Journal of Alzheimer's Disease 38 (2014) 515–531
DOI 10.3233/JAD-130380
IOS Press



# Candidate Blood Proteome Markers of Alzheimer's Disease Onset and Progression: A Systematic Review and Replication Study

Steven J. Kiddle[a,b,1,*], Martina Sattlecker[a,b,1], Petroula Proitsi[a], Andrew Simmons[a,b], Eric Westman[c], Chantal Bazenet[a], Sally K. Nelson[d], Stephen Williams[d], Angela Hodges[a], Caroline Johnston[a,b], Hilkka Soininen[e], Iwona Kłoszewska[f], Patrizia Mecocci[g], Magda Tsolaki[h], Bruno Vellas[i], Stephen Newhouse[a,b], Simon Lovestone[a,b] and Richard J.B. Dobson[a,b,*]

[a]*King's College London, Institute of Psychiatry, London, UK*
[b]*NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK*
[c]*Department of Neurobiology, Karolinska Institutet, Care Sciences and Society, Stockholm, Sweden*
[d]*SomaLogic, Boulder, Colorado, USA*
[e]*Department of Neurology, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland*
[f]*Medical University of Lodz, Lodz, Poland*
[g]*Institute of Gerontology and Geriatrics, University of Perugia, Perugia, Italy*
[h]*3rd Department of Neurology, Aristotle University, Thessaloniki, Greece*
[i]*INSERM University of Toulouse, Toulouse, France*

Handling Associate Editor: Gary Arendash

Accepted 22 July 2013

**Abstract**. A blood-based protein biomarker, or set of protein biomarkers, that could predict onset and progression of Alzheimer's disease (AD) would have great utility; potentially clinically, but also for clinical trials and especially in the selection of subjects for preventative trials. We reviewed a comprehensive list of 21 published discovery or panel-based (>100 proteins) blood proteomics studies of AD, which had identified a total of 163 candidate biomarkers. Few putative blood-based protein biomarkers replicate in independent studies but we found that some proteins do appear in multiple studies; for example, four candidate biomarkers are found to associate with AD-related phenotypes in five independent research cohorts in these 21 studies: α-1-antitrypsin, α-2-macroglobulin, apolipoprotein E, and complement C3. Using SomaLogic's SOMAscan proteomics technology, we were able to conduct a large-scale replication study for 94 of the 163 candidate biomarkers from these 21 published studies in plasma samples from 677 subjects from the AddNeuroMed (ANM) and the Alzheimer's Research UK/Maudsley BRC Dementia Case Registry at King's Health Partners (ARUK/DCR) research cohorts. Nine of the 94 previously reported candidates were found to associate with AD-related phenotypes (False Discovery Rate (FDR) q-value $<0.1$). These proteins show sufficient replication to be considered for further investigation as a biomarker set. Overall, we show that there are some signs of a replicable signal in the range of proteins identified in previous studies and we are able to further replicate some of these. This suggests that AD pathology does affect the blood proteome with some consistency.

Keywords: Alzheimer's disease, biomarkers, blood, magnetic resonance imaging, nucleotide aptamers, proteome

[1]These authors contributed equally to this manuscript.

*Correspondence to: Steven J. Kiddle or Richard Dobson, Box P092, De Crespigny Park, London, SE5 8AF, UK. Tel.: +44 20 7848 0924; Fax: +44 20 7848 0866; E-mails: steven.kiddle@kcl.ac.uk or richard.j.dobson@kcl.ac.uk.

ISSN 1387-2877/14/$27.50 © 2014 – IOS Press and the authors. All rights reserved
This article is published online with Open Access and distributed under the terms of the Creative Commons Attribution Non-Commercial License.

## INTRODUCTION

Alzheimer's disease (AD) is a common and progressive neurodegenerative condition causing considerable burden to both individuals and health economies. No disease modification therapies are available yet and diagnosis can only be made definitively at postmortem. Protein biomarkers such as amyloid-β and tau in cerebrospinal fluid (CSF) can contribute to the diagnosis of AD [1] and in particular may aid early diagnosis and help predict conversion of people from pre-dementia states such as mild cognitive impairment (MCI) to AD [2]. However, lumbar puncture to extract CSF is a relatively invasive procedure. The level of amyloid-β in the brain can also be measured by positron emission tomography, although facilities for such imaging are currently restricted.

