
Publishers page: http://dx.doi.org/10.3233/JAD-200208
<http://dx.doi.org/10.3233/JAD-200208>

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher’s version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.
Dickkopf-1 Overexpression in vitro Nominates Candidate Blood Biomarkers Relating to Alzheimer’s Disease Pathology

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

\textsuperscript{a}Department of Psychiatry, University of Oxford, UK
\textsuperscript{b}King’s College London, Institute of Psychiatry, Psychology and Neuroscience, Maurice Wohl Institute Clinical Neuroscience Institute, London, UK
\textsuperscript{c}Lübeck Interdisciplinary Platform for Genome Analytics, Institutes of Neurogenetics and Cardiogenetics, University of Lübeck, Lübeck, Germany
\textsuperscript{d}Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany
\textsuperscript{e}MRC Social, Genetic and Developmental Psychiatry Centre, King’s College London, UK
\textsuperscript{f}NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK
\textsuperscript{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
\textsuperscript{h}“Victor Babes” National Institute of Pathology, Bucharest, Romania
\textsuperscript{i}Department of Psychiatry and Neurochemistry, Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden
\textsuperscript{j}Wallenberg Centre for Molecular and Translational Medicine, University of Gothenburg, Gothenburg, Sweden
\textsuperscript{k}Dementia Research Institute Cardiff, Cardiff University, Cardiff, UK
\textsuperscript{l}Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Alzheimer Center Limburg, Maastricht University, Maastricht, the Netherlands
\textsuperscript{m}Alzheimer Center, VU University Medical Center, Amsterdam, the Netherlands

Correspondence to: Liu Shi, Department of Psychiatry, University of Oxford, UK. E-mail: liu.shi@psych.ox.ac.uk.

This article is published online with Open Access and distributed under the terms of the Creative Commons Attribution Non-Commercial License (CC BY-NC 4.0).
University Hospital Leuven, Leuven, Belgium
Laboratory for Cognitive Neurology, Department of Neurosciences, KU Leuven, Belgium
Center for Neurosciences, Vrije Universiteit Brussel (VUB), Brussels, Belgium
Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium
Institute Born-Bunge, University of Antwerp, Antwerp, Belgium
Neurodegenerative Brain Diseases Group, Center for Molecular Neurology, VIB, Belgium
Department of Neurology, UZ Brussels, Brussels, Belgium
University of Geneva, Geneva, Switzerland
IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy
AIX Marseille University, INS, Ap-hm, Marseille, France
Neurosciences Therapeutic Area, GlaxoSmithKline R&D, Stevenage, UK
University of Lille, Inserm, CHU Lille, France
Alzheimer’s disease & other cognitive disorders unit, Hospital Clinic, Barcelona, Spain
BarcelonaBeta Brain Research Center, Universitat Pompeu Fabra, Barcelona, Spain
Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK
1st Department of Neurology, AHEPA University Hospital, school of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Makedonia, Greece
Department of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
Department of Psychiatry, University Hospital of Lausanne, Lausanne, Switzerland
Geriatric Psychiatry, Department of Psychiatry, Geneva University Hospitals, and University of Geneva, Geneva, Switzerland
CITA-Alzheimer Foundation, San Sebastian, Spain
Organización Sanitaria Integrada Goierri – Alto Urola, Osakidetza, Spain
Danish Dementia Research Centre, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
Neurochemistry Laboratory, dept of Clinical Chemistry, Amsterdam Neuroscience, Amsterdam University Medical Centers, Vrije Universiteit, Amsterdam, the Netherlands
School of Medical Sciences, Örebro University, Örebro, Sweden
Department of Neurobiology, Care Sciences and Society, Division of Clinical Geriatrics, Karolinska Institute, Stockholm, Sweden
Department of Old Age Psychiatry, Psychology and Neuroscience, King’s College London, UK
Department of Psychiatry, Örebro Universitetssjukhus, Örebro, Sweden
Department of Geriatric Psychiatry, Zentralinstitut für Seelische Gesundheit, University of Heidelberg, Mannheim, Germany
Kings College London, London, UK
The Systems Medicine Group, Steno Diabetes Center Copenhagen, Gentofte, Denmark
Department of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam, The Netherlands
UCL Institutes of Neurology and Healthcare Engineering, London, UK
Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden
Department of Clinical Biochemistry, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
The Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark
Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark
Neurology / Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland
Toulouse Gerontopole University Hospital, Université Paul Sabatier, INSERM U 558, France
Medical University of Lodz, Lodz, Poland
Section of Gerontology and Geriatrics, Department of Medicine, University of Perugia, Perugia, Italy
UK Dementia Research Institute at UCL, London, United Kingdom
Department of Neurodegenerative Disease, UCL Institute of Neurology, London, United Kingdom
UCB, Braine-l’Alleud, Belgium, formerly Janssen R&D, LLC. Beerse, Belgium at the time of study conduct.
Abstract.

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

Objective: We aim to test whether DKK1 induced protein signature obtained in vitro were associated with markers of AD pathology as used in the amyloid/tau/neurodegeneration (ATN) framework as well as with clinical outcomes.

