academy at University of Gothenburg,
- 34 Gothenburg, Sweden
- ⁵ Neurology Department, Hospital Sant Pau, Barcelona, Spain. Centro de Investigación en Red en
- 36 enfermedades neurodegenerativas (CIBERNED).
- 37 ^t University Hospital of Lausanne, Lausanne, Switzerland
- ⁴ Geriatric Psychiatry, Department of Mental Health and Psychiatry, Geneva university Hospitals,
- 39 Geneva, Switzerland
- 40 ^v CITA-Alzheimer Foundation, San Sebastian, Spain
- 41 ^w AC Immune SA, Lausanne, Switzerland, formerly Janssen R&D, LLC. Beerse, Belgium at the time of
- 42 study conduct
- 43 ^{*} Department of Radiology and Nuclear Medicine, Amsterdam UMC, Vrije Universiteit, The
- 44 Netherland
- ⁹ Queen Square Institute of Neurology and Centre for Medical Image Computing, University College
 London, UK
- 47 ^z Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden
- ⁴⁸ ^{aa} Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology,
- 49 Sahlgrenska Academy, University of Gothenburg, Mölndal, Sweden

- 50 ^{bb} UK Dementia Research Institute at UCL, London, United Kingdom
- ⁵¹^{cc} Department of Neurodegenerative Disease, UCL Institute of Neurology, London, United Kingdom
- 52 ^{dd} Janssen Medical (UK), High Wycombe, UK
- ^{ee}Department of Psychology, University of Oslo, Oslo, Norway
- 54
- 55
- 56
- 57 + Contributed equally
- 58 § Senior authors with equal contribution
- 59 ***correspondence**:
- 60 Cristina Legido-Quigley: cristina.legido quigley@kcl.ac.uk
- 61

62 Abstract (150 words limit)

- 63 INTRODUCTION: This study employed an integrative system and causal inference approach to
- 64 explore molecular signatures in blood and CSF, the Amyloid/Tau/Neurodegeneration [AT(N)]
- 65 framework, MCI conversion to AD, and genetic risk for AD.
- 66 METHODS: Using the EMIF-AD MBD cohort, we measured 696 proteins in cerebrospinal fluid
- 67 (n=371), 4001 proteins in plasma (n=972), 611 metabolites in plasma (n=696) and genotyped data in
- 68 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).
- 71 **RESULTS**: AT(N) framework associated key hubs were mostly proteins and few lipids. In MR analyses,
- 72 Proprotein Convertase Subtilisin/Kexin Type 7 showed weak causal associations with AD, and AD was
- 73 causally associated with Reticulocalbin 2 and sphingomyelins.
- 74 DISCUSSION: This study reveals multi-omics networks associated with AT(N) and MCI conversion and
- 75 highlights AD causal candidates.
- 76 Key words: Alzheimer's disease; multi-omics; AT(N) framework; polygenic risk score; Mendelian
- 77 randomization; multimodal biomarker
- 78

79 1. Introduction

80 Alzheimer's disease (AD) is characterised by the presence of β -amyloid (A β) containing plaques, and 81 neurofibrillary tangles composed of modified tau protein together with the progressive loss of 82 synapses and neurons [1]. The National Institute on Aging and Alzheimer's Association (NIA-AA) have 83 proposed to classify AD based on biomarkers of amyloid pathology (A), tau pathology (T), and 84 neurodegeneration (N) (the ATN framework) [2]. Yet, despite their diagnostic utility, these three 85 markers reflect only a portion of the complex pathophysiology of AD. In prodromal stages, the 86 interplay between AT(N) changes, genetic factors and peripheral molecular changes may affect the 87 rate of disease progression.

