Dickkopf-1 overexpression *in vitro* nominates candidate blood biomarkers relating to Alzheimer's pathology

Liu Shi[a]*, Laura M. Winchester[a], Benjamine Y. Liu[a], Richard Killick[b], Elena M. Ribe[a], Sarah Westwood[a], Alison L. Baird[a], Noel J Buckley[a], Shengjun Hong[c], Valerija Dobricic[c], Fabian Kilpert[c], Andre Franke[d], Steven Kiddle[b,e], Martina Sattlecker[b,e], Richard Dobson[jjj,kkk], Antonio Cuadrado[g,h], Abdul Hye[b], Nicholas J. Ashton[b,f,i,j], Angharad R. Morgan[k], Isabelle Bos[l,m], Stephanie J. B. Vos[l], Mara ten Kate[m], Philip Scheltens[m], Rik Vandenberghe[n], Silvy Gabel[n,o] , Karen Meersmans[n,o], Sebastiaan Engelborghs[p,q,r,t], Ellen E. De Roeck[p,q,r], Kristel Sleegers[r,s], Giovanni B. Frisoni[u,v], Olivier Blin[w], Jill C. Richardson[x], Régis Bordet[y], José L. Molinuevo[z,aa], Lorena Rami[aa], Anders Wallin[i,lll], Petronella Kettunen[i,bb], Magda Tsolaki[cc], Frans Verhey[l], Alberto Lleó[dd], Daniel Alcolea[dd], Julius Popp[ee,ff], Gwendoline Peyratout[ee], Pablo Martinez-Lage[gg], Mikel Tainta[gg,hh], Peter Johannsen[ii], Charlotte E. Teunissen[jj], Yvonne Freund-Levi[kk,ll,mm,nn], Lutz Frölich[oo], Cristina Legido-Quigley[pp,qq], Frederik Barkhof[rr,ss], Kaj Blennow[i,tt], Katrine Laura Rasmussen[uu,vv,ww], Børge Grønne Nordestgaard[vv,ww,xx,hhh], Ruth Frikke-Schmidt[uu,vv,ww], Sune Fallgaard Nielsen[vv,ww,xx], Hilkka Soininen[yy], Bruno Vellas[zz], Iwona Kloszewska[aaa], Patrizia Mecocci[bbb], Henrik Zetterberg[i,tt,ccc,ddd], B. Paul Morgan[k], Johannes Streffer[q,eee], Pieter Jelle Visser[l,m,fff], Lars Bertram[c,ggg], Alejo J. Nevado-Holgado[a], Simon Lovestone[a,iii]

[a] Department of Psychiatry, University of Oxford, UK

[b] King's College London, Institute of Psychiatry, Psychology and Neuroscience, Maurice Wohl Institute Clinical Neuroscience Institute, London, UK

[c] Lübeck Interdisciplinary Platform for Genome Analytics, Institutes of Neurogenetics and Cardiogenetics, University of Lübeck, Lübeck, Germany

[d] Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany

[e] MRC Social, Genetic and Developmental Psychiatry Centre, King's College London, UK

[f] NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK

[g] Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Instituto de Investigación Sanitaria La Paz (IdiPaz), Instituto de Investigaciones Biomédicas Alberto Sols UAM-CSIC, and Department of Biochemistry, Faculty of Medicine, Autonomous University of Madrid, Madrid, Spain

[h] "Victor Babes" National Institute of Pathology, Bucharest, Romania

[i] Department of Psychiatry and Neurochemistry, Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden

[j] Wallenberg Centre for Molecular and Translational Medicine, University of Gothenburg, Gothenburg, Sweden

[k] Dementia Research Institute Cardiff, Cardiff University, Cardiff, UK

[l] Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Alzheimer Centrum Limburg, Maastricht University, Maastricht, the Netherlands

[m] Alzheimer Center, VU University Medical Center, Amsterdam, the Netherlands

[n] University Hospital Leuven, Leuven, Belgium

[o] Laboratory for Cognitive Neurology, Department of Neurosciences, KU Leuven, Belgium

[p] Center for Neurosciences, Vrije Universiteit Brussel (VUB), Brussels, Belgium

[q] Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium

[r] Institute Born-Bunge, University of Antwerp, Antwerp, Belgium

[s] Neurodegenerative Brain Diseases Group, Center for Molecular Neurology, VIB, Belgium

[t] Department of Neurology, UZ Brussel, Brussels, Belgium

[u] University of Geneva, Geneva, Switzerland

[v] IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy

[w] AIX Marseille University, INS, Ap-hm, Marseille, France

[x] Neurosciences Therapeutic Area, GlaxoSmithKline R&D, Stevenage, UK

[y] University of Lille, Inserm, CHU Lille, France

[z] Alzheimer's disease & other cognitive disorders unit, Hospital Clínic, Barcelona, Spain

[aa] BarcelonaBeta Brain Research Center, Universitat Pompeu Fabra, Barcelona, Spain

[bb] Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

[cc] 1st Department of Neurology, AHEPA University Hospital, school of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Makedonia, Greece

[dd] Department of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

[ee] Department of Psychiatry, University Hospital of Lausanne, Lausanne, Switzerland

[ff] Geriatric Psychiatry, Department of Psychiatry, Geneva University Hospitals, and University of Geneva, Geneva, Switzerland

[gg] CITA-Alzheimer Foundation, San Sebastian, Spain

[hh] Organización Sanitaria Integrada Goierri – Alto Urola, Osakidetza, Spain

[ii] Danish Dementia Research Centre, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

[jj] Neurochemistry Laboratory, dept of Clinical Chemistry, Amsterdam Neuroscience, Amsterdam University Medical Centers, Vrije Universiteit, the Netherlands

[kk] School of Medical Sciences, Örebro University, Örebro, Sweden

[ll] Department of Neurobiology, Care Sciences and Society, Division of Clinical Geriatrics, Karolinska Institute, Stockholm, Sweden

[mm] Department of Old Age Psychiatry, Psychology and Neuroscience, King's College London, UK

[nn] Department of Psychiatry, Örebro Universitetssjukhus, Örebro, Sweden

[oo] Department of Geriatric Psychiatry, Zentralinstitut für Seelische Gesundheit, University of Heidelberg, Mannheim, Germany

[pp] Kings College London, London, UK

[qq] The Systems Medicine Group, Steno Diabetes Center Copenhagen, Gentofte, Denmark

[rr] Department of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam, The Netherland

[ss] UCL Institutes of Neurology and Healthcare Engineering, London, UK

[tt] Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden

[uu] Department of Clinical Biochemistry, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

[vv] The Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark

[ww] Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[xx] Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark

[yy] Neurology / Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland

[zz] Toulouse Gerontopole University Hospital, Univeriste Paul Sabatier, INSERM U 558, France

[aaa] Medical University of Lodz, Lodz, Poland

[bbb] Section of Gerontology and Geriatrics, Department of Medicine, University of Perugia, Perugia, Italy

[ccc] UK Dementia Research Institute at UCL, London, United Kingdom

[ddd] Department of Neurodegenerative Disease, UCL Institute of Neurology, London, United Kingdom

[eee] UCB, Braine-l'Alleud, Belgium, formerly Janssen R&D, LLC. Beerse, Belgium at the time of study conduct

[fff] Department of Neurobiology, Care Sciences and Society, Division of Neurogeriatrics, Karolinska Institutet, Stockholm Sweden

[ggg] Department of Psychology, University of Oslo, Oslo, Norway

[hhh] The Copenhagen City Heart Study, Frederiksberg Hospital, Copenhagen University Hospital, Frederiksberg, Denmark

[iii] Currently at Janssen-Cilag UK, formerly at Department of Psychiatry, University of Oxford, UK at the time of the study conduct

[jjj] Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

[kkk] Institute of Health Informatics, University College London, London, UK

[lll] Memory Clinic at Department of Neuropsychiatry, Sahlgrenska University Hospital, Mölndal, Sweden

*** Correspondence:** Liu Shi, liu.shi@psych.ox.ac.uk

Abstract

Background: Previous studies suggest that Dickkopf-1 (DKK1), an inhibitor of Wnt signalling, plays a role in amyloid-induced toxicity and hence Alzheimer's disease. However, the effect of DKK1 expression on protein expression, and whether such proteins are altered in disease, is unknown.

Methods: We first overexpressed DKK1 in HEK293A cells and quantified 1128 proteins in cell lysates using aptamer capture arrays (SomaScan) to obtain a protein signature induced by DKK1. We then used the same assay to measure the DKK1-signature proteins in human plasma in two large cohorts, EMIF (n = 785) and ANM (n = 677), to explore the association of DKK1-signature proteins with markers of AD pathology as used in the amyloid/tau/neurodegeneration (ATN) framework as well as with clinical outcomes.

Results: We identified a 100-protein signature induced by DKK1 *in vitro*. In plasma, proteins from this signature set were altered primarily in individuals with markers indicative of amyloid pathology. Subsets of proteins, along with age and *apolipoprotein E* ε4 genotype distinguished amyloid pathology (A+T-N-, A+T+N-, A+T-N+ and A+T+N+) from no AD pathology (A-T-N-) with an area under the curve (AUC) of 0.72, 0.81, 0.88 and 0.85 respectively. In terms of AD clinical outcomes, we found that some signature proteins (e.g. Complement C3 and albumin) were associated with cognitive score ($p < 0.001$ for both proteins) and AD diagnosis ($p < 0.001$ for both proteins) in both cohorts.

Conclusions: Our results add further evidence for a role of DKK regulation of Wnt signalling in AD and suggest that DKK1 induced signature proteins obtained *in vitro* could reflect the ATN framework as well as predict disease severity and progression *in vivo*.

Keywords: ATN framework; Dickkopf-1; Wnt signalling; SOMAScan; replication

Introduction

Alzheimer's disease (AD) is characterised by the presence of β-amyloid (Aβ) containing plaques and neurofibrillary tangles composed of modified tau protein together with the progressive loss of synapses and eventually neurons [1]. Recently, the National Institute on Aging-Alzheimer's Association (NIA-AA) proposed a classification system for disease - the ATN framework [2] - based on three biomarker types where "A" represents amyloid pathology, measured by cortical amyloid positron emission tomography (PET) ligand binding or low cerebrospinal fluid (CSF) Aβ42; "T" represents tau pathology, measured by elevated CSF phosphorylated tau (P-tau) or cortical tau PET ligand binding; and "N" represents neurodegeneration or neuronal injury, measured by elevated CSF total tau (T-tau), 18F-fluoro-deoxyglucose (FDG) PET, or brain atrophy on magnetic resonance imaging (MRI). Dichotomizing these biomarkers as normal or abnormal results in eight ATN profiles; absence of AD related pathology (A-T-N-); the Alzheimer's continuum including indications of amyloid pathology (A+T-N-, A+T+N-, A+T-N+ and A+T+N+); and Suspected Non-Alzheimer Pathology (SNAP), or non-amyloid dementia (A-T-N+, A-T+N- and A-T+N+) [2].

The ATN framework has considerable face validity and has rapidly found wide acceptance in the research field. As clinical trials are increasingly targeting a range of pathologies, the ATN framework helps to inform participant inclusion and potentially also trial outcomes [3]. Moreover, the framework has predictive validity with, for example, people with A+T+N+ showing faster decline than other categories [4, 5]. The ATN framework is limited by biomarkers that are either not yet fully

qualified or are relatively invasive and where access can be difficult. A blood-based version of the ATN framework would be of considerable value and recent progress suggests such an objective is realisable.

Whilst considerable progress has been made in understanding the formation, and to some extent the effects of, the three pathological processes that define the ATN classification, much less progress has been made in determining the mechanistic relationship between amyloid and tau pathologies and their effects in inducing neuronal dysfunction and death. One potential mechanism that has been proposed to link all three processes is Wnt signalling. Activation of Wnt signalling is neuroprotective against the toxicity of Aβ both *in vitro* and *in vivo* [6, 7] and reverses cognitive deficits in rodent models [8, 9]. Inhibition of Wnt signalling could therefore be a factor triggering the onset and progression of AD (reviewed in [10]). In line with this, Dickkopf-1 (DKK1), an inhibitor of Wnt signalling, has been reported to be elevated in human AD brain, as well as in mouse models with Aβ generation and plaque formation [11-13]. In model systems, DKK1 is induced by Aβ, which in turn drives synaptic loss, tau phosphorylation and neuronal death [14-16] and blockade of DKK1 protects synapses from Aβ-mediated neurotoxicity [16, 17].