Methods: We first overexpressed DKK1 in HEK293A cells and quantified 1,128 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).

Results: We identified a 100-protein signature induced by DKK1 in vitro. 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 of 0.72, 0.81, 0.88, and 0.85, respectively. Furthermore, we found that some signature proteins (e.g., Complement C3 and albumin) were associated with cognitive score and AD diagnosis in both cohorts.

Conclusion: Our results add further evidence for a role of DKK regulation of Wnt signaling 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, replication, SomaScan, Wnt signaling

INTRODUCTION

Alzheimer’s disease (AD) is characterized by the presence of amyloid-β (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 realizable.

While 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 signaling. Activation of Wnt signaling 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 signaling 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 signaling, 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 signaling 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 impairments [18]. Given this, we propose that the DKK1 mediated inhibition of Wnt is a critical factor in the Aβ/DKK1 pathway also protected neurons from Aβ-induced toxicity and that in mice, overexpression of DKK1 induced tau pathology and cognitive impairments [18].

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 1,128 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 centers 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 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 Euro-
Fig. 1. Flowchart of study design. 1) Measurement and quantification of 1,128 proteins in total cell lysates of HEK293A cells overexpressing DKK1; 2) The top 100 proteins that constitute the DKK1-induced signature were identified using partial least squares (PLS) regression; 3) Measuring the association of 100 DKK1-induced proteins with ATN framework in vivo and obtaining 23 proteins that were significantly associated with any single ATN abnormal; 4) Measuring the association of AD-related proteins with other AD clinical outcomes; 5) Mendelian randomization to explore the causal relationship between two proteins (complement component 3 [C3] and fibrinogen gamma chain [FGG]) and AD risk, amyloid and tau (both T-tau and P-tau) pathology. AD, Alzheimer’s disease; SNAP, Suspected Non-Alzheimer Pathology; T-tau, total tau; P-tau, phosphorylated tau.

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 e4 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

<table>
<thead>
<tr>
<th>Variable</th>
<th>NPC</th>
<th>AD</th>
<th>SNAP</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (total = 785)</td>
<td>229</td>
<td>105</td>
<td>19</td>
<td>54</td>
</tr>
<tr>
<td>Age (y) (SD)</td>
<td>64 (8.5)</td>
<td>68 (9.7)</td>
<td>72 (6.8)</td>
<td>72 (7.3)</td>
</tr>
<tr>
<td>Male sex N (%)</td>
<td>123 (54)</td>
<td>57 (54)</td>
<td>9 (47)</td>
<td>30 (56)</td>
</tr>
<tr>
<td>APOE e4 carriers N (%)</td>
<td>52 (23)</td>
<td>58 (55)</td>
<td>13 (68)</td>
<td>36 (67)</td>
</tr>
<tr>
<td>MMSE (SD)</td>
<td>27.7 (2.5)</td>
<td>26.5 (3.9)</td>
<td>25.4 (3.9)</td>
<td>24.4 (4.3)</td>
</tr>
</tbody>
</table>

We have a 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 DESCRIPTA, 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 analyzed 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 1,016 and 4,001 proteins measured in ANM and EMIF cohorts, respectively. The $in\ vitro$ experiments were conducted with a version of the assay having 1,128 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 1,128 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 1,128 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, $APOE\ e4$ 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 (MCIs) within 3 years relative to those whose MCI remained stable (MCIs). To analyze the association of proteins with continuous AD phenotypes (i.e., MMSE), we used partial correlation and adjusted for age, $APOE\ e4$ 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 $APOE\ e4$ 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 MR 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 e4 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 analyzed 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 analyzed 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 \(p = 0.7\)) (Fig. 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...
Fig. 2. DKK1 overexpression leads to higher levels of (A) DKK1 and (B) DKK4 expression in HEK293A cell lysate ($n = 5$ per condition). Y axis represents the log transformed of proteins expression abundance measured by Somascan assay.