Conducting unbiased and high-throughput omics-based research in biological fluids and human brain 88 89 tissues provides a data-driven approach to identify the many processes involved in AD pathogenesis 90 and to prioritize links to relevant clinical and neuropathological traits. For example, an increasing 91 number of proteomics studies [3-5], including ours [6-8], have identified AD pathophysiological 92 pathways related to immune response and inflammation, oxidative stress, energy metabolism and 93 mitochondrial function. Metabolomics studies have also identified such pathways related to AD [9-94 11]. A combination of omics, also called multi-omics or deep phenotyping studies, provides an 95 opportunity to explore the molecular interplay with both genotypic and phenotypic variability in AD, 96 bringing in new findings and uncovering novel pathways. Finally, causal inferences approaches allow 97 to scrutinize the causal relationship between molecular markers and AD, highlighting potential 98 interventional targets. Therefore, in this study, we conducted multi-omics analyses with four 99 modalities (cerebrospinal fluid [CSF] proteomics, plasma proteomics, plasma metabolomics and 100 whole blood genetics) from the EMIF-AD MBD study, followed by Mendelian Randomization (MR) 101 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

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

111

112 2. Methods

113 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; <u>http://www.emif.eu/emif-ad-2/</u>); 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

128	pathology (A-T-N-, referring as "A-TN-"), amyloid positive but both T and N negative (A+T-N-,
129	referring as "A+TN-"), amyloid positive and T/N positive (including A+T-N+, A+T+N- and A+T+N+,
130	referring as "A+TN+") and Suspected Non-Alzheimer Pathology (SNAP, including A-T-N+, A-T+N- and
131	A-T+N+). In addition, the following AD-related endophenotypes were also measured for the majority
132	of the subjects: (i) CSF YKL-40; (ii) MRI measures of white matter hyperintensities; (iii) clinical
133	assessments including baseline diagnosis, baseline MMSE score and MCI conversion [12].
134	2.2. Omics analyses
135	We performed multi-omics analyses for these subjects including CSF proteomics, plasma proteomics
136	and metabolomics as well as genome-wide SNP genotyping analyses (Figure 1).
137	CSF proteomics

- 138 We used tandem mass tag (TMT) technique to measure proteins in CSF. More details can be found in
- 139 [13]. We imputed proteins using K-nearest neighbour (K=10) and removed any missing > 70%,

140 leading to a total of 696 proteins in 371 samples for further analysis.

141 Plasma proteomics

142 We used the SOMAscan assay platform (SomaLogic Inc.) to measure proteins in plasma. SOMAscan

143 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 144

145 described previously [14].

146 Plasma metabolomics

147 We measured plasma metabolites using Metabolon platform (Metabolon Inc.). Metabolites with 148 more than 70% missing were excluded and we imputed the missing metabolites using K-nearest 149 neighbour (K=10), resulting in 611 metabolites in 696 subjects for further analysis. More details can 150 be found in [11].

151 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].

158 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

161 of variance (ANOVA) and chi-square tests to compare continuous and binary variables, respectively.

162 Weighted Gene Correlation Network Analysis (WGCNA)

163 We used the R package WGCNA [16] to construct a weighted and unsigned co-expression network 164 for each individual omics layer. This clustering method is based on calculating correlations between 165 paired variables. The resulting modules or groups of co-expressed analytes were used to calculate 166 The module eigenprotein/eigenmetabolite metrics. eigenprotein/eigenmetabolite-based 167 connectivity (kME) value was used to represent the strength of an analyte's correlation with the 168 module. Analytes with high intramodular kME in the top 90th percentile within a module were 169 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.

174 Pathway enrichment analysis

175 Protein WebGestalt pathway enrichment analysis was performed using software 176 (http://www.webgestalt.org/). Briefly, proteins within a module were assembled into a "protein list" 177 and all proteins measured were used as "background". This enrichment analysis was performed on 178 the KEGG database. Metabolite enrichment analysis was performed using the hypergeometric test. 179 The original 60 sub-pathways pre-defined by Metabolon based on the KEGG database were 180 employed as reference [17]. We further performed cell type enrichment analysis for CSF proteins 181 using BEST tool (<u>http://best.psych.ac.cn/#</u>).

182 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.

190 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.

197 Partial correlation network

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'.

204 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.

211

212 3. Results

213 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.

221

222 3.2 Co-expression network analysis of individual omics modalities reveals modules linked to AD

223 endophenotypes

224 We first performed a clustering analysis of the CSF proteome using WGCNA. We found four modules 225 (M) of co-expressed proteins. We ranked modules based on size from largest (M1 turquoise; n= 526 226 proteins) to smallest (M4 yellow; n=51 proteins) (Figure 2A, Table S1). We further investigated the 227 biological significance of proteins in each module and found that three modules (M1 turquoise, M2 228 blue and M4 yellow modules) were enriched with various pathways after FDR correction (Figure 2B). 229 When checking cell type enrichment, we found that all four modules were enriched with endothelial 230 cells. Furthermore, M1 turquoise module was enriched with oligodendrocytes, neurons and 231 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 Ptau 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.

