An Integrative Multi-omics Approach Reveals New Central Nervous System Pathway Alterations in Alzheimer's Disease

Christopher Clark
University of Zurich: Universitat Zurich
Loïc Dayon
Nestlé Institute of Health Sciences SA: Nestle Institute of Health Sciences SA
Mojgan Masoodi
Inselspital University Hospital Bern: Inselspital Universitatsspital Bern
Gene L Bowman
Oregon Health & Science University
Julius Popp (julius.popp@pukzh.ch)
Centre Hospitalier Universitaire Vaudois https://orcid.org/0000-0002-0068-0312

Research

Keywords: Alzheimer's disease, CSF, MOFA, multi-omics, biomarkers

Posted Date: December 21st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-130933/v1

Version of Record: A version of this preprint was published at Alzheimer's Research and Therapy on April 1st, 2021. See the published version at https://doi.org/10.1186/s13195-021-00814-7.

Abstract

Background: Multiple pathophysiological processes have been described in Alzheimer's disease (AD). Their inter-individual variations, complex interrelations, and relevance for clinical manifestation and disease progression remain poorly understood, however. We tested the hypothesis that cerebrospinal fluid (CSF) integrative multi-omics analysis highlights novel interacting pathway alterations in AD.

Methods: We performed multi-level CSF omics in a well-characterized cohort of older adults including subjects with normal cognition, mild cognitive impairment, and mild dementia. Proteomics, metabolomics, lipidomics, one-carbon metabolism, and neuroinflammation related molecules were analysed applying Elastic-net regression and Multi-Omics Factor Analysis followed by pathway enrichment. Multivariate analysis was used to select best predictive models of AD pathology and cognitive decline.

Results: Multi-omics integration identified five major dimensions of heterogenicity explaining the variance within the cohort and differentially associated with AD . Further analysis exposed multiple interactions between single 'omics modalities and distinct multi-omics molecular signatures differentially related to amyloid pathology, neuronal injury, and tau hyperphosphorylation. Enrichment pathway analysis revealed overrepresentation of the hemostasis, immune response and extracellular matrix signalling pathways in association with AD. Further, combinations of four selected molecules significantly improved prediction of both AD (protein 14-3-3 zeta/delta, clusterin, interleukin-15, and transgelin-2) and cognitive decline (protein 14-3-3 zeta/delta, clusterin, cholesteryl ester 27:1 16:0 and monocyte chemoattractant protein-1).

Conclusions: Applying an integrative multi-omics approach we confirmed previously reported associations with AD pathology and report new molecular and pathways alterations. These findings are relevant for the development of personalized diagnosis and treatment approaches in AD.

Background:

Along with amyloid pathology and tau-related neurodegeneration, multiple other molecular alterations and pathway dysregulations have been reported in Alzheimer's disease (AD). However, the contribution and relevance of these alterations for clinical manifestation and progression of the disease as well as their inter-individual variations, and complex interrelations, remain poorly understood. While these processes are generally not considered part of the "core" AD pathology, they may substantially contribute to the development of amyloid pathology and neurodegeneration, and precipitate the manifestation of symptoms. As they may be occurring at early clinical and preclinical disease stages, a better understanding of these processes may be highly relevant for both early diagnosis and prognosis, and the design of targeted interventions to interfere with developing AD pathology and clinical disease progression.

'Omics approaches and technologies have made major progress over the past decade to resolve the complexity of the metabolome, lipidome and proteome (1). As powerful phenotyping technologies, 'omics

significantly accelerate the understanding of mechanisms of pathophysiological alterations that underlie complex diseases such as AD (2, 3). Beyond the potential of identifying altered biofluid molecule profiles that could be used as biomarkers, these technological advances also offer the opportunity to explore different types of molecules in parallel by combining multiple 'omics methods. Recent statistical advances have made it possible to integrate the information from multiple data modalities for a thorough exploration of endophenotype networks, and biological interactions related to disease (4). While multi-omics approaches have recently shown their potential in relation to different other pathological conditions (5–7), these methods still need to be more broadly adapted and applied in AD (8).

Here, we hypothesize that unidentified metabolic network alterations are present in AD and investigate these alterations across multiple biochemical pathways by using a multi-layer dataset acquired by analysis of cerebrospinal fluid (CSF) from a cohort of elder subjects with normal cognition, mild cognitive impairment (MCI) and mild AD dementia. In order to integrate data from different 'omics platforms in an unbiased fashion while considering interactions between modalities, we combine different approaches including single 'omics analysis and Multi-Omics Factor Analysis (MOFA) (9, 10).

Methods:

Study population:

One hundred and twenty community dwelling individuals, aged 55 or older, including subjects with normal cognition, mild cognitive impairment (MCI) or mild AD dementia (referred to as AD group) were enrolled into a brain aging study conducted in the Department of Psychiatry and the Department of Clinical Neurosciences, University Hospital of Lausanne, Switzerland. They were recruited among memory clinic outpatients or through advertisement. An overall clinical, neurological and comprehensive neuropsychological assessment was performed between 2013 and 2016, which included the Mini Mental State Examination (MMSE, (11)), and candidates with unstable medical conditions or with neurological or psychiatric diseases that could interfere with cognitive performance were excluded as previously described (12). All subjects in the AD group had clinical diagnoses of either MCI or mild dementia, and a Clinical Dementia Rating (CDR, (13)) score > 0.5 and displayed memory impairment and/or impairment in other cognitive domains such as executive tasks or language skills (14). Clinical and neuropsychological follow-up evaluations were performed at 18 and 36 months using the same methods and tests. MMSE change (MMSE score at baseline – MMSE at follow-up visit ≥ 2 was used to classify participants with decreased overall cognition.

Study procedures:

Clinical assessment:

We determined Mini-Mental State Examination (MMSE), CDR, and CDR sum of boxes (CDR-SoB), for all participants. CDR-SoB and CDR were based on the information available from the participant and his/her

relative, the clinical examination and comprehensive neuropsychological test performance, as previously described (12).

Biochemical sample collection and handling:

10–12 ml of cerebrospinal fluid (CSF) obtained from lumbar punctures conducted after an overnight fast at participant inclusion, were spun down at 4 ^{o}C, immediately aliquoted, and snap frozen at -80 ^{o}C until assayed (12). Study personnel blinded to clinical data performed biochemical and genetic analyses.

Analyte measurements:

Multiple 'omics data from different pathways and various biological levels were acquired. We used mass spectrometry (MS) for proteomics, one-carbon metabolism and lipidomics, proton nuclear magnetic resonance (^{1}H NMR) for metabonomics, and immunoassays for neuroinflammation. The initial number of analytes measured in CSF, the final number of analytes selected per platform (a total of 891 analytes covered), and quantification method used for each platform are described in Table 1. Further details about the analysis methods and analytes measured by each platform can be found in **Additional File 1**.

Genetic measures:

The *APOE* genotype was determined by PCR as previously described (15). Participants with one or more e4 alleles were classified as carriers.

Table 1
Datasets used in this study

Dataset	Analytes Initial/Final	Quantification Technique	References
Proteomics	791/772	LC-MS/MS	(16, 17)
Neuroinflammation	38/21	Multi-array sandwich immunoassay	(15, 18, 19)
One-carbon metabolism	17/9	LC-MS/MS	(20, 21)
Metabonomics	71/63	^{1}H NMR	(22)
Lipidomics	65/26	MS	(23, 24)
Biomarkers of core AD pathology	3/3	ELISA	(18)

Available datasets from the cohort along with the number of analytes measured in this study and the associated quantification methods. For each dataset the initial number of analytes quantified, the number of measurements remaining after quality control, quantification technique used, and technical references are indicated. LC-MS/MS, liquid chromatography tandem mass spectrometry; ^{1}H NMR, Proton nuclear magnetic resonance; MS, mass spectrometry; ELISA, enzyme-linked immunosorbent assay.

Cerebrospinal fluid AD biomarkers:

CSF beta-amyloid 1–42 ($A\beta_{1-42}$), total-tau (Tau), and tau phosphorylated at threonine 181 (P-Tau) concentrations were measured using commercially available ELISA kits (Fujirebio, Gent, Belgium). AD neuropathology was defined as P-Tau/$A\beta_{1-42}$ ratio > 0.0779 as described previously (12).

Data preparation and transformation:

Lipidomics: 26 high quality intact lipids with less than 5% of null values were selected as continuous numerical markers from 65 original measurements. Numerical lipid marker values were log10-transformed prior to analysis.

Metabolomics: 71 peak integrals were originally measured in CSF. 63 analytes with less than 5% missing values were selected from the obtained spectra. Peak integral values were log10-transformed prior to analysis.

One-carbon metabolomics: 17 analytes were initially measured in CSF. Some analytes could not be measured in the majority of samples and were excluded from the analysis (*i.e.*, homocysteic acid, dimethylglycine, betaine, total homocysteine, pyridoxine and pyridoxamine); taurine and glycine data were inconsistent and were also filtered out resulting in 9 measured analytes (*i.e.*, choline, cystathionine, methionine, riboflavin, S-adenosylhomocysteine, S-adenosylmethionine, serine, cysteine and 5-methyltetrahydrofolate). Numerical values were log10-transformed prior to analysis.

Neuroinflammatory markers: 38 markers were measured in CSF. Calibration curves, batch effects, and lower limit of quantification were controlled. After this quality control, 17 markers were removed, resulting in 21 markers selected. Concentrations were log10-transformed prior to analysis.

Proteomics: Relative quantification data were available for all subject samples as log2 ratios as previously described (16). Analytes with more than 30% missing data points were removed, resulting in 772 proteins measured from an initial number of 791.

Before analysis, outliers (*i.e.*, data points that exceeded the cutoff value of mean ± 3 × standard deviation) were replaced by the cutoff value in all datasets.

Statistical and analytical approaches:

Descriptive statistics for the cohort were performed using t-tests comparing AD and control groups for continuous variables and Chi-square tests for categorical variables. Data was clustered by hierarchical clustering across samples and factors values or loadings.