The motivation for blood-based biomarkers of AD is the desire for a relatively non-invasive marker of AD pathology that could be easily measured in community settings [3]. Despite being a relatively non-invasive and useful biological matrix for diagnostic biomarkers, peripheral blood is a complex tissue containing proteins originating from many organs, thus making it non-specific to markers of brain disorders. For example, plasma levels of amyloid-β do not appear to show clinical utility in AD [4] despite showing great utility in CSF. However, preliminary results from discovery studies suggest that useful signals of AD such as proteins and other relevant biological markers may exist in blood [3]. Proteins represent one source of blood-based biomarkers of AD; other relevant sources and relevant assays are reviewed in Bazenet et al. [3].

Numerous discovery studies for blood-based protein markers of AD have been underway for at least a decade, ranging from proteomics technologies using both gel and non-gel based mass spectrometry (MS) (e.g., [5, 6]) to antibody capture arrays (e.g., [7, 8]). The failure to replicate discoveries has been the fundamental issue surrounding the development of a useful diagnostic panel thus far. The problems of non-replicability may be the result of non-homogeneity of proteomic platforms or research cohorts, over-fitting of data, or technical issues and non-standardized sample collection protocols. In order to explore some degree of concordance among the various discovery efforts, we performed a systematic review of blood-based proteomics AD studies and then applied a modified aptamer-based array to a large sample of subjects in a substantial replication experiment.

## METHODS

### *Candidate blood-based protein markers*

#### *Literature search for candidate markers of AD-related phenotypes*

We included studies of plasma, serum, and leukocyte proteins in our search for general blood protein markers of AD-related phenotypes. In order to be included as a candidate marker, a protein had to have been identified in a discovery, rather than a candidate, based study. Exceptions were made for panel-based studies that included over 100 candidate proteins, as these were considered to be broad enough to be unbiased. We first identified studies fitting these criteria from two recent reviews of blood protein markers of AD by Lista et al. [9] and Zurbig and Jahn [10]. We then added studies which have used the Myriad Rules Based Medicine (RBM) Human Discovery Multi-Analyte Profile (MAP), several of which were published after the reviews mentioned above. The PubMed search term 'Alzheimer blood protein discovery' was used to identify additional studies.

#### *Gene ontology and pathway over-representation analysis*

Gene Ontology (GO, [11]) and Kyoto Encyclopedia of Genes and Genomes (KEGG, [12]) functional annotation over-representation analysis was performed on protein lists identified from the literature review using the Database for Annotation, Visualization and Integrated Discovery (DAVID, http://david.abcc.ncifcrf.gov/, [13, 14]) and the Biological Networks Gene Ontology tool (BINGO, [15]). A background/reference set of the UniProt identifiers of the candidate biomarkers from the literature review detailed above, and the UniProt identifiers from the plasma proteome database (http://www.plasmaproteomedatabase.org/, [16]) was used for the DAVID and BINGO analyses. DAVID analyses with default settings were conducted separately for: GO molecular function, GO cellular component, GO biological process, and KEGG functional annotation terms. The BINGO tool was used with a generic GOslim ontology (a less detailed set of GO annotations) within the cytoscape [17] platform so that results could be visualized.

### *Replication of candidates using plasma samples from the combined AddNeuroMed and ARUK/Maudsley BRC dementia case registry at king's health partners cohorts*

#### *Subjects*

We investigated literature-derived AD related blood proteins in a total of 677 subjects. Of these subjects, 412 (109 controls, 109 MCI [43 of which converted to AD within a year of measurement], 194 AD) were recruited from the EU funded AddNeuroMed (ANM) biomarker project study [18, 19]; 232 subjects (100 controls, 40 MCI [none of which converted within a year of measurement], 92 AD) were recruited from the Alzheimer's Research UK [5]; and 33 AD subjects from the Maudsley Biomedical Research Center (BRC) Dementia Case Registry at King's Health Partners (DCR). It should be noted that these MCI-AD conversion rates are not reflective of the general conversion rates in these cohorts, due to sample depletion by previous studies. The conversion rates are also not reflective of the population level conversion rates; for example, ANM recruited from memory clinics, which will be enriched with MCI subjects with a likely AD-endpoint. The same diagnostic criteria were used in all of these cohorts. Some of these subjects have been used in previous biomarker studies [5, 6, 20, 21]. The relevant ethics board approved the study and informed consent was obtained for all subjects. All subjects were assessed with a standardized assessment protocol including informant interview for diagnosis and the Mini Mental State Examination (MMSE); as detailed in the references above.