239 We used the same approach to analyse plasma proteomics and metabolomics data. We obtained 240 nine modules from plasma proteins (Figure 2C, previously published [26]). Four modules (M2, M3, 241 M4 and M8) had positive correlations with "A", "T" and "N". One (M3) and four (M1, M3, M8 and 242 M9) modules were associated with "V" and "I", respectively. In comparison, most plasma modules 243 were associated with "C" and MCI conversion. Furthermore, such associations were in concordance 244 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, 245 246 they were increased in MCI converters (n=103) compared with MCI non-converters (n=223) (Figure

247 2C). We further investigated the biological significance of proteins in four AT(N) markers-related

248 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).

250 For plasma metabolomics, we obtained seven modules (Figure 2E), among which M4 module was 251 negatively associated with "A", "T" and "N" and M3 module was positively associated with "N". 252 Furthermore, one (M1) and two (M1 and M4) modules were associated with "V" and "I", 253 respectively. Two (M3 and M4) and four (M1, M5, M6 and M7) modules were associated with "C" 254 and MCI conversion respectively. Furthermore, such associations were in concordance with AT(N) 255 markers correlations. We further investigated the biological significance of metabolites in AT(N) 256 markers-related modules (M3 and M4) and found that they were enriched in lipid pathways (Figure 257 2F).

258

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

260 We dichotomized AT(N) biomarkers as normal or abnormal and categorized individuals into one of 261 four groups: A-T-N- (no pathology), A+TN- (amyloid pathology), A+TN+ (Alzheimer pathology) and A-262 TN+ (SNAP). We then assessed the expression of each module eigenprotein/eigenmetabolite across 263 different ATN groups. For CSF protein modules, we found that three modules (M1 turquoise, M2 264 blue and M4 yellow) showed a significant difference across ATN profiles from one-way ANOVA test 265 (Figure 3A-C). Four plasma protein modules (M2 blue, M3 brown, M4 yellow and M8 pink, Figure 266 **3D-G,** adapted from [26]) and three plasma metabolites modules (M4 yellow, M5 green and M3 267 brown, Figure 3H-J) showed a significant difference across ATN profiles.

268

269 3.4 Association between AT(N) framework-related modules and AD PRS

270 We firstly selected AT(N) framework-related modules from each individual omics for further analysis. 271 As a result, we selected three CSF protein modules (M1 turquoise, M2 blue and M4 yellow), four 272 plasma protein modules (M2 blue, M3 brown, M4 yellow and M8 pink) and three plasma metabolite 273 modules (M3 brown, M4 yellow and M5 green). We then analysed the correlations between these 274 ten modules as well as between these modules and AD PRS. When analysing associations between 275 modules, we found that the metabolite M5 green module was negatively correlated with three 276 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 277 278 positive correlation was observed between metabolite M4 yellow module and plasma protein M4 279 yellow module as well as between metabolite M3 brown module and CSF protein M4 yellow module. 280 In addition, CSF protein M4 yellow module was positively associated with plasma protein M3 brown 281 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**).

287

288 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

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**).

299 We also investigated the association of these proteins/metabolites with AD PRS (APOE region 300 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 301 302 included and excluded). Similar trend was also observed for most proteins in M4 yellow module, 303 with only six proteins being positively associated with AD PRS at P_{T} =5x10⁸ (APOE region included), 304 whereas most proteins, except for three, were associated with the $P_T=0.1$ AD PRS (with APOE and 305 without APOE) (Figure 4A). For hub metabolites, three SMs in M3 brown module and three PEs in 306 M5 green module were associated with AD PRS ($P_T=5x10^{-8}$) with and without APOE region 307 respectively. However, such associations did not pass FDR correction (Figure 4A in light red). No 308 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**).

313 **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 $\varepsilon 4$ 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 $\varepsilon 4$ genotype and several metabolites/proteins were selected from LASSO. Of

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).