Collectively, DKK1 mediated inhibition of Wnt signalling might be a critical factor in the Aβ-mediated pathway driving tau pathology and hence neuronal dysfunction and loss. Previously we demonstrated that knockdown of genes on a shared Aβ/DKK1 pathway also protected neurons from Aβ-induced toxicity and that in mice, overexpression of DKK1 induced tau pathology and cognitive impairment [18]. Given this, we propose that the DKK1 induced pathway might reflect the ATN framework in man. However, the molecular signature we previously identified was based on neuronal gene expression and included many master regulators or transcription factors and hence was not readily translatable to human studies. Therefore, in order to explore whether a DKK1 induced signature was recognisable in peripheral fluids in human *in vivo* studies, we first determined a protein signature induced by DKK1 in human cells *in vitr*o and then tested whether this empirically

defined protein signature was associated with the ATN framework and clinical outcomes in human plasma from two large independent cohorts including people clinically defined with AD, with mild cognitive impairment (MCI) and apparently healthy controls (Figure 1).

Materials and Methods

HEK293A cells with overexpression of DKK1

In order to establish a DKK1-induced protein signature, a HIS-tagged *DKK1* cDNA was synthesized (GENEWIZ, UK) and cloned into the mammalian expression construct pcDNA3.1+ (Invitrogen, UK) and validated by sequencing. HEK293A cells (an adherent strain of HEK293) were cultured in DMEM + 10% FCS in 12-well plates until 80% confluent and transfected with the *DKK1* construct or the empty vector control using FuGene 6 according to the manufacturer instructions (Promega, UK). The next day, the FCS content of the media was adjusted to 2% and the cells maintained for a further 6 h. Media was then removed and total cell lysates collected in M-PER (ThermoFisher, UK) for proteomic array studies using the SOMAScan platform (SomaLogic, Boulder, CO), allowing for the simultaneous measurement and quantification of 1128 proteins (n = 5 per condition). All protein data were log-transformed prior to analysis.

Study participants

We used plasma samples recruited from two previously reported cross European studies: AddNeuroMed (ANM) [19] and the European Medical Information Framework for Alzheimer's disease multi-modal biomarker discovery (EMIF-AD MBD) study [20].

ANM sample collection was performed at six different centres across Europe: University of Kuopio, Finland; Aristotle University of Thessaloniki, Greece; King's College London, United Kingdom; University of Lodz, Poland; University of Perugia, Italy; and University of Toulouse, France [19]. We used 677 subjects from the ANM cohort including 319 AD patients, 149 mild cognitive impairment (MCI) individuals and 209 elderly unaffected controls (CTL). General clinical and demographic

information were available for all subjects (including *APOE* ε4 genotype data) (Supplementary Table 1). Furthermore, the majority participants (84%) had Mini Mental State Examination (MMSE) measurement and around 60% of the AD patients had Alzheimer's Disease Assessment Scale - Cognitive subscale (ADAS-Cog) measurement (Supplementary Table 1) [21].

The EMIF-AD MBD is part of the European Medical Information Framework for Alzheimer's disease (EMIF-AD; http://www.emif.eu/), a European wide collaboration to facilitate the re-use of existing healthcare data and the sharing of cohort samples for the benefit of AD research. Overall, the EMIF-AD MBD study collected samples from 11 European cohorts DESCRIPA, EDAR, PharmaCog, Amsterdam, Antwerp, San Sebastian GAP, Gothenburg, Barcelona IDIBAPS, Lausanne, Leuven and Barcelona St Pau [20]. We used 785 subjects from the EMIF study comprising 183 AD patients, 382 MCI and 220 CTL. In addition to general clinical and demographic information, each subject had a measure of both Aβ and tau (including total tau [T-tau] and phosphorylated tau [P-tau]) pathology. The classification of the status (abnormal/normal) of amyloid, T-tau and P-tau has been described previously [20]. Briefly, amyloid pathology was determined using CSF Aβ42 in the majority and PET amyloid in a minority, based on which the individuals were classified into abnormal and normal status [20]. The levels of T-tau and P-tau in CSF were analysed locally and the local cut-off point was used to determine their status (abnormal/normal) [20]. For these subjects, "A" is defined by amyloid status, "T" is by P-tau status and "N" is by T-tau status. Dichotomizing these biomarkers as normal or abnormal results in eight ATN profiles (Table 1). In addition, each subject had MMSE measurement and the majority (over 72%) had other neuropsychological tests for memory, language and executive functioning as previously reported (Supplementary Table 1) [20]. Furthermore, each individual had genome-wide single nucleotide polymorphism (SNP) genotyping. The details of SNP assays and raw data processing were described in [22].

Protein quantification of human plasma in two cohorts

The SOMAScan assay, an aptamer-based assay, (SomaLogic, Boulder, CO) [23], was used to measure plasma proteins in subjects collected from both ANM and EMIF cohorts. Because of an evolving platform, different versions of the SOMAScan assay were used in samples from the two cohorts, with 1016 and 4001 proteins measured in ANM and EMIF cohorts, respectively. The *in vitro* experiments were conducted with a version of the assay having 1128 proteins. The three versions of the assay used here were generated such as to ensure data interoperability. The abundance of each protein was log-transformed for all subsequent analyses.

Statistical analysis

All statistical analyses were completed using R (version 3.3.2). We used Student's t-test to assess the relative levels of DKK family proteins (DKK1, DKK3, DKK4 and DKK-Like 1 [DKKL1]) as measured on the SomaLogic platform, in DKK1 overexpressing and control cells. Partial least squares (PLS) regression was used to obtain a signature consisting of the 100 top, or most altered, proteins as a consequence of DKK1 expression in HEK293A cell lysates. A PLS regression model was fitted to the data using all 1128 proteins as the predictors (X) and the DKK1 or control status as the response variable (Y). We ranked proteins based on the calculated coefficients using two components from the resulting PLS regression model. The coefficients corresponding to each protein in the model are a proxy for how much each protein contributes to the signal.

The top 100 proteins that contributed to this multivariate signature most constitute the 'DKK1-induced signature' for subsequent analysis. The biological significance of the DKK1-induced signature was evaluated using the DAVID Bioinformatics Resource, version 6.8 Functional Annotation tool. Briefly, the 100 proteins were selected as our 'gene list' while all 1128 proteins quantified in the study were selected as our 'background gene list'. The enrichment analysis was performed on the KEGG database.

To compare the association of proteins with the ATN framework, we used logistic regression to compare proteins in different ATN profiles to 'no pathology controls' (A-T-N-), adjusting for age,

APOE ε4 genotype and gender. Logistic regression was also used to compare proteins in different AD diagnostic groups as well as between MCI participants who subsequently converted to dementia (MCIc) within 3 years relative to those whose MCI remained stable (MCIs). To analyse the association of proteins with continuous AD phenotypes (*i.e.* MMSE), we used partial correlation and adjusted for age, *APOE* ε4 genotype and gender. *P* values obtained from both logistic regression and partial correlation were corrected using false discovery rate (FDR) and corrected *p* values were presented in heat map. Furthermore, for visual presentation, we presented the absolute protein expression value in different ATN and AD diagnostic groups in box plots. Student's t-test was used to assess pairwise difference and uncorrected *p* values were presented in the box plots.

Forward stepwise logistic regression was used to find the analyte set that optimally discriminated amyloid pathology (A+T-N-, A+T+N-, A+T-N+ and A+T+N+) from no AD pathology (A-T-N-) in all subjects as well as in only cognitively normal individuals. In both cases, demographic covariates age and *APOE* ε4 genotype were included in models as potential predictors. For each comparison, the data set was randomly split into training (90%) and validation (10%) sets. The training set was used to select variables and fit the model which was then tested on the validation set using receiver operating curve (ROC) analysis. The 95% confidence Intervals of AUC was calculated using the ci.auc function.

Mendelian randomization

Mendelian randomization (MR) was used to investigate the causal relationship between two most promising (see results section) proteins (C3 and FGG) and AD risk, Aβ and tau (T-tau and P-tau) pathology. As genetic variants are passed from parents to child at conception and remain largely unaltered by environment throughout an individual's lifetime, reverse causation and confounding can be limited, making MR a powerful tool to examine causality between the exposure and outcome [24, 25]. The Mendelian randomization approach was based on three assumptions: 1) The genetic variants used as instrumental variables are associated with exposures. 2) The genetic variants are

not associated with any confounders of the exposure-outcome relation. 3) The genetic variants are associated with outcome only through the exposure, namely a lack of pleiotropy [24].

For C3, we selected three SNPs as instrumental variables; rs1065489 [*CFH*], rs429608 [*SKIV2L*] and rs448260 [*C3*]. The association of these SNPs with plasma C3 levels have been validated in 95,442 individuals ($p < 10^{-67}$) [26]. For FGG, we selected 24 SNPs as instrumental variables. These have been shown to be significantly associated with plasma FGG levels in a large Genome wide association studies (GWAS) study including more than 100,000 subjects [27] (Assumption 1). Then we checked whether the SNPs were either in linkage disequilibrium with one another, or were associated with known risk factors for AD (*e.g. APOE* ε4 genotype) (assumption 2). After verifying no direct association with AD, we acquired the summary statistics of each SNP with C3 and FGG from both studies separately as exposure estimates [26, 27]. For AD risk as outcome, we acquired the summary statistics for the association of each SNP with AD risk from a previously published GWAS study; International Genomics of Alzheimer's Project (IGAP) by Lambert *et al.* [28]. One C3 SNP (rs429608 [*SKIV2L*]) was not found in IGAP; we therefore acquired summary statistics of rs429608 in another GWAS study - UK Biobank (UKBB) [29]. For Aβ and tau (T-tau and P-tau) as outcome, we analysed the association of each SNP with Aβ and tau status in EMIF cohort using PLINK (v1.7). Using a two-sample MR approach, the exposure SNP (SNP-C3 and SNP-FGG) and outcome SNP (SNP - AD risk, SNP - Aβ, SNP - T-tau and SNP - P-tau) associations were used to compute estimates of each exposure-outcome association. We then used two MR methods to test the robust causal inference including an inverse-variance weighted (IVW) method [30] the weighted median method [31].

To test the third assumption, we used the MR-Egger method to calculate values of intercepts and their *p* values. If the intercepts do not deviate markedly from zero, it indicates that substantial horizontal pleiotropy of the SNPs is less likely [32]. Furthermore, to determine if there was any single SNP driving the relationship, we performed a leave-one-out analysis where the MR is performed removing a different SNP in each iteration [33].

Results

DKK1 and DKK4 were differentially expressed in DKK1-overexpressing cells compared to controls

In addition to a large number of other proteins representing a wide range of biological processes, both versions of the SOMAScan assay used here include aptamers selected to bind and hence measure relative abundance of DKK proteins including DKK1, DKK3 and DKK4 and the related protein DKK-Like 1 (DKKL1). In order to determine the protein signature induced by DKK1, we used the SOMAScan assay to compare lysates from HEK293A cells engineered to over-express human DKK1 with control cells transfected with empty vector (n=5 in each case). We first determined the ability of the SOMAScan assay to detect DKK1 and to differentiate this protein from other structurally similar DKK isoforms. We found that among the DKK family proteins, both DKK1 and DKK4 registered an increase in DKK1-overexpressing cells compared both to the other DKK isoforms and to all other protein measures (for DKK1 $p = 0.008$, corrected $p = 0.7$; for DKK4 $p = 0.008$, corrected $p = 0.7$) (Figure 2). The most likely explanation for this observation is that the DKK4 read-out on the SomaLogic panel is in fact at least in part a read-out of DKK1 due to cross reactivity and hence we refer subsequently to this as DKK1/4.

DKK1-induced proteomic signature was enriched in AD pathways

Having shown that the SOMAScan assay identifies the overexpression of DKK1 (with concomitant signal in the aptamer raised against DKK4), we used PLS to identify a multivariate proteomic signature that distinguished DKK1 overexpression cell lysates from controls. We ranked the proteins based on PLS coefficients (absolute value), and then selected the 100 proteins with the largest contribution to the DKK1-induced multivariate signature (Supplementary Table 2). As an exploratory study only we then assessed the biological significance of this signature using the DAVID Bioinformatics Resource. Overall, eight pathways were enriched though not passing multiple correction (Supplementary Table 3). Interestingly, AD was nominated as the second highest-ranked pathway ($p = 0.014$, corrected $p = 0.075$) in relation to disease.

Association of DKK1-induced signature with the ATN framework *in vivo*

We then investigated whether the top 100 proteins induced by DKK1, referred to here as the DKK1-induced signature, was associated with the ATN framework in the EMIF cohort (n = 785). Table 1 shows the characteristics of the participants split by ATN framework. No significant difference in sex was found among 8 profiles. Participants within the AD and SNAP groups were older than those in the 'No Pathology Control' (NPC) group except those with an A-T+N- profile. The prevalence of *APOE* ε4 carriers was higher in AD group than those in NPC or SNAP. Furthermore, MMSE was lower in AD group than individuals in NPC or SNAP except those in A+T-N- profile.