Feature selection methods:

Single-modality approaches: To overcome the bias resulting from correlation between variables and thus unreliability and saturation of standard regression techniques, we used Elastic-Net regularization (α = 0.5) for regression analysis. This was performed separately for each individual 'omics platform using R software with custom routines implementing the *glmnet* package (25). Features associated with each pre-specified CSF biomarker endpoint were considered as a continuous dependent variable and selected using the minimal lambda statistic.

Multi-omics factor analysis: This analysis was performed using the MOFA package in R and Python software (9). Latent factors (also referred as LFs) were selected to explain a minimum of 2% variance in at least one data type. The MOFA model was trained over 938 iterations with a convergence threshold of 0.1. Individual analytes were considered strongly associated with a given LF if their normalised absolute loading value was > 0.8. More details about the MOFA method can be found in **Additional File 1**. The trained MOFA model was validated using both a correlation approach and CSF AD biomarker predictions (**Additional File 2**, **Fig. S1** and **Fig. S2**, respectively)

Models for AD pathology and cognitive decline prediction:

Predictions were ran using the *glm* package in R. We first constructed a reference models using age, sex, years of education, baseline MMSE score and APOE4 carrier status for AD pathology prediction and age, sex, years of education, baseline MMSE score, APOE4 carrier status and time to last follow-up for cognitive decline prediction. We then used an iterative approach, first adding all analytes identified by the MOFA model and selected the model displaying the smallest Akaike information criterion (AIC) value, before performing the following iteration with the remaining analytes. Performance of the models was analyzed by comparing area under the curve (AUC) of the resulting ROC curves using the DeLong method. No further improvements to the AUC was observed after five iterations for both predictions. Confusion matrices to assess sensitivity and specificity were calculated for all models.

Pathway enrichment:

Proteins selected by the MOFA model were searched for in the UniProt database (26) and their entry number was then subsequently used within the Reactome database (27). A separate over-representation analysis was performed for each LF. This analysis used hypergeometric distribution to determine which pathways and biological reactions were over-represented within the dataset. Over-represented pathways were then manually grouped into broader ontology-based categories (**Additional File 3**, **Table S1**).

Results:

Cohort description: The clinical and demographical characteristics of the participants included in this study are shown in Table 2.

Table 2
Study cohort

	Control (n = 79)	AD group (n = 41)	P-value
Age (years)	68.42 ± 8.23	74.15 ± 5.7	< 0.001
Sex (%, female)	67.1	58.5	0.354
Education (years)	12.51 ± 2.7	12.10 ± 12.1	0.404
CDR-SoB	0.456 ± 0.9	2.20 ± 2.03	< 0.001
MMSE	27.85 ± 2.28	25.15 ± 3.71	< 0.001
P-Tau/$A\beta_{1-42}$ ratio	0.048 ± 0.127	0.165 ± 0.104	< 0.001

Characteristics of the study cohort. P-value was obtained from t-test for continuous variables or chi-square statistics for sex.

Single-modality feature selection: Elastic-Net regression within each single 'omics modality identified 82 molecules associated with CSF "core" biomarkers of AD pathology (*i.e.*, $A\beta_{1-42}$, Tau and P-Tau, **Additional File 3**, **Table S2**). Strikingly, distinct panels of CSF analytes were associated with either $A\beta_{1-42}$, or Tau and P-Tau, reflecting alterations of different pathways in relation to amyloid pathology, neurodegeneration, and tau pathology, with very little overlap (Fig. 1). Only protein 14-3-3 zeta/delta was associated with all three biomarkers.

Overview of the MOFA model: In parallel, we trained a MOFA model, to identify major dimensions of heterogenicity (latent factors; LFs) responsible for the variance within the cohort. This model identified five LFs that each explained a minimum of 2% variance in at least one of the analysed metabolic levels. Among these factors, LF1 and LF2 were present in most multi-omics modalities, revealing a broad participation to variance within the cohort (Fig. 2). On the other hand, the remaining LFs only captured variance across some modalities (three for LF4 and LF5, two for LF3) and had a smaller contribution to overall variance. Across all LFs, the CSF AD biomarkers accounted for 38.5%, proteins 39.8%, lipids 10.3%, neuroinflammation markers 10.3%, one-carbon metabolites 9% and other metabolites 3.7% of the variance contained within the cohort (Fig. 2). We next produced clustered heatmaps of the weight (*i.e.*, the association of an individual molecule with the LF) of each analyte across different LFs (Fig. 3A-E). This analysis revealed that amongst LFs there are specific contributions of analytes from different 'omics levels. For example, molecules within the one-carbon metabolism were differentially associated with LF1 and LF2 (Fig. 1C). Since only three CSF core AD biomarkers were measured, we did not produce heatmaps to analyse the association of $A\beta_{1-42}$, Tau and P-Tau with these five LFs; but rather we inspected their absolute individual loadings across all LFs (Fig. 4). This revealed that individual CSF AD biomarkers had different contributions across the identified LFs. CSF Tau and P-Tau levels were strongly associated with LF1, LF2 and LF3, while $A\beta_{1-42}$ was the main contributor to variance amongst CSF AD

biomarkers to LF4 and LF5 indicating that these latter LFs were associated with amyloid pathology and the former with tau pathology and related neurodegeneration.

Individual analyte contributions to LFs: We next addressed the contribution of individual molecules to variance within the cohort and how these molecules aligned with CSF AD biomarkers (Table 3a for proteins and Table 3b for other analytes). Using their normalised absolute loadings, we selected 36 proteins, 7 neuroinflammatory markers, 3 one-carbon metabolites, 5 lipids and 7 other metabolites (not counting analytes selected in multiple LFs) that contributed the most to variance within the cohort. Since each individual LF was strongly associated with an individual CSF AD biomarker, we next investigated the relationship between the identified analytes and the expression levels of CSF AD biomarkers in the relevant LFs. Four molecules of interest, manually selected from three different modalities, and their associations with LFs and therefore with either $A\beta_{1-42}$ and Tau (or P-Tau) are presented in Fig. 5. Additionally, 14 of the selected proteins were associated with the presence of cognitive impairment at baseline (**Additional File 3**, **Table S3**).

Table 3a
Proteins associated with latent factors

LF	Analyte	Full Name	Entry#	Previously reported AD association
Proteomics				
1	NRN1	Neuritin isoform 1 precursor	Q9NPD7	
1	SMS	Spermine synthase	P52788	Yes (48)
1	NXPH4	Neurexophilin-4	O95158	
1	LTBP1	Latent-transforming growth factor beta-binding protein 1	Q14766	
1	CLUS	Clusterin	P10909	Yes (49)
1	NPDC1	Neural proliferation differentiation and control protein 1	Q9NQX5	
1	PNOC	Prepronociceptin	Q13519	
1	DYL2	Dynein light chain 2, cytoplasmic	Q96FJ2	
1	PDGFB	Platelet-derived growth factor subunit B	P01127	Yes (50)
1	SAP3	Sphingolipid activator protein 3	P17900	
1	MT1E	Metallothionein-1E	P04732	Yes (51)
1	PCSK1	Neuroendocrine convertase 1	P29120	Yes (52)
1	TAGL	Transgelin-2	P37802	Yes (53)
1	MT3	Metallothionein-3	P25713	Yes (51)
1	LY6H			
2	SAMP	Spindle-associated membrane protein 1	Q5SNT2	
2	VTNC	Vitronectin	P04004	Yes (54)
2	KNG1	Kininogen-1	P01042	
2	FETUA	Alpha-2-HS-glycoprotein	P02765	Yes (55)
2	HELZ	Probable helicase with zinc finger domain	P42694	
2	PLMN	Plasminogen	P00747	Yes (56)
2	PGRP2	N-acetylmuramoyl-L-alanine amidase	Q96PD5	
2	AFAM	Afamin	P43652	Yes (55)
2	ITIH1	Inter-alpha-trypsin inhibitor heavy chain H1	P19827	Yes (57)

LF	Analyte	Full Name	Entry#	Previously reported AD association
2	CO8B	Complement component C8 beta chain	P07358	Yes (58)
2	FIBA	Fibrinogen alpha chain	P02671	Yes (59)
2	CO6	Complement Component C6	P13671	Yes (58)
2	ITIH4	Inter-alpha-trypsin inhibitor heavy chain H4	Q14624	Yes (57)
3	EPDR1	Mammalian ependymin-related protein 1	Q9UM22	
3	SIAE	Sialate O-acetylesterase	Q9HAT2	
4	X1433Z	14-3-3 protein zeta/delta	P63104	Yes (60)
4	S10A6	Protein S100-A6	P06703	Yes (61)
4	PRDX6	Peroxiredoxin-6	P30041	Yes (62)
5	VTM2A	V-set and transmembrane domain-containing protein 2A	Q8TAG5	
5	S10A6	Protein S100-A6	P06703	Yes (61)
5	CMGA	Chromogranin-A	P10645	Yes (63)
5	ZP2	Zona pellucida sperm-binding protein 2	Q05996	
5	SLIK1	SLIT and NTRK-like protein 1	Q96PX8	Yes (64)

Table 3a: CSF proteins significantly associated with the LFs within the MOFA model and whether they have been previously associated with AD. Entry# denotes the protein identifier within the UniProt database.