327 3.7 Causal links of hub proteins/metabolites with AD

328 We finally used a bidirectional two-sample Mendelian randomization to determine whether there 329 was evidence for a causal relationship of MCI conversion related hub proteins/metabolites with 330 Alzheimer's disease. Using Wald ratio estimate, we observed weak associations between PCSK7 and 331 AD as well as between COL15A1 and AD using data from Sun et al. [27]. In sensitivity analyses, the 332 causal relationship between PCSK7 and AD was replicated using summary data from an independent 333 protein GWA study by Suhre et al. [28] (Table 2). Further support for causal effects for the 334 association of PCSK7 with AD came from multiple-cis instrument MR (p<0.001 for IVW, 95% CI = 0.8 335 to 0.9, N_{SNPS} = 4, Figure S4), although this was not the case for COL15A1 (Table S10). Multiple-cis 336 instrument MR robust methods (MR-Egger and Weighted-median MR) and sensitivity analyses 337 estimates for PCSK7 were consistent with Wald ratio estimates in direction and magnitude, and 338 showed no horizontal pleiotropy or evidence of heterogeneity, further supporting the validity of the 339 MR assumptions (Table S10). In reverse MR analysis, we identified a causal association between 340 Alzheimer's disease, RCN2 and SM (Table 2, Figure S1-3). Robust methods and sensitivity analyses 341 provided additional support for such causal effects (Table S10).

342

343 **4. Discussion (words:1161)**

344 Alzheimer's disease is characterized by non-linear and heterogeneous biological alterations. Multi-345 level biological networks underlie AD pathophysiology, including but not limited to proteostasis 346 (amyloid- β and tau), synaptic homeostasis, inflammatory and immune responses, lipid and energy 347 metabolism, and oxidative stress [30]. Therefore, a systems-level approach is needed to fully capture 348 AD multifaceted pathophysiology. Here we used unbiased and high throughput multi-omics profiling 349 of AD. We applied correlation network analysis to identify modules linked to a variety of AD 350 endophenotypes including "A", "T", "N", "V", "I" and "C". We found that four modules obtained 351 from CSF proteins were associated with at least one pathology marker of "T" (P-tau), "N" (T-tau) and 352 "I" (YKL-40). Furthermore, the three "I" related modules (M1 turquoise, M2 blue and M4 yellow) 353 were enriched with either microglia or astrocytes, which are key cellular drivers and regulators of 354 neuroinflammation [31], further indicating the consistency between correlation network analysis 355 and cell type enrichment analysis. In addition, of the four modules, three were enriched with various 356 pathways which have been reported being associated with Alzheimer's such as Ras signalling 357 pathway [32], axon guidance [33], cell adhesion molecules (CAMs) [34], and lysosome pathway [35], 358 further demonstrating the relatedness of these proteins with AD.

359 From plasma metabolomics, we found that the M3 brown module was associated with "N" (T-tau) 360 and "C" (cognition) and enriched with sphingolipid and ceramide metabolism. These findings align 361 with literature report as the lipids within this module have been reported being associated with 362 cognitive progression [36] and hippocampal atrophy [37]. In addition, M4 yellow module was 363 associated with five AD pathology markers ("A", "T", "N", "I" and "C") and enriched with three 364 pathways including gamma-glutamyl amino acid, plasmalogen, and polyamine metabolism. The 365 findings are also consistent with previous reports showing that these pathways were associated with 366 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 PRS at 5×10^{-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 5×10^{-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 = 5×10^{-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.

381 We finally investigated the causal relationship between A/T/N hubs associated with MCI conversion 382 status and AD. Our MR analyses highlighted a potential weak causal relationship between PCSK7 and 383 AD which was robust in both single and multiple cis instruments MR analyses and was replicated 384 using an independent pQTL dataset. We also found a causal relationship in the opposite direction, 385 whereby AD status is potentially causally linked to RCN2 that has been proposed as a therapeutic 386 target for atherosclerosis [47]. Finally, although we didn't have GWA summary data for the SM and 387 PE hubs examined in this study, our MR analyses showed that AD was causally linked to SM levels, as 388 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].

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.