#### *Protein measures*

Proteins were measured using a Slow Off-rate Modified Aptamer (SOMAmer)-based capture array called 'SOMAscan' (SomaLogic, Inc, Boulder, Colorado). This approach uses chemically modified nucleotides to transform a protein signal to a nucleotide signal that can be quantified using relative florescence on microarrays. This assay has been shown to have a median intra- and inter-run coefficient of variation of ~5%. The median lower and upper limits of quantification were ~1 pM and ~1.5 nM in buffer, and ~2.95 pM and ~1.5 nM for a subset of the somamers in plasma (full details are given in Gold et al. [22]).

Quality control is performed at the sample and SOMAmer level, and involves the use of control SOMAmers on the microarray and calibration samples.

At the sample level, hybridization controls on the microarray are used to monitor sample-by-sample variability in hybridization, while the median signal over all SOMAmers is used to monitor overall technical variability. The resulting hybridization scale factor and median scale factor are used to normalize data across samples. The acceptance criteria for these values are 0.4–2.5, based on historical trends in these values. Somamer-by-somamer calibration occurs through the repeated measurement of calibration samples, these samples are of the same matrix as the study samples, and are used to monitor repeatability and batch to batch variability. Historical values for these calibrator samples for each SOMAmer are used to generate a calibration scale factor. The acceptance criteria for calibrator scale factors is that 95% of SOMAmers must have a calibration scale factor within ±0.4 of the median.

The assay required 8 μL of plasma from each sample. A single assay was used per plasma sample, i.e., no technical replicates were performed. Additionally, the samples were run in two batches ensuring an even mix of diagnosis groups in each batch. Seven outliers, identified using principal component analysis, were removed from the downstream analysis. Principal component analysis also showed that protein measures were affected by study center and thus we either added center as a covariate or adjusted for center using linear regression in all downstream analysis. All remaining samples were log2 transformed.

The assay measures the level of 1,001 human proteins representing a range of biomedically relevant molecular pathways and gene families. For each of the 94 literature-derived AD candidate protein markers targeted by the SOMAscan assay, a single measure was selected where possible. As such, for 88 candidate proteins, a SOMAscan probe was chosen that targets that protein alone. For two candidates, the probes that targeted them also targeted an additional protein: the probe targeting Complement C4-A also targeted Complement C4-B and the probe targeting Interleukin-12 subunit β also targeted Interleukin-12 subunit α. Another four probes were chosen as they targeted multiple candidate proteins not targeted by other probes: complement C4A/B, creatine kinase m- and k-type, complement C8 α/β/γ, and fibrinogen α/β/γ chain. The candidate proteins represented by the 94 SomaLogic measures are indicated in Supplementary Table 1.

Table 1
Summary of 21 literature studies of blood-based protein biomarkers of AD, limited to discovery or panel-based (>100 proteins assayed) studies. CTL, control