We used logistic regression to compare proteins in different ATN framework groups to the NPC group (A-T-N-), adjusting for age, *APOE* ε4 genotype and gender. Of the 100 proteins, the levels of 23 proteins significantly altered in participants with at least one abnormal ATN biomarker, *i.e.* either A+, T+ or N+ after FDR correction (Figure 3A). Furthermore, the majority of proteins were associated with amyloid-related pathology rather than with non-amyloid pathology (SNAP). In subsequent analyses we therefore focused on comparisons within the AD (A+) group to the NPC (A-T-N-) group, omitting the SNAP group (A- but T+ or N+). Based on their expression, these proteins could be divided into 3 groups: (1) those influenced only by amyloid pathology (A+) and independent of P-tau or T-tau status. For example, FGG increased in all individuals with abnormal amyloid pathology although it did not achieve significance in A+T+N- subjects (Figure 3B) perhaps due to a low number of only 19 subjects with this profile. Other proteins belonging to this group were BRF-1, Coagulation Factor VII, CKAP2, HMG-1, CAMK2D, AURKB, BFL1, C3 and albumin; (2) those influenced by both amyloid and T-tau (A+N+). For example, DKK1/4 increased in individuals with A+T-N+ and A+T+N+ profiles (Figure 3C). Three other proteins also belonged to this group: eotaxin, coactosin-like protein and annexin I; (3) those influenced by amyloid, P-tau and T-tau (A+T+N+), resulting changed levels in only A+T+N+ individuals, *e.g.*, DKK1 (Figure 3D), CHST2, MK01, CONA1, DLRB1, FN1.4, Cytochrome c and SHP-2.

We then used forward stepwise logistic regression to identify optimal analyte sets to distinguish the NPC group from different ATN profiles within the AD (A+) group. We used AD related proteins as well as age and *APOE* ε4 genotype as input features. Results showed that a model containing HMG-1 as well as age and *APOE* ε4 genotype best discriminate A+T-N- from NPC group with an area under the curve (AUC) of 0.72 (Figure 3E). The optimal model to differentiate A+T+N- from NPC group contained 6 features including SHP-2, FN1.4, CKAP2 and CHST2 as well as age and *APOE* ε4 genotype. For differentiating A+T-N+ from NPC group, a model containing 7 proteins (annexin I, Albumin, Cytochrome c, Eotaxin, DKK1, CONA1 and AURKB) with age and *APOE* ε4 genotype was selected. The optimal model to differentiate A+T+N+ from NPC group included 7 features including C3, AURKB, Eotaxin, BRF-1, albumin as well as age and *APOE* ε4 genotype (Figure 3E). The comparison between each optimal model and the combination of age and *APOE* ε4 genotype were shown in Supplementary Table 4.

We also set out to determine the optimal analyte sets to differentiate the NPC group from the amyloid positive ATN profiles (A+) groups in only cognitively normal individuals (Figure 3F). The optimal model that separated the A+T+N+ from the NPC group included 4 features including CKAP2, C3 as well as age and *APOE* ε4 genotype. A model consisting of four proteins (Coagulation Factor VII, SHP-2, FN1.4 and DKK4) best discriminated the A+T-N- group from the NPC group with an AUC of 0.92 and a model consisting of 3 proteins (CKAP2, FN1.4 and Cytochrome c) together with age and *APOE* differentiated the A+T-N+ group from the NPC with an AUC of 0.94 (Figure 3F) and (Supplementary Table 4). The A+T+N- group, with only three members, was too small to test for differentiation from the NPC group.

Association between ATN related proteins and AD clinical outcomes

We then determined the relationship between ATN framework-related proteins and AD clinical outcomes in two large independent cohorts: EMIF (183 AD, 382 MCI and 220 CTL) and ANM (319 AD, 149 MCI and 209 CTL). While we found many proteins to be associated with AD clinical outcomes in

the EMIF cohort, few of these replicated in the ANM cohort (Supplementary Figure 1). However, we did find robust replication for two protein associations with clinical features; C3 and albumin. In the EMIF cohort, C3 was significantly decreased in AD compared to CTL and MCI individuals, a change also observed in the ANM cohort (Figure 4A). Furthermore, C3 was significantly decreased in MCIc (EMIF n = 100; ANM n = 43) relative to MCIs (EMIF n = 219; ANM n = 106) in both two cohorts (Figure 4B). C3 had consistent protective effects on cognition as it was positively associated with MMSE score in both EMIF ($R^2 = 0.29$, $p < 0.001$) and ANM cohorts ($R^2 = 0.08$, $p < 0.001$) although the association was weak in the ANM cohort (Figure 4C). Albumin was also significantly decreased in AD compared to CTL in both cohorts (Supplementary Figure 2A). Furthermore, it was positively associated with baseline MMSE score in both EMIF ($R^2 = 0.25$, $p < 0.001$) and ANM cohorts ($R^2 = 0.17$, $p < 0.001$) (Supplementary Figure 2B).

Causal relationship between C3 and AD risk, amyloid and tau pathology

We then sought to investigate the causal relationship between two proteins (C3 and FGG) with AD risk and Aβ and tau pathology (T-tau and P-tau) using two sample MR. A prerequisite for MR is evidence of genetic variations associated with the exposure variable - in this case levels of the proteins in plasma. To identify such variants, we interrogated the GWAS catalogue for all genes encoding the proteins in the panel and found that four proteins with at least two SNPs in their encoding gene significantly associated with their levels in blood ($p < 10^{-8}$). These proteins were C3, FGG, CONA1 and coagulation factor VII. Given that C3 and FGG were associated with AD from previous biomarker studies from our group and others [26, 34-36], we selected these two proteins to further explore their causal relationship with AD. For exposure estimates, we selected three C3 SNPs [26] and 24 FGG SNPs [27] as instruments for MR analysis. For AD risk as outcome, we acquired the summary statistics for the association of each SNP with AD risk from IGAP and UK Biobank [28, 29]. For Aβ and tau pathologies as outcomes, we obtained the association of each SNP with biomarker-based Aβ and tau status from the EMIF cohort.

We first confirmed no pleiotropic effects for these genes given that intercept of genetic variants from MR-Egger regression was close to zero. We then performed MR analysis using both weighted median (WM) and inverse-variance weighted (IVW) approaches. Results showed that lower C3 was likely to be causally related to high AD risk using both WM (effect size [β] = 0.75, standard error of the effect size [se] = 0.44, 95% CI [-0.14, 1.63], p = 0.09) and IVW (β = 0.72, se = 0.40, 95% CI [-0.05, 1.50], p = 0.06) methods (Figure 5). In contrast, such a relationship was not found between C3 with either Aβ or tau (both T-tau and P-tau) status (Figure 5). Furthermore, results from the leave-one-out analysis demonstrated that no single SNP was driving the majority of the association signal between C3 and AD risk (Supplementary Figure 3). For FGG, no causal relationship was found between FGG and AD risk, or FGG and Aβ or FGG and tau (Supplementary Figure 4 and Supplementary Table 5).

Discussion

Identification of biomarkers tends to fall into two different designs; either hypothesis-driven targeted measures of features known to be associated with the disease in question or data-driven high dimensionality agnostic platform approaches ("omics"). Here we utilise a novel approach with a hybrid design where first we used agnostic high dimensionality proteomics in an *in vitro* model of a hypothesised driver of disease mechanism and then used the derived signature in a targeted study in human samples. As Wnt signalling has been proposed to be protective and an increase in the wnt inhibitor DKK1 has been found to be increased in AD [10, 37-39], and hence a possible driver of disease mechanisms, we first empirically identified a DKK1-induced signature from *in vitro* human cell models of DKK1 over-expression and then determined the association of this signature with AD pathology in two large independent cohorts; EMIF (n = 785) and ANM (n = 677). From analysis of high dimensionality proteomics of over 1000 proteins, we determined a 100-protein signature induced by DKK1 and found that this protein set was enriched in molecular pathways known to be

423 associated with AD, adding further evidence to the relevance and possible importance of this
424 mechanism or pathway in disease.

425 We then explored the relationship of the identified signature to the biomarker based AD
426 classification, *i.e.* ATN framework. The role of DKK1 and Wnt signalling in AD is suggested, from
427 multiple lines of evidence mentioned above, to somehow transmit a signal from amyloid to tau
428 pathology and hence neurotoxicity [14-16]. The mechanism of such a transmission is unknown
429 although might be through the canonical Wnt pathway regulation of the tau kinase glycogen
430 synthase kinase-3 (GSK3) or through the non-canonical Wnt pathways that include Rho/Roc and
431 their effects on synaptic resilience [10, 40, 41]. If DKK1 is either a direct or indirect link between
432 amyloid and tau pathology with subsequent effects on neurotoxicity, one would expect that a DKK1
433 induced signature might be associated with these different components of the AD pathway. We
434 indeed found this to be the case with a strong association of many of the DKK1-signature proteins
435 being associated with ATN classifications including amyloid (*i.e.* AD group, A+) but less association
436 with the non-amyloid group (SNAP, A-). These results offer confirmation of DKK inhibition of Wnt
437 signalling as a factor in Alzheimer pathology and specifically add weight to data from *in vitro* and *in*
438 *vivo* models and from human brain studies that DKK1 is increased in response to amyloid and as a
439 consequence increases risk of tau pathology and neurodegeneration.

440 Based on this, we further investigated the predictive value of DKK1-induced proteins in
441 discriminating amyloid pathology (A+T-N-, A+T+N-, A+T-N+ and A+T+N+) from no AD pathology in all
442 subjects as well as in cognitively normal individuals. We found the combination of different subsets
443 of proteins along with age and APOE ε4 genotype was able to differentiate the different ATN profiles
444 with a high AUC, especially in normal individuals (AUC > 0.9). It should be noted that forward
445 stepwise regression makes an arbitrary decision as to select highly correlated proteins. Therefore,
446 other proteins that were highly correlated to those selected proteins could equally function as
447 biomarkers for ATN classification. For example, HMG-1 was selected in the model to discriminate

A+T-N- from NPC in all subjects. As shown in Supplementary Figure 5, it was highly related to BRF.1, indicating that BRF.1 could also be a useful marker to discriminate A+T-N- from NPC.

Our study is the largest we are aware of to report a plasma biomarker indicative of the ATN framework both in terms of the number of proteins assayed and in sample size. As the data was derived from a biomarker platform with claims to have value in other clinical settings (see for example Somalogic.com/somasignals) then the identification of a signature indicative of AD pathological processes might have value in screening from existing data for possible suitability for clinical trials or, when therapies become available, possibly for early intervention, including those related to this particular target. As an approach to precision medicine, this model of biomarker discovery might have wider applicability.

In addition to the ATN framework, we also found the DKK1-derived protein signature associated with AD clinical outcomes (*i.e.* MMSE score and MCI conversion). We observed a robust replication especially for the association of C3 with clinical features, in line with some recent genetic and biomarkers studies [34, 36, 42]. For example, genetic studies strongly implicate complement signalling with AD pathogenesis with increasing attention being paid especially to the complement pathway node of C3 and C5 metabolism, regulated by CR1, one of the common variants most strongly associated with AD [42]. Biomarker studies too suggest that complement signalling is critically altered in AD with a large range of complement proteins, including C3 being repeatedly nominated in agnostic proteomic studies [34, 36]. However, neither genetic nor biomarker studies can alone demonstrate directional causality – in other words, although the genetics strongly implicate complement as a causative biological process neither they nor the biomarker studies are able to say whether any given complement protein is exacerbating or protecting against disease.

Empirical studies have attempted to address this question, critically important in drug development, using model systems. However, the results of such studies are less than clear. In some preclinical *in vivo* studies, C3 knock out offers synaptic protection and reduces amyloid burden in a range of

models [43, 44] although in other studies increased amyloid accumulation and neurodegeneration has been reported [45, 46]. Therefore, to further explore causality, we employed a Mendelian randomization approach with this protein data combined with knowledge of C3 quantitative trait locus (QTL) SNPs associated with protein levels in this data set. We were able to substantiate the association between low levels of C3 and high risk of AD, suggesting a causal influence of C3 on risk of AD, further supporting findings from Rasmussen et al. [26] who studied more than 95,000 individuals from the general population.

We acknowledge that the sample distribution within the ATN framework is a limitation of this study. As shown in Table 1, A+T+N- profile only included 19 subjects. The small sample size might explain why proteins did not reach significance in A+T+N- profile when comparing to A-T-N- profile. It might also lead to the fact that we have not obtained a subset of proteins which could reflect A+T+ profile. Further research in large, well-characterized cohorts to replicate, validate, and extend these findings is needed.

In conclusion, our results add to the evidence base indicating a role for DKK1 and the Wnt pathway in AD pathogenesis. It suggests that a protein signature derived from a human cell model of DKK1 over-expression, when measured in human plasma, is significantly associated with the staging according to the ATN framework and could discriminate different amyloid-positive classes from No Pathology Controls. Furthermore, a subset of the DKK1-signature proteins are strongly associated with disease severity and progression and these association can be replicated in two large independent cohorts. Taken together, our results indicate that this novel, empirically generated approach can help identify biomarkers of utility for the selection of participants for clinical trials as well as for monitoring trial outcomes.