407

408 **Conflict of Interest**

409 SL is named as an inventor on biomarker intellectual property protected by Proteome Sciences and 410 Kings College London unrelated to the current study and within the past five years has advised for 411 Optum labs, Merck, SomaLogic and been the recipient of funding from AstraZeneca and other 412 companies via the IMI funding scheme. HZ has served at scientific advisory boards and/or as a 413 consultant for Abbvie, Alector, ALZPath, Annexon, Apellis, Artery Therapeutics, AZTherapies, CogRx, 414 Denali, Eisai, Nervgen, Novo Nordisk, Pinteon Therapeutics, Red Abbey Labs, reMYND, Passage Bio, 415 Roche, Samumed, Siemens Healthineers, Triplet Therapeutics, and Wave, has given lectures in 416 symposia sponsored by Cellectricon, Fujirebio, Alzecure, Biogen, and Roche, and is a co-founder of

417 Brain Biomarker Solutions in Gothenburg AB (BBS), which is a part of the GU Ventures Incubator 418 Program (outside submitted work). AL has served at scientific advisory boards of Fujirebio Europe, Eli 419 Lilly, Novartis, Nutricia and Otsuka and is the inventor of a patent on synaptic markers in CSF (all 420 unrelated to this study). JP has served at scientific advisory boards of Fujirebio Europe, Eli Lilly and 421 Nestlé Institute of Health Sciences, all unrelated to this study. SE has received unrestricted research 422 grants from Janssen Pharmaceutica and ADx Neurosciences and has served at scientific advisory 423 boards of Biogen, Eisai, icometrix, Novartis, Nutricia / Danone, Pfizer, Roche, all unrelated to this 424 study. FB is a steering committee or iDMC member for Biogen, Merck, Roche, EISAI and Prothena. 425 Consultant for Roche, Biogen, Merck, IXICO, Jansen, Combinostics. Research agreements with Merck, 426 Biogen, GE Healthcare, Roche. Co-founder and shareholder of Queen Square Analytics LTD, all 427 unrelated to this study.

428 Acknowledgements

429 This research was conducted as part of the EMIF-AD MBD project which has received support from 430 the Innovative Medicines Initiative Joint Undertaking under EMIF grant agreement no 115372, 431 resources of which are composed of financial contribution from the European Union's Seventh 432 Framework Programme (FP7/2007-2013) and EFPIA companies' in-kind contribution. The DESCRIPA 433 study was funded by the European Commission within the 5th framework program (QLRT-2001-434 2455). The EDAR study was funded by the European Commission within the 5th framework program 435 (contract # 37670). The Leuven cohort was funded by the Stichting voor Alzheimer Onderzoek (grant 436 numbers #11020, #13007 and #15005). RV is a senior clinical investigator of the Flemish Research 437 Foundation (FWO). The San Sebastian GAP study is partially funded by the Department of Health of 438 the Basque Government (allocation 17.0.1.08.12.0000.2.454.01.41142.001.H). We acknowledge the 439 contribution of the personnel of the Genomic Service Facility at the VIB-U Antwerp Center for 440 Molecular Neurology. The research at VIB-CMN is funded in part by the University of Antwerp 441 Research Fund. HZ is a Wallenberg Scholar supported by grants from the Swedish Research Council

442	(#2018-02532), the European Research Council (#681712), Swedish State Support for Clinical
443	Research (#ALFGBG-720931), the Alzheimer Drug Discovery Foundation (ADDF), USA (#201809-
444	2016862), and the UK Dementia Research Institute at UCL. FB is supported by the NIHR biomedical
445	research centre at UCLH. LS is funded by the Virtual Brain Cloud from European commission (grant
446	no. H2020-SC1-DTH-2018-1). R.G. was supported by the National Institute for Health Research (NIHR)
447	Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College
448	London. This paper represents independent research part-funded by the National Institute for
449	Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation
450	Trust and King's College London. The views expressed are those of the author(s) and not necessarily
451	those of the NHS, the NIHR or the Department of Health and Social Care. JX and CLQ thank Lundbeck
452	Fonden for the support (grant no. R344-2020-989).