| Cohort | Study | Sample size | Tissue | Proteomic approach | Outcome variable(s) |
|---|---|---|---|---|---|
| ANM | Thambisetty et al. [21] | 91 | Plasma | 2D-Gel Electrophoresis (GE) Liquid Chromatography (LC)/MS/MS | Hippocampal atrophy and rapid clinical progression |
| ADNI | Soares et al. [39] | 566 | Plasma | Myriad RBM - Luminex xMAP | (AD versus CTL) and (MCI versus CTL) |
| ADNI | Kiddle et al. [42] | 71 | Plasma | Myriad RBM - Luminex xMAP | Brain amyloid burden |
| AIBL | Doecke et al. [40] | 961 | Plasma | Myriad RBM - Luminex xMAP | AD versus CTL |
| Akuffo | Akuffo et al. [34] | 193 | Plasma | 2D-GE LC/MS/MS | Markers of AD drug efficacy |
| ARUK | Hye et al. [5] | 100 | Plasma | 2D-GE LC/MS/MS | AD versus CTL |
| ARUK | Thambisetty et al. [20] | 26 | Plasma | 2D-GE | AD versus MCI |
| ARUK | Guntert et al. [6] | 45 | Plasma | Tandem Mass Tag labeled LC/MS/MS | Predicting CTL, slow decliners and fast decliners |
| BLSA | Thambisetty et al. [43] | 57 | Plasma | 2D-GE | Brain amyloid burden |
| Choi | Choi et al. [29] | 18 | Plasma | 1D-GE Matrix-assisted laser desorption/ionization (MALDI) – Time of Flight (TOF)/MS of oxidized proteins | AD versus CTL |
| genADA | Cutler et al. [35] | 94 | Plasma | 2D-GE LC/MS/MS | AD versus CTL |
| Henkel | Henkel et al. [38] | 14 | Plasma | Immunodepletion and Difference GE | AD versus CTL |
| Hu | Hu et al. [41] | 230 | Plasma | Myriad RBM - Luminex xMAP | Associated with mild dementia/MCI/AD |
| Liao | Liao et al. [33] | 20 | Plasma | 2D-GE MS | AD versus CTL |
| Liu | Liu et al. [32] | 133 | Serum | 2D-GE MALDI-TOF MS | AD versus CTL |
| Mhyre | Mhyre et al. [36] | 15 | Plasma | 2D-GE MALDI-TOF MS | Markers of AD drug efficacy |
| Ray | Ray et al. [7] | 259 | Leukocytes | Arrayed sandwich Enzyme Linked Immunoassays (ELISAs) | AD versus CTL |
| Rotterdam Scan Study | Ijesselstijn et al. [37] | 86 | Serum | Nano LC Orbitrap MS | Pre-symptomatic AD versus CTL |
| TARC | O'Bryant et al. [8] | 400 | Serum | Myriad RBM - Luminex xMAP | AD versus CTL |
| Yu | Yu et al. [30] | 19 | Plasma | 1D and 2D-GE MALDI-TOF MS | AD versus CTL |
| Zhang | Zhang et al. [31] | 71 | Serum | 2D-GE MALDI-TOF MS and ELISA | AD versus CTL |

*Structural magnetic resonance imaging*

Volumes of the whole brain, hippocampi, and entorhinal cortices were obtained using FreeSurfer 5.1.0 from 274 ANM subjects who had undergone structural magnetic resonance imaging (sMRI). These regions were selected as they are known to be related to early AD pathology, and were normalized by intracranial volume (ICV) [23]. Detailed information about data acquisition, pre-processing, and quality control assessment have been described for this cohort in detail elsewhere [19, 24–27].

*Rate of cognitive decline*

The rate of cognitive decline in 329 AD patients (214 ANM, 87 ARUK, and 28 DCR) was calculated based on longitudinal MMSE assessments. For the ANM cohort, MMSE scores were gathered at five visits, where visits were three months apart. For ARUK and DCR, MMSE scores were obtained annually. To estimate the rate of cognitive decline, only samples with at least three MMSE measures were included. Linear mixed effect models were generated using the package 'nlme' in R. This was done separately for ANM, and for DCR and ARUK together, due to the differences in assessment windows between the cohorts. Samples and center were added as random effect to the model. Further covariates, including age of onset, disease duration at baseline, gender, $APOE$ ε4 allele presence, living in a nursing home, and years of education were investigated for their effect on the rate of decline. We found that age of onset, living in a nursing home, and education had a significant effect on the rate of decline ($p < 0.05$) and thus were included as fixed effects in the final model. The slope coefficient obtained from the final model was then used as the rate of cognitive decline, defined as the change in MMSE per day. We had plasma protein measures for 239 (173 ANM, 44 ARUK, and 22 DCR) of the 329 AD patients with rate of cognitive decline slopes.