Table 3b

Other molecules associated with latent factors

LF	Analyte	Full Name	Entry#	Previously reported AD association
Neuroinflammation				
1	sVCAM-1	Circulating vascular cell adhesion molecule-1	P19320	Yes (65)
1	IL-15	Interleukin-15	P40933	Yes (65)
1	sICAM-1	Soluble Intracellular adhesion molecule-1	P05362	Yes (65)
2	SAA	Serum amyloid A	P0DJI8	Yes (66)
2	PIGF_1R	Insulin-like growth factor 1 receptor	P08069	Yes (67)
3	PIGF_1R	Insulin-like growth factor 1 receptor	P08069	Yes (67)
4	IL-16	Interleukin-16	Q14005	Yes (68)
5	MCP-1	Monocyte chemoattractant protein-1	P13500	Yes (65)
5	PIGF_1R	Insulin-like growth factor 1 receptor	P08069	Yes (67)
One-carbon metabolism				
1	MTHF	5-methyltetrahydrofolate	20612	Yes (69)
1	SAH	S-adenosyl-L-homocysteine	16680	Yes (70)
2	CYST	Total Cysteine	15356	Yes(71)
3	SAH	S-adenosyl-L-homocysteine	16680	Yes (70)
4	CYST	Total Cysteine	15356	Yes (71)
5	CYST	Total Cysteine	15356	Yes (71)
Metabonomics				
1	N/A	Glycoproteins	17089	Yes (72)
2	N/A	Alanine	16449	Yes (73)
2	N/A	Valine	27266	
2	N/A	Glycoproteins	17089	Yes (72)
3	N/A	Inositol	24848	
4	N/A	Glycoproteins	17089	Yes (72)
5	N/A	Formic acid	30751	
5	S69	Unidentified metabolite	N/A	N/A

5	N/A	Acetoacetic acid	15344	Yes (74)
Lipidomics				
1	PC 32:0	1,2-dihexadecanoyl-sn-glycero-3-phosphocholine	N/A	(41)
2	SE 27:1 18:2	Cholesteryl ester	N/A	
3	PC 32:0	1,2-dihexadecanoyl-sn-glycero-3-phosphocholine	N/A	(41)
4	SE 27:1 18:2	Cholesteryl ester	N/A	
4	SE 27:1 20:4	Cholesteryl ester	N/A	
4	SE 27:1 16:0	Cholesteryl ester	N/A	
5	LPG 20:1	1-(11Z-eicosenoyl)-glycero-3-phospho-(1'-sn-glycerol)	N/A	

Table 3b: CSF biomolecules significantly associated with the LFs within the MOFA model and whether they have been previously associated with AD. Entry# denotes the analyte identifier within the UniProt database (for neuroinflammation) or ChEBI database (for other analytes).

Prediction of AD pathology and cognitive decline using MOFA-selected molecules: To confirm *a posteriori* the importance of the molecules selected by our MOFA model in clinical practice, we analysed their contribution to prediction of both AD pathology and global cognitive decline (see **Methods**). The addition of the four following analytes: protein 14-3-3 zeta/delta, clusterin, interleukin-15, and transgelin-2, improved the AUC of the ROC curve when compared to the reference model for AD prediction (Fig. 6A, p-value < 0.001). The addition of protein 14-3-3 zeta/delta, clusterin, cholesteryl ester 27:1 16:0 and monocyte chemoattractant protein-1, improved prediction on cognitive decline (Fig. 6B, p-value = 0.047)

Metabolic pathway enrichment: Using the Reactome database and coarse-grain ontological categories (See **Methods** and **Additional File 3, Table S1**), we investigated which biological pathways were over-represented within each LF for the proteomic modality. Other modalities were not analysed in this fashion since they were selected *a priori* to represent distinct metabolic pathways (one-carbon metabolism and inflammatory markers) or did not contain enough molecules to conduct pathway analysis. Lipids were also excluded from this analysis since our quantification method did not allow to dissociate between different isoforms of compounds with the same chemical formula. This approach revealed an overrepresentation of the hemostasis (28.8%), immune response (20.8%) and extracellular matrix signalling pathways (8.8%) (Fig. 7).

Discussion:

Here, we applied a multi-layered integrative approach to disentangle sources of variance within a cohort of elderly participants with normal cognition, mild cognitive impairment or mild AD dementia. We identified five major dimensions of heterogenicity that together comprehensively explained the variance within the cohort and were associated with core AD pathology. Further analysis revealed multiple interactions between single 'omics modalities, distinct multi-omics molecular patterns differentially associated with amyloid aggregation, neurodegeneration, and tau hyperphosphorylation, and novel molecules associated with cognitive impairment. Specific signatures of four molecules improved the accuracy of both AD and cognitive decline prediction. Additionally, pathway enrichment showed over-representation of the hemostasis, immune response and extracellular matrix signalling pathways in association with AD.

Single modality feature selection: We first used Elastic-net regression, to identify molecules associated with individual biomarkers of CSF AD pathology without considering any possible interactions between different 'omics modalities. This approach identified several proteins (SPARC-related modular calcium-binding protein 1, brain acid soluble protein 1, neuromodulin, pyruvate kinase PKM, thymosin beta-10, 14-3-3 protein zeta/delta, and fructose-bisphosphate aldolase A) in strong accordance with recent studies of the AD CSF proteome (16, 28). The zeta/delta isoform of protein 14-3-3 was associated with Aβ1–42, Tau, and P-Tau levels. This apoptosis inhibitor, one of the most abundant proteins in the brain, was previously found to exhibit altered levels in AD and modulate AD risk (29, 30). We also identified associations of neurofilament medium polypeptide with Tau levels and of reelin with $A\beta_{1-42}$ and Tau levels. Both these molecules have previously been associated with AD (31–33). Regarding neuroinflammatory molecules, C-reactive protein and monocyte chemoattractant protein-1 have previously been associated with AD, albeit in plasma (34). In addition, we have also previously shown that soluble intracellular cell adhesion molecule-1 in CSF is associated with AD (18). At metabolite level, we identified 10 molecules in CSF associated with Tau and P-Tau, which differ from the blood biomarkers associated with AD identified in a recent study in a large sample (35). Overall, our approach identified more molecules associated with AD pathology as compared to previous studies. A likely source of differences is the use of Elastic-Net regression in the current study which eliminates saturation of the regression and could therefore identify more associations.

Heterogenicity within the cohort: An important strength of our study is to consider all interactions between multiple biological levels and their associations with the heterogenicity within the cohort. This was achieved by training a MOFA model on the multi-omics dataset which has the advantage of not giving any additional analysis weight to the established CSF biomarkers of core AD pathology while also reducing the complexity of the data to better depict the sources of variation. This revealed proteomic measures and CSF core AD biomarkers as the main contributors to the variance. This was to be expected since i) protein expression levels do not only reveal changes related to AD pathology, but also reflect the effects of different environments, life style, health conditions, and genetic backgrounds; all factors potentially affecting protein expression and regulation (36); and ii) our sample contains a large proportion of participants with AD, each displaying CSF AD biomarkers significantly different from

subjects without AD. Nonetheless, this approach identified 21 proteins with previously reported association to AD, suggesting the MOFA approach can accurately disentangle the inter-individual heterogeneity driven by AD pathology and differentiate between individual (*i.e.*, not repeated in the dataset) and cohort heterogeneity (*i.e.*, underlying changes in many participants). Conversely, the metabolomic dataset was only responsible for a small amount of the cohort heterogeneity (3.7%), a possible explanation being that it represents individual heterogeneity for the most part caused by the environment, disease processes or nutritional habits. This low contribution of metabolomics to variance could also result from the lower dimensionality of the metabolomics dataset as molecules within had lower concentrations compared to molecules in the other modalities. Yet, despite this low level of variance, our model was able to correctly retrieve metabolites previously reported in association with AD, underlining the sensitivity of the model. This is further supported by the ability of our model to identify a four-molecule signature that improves the prediction of the occurrence of AD pathology, confirming that the identified LFs and molecules reflect metabolic differences resulting from the presence of AD pathology rather than from other factors. This also confirms the clinical and diagnosis relevance of the identified molecules.

Associations between LF and specific aspects of AD pathology: We next investigated how individual LFs related to specific aspects of AD pathology by comparing which CSF AD biomarkers were most strongly associated with each of them. This revealed that LFs 1–3 were primarily associated with CSF Tau levels suggesting relationships with neuronal injury while LF4 and LF5 were mainly associated with CSF $A\beta_{1-42}$ suggesting implication in the development of amyloid pathology. In LF3 and LF4, the association of $A\beta_{1-42}$ was opposite to those of Tau and P-Tau. We speculate that the molecules within and the associated alterations could play a role in both amyloid aggregation and tau-related neurodegeneration or represent a consequence of developing cerebral AD pathology.

Interactions between LFs and 'omics modalities: Besides the identification of molecular profiles and metabolic pathways alterations associated with AD, our approach also disentangled how components of individual LFs interact with each other to explain variance within the cohort. In other words, the contribution of individual LFs to total variance results from a specific combination of the different 'omics modalities. Indeed, while the variance explained by LF1 and LF2 was associated with all modalities, other LFs only contained a subset of these (one-carbon metabolism and metabolomics were only very weakly associated with LF3 and LF4, whereas lipidomics was nearly absent from LF3 and LF5); revealing specific interactions between a subset of molecules and particular metabolic pathways. Individual molecules also presented different patterns of association across LFs. For example, a subset of lipids, including PC 32:0, PC 34:1, LPA 18:3 and TAG 54:3, had a strong positive association with LF2 and a weak negative association with LF4. Since LF2 was associated with all tested modalities, this suggests these analytes interact within multiple biological pathways and could be within a hub of metabolic changes. LF2 is associated with both Tau and P-Tau; neurodegeneration and tau pathology could therefore relate to a more general metabolic alteration. The association of PC 32:0 with tau pathology in single 'omics approaches supports this assumption. In contrast, LF4 is strongly associated with amyloid

pathology and it is only associated with changes in lipids and proteins (in addition to CSF AD biomarkers). Therefore, only a subset of lipids appears to interact directly with amyloid pathology.