453 Ethics Statement

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

456 Data Availability

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

460 References

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

469	5.	Butterfield, D.A. and D. Boyd-Kimball, Oxidative Stress, Amyloid-6 Peptide, and Altered Key
470		Molecular Pathways in the Pathogenesis and Progression of Alzheimer's Disease. J
471		Alzheimers Dis, 2018. 62 (3): p. 1345-1367.
472	6.	Ashton, N.J., et al., Blood protein predictors of brain amyloid for enrichment in clinical trials?
473	_	Alzheimers Dement (Amst), 2015. 1(1): p. 48-60.
474	7.	Hye, A., et al., <i>Plasma proteins predict conversion to dementia from prodromal disease.</i>
475	•	Alzheimers Dement, 2014. 10(6): p. 799-807.e2.
476	8.	Shi, L., et al., Discovery and validation of plasma proteomic biomarkers relating to brain
4//	0	amyloid burden by SOWAscan assay. Alzheimers Dement, 2019. 15 (11): p. 1478-1488.
478	9.	Wilkins, J.W. and E. Trushina, Application of Wetabolomics in Alzheimer's Disease. Front
479	10	Neuror, 2017. 8. p. 719. Tänning E and E Truching, Oxidative Stress Supertin Dusfunction, and Alzheimer's Disease L
400 401	10.	Alzheimers Dis 2017 57 (A): p. 1105 1121
401	11	Aizhenners Dis, 2017. 57 (4). p. 1105-1121. Kim M. et al. Brimary fatty amides in plasma associated with brain amyleid byrden.
402	11.	kin, W., et al., Frindry Jully annues in plasma associated with brain anytoid burden,
485		Alzheimer's Disease biomarker discovery cohort Alzheimers Dement 2019 15(6): n 817-827
404	12	Bos Let al. The EMIE-AD Multimodal Biomarker Discovery study: design methods and
486	12.	cohort characteristics Alzheimers Res Ther 2018 10 (1): n 64
487	13	Tims B M et al. Pathonhysiological subtypes of Alzheimer's disease based on cerebrospingl
488	10.	fluid proteomics Brain 2020 143(12) p 3776-3792
489	14.	Shi, L., et al., Discovery and validation of plasma proteomic biomarkers relating to brain
490		amyloid burden by SOMAscan assay. Alzheimer's & Dementia. 2019. 15 (11): p. 1478-1488.
491	15.	Hong, S., et al., Genome-wide association study of Alzheimer's disease CSF biomarkers in the
492		EMIF-AD Multimodal Biomarker Discovery dataset. Transl Psychiatry, 2020. 10 (1): p. 403.
493	16.	Langfelder, P. and S. Horvath, WGCNA: an R package for weighted correlation network
494		analysis. BMC Bioinformatics, 2008. 9 : p. 559.
495	17.	Xu, J., et al., Sex-Specific Metabolic Pathways Were Associated with Alzheimer's Disease (AD)
496		Endophenotypes in the European Medical Information Framework for AD Multimodal
497		Biomarker Discovery Cohort. Biomedicines, 2021. 9 (11).
498	18.	Kunkle, B.W., et al., Genetic meta-analysis of diagnosed Alzheimer's disease identifies new
499		risk loci and implicates Aβ, tau, immunity and lipid processing. Nat Genet, 2019. 51 (3): p.
500		414-430.
501	19.	Euesden, J., C.M. Lewis, and P.F. O'Reilly, <i>PRSice: Polygenic Risk Score software.</i>
502		Bioinformatics, 2015. 31 (9): p. 1466-8.
503	20.	Zettergren, A., et al., Association between polygenic risk score of Alzheimer's disease and
504		plasma phosphorylated tau in individuals from the Alzheimer's Disease Neuroimaging
505		Initiative. Alzheimers Res Ther, 2021. 13 (1): p. 17.
506	21.	Leonenko, G., et al., Identifying individuals with high risk of Alzheimer's disease using
507		polygenic risk scores. Nat Commun, 2021. 12(1): p. 4506.
508	22.	Hong, S., et al., IMEM106B and CPOX are genetic determinants of cerebrospinal fluid
509	22	Alzheimer's disease biomarker levels. Alzheimer's Dement, 2021. 17(10): p. 1628-1640.
510	23.	Tofte, N., et al., Metabolomic Assessment Reveals Alteration in Polyois and Branched Chain
511		Amino Acids Associated with Present and Future Renai Impairment in a Discovery Conort of
512	24	Borner Sons with Type 1 Diabetes. From Endocrinol (Lausanne), 2019. 10: p. 818.
515	24.	human phenome. Elife. 2018. 7
515	25	Vavorska, O, O, and S, Burgess, Mendelian Randomization: an R nackage for performing
516	2٦.	Mendelian randomization analyses using summarized data Int Enidemiol 2017 46 (6), n
517		1734-1739.
518	26.	Shi, L., et al., Replication study of plasma proteins relating to Alzheimer's pathology
519		Alzheimers Dement, 2021. 17 (9): p. 1452-1464.