*Statistical analysis*

All proteins were analyzed for their association with presence of at least one $APOE$ ε4 allele using

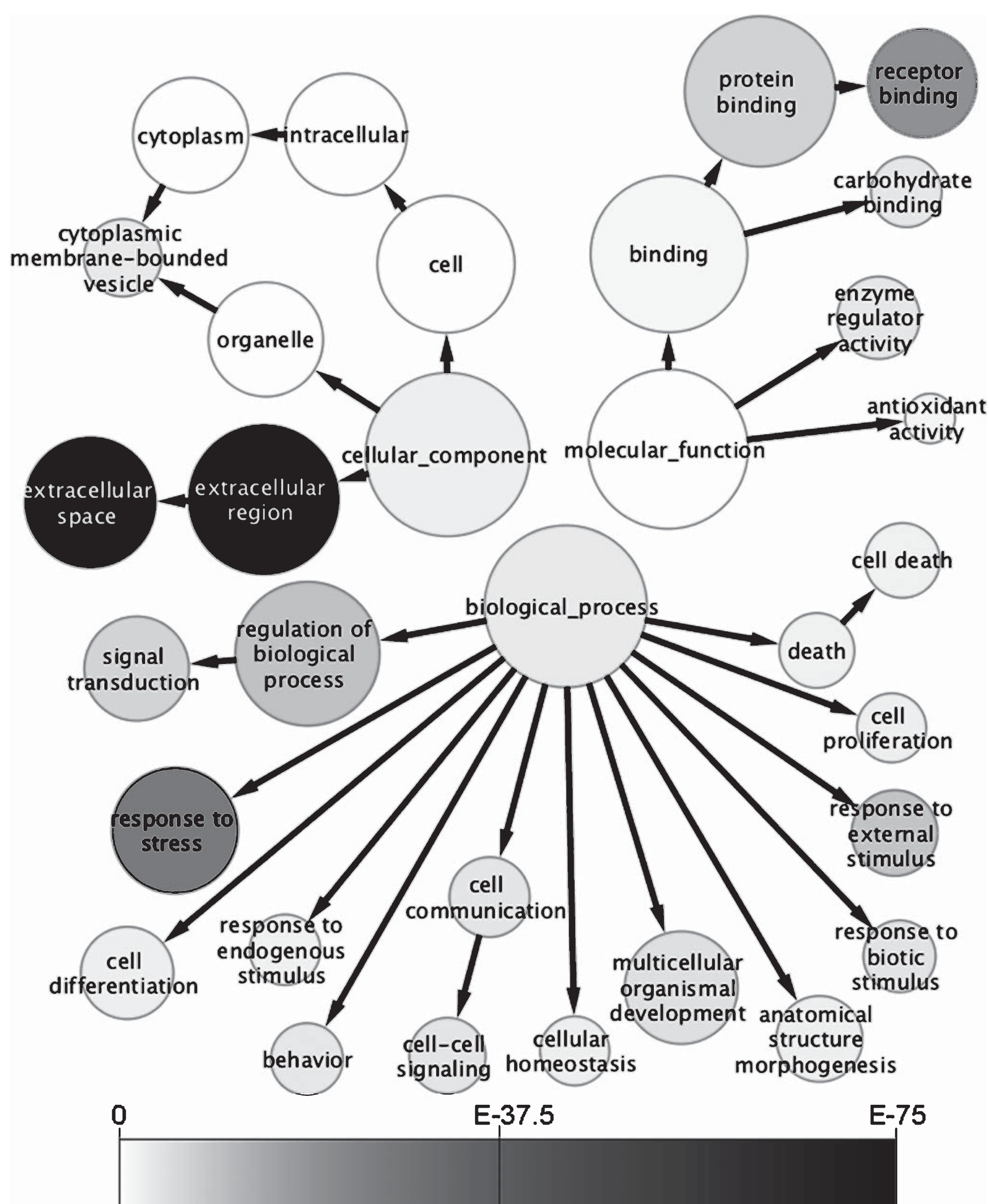


Fig. 1. BINGO analysis of over-represented GOslim terms in list of 163 candidate plasma protein biomarkers of AD in comparison to all proteins known to be present in plasma. GO terms are shown as nodes, with term hierarchies shown as arrows between nodes/GO terms. The *p*-value of the over-representation of each GO term is illustrated by the color of each node/GO term.

the Mann-Whitney-Wilcoxon test using the function Wilcox.test in R. All proteins were also analyzed individually for their association with AD phenotypes: disease status (AD versus CTL, MCI stable versus MCI converter), sMRI imaging measures (whole brain volume, left and right entorhinal cortex/hippocampal volume), and rate of cognitive decline (MMSE). Since disease status is a categorical predictor, logistic regression was employed for each protein by adjusting for age at sample acquisition, gender, $APOE$ ε4 presence, and research center. Logistic regression models were developed for AD versus CTL and MCI stable versus MCI converter subjects separately.