Novel associations uncovered by the MOFA model: The MOFA model uncovered additional relationships not revealed by single 'omics exploration paradigms, such as the association of S-adenosylhomocysteine and glycoproteins associated with CSF $A\beta_{1-42}$, Tau and P-Tau; and of total cysteine associated with $A\beta_{1-42}$. Indeed, since the trained MOFA model did not only consider molecules from one modality but the whole dataset from different 'omics, it was able to reveal additional associations resulting from the downstream effects of these molecules or from interactions with other modalities. While several analytes identified by the trained model have previously been associated with AD (see **Additional File 1**), it also uncovered novel associations, such as dynein light-chain 2, cytoplasmic (DYL2) and neurexophilin-4 (NXPH4). Both were associated with LF1 and with cognitive impairment (**Additional File 3, Table S3**). DYL2 is thought to regulate dynein function (37) and maintain cytoskeletal structure, therefore regulating synaptic function (38). NXPH4 structurally resembles neurexophilin-1, an α-neurexin ligand, which promotes adhesion between dendrites and axons and modulates specific cerebellar synapses and motor functions (39). Altered levels of these proteins may therefore be associated with neurodegeneration processes and related to cognitive impairment in AD. Another novel analyte we identified is the cholesteryl ester SE 27:1 16:0. While links between phosphatidylcholine metabolism and AD in general (40) and PC 32:0 in particular (41) have been previously reported, to our knowledge cholesteryl esters have not previously been associated with AD pathology. In our MOFA model, this cholesteryl ester was strongly correlated to LF4, suggesting a role in amyloid pathology. These molecules were also associated with cognitive performance as measured by MMSE (**Additional File 3, Table S4**). Together, these results demonstrate the capacity of integrative multi-omics to provide additional insights into the relationship of molecular alterations with specific aspects of the AD pathology.

Prediction of AD pathology and cognitive decline using MOFA-selected molecules: Molecular signatures associated with AD or predictive of cognitive decline were derived from our model. Both signatures contain four molecules each, taken from multiple biological levels, and share two common molecules, protein 14-3-3 zeta/delta and clusterin, suggesting common biological pathways associated with AD and cognitive decline. Both signatures also significantly improved the prediction performance when added to reference models. These findings demonstrate the ability of our model to identify molecules reflecting metabolic differences related to AD pathology or cognitive decline rather than to other factors. While these results need validation in an independent cohort, they already demonstrate the ability of our model to identify biomarker combinations that may be used in clinical practice.

Infer pathway relationships with AD pathology: One important strength of the MOFA approach is that it enables addressing the relationship between multiple biological pathways and associate them with sources of variance (*i.e.*, LFs). Using over-representation of metabolic pathways, we were able to show that individual LFs, and the main related pathological aspects of AD (*i.e.*, amyloid aggregation, neurodegeneration and tau pathology) are associated with distinct pathways. Hemostasis and immune response were the most over-represented. Only the immune response was associated with all LFs in

which individual pathways could be identified. LF1 and LF2 presented a significant enrichment in biomolecules implicated in hemostasis, suggesting an association between this pathway and neuronal injury, and tau pathology. While an association between hemostasis and amyloid pathology pathway was previously described (42), in particular related to expression of amyloid precursor protein and release of Aβ (43), there have also been recent reports of an association between Tau and hemostasis (44). Molecules involved in the extracellular matrix were significantly enriched in LF2, also suggesting an association with tau-related pathology, in line with previous reports (45). However, this pathway was not detected within LF1 or other LFs. We therefore hypothesise that the molecules involved are those presenting a specific pattern of association with LF2, such as PC 32:0, PC 34:1, LPA 18:3 and TAG 54:3. Neuronal function was confined to associations with LF5, suggesting little variation and differences in signal transmission and synaptic function across the cohort since this LF only explained 8% of the variance. Nonetheless, this result suggests an association with amyloid pathology, which is in accordance with previous findings of amyloid being released in an activity-dependent fashion from neurons and modulating synaptic function and plasticity (46, 47). Overall, the enriched metabolic pathways suggest that AD pathology affects not only pathways related to neuronal biological systems but is linked to a broader spectrum of metabolic dysfunctions.

Limitations: The inclusion of some targeted analysis results in the multi-omics models may be considered as a limitation. While the proteomic and lipidomic dataset are hypothesis-free measurements and the study could be limited to this data, we chose to include further available modalities. In particular, we considered neuroinflammation and one-carbon metabolism given their previously reported associations with AD and relevance for brain metabolism. The replication of these and other previously reported associations in our MOFA model supports the validity of the new findings revealed in the present study. Our findings, in particular the identified biomarker combinations need validation in independent cohorts.

Conclusions:

Here, applying integrative multi-omics in AD, we have identified five axes of variation within a cohort of individuals with or without cognitive impairment and AD pathology. These five LFs were associated with different aspects of the core AD pathology. We confirmed several previously reported associations with AD core pathology and identified new molecular patterns interrelated within each LF. Additionally, we identified molecular biomarker signatures improving the diagnosis of AD pathology and the prediction of future cognitive decline. Furthermore, using pathway enrichment analysis, we have revealed metabolic pathways represented within single LFs and explored specific relationships with markers of amyloid pathology, neuronal injury, and tau hyperphosphorylation. These findings demonstrate the added value of integrative multi-omics analysis to uncover interrelated pathway alterations in AD and its ability to identify biomarker combinations that and may be used in clinical practice. This is relevant for the development of both personalized diagnosis and tailored therapeutic interventions in AD.

Abbreviations:

$A\beta_{1\text{-}42}$: CSF beta-amyloid 1-42; AD: Alzheimer's disease; AIC: Akaike information criterion; AUC: area under the curve; CDR: clinical dementia rating; CDR-SoB: CDR sum of boxes; CSF: cerebrospinal fluid; ^{1}H NMR: proton nuclear magnetic resonance; LF: latent factor; MCI: mild cognitive impairment; MMSE: Mini mental state examination; MOFA: Multi-omics factor analysis; MS: mass spectrometry; P-Tau: tau phosphorylated at threonine 181; Tau: total-tau

Declarations:

Ethics approval and consent to participate:

The local ethics committee of Vaud, Switzerland approved this study (No. 171/2013), and all participants or their legal representatives provided written informed consent.

Consent for publication:

Not applicable

Availability of data and materials:

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request

Competing interests:

JP received consultation honoraria from Nestle Institute of Health Sciences, Innovation Campus, EPFL, Lausanne, Switzerland, Ono Pharma, and from Fujirebio Europe. The other authors declare no potential conflicts of interest.

Funding:

This work was supported by grants from the Swiss National Research Foundation (to JP, SNF 320030_141179), Synapsis Foundation - Alzheimer Research Switzerland (to JP, grant number 2017-PI01) and funding from the Nestlé Institute of Health Sciences.

Authors' contributions:

CC and JP designed the study. LD, MM and GB designed the experiments and analyzed multi-omics data. CC performed data processing, bioinformatics analysis, and generated text and figures for the manuscript. CC, LD, MM, GB and JP designed and established the study cohorts and contributed to writing the paper. JP supervised and guided the project. CC and JP wrote the manuscript. All authors revised and approved the final draft.

Acknowledgements:

We thank Antonio Nuñez Galindo, John Corthésy, Ornella Cominetti, Delphine Morin-Rivron, Seu-Ping Guiraud, Laeticia Da Siva, François-Pierre Martin, Sofia Moço, Sebastiano Collino, Martin Kussmann, Sylviane Métairon, Julien Marquis, Patrick Descombes and Richard Kirkland for biological and genetic analyses.

References:

1. Sancesario GM, Bernardini S. Alzheimer's disease in the omics era. Clin Biochem 2018; 59:9–16.
2. Dayon L, Cominetti O, Wojcik J, Galindo AN, Oikonomidi A, Henry H et al. Proteomes of Paired Human Cerebrospinal Fluid and Plasma: Relation to Blood-Brain Barrier Permeability in Older Adults. J Proteome Res 2019; 18(3):1162–74.
3. van der Velpen V, Teav T, Gallart-Ayala H, Mehl F, Konz I, Clark C et al. Systemic and central nervous system metabolic alterations in Alzheimer's disease. Alzheimers Res Ther 2019; 11(1):93.
4. Huang S, Chaudhary K, Garmire LX. More Is Better: Recent Progress in Multi-Omics Data Integration Methods. Frontiers in genetics 2017; 8:84.
5. Su M-W, Chang C-K, Lin C-W, Ling S-J, Hsiung C-N, Chu H-W et al. Blood multiomics reveal insights into population clusters with low prevalence of diabetes, dyslipidemia and hypertension. PLoS ONE 2020; 15(3):e0229922.
6. Chalikiopoulou C, Jenko Bizjan B, Leventopoulos G, Smaili K, Blagus T, Menti A et al. Multiomics Analysis Coupled with Text Mining Identify Novel Biomarker Candidates for Recurrent Cardiovascular Events. OMICS 2020.
7. Wang B, Lunetta KL, Dupuis J, Lubitz SA, Trinquart L, Yao L et al. Integrative Omics Approach to Identifying Genes Associated With Atrial Fibrillation. Circ Res 2020; 126(3):350–60.
8. Badhwar A, McFall GP, Sapkota S, Black SE, Chertkow H, Duchesne S et al. A multiomics approach to heterogeneity in Alzheimer's disease: focused review and roadmap. Brain 2019.
9. Argelaguet R, Velten B, Arnol D, Dietrich S, Zenz T, Marioni JC et al. Multi-Omics Factor Analysis-a framework for unsupervised integration of multi-omics data sets. Molecular systems biology 2018; 14(6):e8124.
10. Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine Learning and Integrative Analysis of Biomedical Big Data. Genes 2019; 10(2).
11. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". Journal of psychiatric research 1975; 12(3):189–98. Available from: URL: https://pubmed.ncbi.nlm.nih.gov/1202204/.
12. Tautvydaite D, Antonietti JP, Henry H, Gunten A von, Popp J. Relations between personality changes and cerebrospinal fluid biomarkers of Alzheimer's disease pathology. Journal of psychiatric research 2017; 90:12–20.

13. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993; 43(11):2412–4.
14. Mathys J, Gholamrezaee M, Henry H, Gunten A von, Popp J. Decreasing body mass index is associated with cerebrospinal fluid markers of Alzheimer's pathology in MCI and mild dementia. Exp Gerontol 2017; 100:45–53.
15. Bowman GL, Dayon L, Kirkland R, Wojcik J, Peyratout G, Severin IC et al. Blood-brain barrier breakdown, neuroinflammation, and cognitive decline in older adults. Alzheimer's & dementia : the journal of the Alzheimer's Association 2018; 14(12):1640–50.
16. Dayon L, Nunez Galindo A, Wojcik J, Cominetti O, Corthesy J, Oikonomidi A et al. Alzheimer disease pathology and the cerebrospinal fluid proteome. Alzheimer's research & therapy 2018; 10(1):66.
17. Dayon L, Wojcik J, Nunez Galindo A, Corthesy J, Cominetti O, Oikonomidi A et al. Plasma Proteomic Profiles of Cerebrospinal Fluid-Defined Alzheimer's Disease Pathology in Older Adults. Journal of Alzheimer's disease : JAD 2017; 60(4):1641–52.
18. Popp J, Oikonomidi A, Tautvydaite D, Dayon L, Bacher M, Migliavacca E et al. Markers of neuroinflammation associated with Alzheimer's disease pathology in older adults. Brain, behavior, and immunity 2017; 62:203–11.
19. Oikonomidi A, Tautvydaite D, Gholamrezaee MM, Henry H, Bacher M, Popp J. Macrophage Migration Inhibitory Factor is Associated with Biomarkers of Alzheimer's Disease Pathology and Predicts Cognitive Decline in Mild Cognitive Impairment and Mild Dementia. Journal of Alzheimer's disease : JAD 2017; 60(1):273–81.
20. Oikonomidi A, Lewczuk P, Kornhuber J, Smulders Y, Linnebank M, Semmler A et al. Homocysteine metabolism is associated with cerebrospinal fluid levels of soluble amyloid precursor protein and amyloid beta. Journal of neurochemistry 2016; 139(2):324–32.
21. Guiraud SP, Montoliu I, Da Silva L, Dayon L, Galindo AN, Corthesy J et al. High-throughput and simultaneous quantitative analysis of homocysteine-methionine cycle metabolites and co-factors in blood plasma and cerebrospinal fluid by isotope dilution LC-MS/MS. Analytical and bioanalytical chemistry 2017; 409(1):295–305.
22. Lauria M, Persico M, Dordevic N, Cominetti O, Matone A, Hosking J et al. Consensus Clustering of temporal profiles for the identification of metabolic markers of pre-diabetes in childhood (EarlyBird 73). Scientific reports 2018; 8(1):1393.
23. Surma MA, Herzog R, Vasilj A, Klose C, Christinat N, Morin-Rivron D et al. An automated shotgun lipidomics platform for high throughput, comprehensive, and quantitative analysis of blood plasma intact lipids. European journal of lipid science and technology : EJLST 2015; 117(10):1540–9.
24. Chakrabarti A, Eiden M, Morin-Rivron D, Christinat N, Monteiro JP, Kaput J et al. Impact of multi-micronutrient supplementation on lipidemia of children and adolescents. Clin Nutr 2019.
25. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J. Stat. Soft. 2010; 33(1).
26. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 2019; 47(D1):D506-D515.

27. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A et al. The reactome pathway knowledgebase. Nucleic acids research 2019.
28. Zhou M, Haque RU, Dammer EB, Duong DM, Ping L, Johnson ECB et al. Targeted mass spectrometry to quantify brain-derived cerebrospinal fluid biomarkers in Alzheimer's disease. Clin Proteom 2020; 17(1):1–14.
29. Mateo I, Llorca J, Infante J, Rodríguez-Rodríguez E, Berciano J, Combarros O. Gene-gene interaction between 14-3-3 zeta and butyrylcholinesterase modulates Alzheimer's disease risk. Eur J Neurol 2008; 15(3):219–22.
30. Bader JM, Geyer PE, Müller JB, Strauss MT, Koch M, Leypoldt F et al. Proteome profiling in cerebrospinal fluid reveals novel biomarkers of Alzheimer's disease. Molecular systems biology 2020; 16(6):e9356.
31. George C, Gontier G, Lacube P, François J-C, Holzenberger M, Aïd S. The Alzheimer's disease transcriptome mimics the neuroprotective signature of IGF-1 receptor-deficient neurons. Brain 2017; 140(7):2012–27.
32. Deng Y, Li B, Liu F, Iqbal K, Grundke-Iqbal I, Brandt R et al. Regulation between O-GlcNAcylation and phosphorylation of neurofilament-M and their dysregulation in Alzheimer disease. FASEB J 2008; 22(1):138–45.
33. Cuchillo-Ibañez I, Balmaceda V, Mata-Balaguer T, Lopez-Font I, Sáez-Valero J. Reelin in Alzheimer's Disease, Increased Levels but Impaired Signaling: When More is Less. J Alzheimers Dis 2016; 52(2):403–16.
34. Morgan AR, Touchard S, Leckey C, O'Hagan C, Nevado-Holgado AJ, Barkhof F et al. Inflammatory biomarkers in Alzheimer's disease plasma. Alzheimers Dement 2019; 15(6):776–87.
35. Stamate D, Kim M, Proitsi P, Westwood S, Baird A, Nevado-Holgado A et al. A metabolite-based machine learning approach to diagnose Alzheimer-type dementia in blood: Results from the European Medical Information Framework for Alzheimer disease biomarker discovery cohort. Alzheimers Dement (N Y) 2019; 5:933–8.
36. van Eyk JE, Snyder MP. Precision Medicine: Role of Proteomics in Changing Clinical Management and Care. J Proteome Res 2019; 18(1):1–6.
37. Gaudet P, Livstone MS, Lewis SE, Thomas PD. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. Brief Bioinformatics 2011; 12(5):449–62.
38. Fejtova A, Davydova D, Bischof F, Lazarevic V, Altrock WD, Romorini S et al. Dynein light chain regulates axonal trafficking and synaptic levels of Bassoon. J Cell Biol 2009; 185(2):341–55.
39. Meng X, McGraw CM, Wang W, Jing J, Yeh S-Y, Wang L et al. Neurexophilin4 is a selectively expressed α-neurexin ligand that modulates specific cerebellar synapses and motor functions. Elife 2019; 8.
40. Whiley L, Sen A, Heaton J, Proitsi P, García-Gómez D, Leung R et al. Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. Neurobiol Aging 2014; 35(2):271–8.

41. Orešič M, Hyötyläinen T, Herukka S-K, Sysi-Aho M, Mattila I, Seppänan-Laakso T et al. Metabolome in progression to Alzheimer's disease. Transl Psychiatry 2011; 1:e57.
42. Cortes-Canteli M, Zamolodchikov D, Ahn HJ, Strickland S, Norris EH. Fibrinogen and altered hemostasis in Alzheimer's disease. J Alzheimers Dis 2012; 32(3):599–608.
43. Catricala S, Torti M, Ricevuti G. Alzheimer disease and platelets: how's that relevant. Immun Ageing 2012; 9(1):20.
44. Wesenhagen KEJ, Teunissen CE, Visser PJ, Tijms BM. Cerebrospinal fluid proteomics and biological heterogeneity in Alzheimer's disease: A literature review. Crit Rev Clin Lab Sci 2019:1–13.
45. Huynh MB, Ouidja MO, Chantepie S, Carpentier G, Maïza A, Zhang G et al. Glycosaminoglycans from Alzheimer's disease hippocampus have altered capacities to bind and regulate growth factors activities and to bind tau. PLoS ONE 2019; 14(1):e0209573.
46. Ovsepian SV, O'Leary VB. Neuronal activity and amyloid plaque pathology: an update. J Alzheimers Dis 2016; 49(1):13–9.
47. Peña F, Gutiérrez-Lerma A, Quiroz-Baez R, Arias C. The role of beta-amyloid protein in synaptic function: implications for Alzheimer's disease therapy. Curr Neuropharmacol 2006; 4(2):149–63.
48. Liu P, Fleete MS, Jing Y, Collie ND, Curtis MA, Waldvogel HJ et al. Altered arginine metabolism in Alzheimer's disease brains. Neurobiology of aging 2014; 35(9):1992–2003.
49. Miners JS, Clarke P, Love S. Clusterin levels are increased in Alzheimer's disease and influence the regional distribution of Abeta. Brain pathology (Zurich, Switzerland) 2017; 27(3):305–13.
50. Bjorkqvist M, Ohlsson M, Minthon L, Hansson O. Evaluation of a previously suggested plasma biomarker panel to identify Alzheimer's disease. PloS one 2012; 7(1):e29868.
51. Juarez-Rebollar D, Rios C, Nava-Ruiz C, Mendez-Armenta M. Metallothionein in Brain Disorders. Oxidative medicine and cellular longevity 2017; 2017:5828056.
52. Hokama M, Oka S, Leon J, Ninomiya T, Honda H, Sasaki K et al. Altered expression of diabetes-related genes in Alzheimer's disease brains: the Hisayama study. Cerebral cortex (New York, N.Y. : 1991) 2014; 24(9):2476–88.
53. Muller T, Concannon CG, Ward MW, Walsh CM, Tirniceriu AL, Tribl F et al. Modulation of gene expression and cytoskeletal dynamics by the amyloid precursor protein intracellular domain (AICD). Molecular biology of the cell 2007; 18(1):201–10.
54. Kiddle SJ, Thambisetty M, Simmons A, Riddoch-Contreras J, Hye A, Westman E et al. Plasma based markers of [11C] PiB-PET brain amyloid burden. PloS one 2012; 7(9):e44260.
55. Kitamura Y, Usami R, Ichihara S, Kida H, Satoh M, Tomimoto H et al. Plasma protein profiling for potential biomarkers in the early diagnosis of Alzheimer's disease. Neurological research 2017; 39(3):231–8.
56. Baker SK, Chen ZL, Norris EH, Revenko AS, MacLeod AR, Strickland S. Blood-derived plasminogen drives brain inflammation and plaque deposition in a mouse model of Alzheimer's disease.

Proceedings of the National Academy of Sciences of the United States of America 2018; 115(41):E9687-e9696.