520 521	27.	Sun, B.B., et al., <i>Genomic atlas of the human plasma proteome</i> . Nature, 2018. 558 (7708): p.
521	20	75-75. Subro K, at al. Connecting ganatic rick to disags and points through the human blood
522 523	20.	plasma proteome. Nat Commun, 2017. 8 : p. 14357.
524	29.	Borges, M.C., et al., Circulating Fatty Acids and Risk of Coronary Heart Disease and Stroke:
525		Individual Participant Data Meta-Analysis in Up to 16 126 Participants. J Am Heart Assoc,
526		2020. 9 (5): p. e013131.
527	30.	Hampel, H., et al., Omics sciences for systems biology in Alzheimer's disease: State-of-the-art
528		of the evidence. Ageing Res Rev, 2021. 69 : p. 101346.
529	31.	Hampel, H., et al., A Path Toward Precision Medicine for Neuroinflammatory Mechanisms in
530		Alzheimer's Disease. Front Immunol, 2020. 11 : p. 456.
531	32.	Kirouac, L., et al., Activation of Ras-ERK Signaling and GSK-3 by Amyloid Precursor Protein
532		and Amyloid Beta Facilitates Neurodegeneration in Alzheimer's Disease. eNeuro, 2017. 4(2).
533	33.	Zhang, L., et al., Roles and Mechanisms of Axon-Guidance Molecules in Alzheimer's Disease.
534		Mol Neurobiol, 2021. 58 (7): p. 3290-3307.
535	34.	Nagappan-Chettiar, S., E.M. Johnson-Venkatesh, and H. Umemori, Activity-dependent
536		proteolytic cleavage of cell adhesion molecules regulates excitatory synaptic development
537		<i>and function.</i> Neurosci Res, 2017. 116 : p. 60-69.
538	35.	Zhang, Y., et al., The role of ubiquitin proteasomal system and autophagy-lysosome pathway
539		in Alzheimer's disease. Rev Neurosci, 2017. 28 (8): p. 861-868.
540	36.	Kalinichenko, L.S., et al., Sphingolipid control of cognitive functions in health and disease.
541		Prog Lipid Res, 2022. 86 : p. 101162.
542	37.	Kim, M., et al., Association between Plasma Ceramides and Phosphatidylcholines and
543		Hippocampal Brain Volume in Late Onset Alzheimer's Disease. J Alzheimers Dis, 2017. 60 (3):
544		р. 809-817.
545	38.	Mahajan, U.V., et al., Dysregulation of multiple metabolic networks related to brain
546		transmethylation and polyamine pathways in Alzheimer disease: A targeted metabolomic
547		and transcriptomic study. PLoS Med, 2020. 17 (1): p. e1003012.
548	39.	Bozelli, J.C., Jr., S. Azher, and R.M. Epand, <i>Plasmalogens and Chronic Inflammatory Diseases</i> .
549	40	Front Physiol, 2021. 12: p. 730829.
550	40.	Liu, Y., et al., Plasma liplaome is aysregulated in Alzheimer's disease and is associated with
551	4.1	alsease risk genes. Transi Psychiatry, 2021. 11(1): p. 344.
552	41.	Mierke, M.M., et al., Plasma springomyelins are associated with cognitive progression in
553	40	Alzheimer's disease. J Alzheimer's Dis, 2011. 27 (2): p. 259-69.
554	42.	He, X., et al., Deregulation of springolipia metabolism in Alzheimer's disease. Neurobiol
555	12	Aging, 2010. 31 (5), p. 596-406. Morrow A.B. et al. CSE sphinger under metabolites in Alzheimer's disease
550	45.	nourodogonarction, and nouroinflammation. Alphoimer's & Domontic. 2021, 17: p. o052200
557	4.4	Hano D.A. and V. Dovanaravan, Serum Phoenhetidulethanolamina and
550	44.	Liano, D.A. and V. Devanarayan, serum Phosphuluyiethunorumine und
560		Predict Progression from Mild Cognitive Imngirment 1 Alzheimers Dis 2021 80 (1): p. 311-
561		210
562	45	Calzada E. O. Onguka and S.M. Claynool Phosphatidylethanolamine Metabolism in Health
563	45.	and Disease Int Rev Cell Mol Biol 2016 321 n 29-88
564	46	Ma X et al Metabolic Reprogramming of Microglig Enhances Proinflammatory Cytokine
565	40.	Release through EnhA2/n38 MAPK Pathway in Alzheimer's Disease Alzheimers Dis 2022
566	47.	Li, J., et al., Reticulocalbin 2 as a Potential Biomarker and Therapeutic Taraet for
567		Atherosclerosis. Cells. 2022. 11 (7).
568	48.	Compton, H., et al., Effects of genetic liability to Alzheimer's disease on circulatina
569		metabolites across the life course. medRxiv, 2022: p. 2022.03.24.22272867.