Acknowledgements

This research was conducted as part of the EMIF-AD project which has received support from the Innovative Medicines Initiative Joint Undertaking under EMIF grant agreement n° 115372, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in-kind contribution. The DESCRIPA study was funded by the European Commission within the 5th framework program (QLRT-2001-2455). The EDAR study was funded by the European Commission within the 5th framework program (contract # 37670). The Leuven cohort was funded by the Stichting voor Alzheimer Onderzoek (grant numbers #11020, #13007 and #15005). The Lausanne cohort study was supported by a grant from the Swiss National Research Foundation to JP (SNF 320030_141179). RV is a senior clinical investigator of the Flemish Research Foundation (FWO). JS is currently an employee of UCB, Braine-l'Alleud, Belgium. The San Sebastian GAP study is partially funded by the Department of Health of the Basque Government (allocation 17.0.1.08.12.0000.2.454.01.41142.001.H). We acknowledge the contribution of the personnel of the Genomic Service Facility at the VIB-U Antwerp Center for Molecular Neurology. The research at VIB-CMN is funded in part by the University of Antwerp Research Fund. FB is supported by the NIHR biomedical research centre at UCLH. RD is supported by Health Data Research UK, the National Institute for Health Research (NIHR) Biomedical Research Centre at South London, Maudsley NHS Foundation Trust, King's College London and the NIHR University College London Hospitals Biomedical Research Centre. LS is funded by DPUK through MRC (grant no. MR/L023784/2) and the UK Medical Research Council Award to the University of Oxford (grant no. MC_PC_17215). Also support was received from the NIHR Biomedical Research Centre at Oxford Health NHS Foundation Trust.

Data availability: The data sets generated and analysed during the present study are available online. Briefly, access to EMIF Somascan data could be obtained via the EMIF-AD Catalogue (http://www.emif.eu/emif-catalogue/). Access to ANM Somascan data could be via AData (Viewer) (https://adata.scai.fraunhofer.de/). Research questions need to be submitted when applying for the access. *In vitro* Somascan data are available upon request.

Conflict of Interest

SL has been named as an inventor on biomarker intellectual property protected by Proteome Sciences and Kings College London; in both cases not the subject of the current study. He has served on the advisory board to SomaLogic Inc. and has provided consultancy for Merck and Eisai within the past three years. He is currently an employee of Janssen-Cilag, UK. HZ has served at scientific advisory boards for Roche Diagnostics, Wave, Samumed and CogRx, has given lectures in symposia sponsored by Biogen and Alzecure, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB, a GU Ventures-based platform company at the University of Gothenburg (all outside submitted work). JLM has served at scientific advisory boards for Eli Lilly, Roche Diagnostics, Roche, Biogen, Lundbeck, Novartis, IBL, Axovant, Oryzon, Merck (all unrelated to this study). AL has served at scientific advisory boards of Fujirebio Europe, Eli Lilly, Novartis, Nutricia and Otsuka and is the inventor of a patent on synaptic markers in CSF (all unrelated to this study). LF has received research funding, consultancy fees or speech honoraria from Allergan, Avid-Eli Lilly, Avanir, Avraham Pharmaceuticals, Axon Neuroscience, Axovant, Biogen, Boehringer Ingelheim, Eisai, Functional Neuromodulation, GE Health Care, Lundbeck, MerckSharpe&Dohme, Novartis, Pfizer, Piramal Imaging, Roche, Schwabe Pharma. JP served as a consultant for Fujirebio and Nestlé Institute of Health Sciences, all unrelated to this study. HS and PJ have served at advisory boards for ACImmune and MERCK outside the submitted work. SE has served as a consultant for Nutricia, Novartis and Eisai and has received research funding of Janssen Pharmaceutica and ADx Neurosciences (paid to institution; all outside submitted work). SK has served on an advisory board for Roche Diagnostics, and as a consultant to DIADEM (all outside of submitted work).

References

[1] Lane CA, Hardy J, Schott JM (2018) Alzheimer's disease. *Eur J Neurol* **25**, 59-70.

[2] Jack CR, Jr., Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Holtzman DM, Jagust W, Jessen F, Karlawish J, Liu E, Molinuevo JL, Montine T, Phelps C, Rankin KP, Rowe CC, Scheltens P, Siemers E, Snyder HM, Sperling R (2018) NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* **14**, 535-562.

[3] Cummings J (2019) The National Institute on Aging-Alzheimer's Association Framework on Alzheimer's disease: Application to clinical trials. *Alzheimers Dement* **15**, 172-178.

[4] Yu JT, Li JQ, Suckling J, Feng L, Pan A, Wang YJ, Song B, Zhu SL, Li DH, Wang HF, Tan CC, Dong Q, Tan L, Mok V, Aisen PS, Weiner MM (2019) Frequency and longitudinal clinical outcomes of Alzheimer's AT(N) biomarker profiles: A longitudinal study. *Alzheimers Dement* **15**, 1208-1217.

[5] Soldan A, Pettigrew C, Fagan AM, Schindler SE, Moghekar A, Fowler C, Li QX, Collins SJ, Carlsson C, Asthana S, Masters CL, Johnson S, Morris JC, Albert M, Gross AL (2019) ATN profiles among cognitively normal individuals and longitudinal cognitive outcomes. *Neurology* **92**, e1567-e1579.

[6] De Ferrari GV, Chacon MA, Barria MI, Garrido JL, Godoy JA, Olivares G, Reyes AE, Alvarez A, Bronfman M, Inestrosa NC (2003) Activation of Wnt signaling rescues neurodegeneration and behavioral impairments induced by beta-amyloid fibrils. *Mol Psychiatry* **8**, 195-208.

[7] Alvarez AR, Godoy JA, Mullendorff K, Olivares GH, Bronfman M, Inestrosa NC (2004) Wnt-3a overcomes beta-amyloid toxicity in rat hippocampal neurons. *Exp Cell Res* **297**, 186-196.

[8] Vargas JY, Fuenzalida M, Inestrosa NC (2014) In vivo activation of Wnt signaling pathway enhances cognitive function of adult mice and reverses cognitive deficits in an Alzheimer's disease model. *Journal of Neuroscience* **34**, 2191-2202.

[9] Toledo EM, Inestrosa NC (2010) Activation of Wnt signaling by lithium and rosiglitazone reduced spatial memory impairment and neurodegeneration in brains of an APPswe/PSEN1DeltaE9 mouse model of Alzheimer's disease. *Mol Psychiatry* **15**, 272-285, 228.

[10] Tapia-Rojas C, Inestrosa NC (2018) Loss of canonical Wnt signaling is involved in the pathogenesis of Alzheimer's disease. *Neural Regen Res* **13**, 1705-1710.

[11] Caricasole A, Copani A, Caraci F, Aronica E, Rozemuller AJ, Caruso A, Storto M, Gaviraghi G, Terstappen GC, Nicoletti F (2004) Induction of Dickkopf-1, a negative modulator of the Wnt pathway, is associated with neuronal degeneration in Alzheimer's brain. *J Neurosci* **24**, 6021-6027.

[12] Rosi MC, Luccarini I, Grossi C, Fiorentini A, Spillantini MG, Prisco A, Scali C, Gianfriddo M, Caricasole A, Terstappen GC, Casamenti F (2010) Increased Dickkopf-1 expression in transgenic mouse models of neurodegenerative disease. *J Neurochem* **112**, 1539-1551.

[13] Bayod S, Felice P, Andres P, Rosa P, Camins A, Pallas M, Canudas AM (2015) Downregulation of canonical Wnt signaling in hippocampus of SAMP8 mice. *Neurobiol Aging* **36**, 720-729.

[14] Galli S, Lopes DM, Ammari R, Kopra J, Millar SE, Gibb A, Salinas PC (2014) Deficient Wnt signalling triggers striatal synaptic degeneration and impaired motor behaviour in adult mice. *Nat Commun* **5**, 4992.

[15] Sellers KJ, Elliott C, Jackson J, Ghosh A, Ribe E, Rojo AI, Jarosz-Griffiths HH, Watson IA, Xia W, Semenov M, Morin P, Hooper NM, Porter R, Preston J, Al-Shawi R, Baillie G, Lovestone S, Cuadrado A, Harte M, Simons P, Srivastava DP, Killick R (2018) Amyloid beta synaptotoxicity is Wnt-PCP dependent and blocked by fasudil. *Alzheimers Dement* **14**, 306-317.

[16] Purro SA, Dickins EM, Salinas PC (2012) The secreted Wnt antagonist Dickkopf-1 is required for amyloid beta-mediated synaptic loss. *J Neurosci* **32**, 3492-3498.

[17] Marzo A, Galli S, Lopes D, McLeod F, Podpolny M, Segovia-Roldan M, Ciani L, Purro S, Cacucci F, Gibb A, Salinas PC (2016) Reversal of Synapse Degeneration by Restoring Wnt Signaling in the Adult Hippocampus. *Curr Biol* **26**, 2551-2561.

[18] Killick R, Ribe EM, Al-Shawi R, Malik B, Hooper C, Fernandes C, Dobson R, Nolan PM, Lourdusamy A, Furney S, Lin K, Breen G, Wroe R, To AW, Leroy K, Causevic M, Usardi A, Robinson M, Noble W, Williamson R, Lunnon K, Kellie S, Reynolds CH, Bazenet C, Hodges A,

Brion JP, Stephenson J, Simons JP, Lovestone S (2014) Clusterin regulates beta-amyloid toxicity via Dickkopf-1-driven induction of the wnt-PCP-JNK pathway. *Mol Psychiatry* **19**, 88-98.

[19] Lovestone S, Francis P, Kloszewska I, Mecocci P, Simmons A, Soininen H, Spenger C, Tsolaki M, Vellas B, Wahlund LO, Ward M, AddNeuroMed C (2009) AddNeuroMed--the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. *Ann N Y Acad Sci* **1180**, 36-46.

[20] Bos I, Vos S, Vandenberghe R, Scheltens P, Engelborghs S, Frisoni G, Molinuevo JL, Wallin A, Lleo A, Popp J, Martinez-Lage P, Baird A, Dobson R, Legido-Quigley C, Sleegers K, Van Broeckhoven C, Bertram L, Ten Kate M, Barkhof F, Zetterberg H, Lovestone S, Streffer J, Visser PJ (2018) The EMIF-AD Multimodal Biomarker Discovery study: design, methods and cohort characteristics. *Alzheimers Res Ther* **10**, 64.

[21] Westman E, Simmons A, Muehlboeck JS, Mecocci P, Vellas B, Tsolaki M, Kloszewska I, Soininen H, Weiner MW, Lovestone S, Spenger C, Wahlund LO (2011) AddNeuroMed and ADNI: similar patterns of Alzheimer's atrophy and automated MRI classification accuracy in Europe and North America. *Neuroimage* **58**, 818-828.

[22] Hong S, Prokopenko D, Dobricic V, Kilpert F, Bos I, Vos SJ, Tijms BM, Andreasson U, Blennow K, Vandenberghe R (2019) Genome-wide association study of Alzheimer's disease CSF biomarkers in the EMIF-AD Multimodal Biomarker Discovery dataset. *bioRxiv*, 774554.

[23] Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, Carter J, Dalby AB, Eaton BE, Fitzwater T, Flather D, Forbes A, Foreman T, Fowler C, Gawande B, Goss M, Gunn M, Gupta S, Halladay D, Heil J, Heilig J, Hicke B, Husar G, Janjic N, Jarvis T, Jennings S, Katilius E, Keeney TR, Kim N, Koch TH, Kraemer S, Kroiss L, Le N, Levine D, Lindsey W, Lollo B, Mayfield W, Mehan M, Mehler R, Nelson SK, Nelson M, Nieuwlandt D, Nikrad M, Ochsner U, Ostroff RM, Otis M, Parker T, Pietrasiewicz S, Resnicow DI, Rohloff J, Sanders G, Sattin S, Schneider D, Singer B, Stanton M, Sterkel A, Stewart A, Stratford S, Vaught JD, Vrkljan M, Walker JJ, Watrobka M,

Waugh S, Weiss A, Wilcox SK, Wolfson A, Wolk SK, Zhang C, Zichi D (2010) Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One* **5**, e15004.

[24] Grover S, Del Greco MF, Stein CM, Ziegler A (2017) Mendelian Randomization. *Methods Mol Biol* **1666**, 581-628.

[25] Davey Smith G, Hemani G (2014) Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* **23**, R89-98.

[26] Rasmussen KL, Nordestgaard BG, Frikke-Schmidt R, Nielsen SF (2018) An updated Alzheimer hypothesis: Complement C3 and risk of Alzheimer's disease-A cohort study of 95,442 individuals. *Alzheimers Dement* **14**, 1589-1601.