57. Song F, Poljak A, Kochan NA, Raftery M, Brodaty H, Smythe GA et al. Plasma protein profiling of Mild Cognitive Impairment and Alzheimer's disease using iTRAQ quantitative proteomics. Proteome science 2014; 12(1):5.
58. Muenchhoff J, Poljak A, Thalamuthu A, Gupta VB, Chatterjee P, Raftery M et al. Changes in the plasma proteome at asymptomatic and symptomatic stages of autosomal dominant Alzheimer's disease. Scientific reports 2016; 6:29078.
59. Noguchi M, Sato T, Nagai K, Utagawa I, Suzuki I, Arito M et al. Roles of serum fibrinogen alpha chain-derived peptides in Alzheimer's disease. International journal of geriatric psychiatry 2014; 29(8):808–18.
60. Foote M, Zhou Y. 14-3-3 proteins in neurological disorders. International journal of biochemistry and molecular biology 2012; 3(2):152–64.
61. Boom A, Pochet R, Authelet M, Pradier L, Borghgraef P, van Leuven F et al. Astrocytic calcium/zinc binding protein S100A6 over expression in Alzheimer's disease and in PS1/APP transgenic mice models. Biochimica et biophysica acta 2004; 1742(1-3):161–8.
62. Power JH, Asad S, Chataway TK, Chegini F, Manavis J, Temlett JA et al. Peroxiredoxin 6 in human brain: molecular forms, cellular distribution and association with Alzheimer's disease pathology. Acta neuropathologica 2008; 115(6):611–22.
63. Duits FH, Brinkmalm G, Teunissen CE, Brinkmalm A, Scheltens P, van der Flier, W. M. et al. Synaptic proteins in CSF as potential novel biomarkers for prognosis in prodromal Alzheimer's disease. Alzheimer's research & therapy 2018; 10(1):5.
64. Devanarayan P, Devanarayan V, Llano DA. Identification of a Simple and Novel Cut-Point Based Cerebrospinal Fluid and MRI Signature for Predicting Alzheimer's Disease Progression that Reinforces the 2018 NIA-AA Research Framework. J Alzheimers Dis 2019; 68(2):537–50.
65. Janelidze S, Mattsson N, Stomrud E, Lindberg O, Palmqvist S, Zetterberg H et al. CSF biomarkers of neuroinflammation and cerebrovascular dysfunction in early Alzheimer disease. Neurology 2018; 91(9):e867-e877.
66. Shang J, Yamashita T, Fukui Y, Song D, Li X, Zhai Y et al. Different Associations of Plasma Biomarkers in Alzheimer's Disease, Mild Cognitive Impairment, Vascular Dementia, and Ischemic Stroke. Journal of clinical neurology (Seoul, Korea) 2018; 14(1):29–34.
67. Zemva J, Schubert M. The role of neuronal insulin/insulin-like growth factor-1 signaling for the pathogenesis of Alzheimer's disease: possible therapeutic implications. CNS & neurological disorders drug targets 2014; 13(2):322–37.
68. Anvar NE, Saliminejad K, Ohadi M, Kamali K, Daneshmand P, Khorshid HR. Association between polymorphisms in Interleukin-16 gene and risk of late-onset Alzheimer's disease. Journal of the neurological sciences 2015; 358(1-2):324–7.

69. Farkas M, Keskitalo S, Smith DEC, Bain N, Semmler A, Ineichen B et al. Hyperhomocysteinemia in Alzheimer's disease: the hen and the egg? J Alzheimers Dis 2013; 33(4):1097–104.
70. Selley ML. A metabolic link between S-adenosylhomocysteine and polyunsaturated fatty acid metabolism in Alzheimer's disease. Neurobiol Aging 2007; 28(12):1834–9.
71. Czech C, Berndt P, Busch K, Schmitz O, Wiemer J, Most V et al. Metabolite profiling of Alzheimer's disease cerebrospinal fluid. PLoS ONE 2012; 7(2):e31501.
72. Frenkel-Pinter M, Shmueli MD, Raz C, Yanku M, Zilberzwige S, Gazit E et al. Interplay between protein glycosylation pathways in Alzheimer's disease. Sci Adv 2017; 3(9):e1601576.
73. Lin CH, Yang HT, Lane HY. D-glutamate, D-serine, and D-alanine differ in their roles in cognitive decline in patients with Alzheimer's disease or mild cognitive impairment. Pharmacology, biochemistry, and behavior 2019; 185:172760.
74. Yin JX, Maalouf M, Han P, Zhao M, Gao M, Dharshaun T et al. Ketones block amyloid entry and improve cognition in an Alzheimer's model. Neurobiology of aging 2016; 39:25–37.

Figures

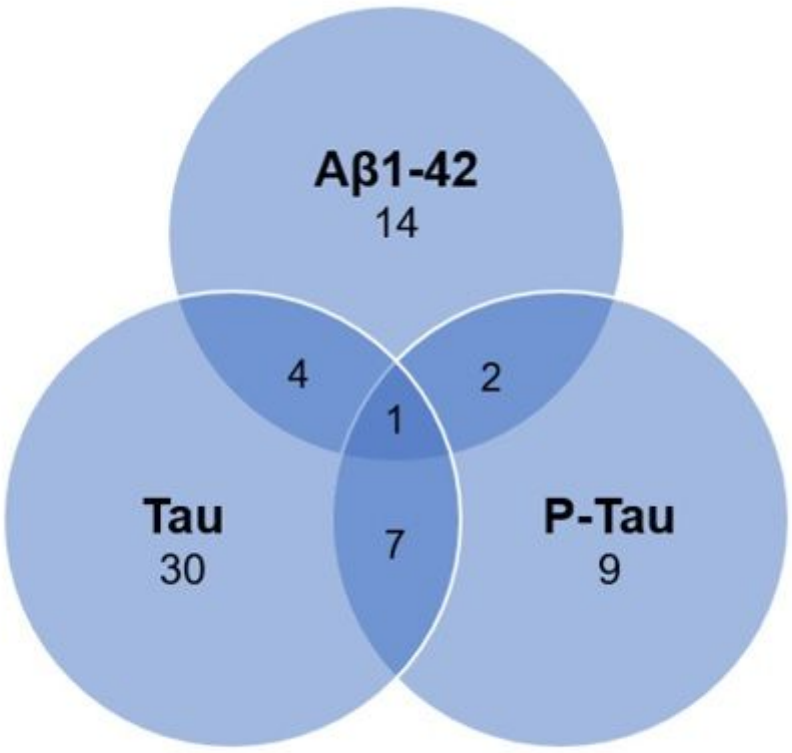

Figure 1

Venn diagram of associations with CSF core AD biomarkers. Venn diagram of associations of analytes obtained by regression models with CSF core AD biomarkers. Number of molecules identified as well as those shared between biomarkers is shown. The full list of associated molecules is presented in Additional File 3, Table S1.

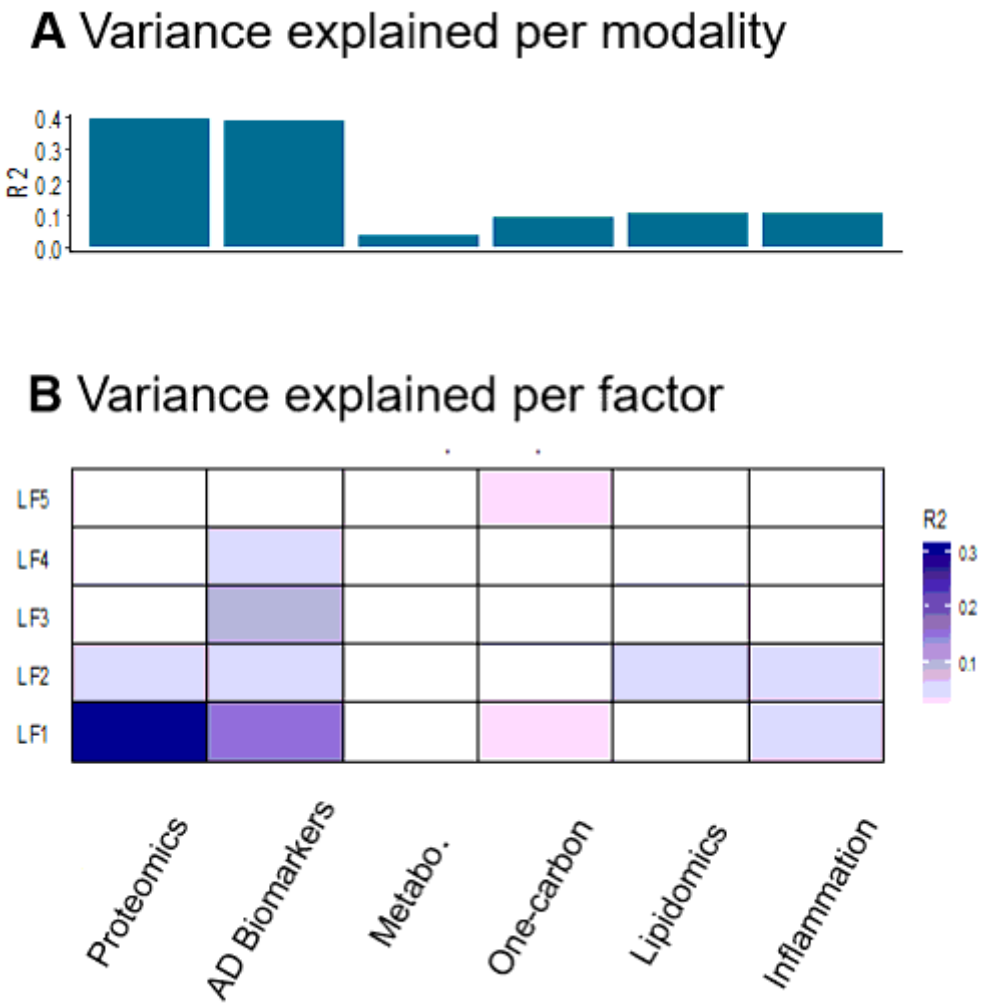

Figure 2

Overview of the MOFA model. Overview of the trained MOFA model showing variance (R2) within the cohort explained by each modality (top) and latent factors (LFs, bottom) from the trained MOFA model.