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 $APOE \varepsilon 4$ 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, $APOE \varepsilon 4$ 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 (Fig. 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 three 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 (Fig. 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 (Fig. 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 (Fig. 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 $APOE \varepsilon 4$ genotype as input features. Results showed that a model containing HMG-1 as well as age and $APOE \varepsilon 4$ genotype best discriminate A+T-N- from NPC group with an area under the curve (AUC) of 0.72 (Fig. 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 $APOE \varepsilon 4$ genotype. For differentiating A+T–N+ from NPC group, a model containing 7 proteins (annexin I, Albumin, Cytochrome c, Eotaxin, DKK1,
Fig. 3. A) Association of 23 DKK1-induced signature with 7 ATN profiles compared to A–T–N–. B–D) Comparison of proteins between A–T–N– (n = 229) and amyloid-positive individuals including A + T–N– (n = 105), A + T + N– (n = 19), A + T–N+ (n = 54), and A + T + N+ (n = 289). E, F) AUC of using proteins along with age and APOE e4 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 Somason assay. *p < 0.05; **p < 0.001; NS., not significant; AUC, area under the curve; CTL, controls.
CONA1, and AURKB) with age and APOE e4 genotype was selected. The optimal model to differentiate A + T + N+ from NPC group included 7 features including C3, AURKB, Eotaxin, BRF-1, and albumin as well as age and APOE e4 genotype (Fig. 3E). The comparison between each optimal model and the combination of age and APOE e4 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 (Fig. 3F). The optimal model that separated the A + T + N+ from the NPC group included four features including CKAP2, C3, age, and APOE e4 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 with an AUC of 0.92 and a model consisting of three 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 (Fig. 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 (Fig. 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 (Fig. 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 (Fig. 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 $\beta = 0.75$, standard error of the effect size $se = 0.44$, 95% CI $[-0.14, 1.63]$, $p = 0.09$) and IVW ($\beta = 0.72$, se $= 0.40$, 95% CI $[-0.05, 1.50]$, $p = 0.06$) methods (Fig. 5). In contrast, such a relationship was not found between C3 with either Aβ or tau (both T-tau and P-tau) status (Fig. 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 tau status from the EMIF cohort.

**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 utilize a novel approach with a hybrid design where first we used agnostic high dimensionality proteomics in an in vitro model of a hypothesized driver of disease mechanism and then used the derived signature in a targeted study in human samples. As Wnt signaling 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 overexpression 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 1,000 proteins, we determined a 100-protein signature induced by DKK1 and found that this protein set was enriched in molecular pathways known to be associated with AD, adding further evidence to the relevance and possible importance of this mechanism or pathway in disease.

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

Based on this, we further investigated the predictive value of DKK1-induced proteins in discriminating amyloid pathology (A+T–N–, A+T+N−, A+T–N+, and A+T+N+) from no AD pathology in all subjects as well as in cognitively normal individuals. We found the combination of different subsets of proteins along with age and APOE e4 genotype was able to differentiate the different ATN profiles with a high AUC, especially in normal individuals (AUC > 0.9). It should be noted that forward stepwise regression makes an arbitrary decision as to select highly correlated proteins. Therefore, other proteins that were highly correlated to those selected proteins
Table 1. Results of Mendelian randomization estimates for associations between genetic variants and AD risk, Aβ, and tau status (T-tau and P-tau).

<table>
<thead>
<tr>
<th>No./Total No.</th>
<th>Study</th>
<th>Beta (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1065489 17008/54162</td>
<td>Lambert et al. [30]</td>
<td>0.98 (-0.19-2.15)</td>
<td>0.10</td>
</tr>
<tr>
<td>rs429608 42034/314278</td>
<td>Marioni et al. [31]</td>
<td>0.65 (-0.64-1.93)</td>
<td>0.33</td>
</tr>
<tr>
<td>rs448260 17008/54162</td>
<td>Lambert et al. [30]</td>
<td>0.28 (-1.47-2.03)</td>
<td>0.76</td>
</tr>
<tr>
<td>WM</td>
<td></td>
<td>0.75 (-0.14-1.63)</td>
<td>0.09</td>
</tr>
<tr>
<td>IVW</td>
<td></td>
<td>0.72 (-0.05-1.50)</td>
<td>0.06</td>
</tr>
</tbody>
</table>

**Amyloid**
- rs1065489 405/656 EMIF: -0.91 (-4.53-2.72) 0.62
- rs429608 405/656 EMIF: 0.70 (-4.67-6.07) 0.79
- rs448260* 405/656 EMIF: -2.02 (-7.43-3.39) 0.46
- WM EMIF: -0.89 (-3.78-1.99) 0.54
- IVW EMIF: -0.78 (-3.41-1.84) 0.60

**T-tau**
- rs1065489 328/568 EMIF: 0.89 (-5.11-6.90) 0.77
- rs429608 328/568 EMIF: 3.74 (-5.28-12.76) 0.41
- rs448260* 328/568 EMIF: 1.20 (-7.70-10.10) 0.79
- WM EMIF: 1.08 (-3.72-5.88) 0.66
- IVW EMIF: 1.63 (-2.73-5.99) 0.46

**P-tau**
- rs1065489 304/570 EMIF: 2.75 (-2.51-8.01) 0.31
- rs429608 304/570 EMIF: 4.86 (-2.98-12.70) 0.22
- rs448260* 304/570 EMIF: -1.83 (-9.62-5.97) 0.65
- WM EMIF: 2.73 (-1.72-7.18) 0.23
- IVW EMIF: 2.15 (-1.66-5.96) 0.27

Fig. 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 (r² = 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.

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 signaling 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 signaling 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 overexpression, 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.

ACKNOWLEDGMENTS

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. MC_PC_17215). Also support was received from the NIHR Biomedical Research Centre at Oxford Health NHS Foundation Trust.

Authors’ disclosures available online (https://www.j-alz.com/manuscript-disclosures/20-0208r3).

DATA AVAILABILITY

The data sets generated and analyzed 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.

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: https://dx.doi.org/10.3233/JAD-200208.

REFERENCES