[27] Sabater-Lleal M, Huang J, Chasman D, Naitza S, Dehghan A, Johnson AD, Teumer A, Reiner AP, Folkersen L, Basu S, Rudnicka AR, Trompet S, Malarstig A, Baumert J, Bis JC, Guo X, Hottenga JJ, Shin SY, Lopez LM, Lahti J, Tanaka T, Yanek LR, Oudot-Mellakh T, Wilson JF, Navarro P, Huffman JE, Zemunik T, Redline S, Mehra R, Pulanic D, Rudan I, Wright AF, Kolcic I, Polasek O, Wild SH, Campbell H, Curb JD, Wallace R, Liu S, Eaton CB, Becker DM, Becker LC, Bandinelli S, Raikkonen K, Widen E, Palotie A, Fornage M, Green D, Gross M, Davies G, Harris SE, Liewald DC, Starr JM, Williams FM, Grant PJ, Spector TD, Strawbridge RJ, Silveira A, Sennblad B, Rivadeneira F, Uitterlinden AG, Franco OH, Hofman A, van Dongen J, Willemsen G, Boomsma DI, Yao J, Swords Jenny N, Haritunians T, McKnight B, Lumley T, Taylor KD, Rotter JI, Psaty BM, Peters A, Gieger C, Illig T, Grotevendt A, Homuth G, Volzke H, Kocher T, Goel A, Franzosi MG, Seedorf U, Clarke R, Steri M, Tarasov KV, Sanna S, Schlessinger D, Stott DJ, Sattar N, Buckley BM, Rumley A, Lowe GD, McArdle WL, Chen MH, Tofler GH, Song J, Boerwinkle E, Folsom AR, Rose LM, Franco-Cereceda A, Teichert M, Ikram MA, Mosley TH, Bevan S, Dichgans M, Rothwell PM, Sudlow CL, Hopewell JC, Chambers JC, Saleheen D, Kooner JS, Danesh J, Nelson CP, Erdmann J, Reilly MP, Kathiresan S, Schunkert H, Morange PE, Ferrucci L, Eriksson JG, Jacobs D, Deary IJ, Soranzo N, Witteman JC, de Geus EJ, Tracy RP, Hayward C, Koenig W, Cucca F, Jukema JW, Eriksson P, Seshadri S, Markus HS, Watkins H,

Samani NJ, Wallaschofski H, Smith NL, Tregouet D, Ridker PM, Tang W, Strachan DP, Hamsten A, O'Donnell CJ (2013) Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated Loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease. *Circulation* **128**, 1310-1324.

[28] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thorton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau MT, Choi SH, Reitz C, Pasquier F, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Moron FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fievet N, Huentelman MW, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuiness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogaeva E, Gallacher J, St George-Hyslop P, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossu P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Deniz Naranjo MC, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannefelt L, Hakonarson H, Pichler S, Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley TH, Jr., Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RF, Passmore P, Montine TJ, Bettens K, Rotter JI, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell JF, Nalls MA, Ritchie K, Lunetta KL, Kauwe JS, Boerwinkle E, Riemenschneider M, Boada M, Hiltuenen M, Martin ER, Schmidt R, Rujescu D, Wang LS, Dartigues JF, Mayeux R, Tzourio C, Hofman A, Nothen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van

677 Duijn CM, Van Broeckhoven C, Moskvina V, Seshadri S, Williams J, Schellenberg GD, Amouyel
678 P (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for
679 Alzheimer's disease. *Nat Genet* **45**, 1452-1458.
680 [29] Marioni RE, Harris SE, Zhang Q, McRae AF, Hagenaars SP, Hill WD, Davies G, Ritchie CW, Gale
681 CR, Starr JM, Goate AM, Porteous DJ, Yang J, Evans KL, Deary IJ, Wray NR, Visscher PM
682 (2018) GWAS on family history of Alzheimer's disease. *Transl Psychiatry* **8**, 99.
683 [30] Borenstein M, Hedges LV, Higgins JP, Rothstein HR (2010) A basic introduction to fixed-effect
684 and random-effects models for meta-analysis. *Res Synth Methods* **1**, 97-111.
685 [31] Bowden J, Davey Smith G, Haycock PC, Burgess S (2016) Consistent Estimation in Mendelian
686 Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet*
687 *Epidemiol* **40**, 304-314.
688 [32] Bowden J, Davey Smith G, Burgess S (2015) Mendelian randomization with invalid
689 instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*
690 **44**, 512-525.
691 [33] Hemani G, Zheng J, Wade KH, Laurin C, Elsworth B, Burgess S, Bowden J, Langdon R, Tan V,
692 Yarmolinsky J (2016) MR-Base: a platform for systematic causal inference across the
693 phenome using billions of genetic associations. *BioRxiv*, 078972.
694 [34] Shi L, Baird AL, Westwood S, Hye A, Dobson R, Thambisetty M, Lovestone S (2018) A Decade
695 of Blood Biomarkers for Alzheimer's Disease Research: An Evolving Field, Improving Study
696 Designs, and the Challenge of Replication. *J Alzheimers Dis* **62**, 1181-1198.
697 [35] Baird AL, Westwood S, Lovestone S (2015) Blood-Based Proteomic Biomarkers of Alzheimer's
698 Disease Pathology. *Front Neurol* **6**, 236.
699 [36] Ashton NJ, Kiddle SJ, Graf J, Ward M, Baird AL, Hye A, Westwood S, Wong KV, Dobson RJ,
700 Rabinovici GD, Miller BL, Rosen HJ, Torres A, Zhang Z, Thurfjell L, Covin A, Hehir CT, Baker D,
701 Bazenet C, Lovestone S (2015) Blood protein predictors of brain amyloid for enrichment in
702 clinical trials? *Alzheimers Dement (Amst)* **1**, 48-60.

[37] Palomer E, Buechler J, Salinas PC (2019) Wnt Signaling Deregulation in the Aging and Alzheimer's Brain. *Front Cell Neurosci* **13**, 227.

[38] Ren C, Gu X, Li H, Lei S, Wang Z, Wang J, Yin P, Zhang C, Wang F, Liu C (2019) The role of DKK1 in Alzheimer's disease: A potential intervention point of brain damage prevention? *Pharmacol Res* **144**, 331-335.

[39] Anderton BH, Dayanandan R, Killick R, Lovestone S (2000) Does dysregulation of the Notch and wingless/Wnt pathways underlie the pathogenesis of Alzheimer's disease? *Mol Med Today* **6**, 54-59.

[40] Libro R, Bramanti P, Mazzon E (2016) The role of the Wnt canonical signaling in neurodegenerative diseases. *Life Sci* **158**, 78-88.

[41] Inestrosa NC, Varela-Nallar L (2014) Wnt signaling in the nervous system and in Alzheimer's disease. *J Mol Cell Biol* **6**, 64-74.

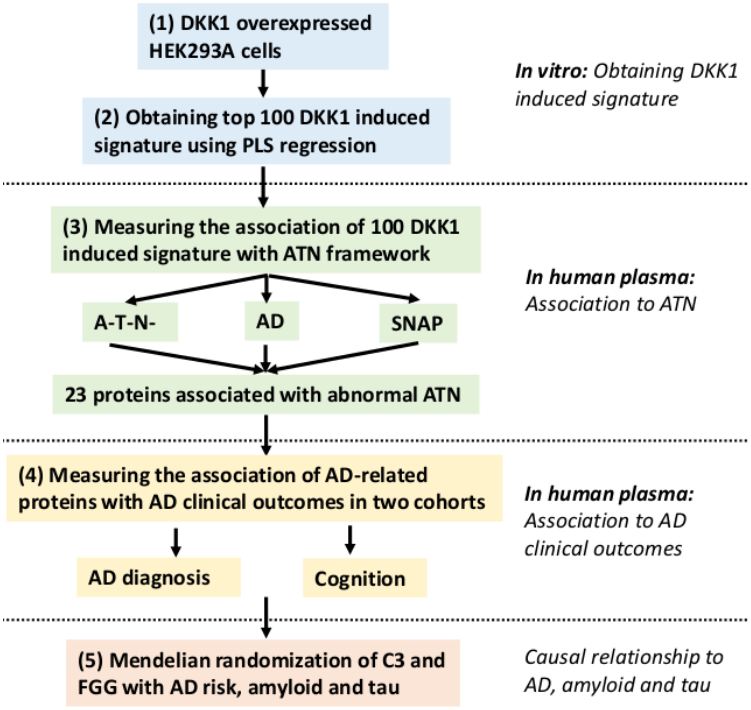
[42] Kucukkilic E, Brookes K, Barber I, Guetta-Baranes T, Morgan K, Hollox EJ (2018) Complement receptor 1 gene (CR1) intragenic duplication and risk of Alzheimer's disease. *Hum Genet* **137**, 305-314.

[43] Counts SE, Ikonomovic MD, Mercado N, Vega IE, Mufson EJ (2017) Biomarkers for the Early Detection and Progression of Alzheimer's Disease. *Neurotherapeutics* **14**, 35-53.

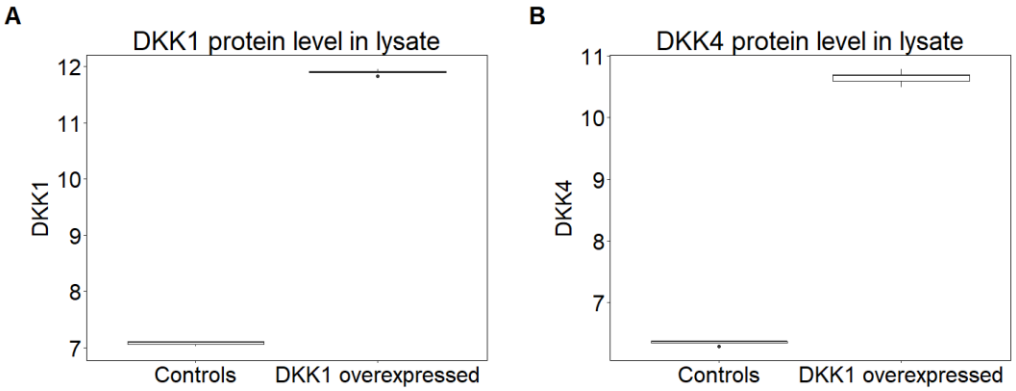
[44] Keshavan A, Heslegrave A, Zetterberg H, Schott JM (2017) Blood Biomarkers for Alzheimer's Disease: Much Promise, Cautious Progress. *Mol Diagn Ther* **21**, 13-22.

[45] Wyss-Coray T, Yan F, Lin AH, Lambris JD, Alexander JJ, Quigg RJ, Masliah E (2002) Prominent neurodegeneration and increased plaque formation in complement-inhibited Alzheimer's mice. *Proc Natl Acad Sci U S A* **99**, 10837-10842.

[46] Maier M, Peng Y, Jiang L, Seabrook TJ, Carroll MC, Lemere CA (2008) Complement C3 deficiency leads to accelerated amyloid beta plaque deposition and neurodegeneration and modulation of the microglia/macrophage phenotype in amyloid precursor protein transgenic mice. *J Neurosci* **28**, 6333-6341.



729 **Figure 1** Flowchart of Study Design. (1) Measurement and quantification of 1128 proteins in total cell
730 lysates of HEK293A cells overexpressing DKK1; (2) The top 100 proteins that constitute the DKK1-
731 induced signature were identified using partial least squares (PLS) regression; (3) Measuring the
732 association of 100 DKK1-induced proteins with ATN framework *in vivo* and obtaining 23 proteins that
733 were significantly associated with any single ATN abnormal; (4) Measuring the association of AD
734 related proteins with other AD clinical outcomes; (5) Mendelian randomization to explore the causal
735 relationship between two proteins (complement component 3 [C3] and fibrinogen gamma chain
736 [FGG]) and AD risk, amyloid and tau (both T-tau and P-tau) pathology. AD, Alzheimer's disease;
737 SNAP, Suspected Non-Alzheimer Pathology; T-tau, total tau; P-tau, phosphorylated tau.



738 **Figure 2** DKK1 overexpression leads to higher levels of (A) DKK1 and (B) DKK4 expression in HEK293A
739 cell lysate (n = 5 per condition). Y axis represents the log transformed of proteins expression
740 abundance measured by Somascan assay.