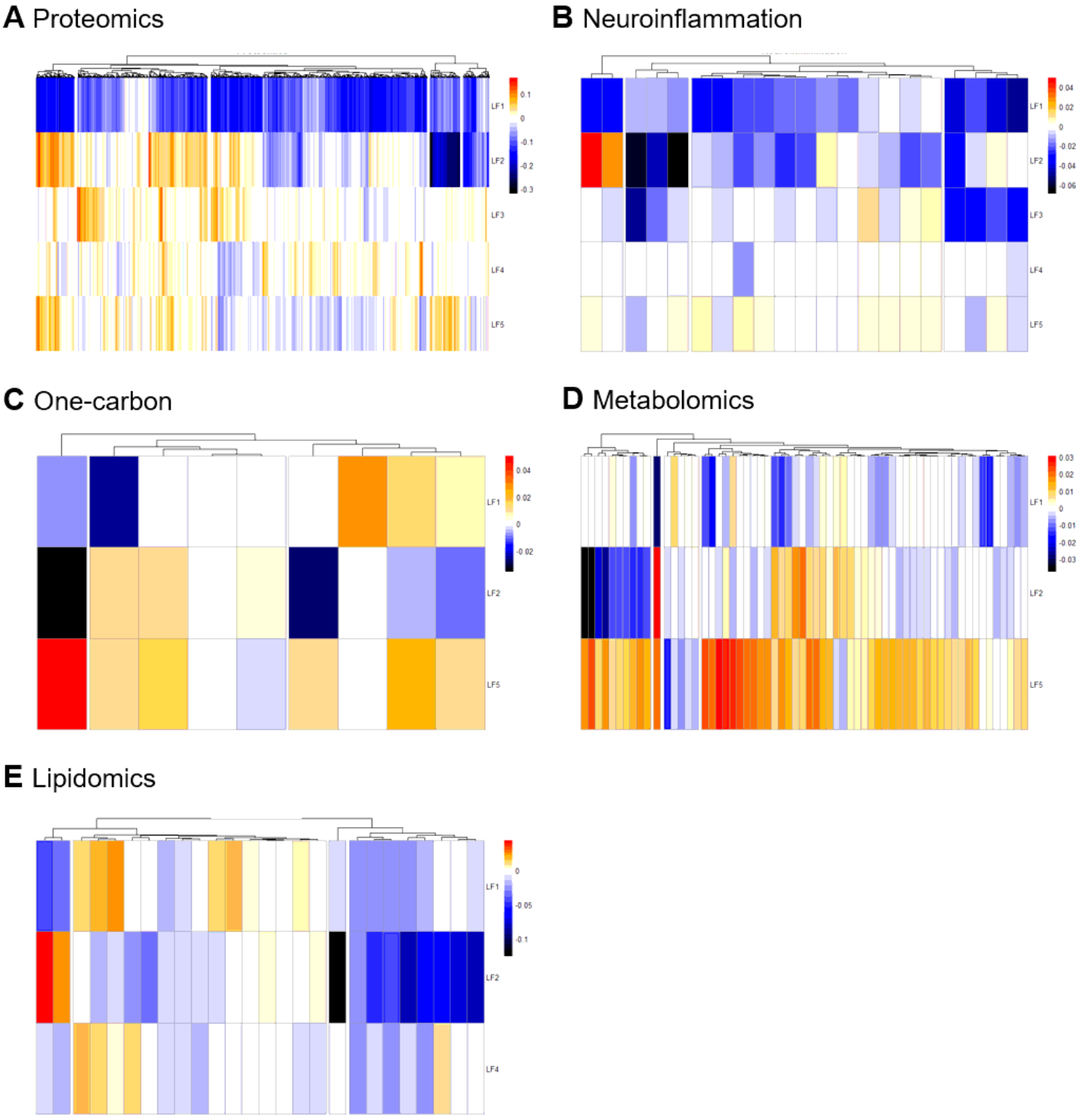

Figure 3

Clustering of loadings across latent factors. Heatmaps of hierarchical clustering of the measured loadings across in LFs for data obtained from proteomics (A), neuroinflammation markers (B), one-carbon metabolism (C), metabonomics (D) and lipidomics (E). Note the distinct pattern within each LF.

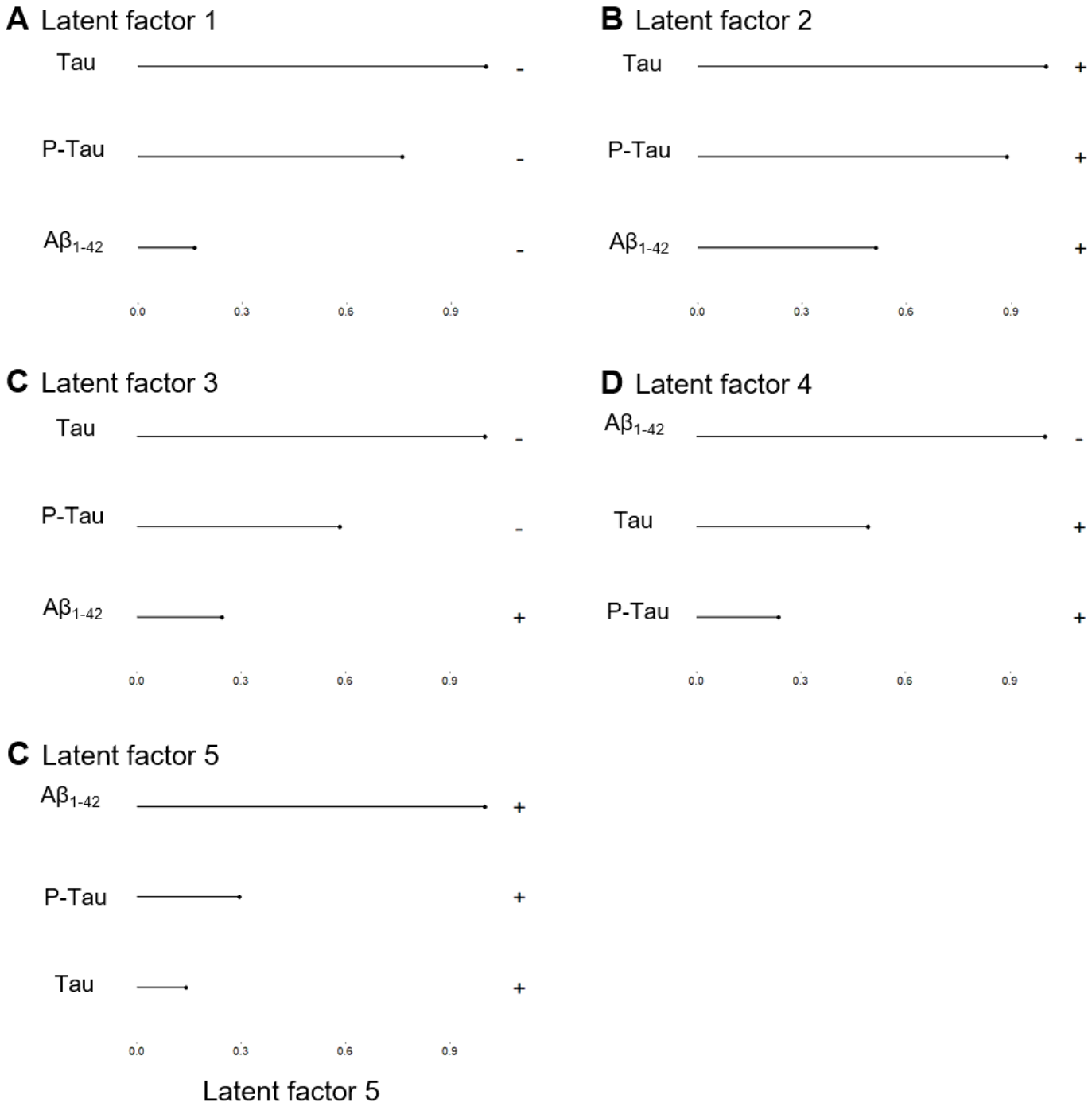

Figure 4

Loadings of CSF AD biomarkers. Normalised loadings of CSF AD biomarkers across the five latent factors of the trained MOFA model. Signs are relative within the model and not absolute.

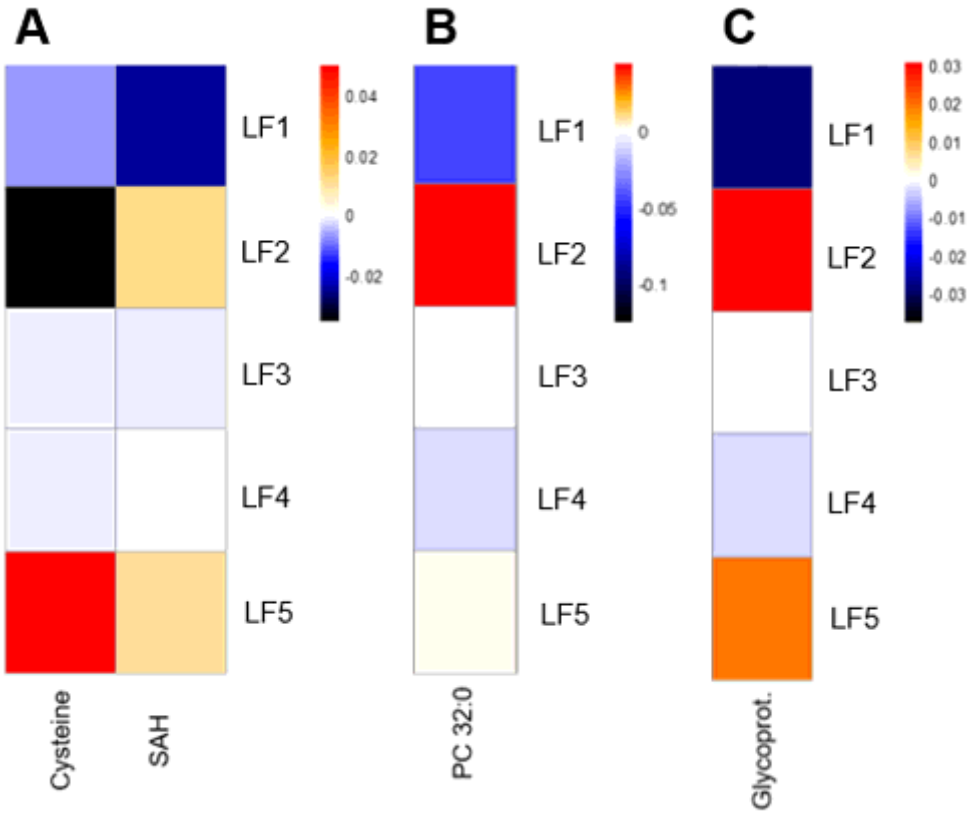

Figure 5

Relationship between selected CSF molecules and latent factors. Relationship between selected CSF molecules and latent factors, indicating associations with specific CSF AD biomarkers. Example analytes that show a specific pattern of association with LFs allowing to identify with which CSF AD biomarker they are associated. A) Measured loadings of total cysteine and S-adenosyl-L-homocysteine within each LF. B) Loadings of PC 32:0 within individual LFs. C) Measured loadings of glycoproteins within individual LFs. SAH, S-adenosyl-L-homocysteine.