570 49. Damotte, V., et al., *Plasma amyloid & levels are driven by genetic variants near APOE, BACE1,*

- 571 APP, PSEN2: A genome-wide association study in over 12,000 non-demented participants.
- 572 Alzheimers Dement, 2021. **17**(10): p. 1663-1674.

573

574

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 E4 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
<i>APOE</i> 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
<i>APOE</i> 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
Amy∣oid + 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
<i>APOE</i> 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; MCl, 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.

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	

582 **Table 2. Examination of the causal relationship between hub proteins/metabolites and** 583 **Alzheimer's using bidirectional Mendelian randomization**

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

586 ** Cochran's Q p<0.001

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

588

589 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.

593 Figure 2. Individual omics modules correlating to AD endophenotypes. (A) Weighted gene 594 correlation network analysis (WGCNA) of the CSF proteomics and cell type enrichment analysis of 595 modules; (B) Enriched KEGG pathways of three modules in CSF proteins; (C) WGCNA of plasma 596 proteomics; (D) Enriched KEGG pathways of three modules in plasma proteins; (E) WGCNA of plasma 597 metabolomics; (F) Enriched KEGG pathways of two modules in plasma metabolites. * and ** denote 598 significant correlations p < 0.05 and p < 0.001 after false discovery rate (FDR) correction respectively; 599 CSF, cerebrospinal fluid; "A", amyloid; "T", tau; "N", neurodegeneration; "V", vascular; "I", 600 inflammation; "C", cognition; +, abnormality; P-tau, phosphorylated tau; T-tau, total tau; WMH, 601 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=5x10⁸ & 0.1); red and blue links denoted positive and negative correlations, respectively. CSF, cerebrospinal fluid; SNAP, Suspected Non-Alzheimer Pathology.

608 Figure 4. (A) Relation of hub proteins/metabolites with AT(N) markers and PRS; hub 609 proteins/metabolites from three CSF protein modules (M1 turquoise, M2 blue and M4 yellow), four 610 plasma protein modules (M2 blue, M3 brown, M4 yellow and M8 pink), and three plasma metabolite 611 modules (M4 yellow, M5 green and M3 brown); red, blue, light red and light blue squares denoted 612 positive association at FDR level (*pFDR*<0.05), negative association at FDR level (*pFDR*<0.05), positive 613 association at nominal level (p<0.05, pFDR>0.05), and negative association at nominal level (p<0.05, 614 pFDR>0.05), respectively. (B) Partial correlation network selected for hub metabolites/proteins, 615 genetic factors, and MCI conversion. PRS, polygenic risk score; MCI, mild cognitive impairment.



Α



В

Turquoise module



С

D

"A"

"T"

"N"

"V"

"]"

"C"

Height



Blue module

Rap1 signaling pathway -Ras signaling pathway Focal adhesion -PI3K-Akt signaling pathway Human papillomavirus infection MAPK signaling pathway Pathways in cancer JAK-STAT signaling pathway Metabolic pathways -