741

742

743

744

745

746

747

748

749

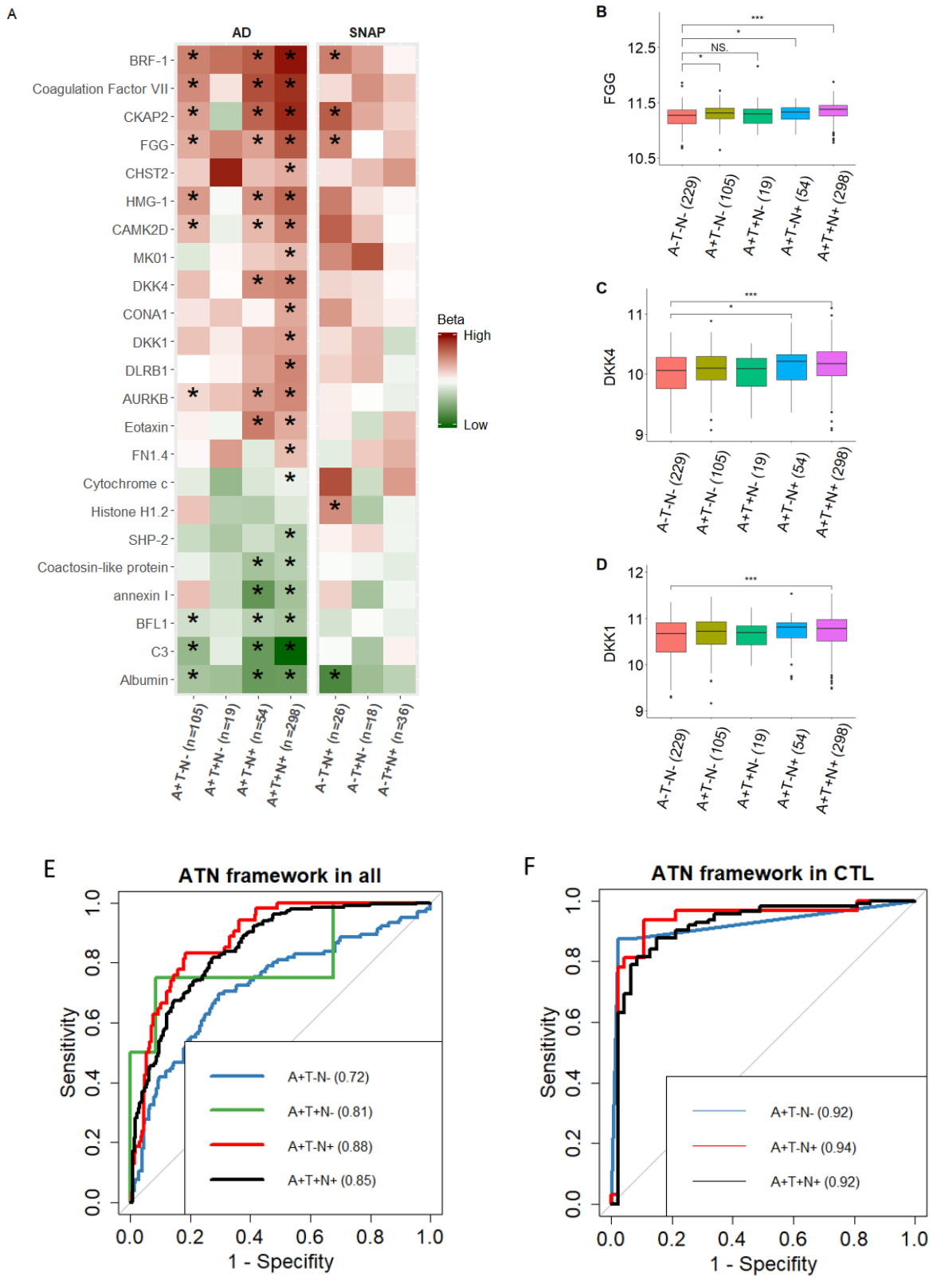
750

751

752

Variable	NPC	AD				SNAP			*P* value
	A-T-N-	A+T-N-	A+T+N-	A+T-N+	A+T+N+	A-T-N+	A-T+N-	A-T+N+	
N (total = 785)	229	105	19	54	298	26	18	36	NA
Age (yrs.) (SD)	64 (8.5)	68 (9.7)	72 (6.8)	72 (7.3)	70 (8.0)	72 (7.1)	64 (6.3)	71 (8.2)	< 0.001
Male sex N (%)	123 (54)	57 (54)	9 (47)	30 (56)	170 (57)	14 (54)	10 (56)	23 (64)	0.95
APOE ε4 carriers N (%)	52 (23)	58 (55)	13 (68)	36 (67)	193 (65)	5 (19)	5 (28)	9 (25)	< 0.001
MMSE (SD)	27.7 (2.5)	26.5 (3.9)	25.4 (3.9)	24.4 (4.3)	24.0 (4.4)	26.7 (3.6)	27.8 (1.7)	26.9 (2.8)	< 0.001

Table 1 Characteristics of 785 participants split by ATN framework. Standard deviation is shown in brackets for age and MMSE in each category. Percentage of cases is shown in brackets for male sex as well as *APOE* ε4 carriers. *P*-values compare each demographic across 8 categories. NPC, No Pathology Control; SNAP, Suspected Non-Alzheimer Pathology; SD, Standard deviation; MMSE, mini mental state examination.

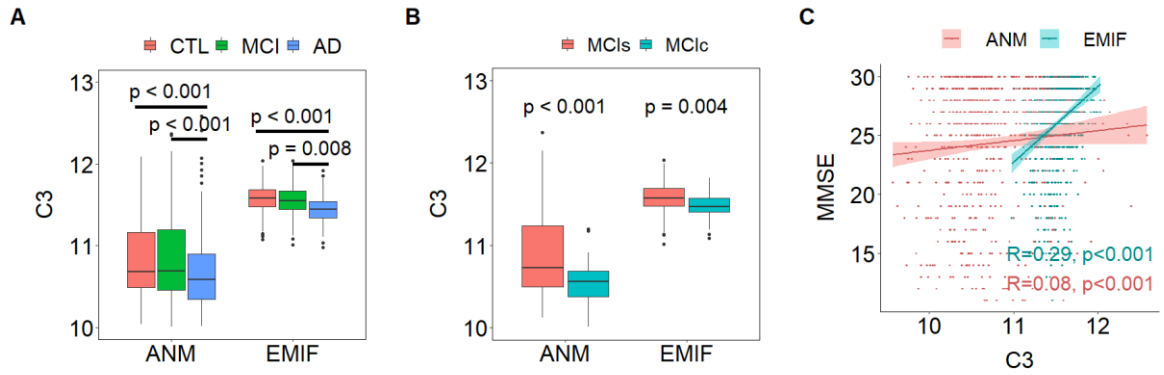


758 **Figure 3** (A) Association of 23 DKK1-induced signature with 7 ATN profiles compared to A-T-N-; (B)

759 (C) and (D) comparison of proteins between A-T-N- (n = 229) and amyloid-positive individuals

760 including A+T-N- (n = 105), A+T+N- (n = 19), A+T-N+ (n = 54) and A+T+N+ (n = 289); (E) and (F) AUC of

using proteins along with age and *APOE* ε4 genotype to differentiate A-T-N- from amyloid-positive individuals in all individuals and healthy controls respectively. High and low beta indicate positive and negative coefficients respectively. SNAP, Suspected Non-Alzheimer Pathology; FGG, fibrinogen gamma chain. In B, C and D, Y axis represents the log transformed of proteins expression abundance measured by Somascan assay. * $p < 0.05$; *** $p < 0.001$; NS., not significant; AUC, area under the curve; CTL, controls.



778 **Figure 4** (A) Comparison of C3 in different AD diagnostic groups in both EMIF and ANM cohorts; (B)
779 comparison of C3 in MCI who subsequently converted to dementia (MCIc) to those whose MCI
780 remained stable (MCIs) in both cohorts; (C) Correlation of MMSE with C3 in both cohorts. Y axis in A
781 & B and X axis of C represent the log transformed of proteins expression abundance measured by
782 Somascan assay. C3, complement component 3; ANM, AddNeuroMed; EMIF, European Medical
783 Information Framework.

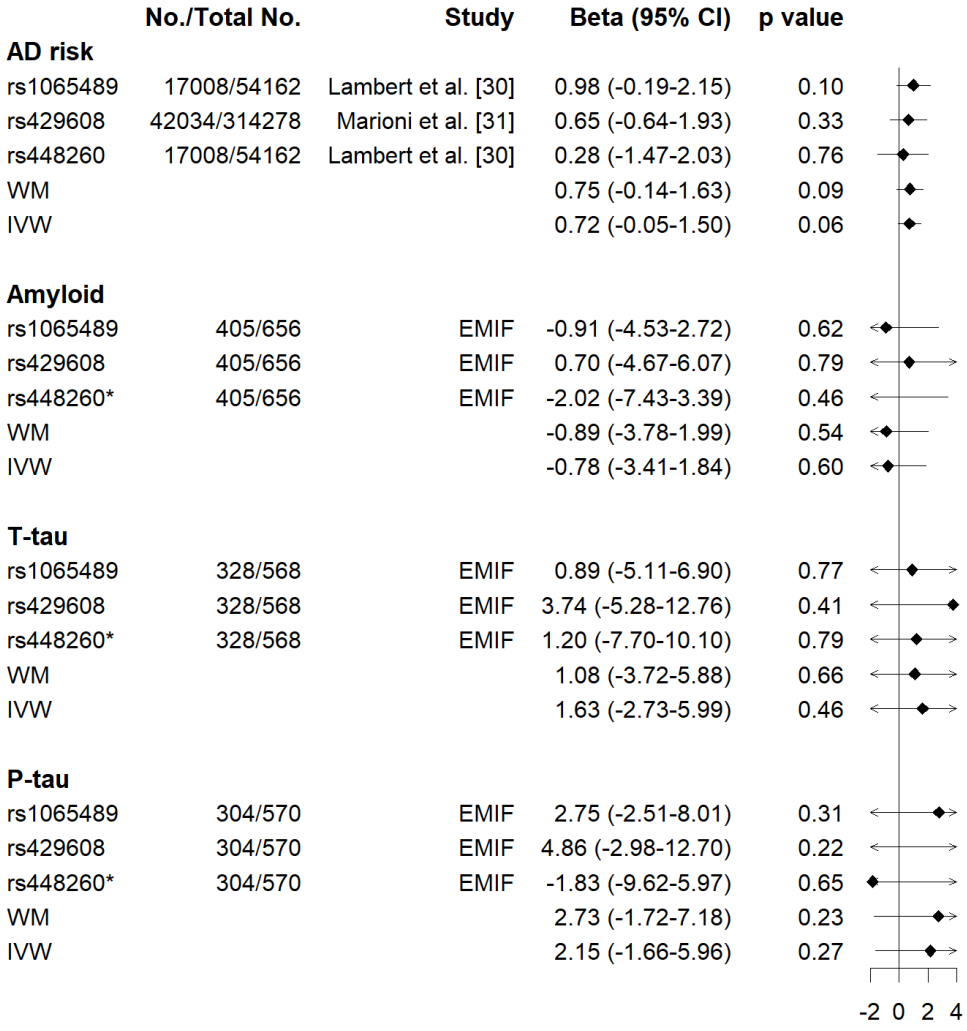


Figure 5 Forest plot of Mendelian randomization estimates the effects of C3 on AD risk, Aβ and tau status (T-tau and P-tau). Lower C3 is likely to be causally related to high AD risk from both inverse-variance weighted (IVW) and weighted median (WM) methods, but such relationship was not found between C3 and amyloid or C3 and tau (T-tau and P-tau). *rs448260 was not found in EMIF data, therefore its proxy rs2287848 (r^2 = 0.93) was used to obtain its association with amyloid and tau. EMIF, European Medical Information Framework; T-tau, total tau; P-tau, phosphorylated tau.

Supplementary Tables and Figures

Characteristics	Sample size	CTL	MCI	AD	*P* value
ANM cohort					
Size	677	209	149	319	NA
Male sex N (%)	677	102 (49)	57 (38)	98 (31)	< 0.001 *
Age	677	74.4 (6.3)	76.1 (6.8)	79.0 (7.0)	< 0.001 †
APOE ε4+ N (%)	677	56 (27)	59 (40)	180 (56)	< 0.001 *
MMSE	571	29.1 (1.0)	26.9 (2.0)	19.6 (5.0)	< 0.001 †
ADAS-Cog	191	NA	NA	24.4 (10.1)	NA
EMIF cohort					
Size	785	220	382	183	NA
Male sex N (%)	785	136 (62)	203 (53)	97 (53)	0.09 *
Age	785	63.3 (7.8)	70.1 (8.1)	70.3 (8.8)	< 0.001 †
APOE ε4+ N (%)	785	75 (34)	185 (48)	111 (61)	< 0.001 *
MMSE	782	28.8 (1.3)	26.2 (2.7)	21.4 (4.8)	< 0.001 †
Executive z score	564	0.25 (1.33)	-0.92 (1.88)	-2.47 (2.30)	< 0.001 †
Language z score	751	-0.32 (0.96)	-0.83 (1.29)	-2.12 (1.19)	< 0.001 †
Memory delayed z score	624	-0.41 (1.13)	-1.21 (1.38)	-2.42 (1.09)	< 0.001 †
Memory immediate z score	694	-0.91 (1.90)	-1.31 (1.42)	-2.48 (1.24)	< 0.001 †

Supplementary Table 1 Sample characteristics of EMIF and ANM cohorts by clinical diagnosis. *chi-square test; †One-way analysis of variance (ANOVA) test; ANM, AddNeuroMed; EMIF, European Medical Information Framework; CTL, control; MCI, mild cognitive impairment; AD, Alzheimer's disease; MMSE, mini mental state examination; ADAS-Cog, Alzheimer's Disease Assessment Scale - Cognitive subscale.