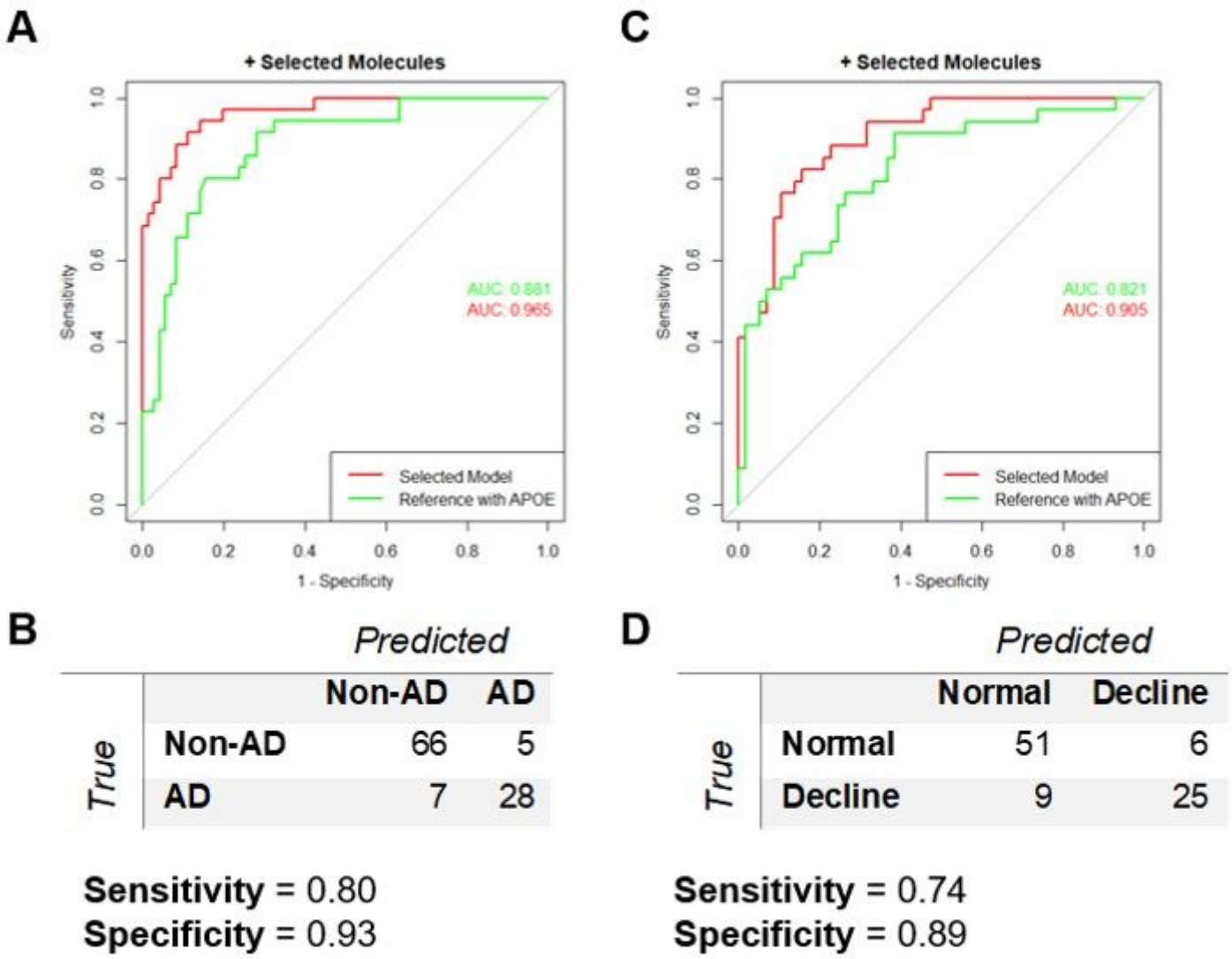

Figure 6

Clinical predictions. Binary logistic regression models to improve clinical predictions. A) ROC curves and AUCs for the reference model (green) and the final prediction model of AD pathology (red) obtained after addition of four analytes (14-3-3 zeta/delta, clusterin, interleukin-15, and transgelin-2) selected by the MOFA model. B) Confusion matrix of the final prediction model of AD. C) ROC curves and AUCs for the reference model (green) and the final prediction model of cognitive decline (red) obtained after addition of four analytes (14-3-3 zeta/delta, clusterin, cholesteryl ester 27:1 16:0 and monocyte chemoattractant protein-1) selected by the MOFA model. D) Confusion matrix of the final prediction model of cognitive decline.

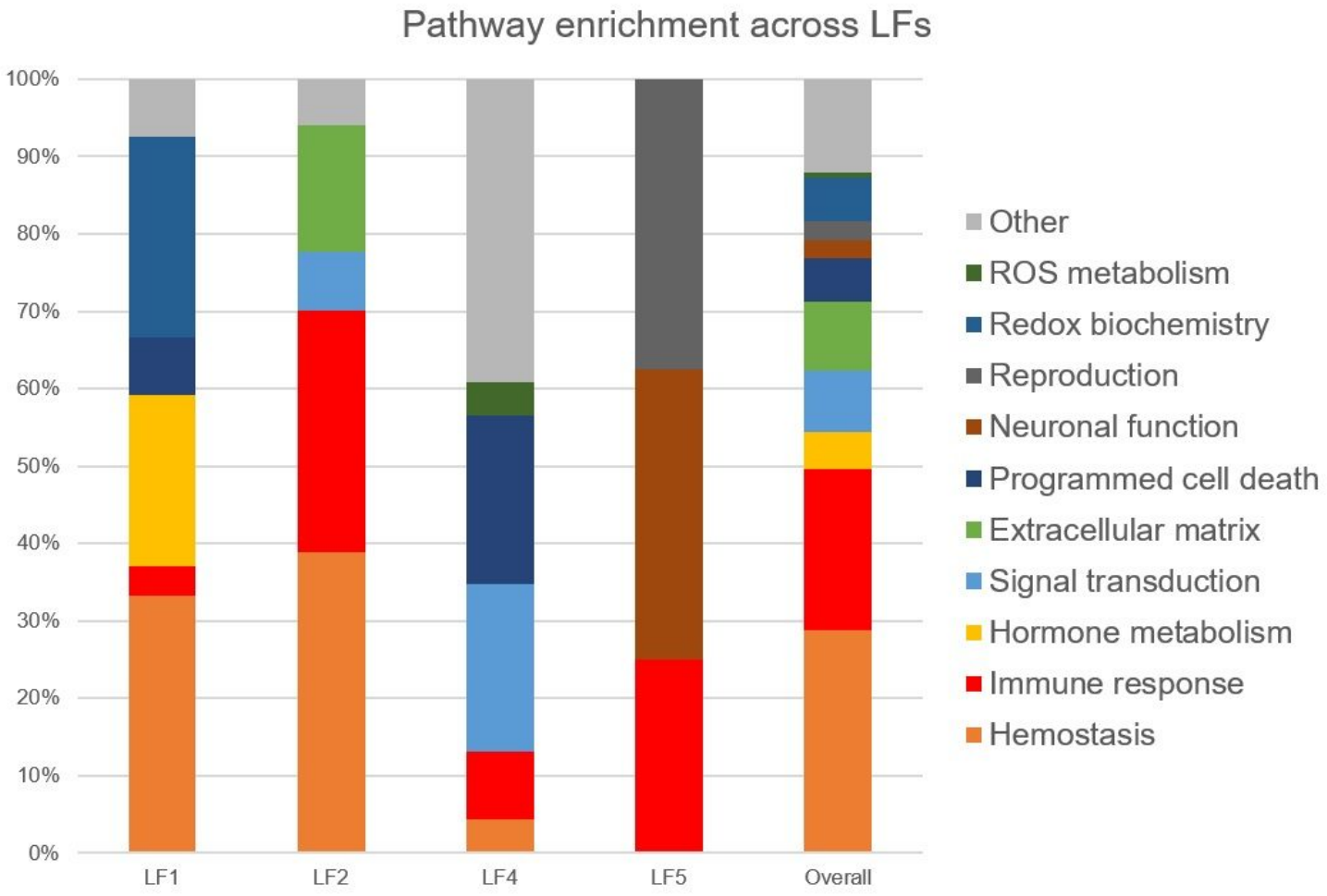

Figure 7

Pathway enrichment. Pathway enrichment analysis of identified proteins across LFs and overall. Top analytes from each LF and identified by the Reactome database were assigned to coarse-grain categories (Fig. S3). The number of entities found within each pathway for each LF and overall (expressed as a percentage) is represented. NB: the low number of analytes associated with LF3 did not allow for an enrichment analysis.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile1Clark.docx
- AdditionalFile2Clark.docx
- AdditionalFile3Clark.docx