Uniprot	Protein Name	Regression coefficients
O94907	Dickkopf-related protein 1	-0.236
Q9UBT3	Dickkopf-related protein 4	-0.210
P01024	C3a anaphylatoxin des Arginine	0.031
P49862	Kallikrein-7	0.022
P01024	Complement C3	0.020
P02671	Fibrinogen	0.018
O15264	Mitogen-activated protein kinase 13	0.017
Q08188	Protein-glutamine gamma-glutamyltransferase E	0.016
Q9H422	Homeodomain-interacting protein kinase 3	0.016
P02775	Connective tissue-activating peptide III	0.016
P09429	High mobility group protein B1	0.016
P02788	Lactotransferrin	0.015
P41240	Tyrosine-protein kinase CSK	0.014
P25098	beta-adrenergic receptor kinase 1	0.013
O14929	Histone acetyltransferase type B catalytic subunit	-0.013
P04406	Glyceraldehyde-3-phosphate dehydrogenase	0.013
P05186	Alkaline phosphatase, tissue-nonspecific isozyme	0.013
P49840	Glycogen synthase kinase-3 alpha/beta	0.013
P63241	Eukaryotic translation initiation factor 5A-1	0.013
Q05513	Protein kinase C zeta type	0.013
P12268	Inosine-5'-monophosphate dehydrogenase 2	0.012
P51665	26S proteasome non-ATPase regulatory subunit 7	0.012
O96017	Serine/threonine-protein kinase Chk2	0.012
P61626	Lysozyme C	0.012
P99999	Cytochrome c	0.012

Q92994	Transcription factor IIIB 90 kDa subunit	0.010
P35354	Prostaglandin G/H synthase 2	0.010
P16403	Histone H1.2	0.009
P12259	Coagulation Factor V	0.009
P52823	Stanniocalcin-1	0.009
Q06124	Tyrosine-protein phosphatase non-receptor type 11	0.009
P16591	Tyrosine-protein kinase Fer	0.009
P56211	cAMP-regulated phosphoprotein 19	0.009
Q05397	Focal adhesion kinase 1	0.009
P11387	DNA topoisomerase 1	-0.009
P25685	DnaJ homolog subfamily B member 1	0.008
P20226	TATA-box-binding protein	-0.008
P50579	Methionine aminopeptidase 2	0.008
P30533	alpha-2-macroglobulin receptor-associated protein	0.008
P12956	X-ray repair cross-complementing protein 6	0.008
P02679	Fibrinogen gamma chain	0.008
P04141	Granulocyte-macrophage colony-stimulating factor	0.008
P14618	Pyruvate kinase PKM	-0.008
Q08345	Epithelial discoidin domain-containing receptor 1	0.008
P01833	Polymeric immunoglobulin receptor	0.007
Q9UQ80	Proliferation-associated protein 2G4	0.007
P02768	Serum albumin	0.007
P40121	Macrophage-capping protein	0.007
P28482	Mitogen-activated protein kinase 1	0.007
Q9Y337	Kallikrein-5	0.007

O00548	Delta-like protein 1	0.007
Q9Y4C5	Carbohydrate sulfotransferase 2	-0.007
P04083	Annexin A1	0.007
P62826	GTP-binding nuclear protein Ran	0.007
P08709	Coagulation Factor VII	0.006
P07996	Thrombospondin-1	0.006
P63279	SUMO-conjugating enzyme UBC9	0.006
P51671	Eotaxin	0.006
P07355	Annexin A2	0.006
Q13557	Calcium/calmodulin-dependent protein kinase type II subunit delta	0.006
Q9Y3C8	Ubiquitin-fold modifier-conjugating enzyme 1	0.006
P27361	Mitogen-activated protein kinase 3	0.006
P05121	Plasminogen activator inhibitor 1	-0.006
P0C0S5	Histone H2A.z	-0.006
P09228	Cystatin-SA	0.006
Q86Y22	Collagen alpha-1(XXIII) chain	-0.006
Q07021	Complement component 1 Q subcomponent-binding protein, mitochondrial	0.006
Q8N3X6	Ligand-dependent nuclear receptor corepressor-like protein	0.006
P01011	Alpha-1-antichymotrypsin	0.006
Q14126	Desmoglein-2	0.006
P19876	Gro-beta/gamma	-0.006
O95219	Sorting nexin-4	0.006
P02751	Fibronectin Fragment 4	-0.006
O00299	Chloride intracellular channel protein 1	0.005
P01033	Metalloproteinase inhibitor 1	0.005

P07288	PSA:alpha-1-antichymotrypsin complex	0.005
Q8N474	Secreted frizzled-related protein 1	0.005
Q08209	Calcineurin	-0.005
Q9UBN6	Tumor necrosis factor receptor superfamily member 10D	0.005
P24941	Cyclin-dependent kinase 2:Cyclin-A2 complex	0.005
Q8WWK9	Cytoskeleton-associated protein 2	0.005
Q9NP97	Dynein light chain roadblock-type 1	-0.005
Q9BQ51	Programmed cell death 1 ligand 2	-0.005
P16333	Cytoplasmic protein NCK1	0.005
P10909	Clusterin	0.005
P03971	Muellerian-inhibiting factor	0.005
P46527	Cyclin-dependent kinase inhibitor 1B	0.005
Q96MS0	Roundabout homolog 3	-0.005
Q96GD4	Aurora kinase B	-0.005
Q13765	Nascent polypeptide-associated complex subunit alpha	0.005
Q14019	Coactosin-like protein	0.005
Q16548	Bcl-2-related protein A1	0.005
P00558	Phosphoglycerate kinase 1	0.005
O00220	Tumor necrosis factor receptor superfamily member 10A	-0.005
O95998	Interleukin-18-binding protein	0.005
P07333	Macrophage colony-stimulating factor 1 receptor	-0.005
P17706	Tyrosine-protein phosphatase non-receptor type 2	0.005
P05231	Interleukin-6	-0.005
P30419	Glycylpeptide N-tetradecanoyltransferase 1	0.005
Q99714	3-hydroxyacyl-CoA dehydrogenase type-2	0.005

Supplementary Table 2 Top 100 proteins with the largest contribution to the Dickkopf-related protein 1-induced multivariate signature using partial least squares regression. The rank was based on their contribution to the DKK1-induced multivariate signature from the largest to smallest.

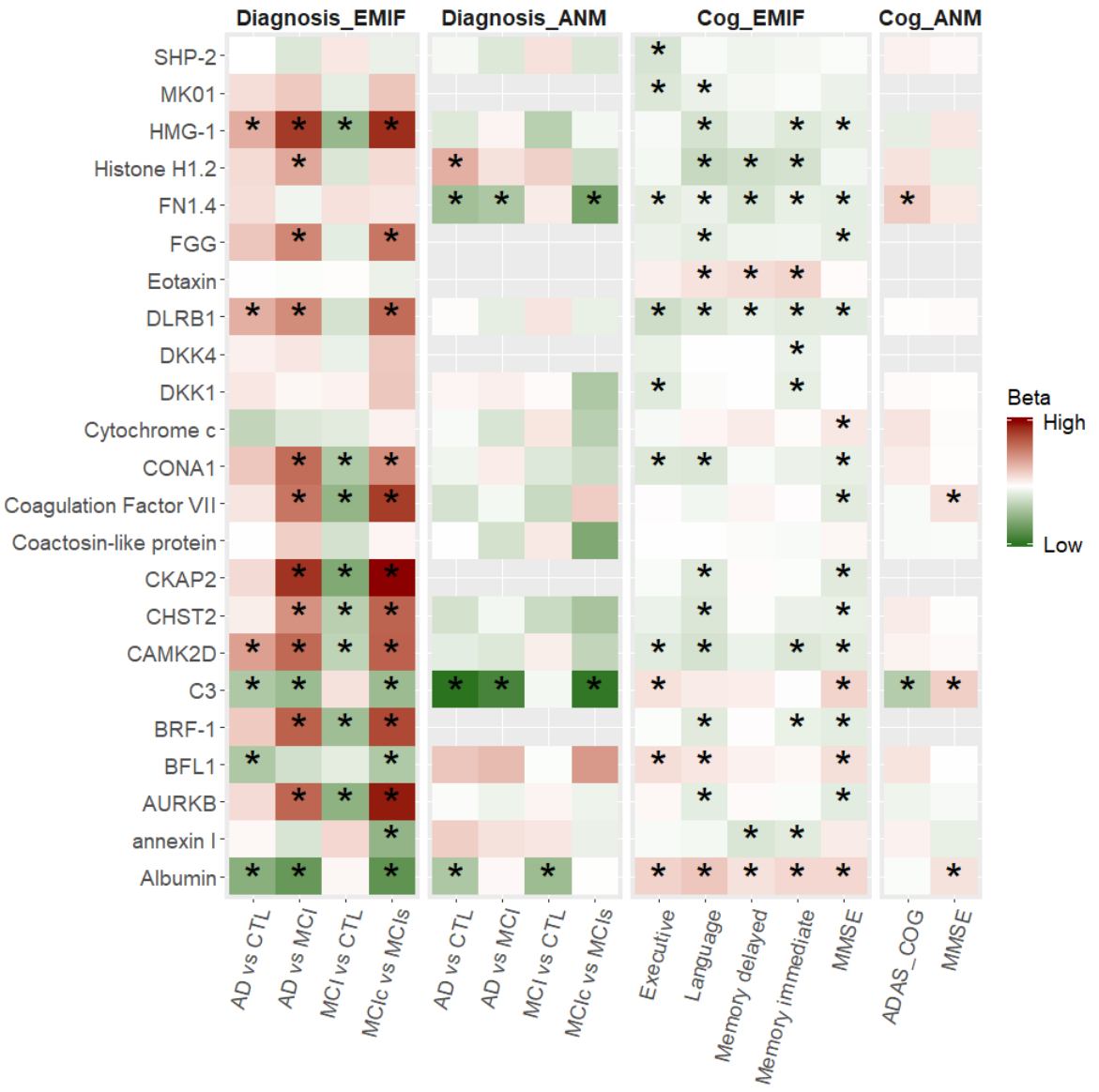
KEGG pathway	*P* value (uncorrected)	FDR corrected *p* value
Small cell lung cancer (P24941, P46527, P99999, P02751, P35354, Q05397)	0.011	0.075
Alzheimer's disease (P99999, P04406, Q99714, P28482, P27361, Q08209)	0.014	0.075
Viral carcinogenesis (P20226, P01024, P24941, P46527, P28482, P27361, P14618)	0.018	0.075
p53 signaling pathway (O96017, P24941, P99999, P05121, P07996)	0.025	0.075
Platelet activation (P02671, P02679, P28482, O15264, P27361, Q05513)	0.033	0.075
Glutamatergic synapse (P25098, P28482, P27361, Q08209)	0.034	0.075
Salmonella infection (P19876, P04141, P05231, P28482, O15264, P27361)	0.039	0.075
HIF-1 signaling pathway (P01033, Q13557, P46527, P04406, P05231, P28482, P27361, P05121)	0.040	0.075

821 **Supplementary Table 3** Enriched pathway of DKK1-induced multivariate signature from KEGG
822 pathway analysis. FDR, false discovery rate.

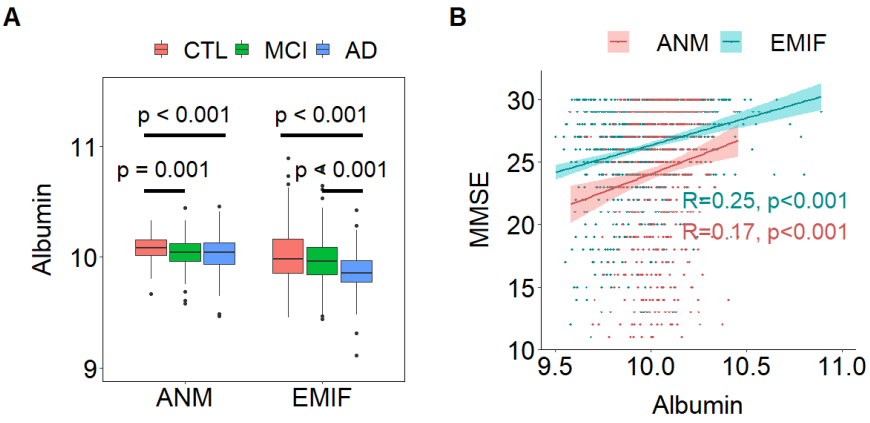
823

ATN profile	AUC of proteins, age and *APOE* ε4 (95% CI)	AUC of age (95% CI)	AUC of age and *APOE* ε4 (95% CI)
In all samples			
A-T-N- (n = 229) vs A+T-N- (n = 105)	0.72 (0.66, 0.78)	0.60 (0.52, 0.65)	0.70 (0.65, 0.78)
A-T-N- (n = 229) vs A+T+N- (n =19)	0.81 (0.74, 0.86)	0.78 (0.66, 0.85)	0.80 (0.71, 0.85)
A-T-N- (n = 229) vs A+T-N+ (n = 54)	0.88 (0.85, 0.92)	0.75 (0.67, 0.80)	0.84 (0.79, 0.88)
A-T-N- (n = 229) vs A+T+N+ (n = 298)	0.85 (0.83, 0.90)	0.70 (0.65, 0.74)	0.80 (0.76, 0.83)
In normal individuals			
A-T-N- (n = 47) vs A+T-N- (n = 16)	0.92 (0.89, 0.99)	0.77 (0.57, 0.84)	0.83 (0.67, 0.91)
A-T-N- (n = 47) vs A+T-N+ (n = 32)	0.94 (0.89, 0.99)	0.80 (0.69, 0.88)	0.91 (0.87, 0.98)
A-T-N- (n = 47) vs A+T+N+ (n = 114)	0.92 (0.86, 0.97)	0.75 (0.65, 0.81)	0.85 (0.78, 0.91)

Supplementary Table 4 AUC of using proteins, age and *APOE* ε4 to differentiate amyloid pathology (A+T-N-, A+T+N-, A+T-N+ and A+T+N+) from no AD pathology (A-T-N-) in all samples and cognitively normal individuals. CI, confidence interval.



827 **Supplementary Figure 1** Heat map displaying the association between 23 proteins and AD clinical
828 diagnosis as well as cognition in both EMIF and ANM cohort. *The association corrected $p < 0.05$. High
829 and low beta indicate positive and negative coefficients respectively. ANM, AddNeuroMed; EMIF,
830 European Medical Information Framework; Cog, cognition; AD, Alzheimer's disease; CTL, control; MCI,
831 mild cognitive impairment; MCIc, MCI conversion; MCIs, MCI stable; MMSE, Mini Mental State
832 Examination; ADAS-Cog, Alzheimer's Disease Assessment Scale - Cognitive subscale.



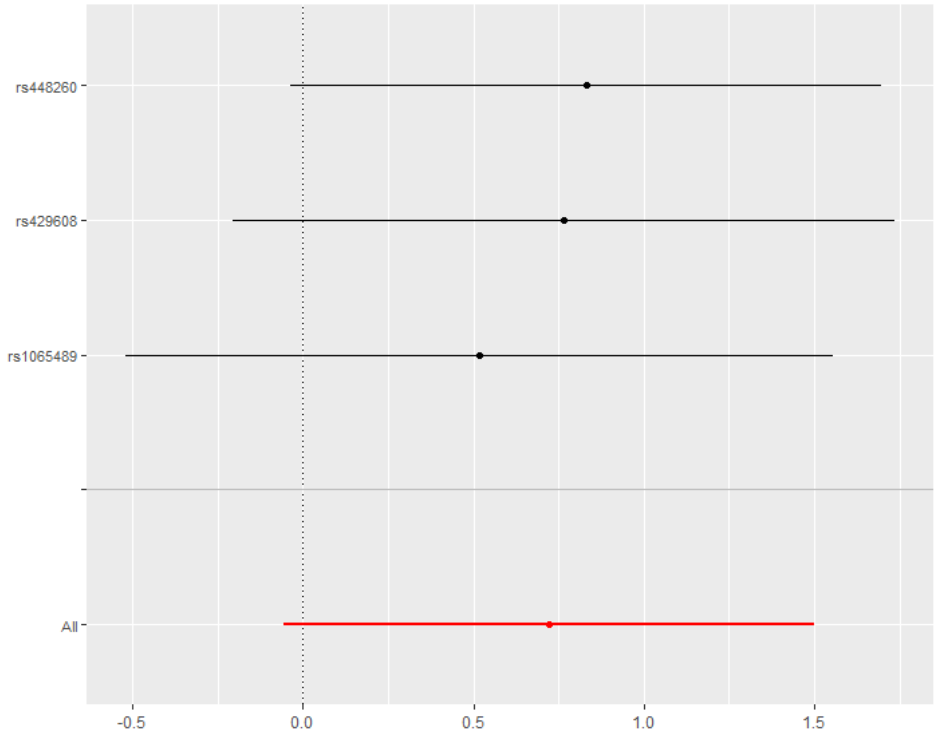
833 **Supplementary Figure 2** (A) Comparison of albumin in different AD diagnostic groups in both EMIF
834 and ANM cohorts; (B) Correlation of MMSE with albumin in both ANM and EMIF cohort. ANM,
835 AddNeuroMed; EMIF, European Medical Information Framework; AD, Alzheimer's disease; CTL,
836 control; MCI, mild cognitive impairment; MMSE, Mini Mental State Examination.

837

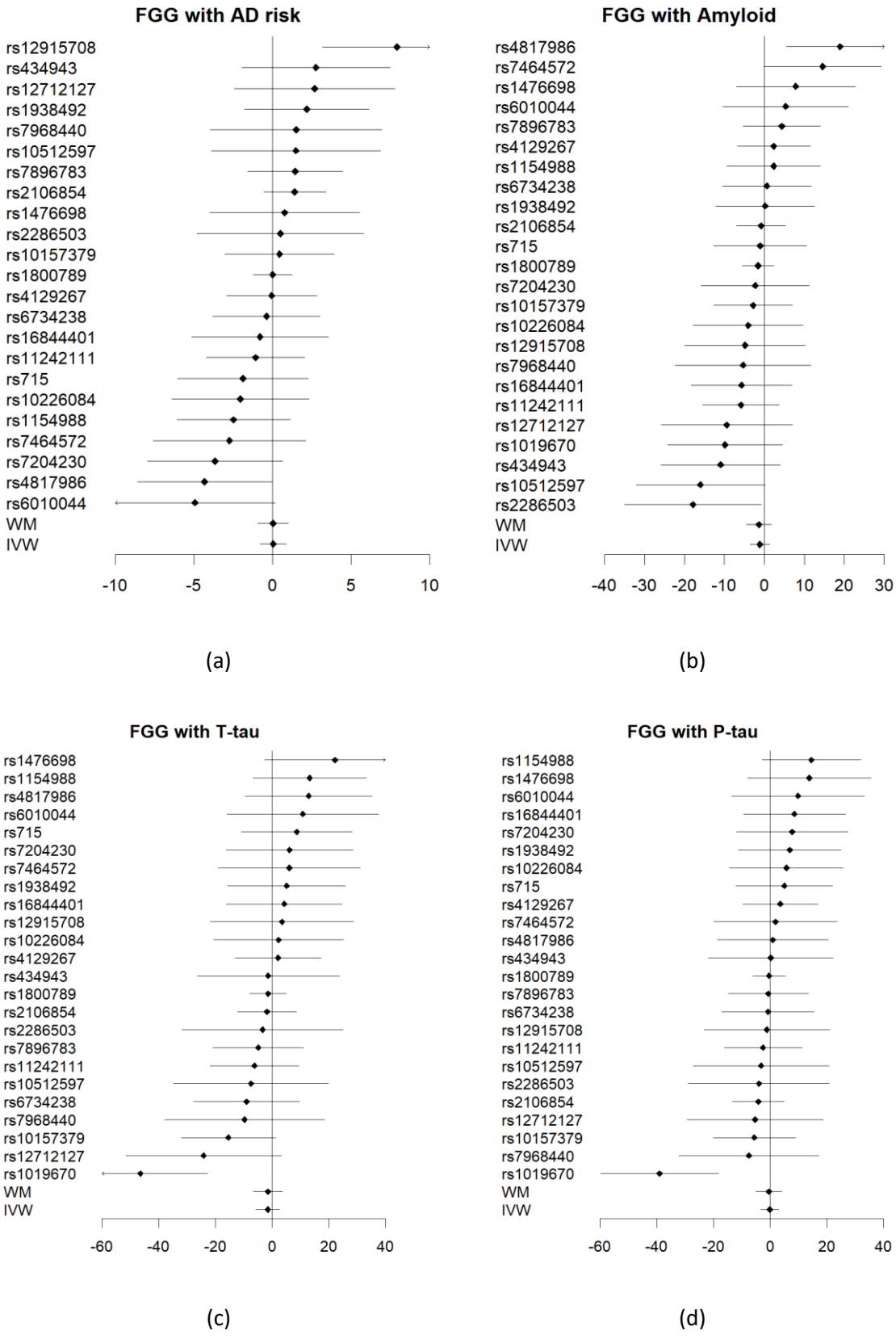
838

839

840



841 **Supplementary Figure 3** Leave-one-out Mendelian randomization estimates the association between
842 plasma complement C3 and AD risk by sequentially removing each single-nucleotide polymorphism
843 (SNP) from the analysis. No single SNP drove the majority of the association signal between C3 and AD
844 risk.



845 **Supplementary Figure 4** Forest plot of Mendelian randomization estimates the effects of (a) FGG on

846 AD risk, (b) FGG on amyloid, (c) FGG on T-tau and (d) FGG on P-tau. The estimated effect size of both

847 weighted median (WM) and inverse-variance weighted (IVW) method showing that no causal
848 relationship was found between FGG and AD risk, or FGG and amyloid or FGG and tau. FGG, fibrinogen
849 gamma chain; AD, Alzheimer's disease; T-tau, total tau; P-tau, phosphorylated tau.

850

851

852

853

854

855

856

857

858

859

860

861

862

863

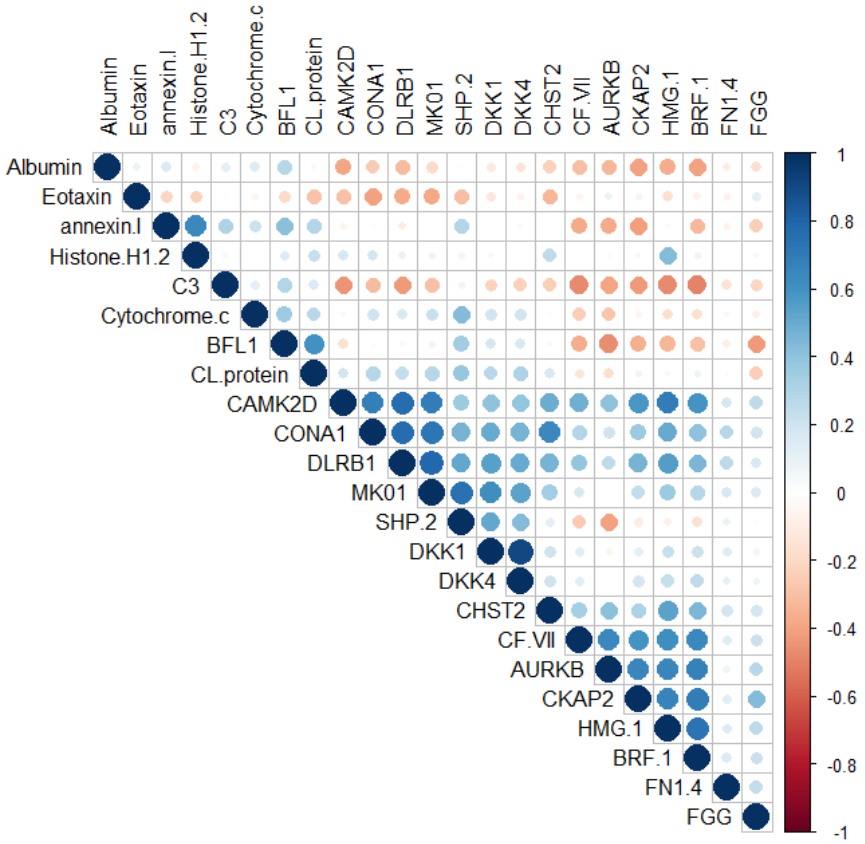
864

865

866

	Methods	β	se	95% CI	P value
FGG with AD risk	WM	0.04	0.50	-0.93 to 1.01	0.94
	IVW	0.04	0.42	-0.79 to 0.88	0.92
FGG with amyloid	WM	-1.35	1.59	-4.46 to 1.77	0.40
	IVW	-1.13	1.24	-3.56 to 1.30	0.36
FGG with T-tau	WM	-1.44	2.63	-6.58 to 3.71	0.58
	IVW	-1.50	2.10	-5.62 to 2.62	0.48
FGG with P-tau	WM	-0.44	2.33	-4.99 to 4.12	0.85
	IVW	-0.10	1.59	-3.22 to 3.01	0.95

867 **Supplementary Table 5** Mendelian randomization estimates of the causal effect of FGG on AD risk,
868 amyloid, T-tau and P-tau using both weighted median (WM) and inverse-variance weighted (IVW)
869 methods.β, beta coefficient; se, standard error of the effect size; CI, confidence interval; FGG,
870 fibrinogen gamma chain; AD, Alzheimer's disease; T-tau, total tau; P-tau, phosphorylated tau.
871



872 **Supplementary Figure 5** Correlation matrix of 23 proteins proteins which significantly altered in

873 participants with at least one abnormal ATN biomarker.