The correlation between sMRI imaging measures and protein measures was investigated using partial Spearman's Rank Correlation (SRC) using the function pcor.test from the R bioconductor package 'ppcor'. Subject age at sampling, gender, research center, and

number of APOE ε4 alleles were taken into account as co-variates in this analysis. Correlation of rate of cognitive decline with protein measures was analyzed using SRC. Covariates were not used for this correlation since they were already taken account of when calculating the rate of cognitive decline.

False discovery rate (FDR) multiple testing corrections were applied to the resulting *p*-values; with both a strict (FDR *q*-value of 0.05) and less strict significance level used (FDR *q*-value of 0.1). This was used to allow both strong and promising results to be detected respectively.

*Classification*

Control and AD subjects were allocated at random to either the training or test set with a roughly 75%:25% split respectively. The training set therefore consisted of 395 subjects (156 controls and 239 AD), whereas the test set consisted of 133 subjects (53 controls and 80 AD). For classification, the protein data was adjusted for research center by linear regression. Over-sampling was performed at random for control subjects to balance the training set. All classification was performed in R using random forest classification using the package 'randomForest'. Additionally, the R package 'ROCR' was used to generate Receiver Operator Characteristic (ROC) curves and to measure the Area Under the Curve (AUC) [28]. Four different sets of predictors were used: 1) co-variates only (age/gender/presence of APOE ε4), 2) co-variates plus the four literature-based candidate markers of AD-related phenotypes seen in five independent cohorts, 3) co-variates plus all the literature-based candidate markers, and 4) co-variates plus literature-based candidate markers significantly different between AD and control subjects in the training set at the Bonferroni corrected 0.05 significance level in Mann-Whitney-Wilcoxon tests.

## RESULTS

*Systematic review of candidate blood protein markers from discovery/panel-based studies*

Thirteen studies fitting the inclusion criteria (discovery or panel-based with >100 analytes) were identified from Lista et al. [9] [5, 7, 8, 20, 21, 29–36], and two studies were identified from Zurbig and Jahn et al. [10] [37, 38]. Four more recently published papers which used the RBM MAP panel were added [39–42]. Finally, we searched PubMed for additional studies, two of which fit our inclusion criteria [6, 43].

In total, we identified 21 discovery or panel-based (>100 analytes) blood proteomics studies (Table 1), comprising 163 separate candidate blood protein markers of AD-related phenotypes (Supplementary Table 1). These 21 studies used blood samples from a total of 18 independent cohorts. Of these 163 candidate blood-based biomarker proteins, ∼61% are in the high-confidence Human Plasma Proteome Project (HPPP) reference set [44] and ∼88% are in the less stringent but more comprehensive Plasma Proteome database [16]. The 21 identified studies included the following cohorts: AddNeuroMed (ANM; [18]), Alzheimer's Disease Neuroimaging Initiative (ADNI; [45]), Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing (AIBL; [46]), Alzheimer's Research UK (ARUK; previously known as Alzheimer's Research Trust, ART; [5]), Baltimore Longitudinal Study of Aging (BLSA; [47]), Canadian Genotype-Phenotype Alzheimer's Disease Associations (genADA; [48]), and Texas Alzheimer's Research Consortium (TARC; [49]).

*Gene ontology and pathway analysis of candidate markers*

In order to examine possible roles of these putative biomarkers in disease processes we used DAVID [13, 14] and BINGO [15] to examine over-represented functional annotations assigned to the 163 candidate markers. UniProt identifiers for the candidates and the reference set are given in Supplementary Table 2. Over-represented GO terms from the BINGO analysis are shown in Fig. 1. It can be seen that 'signal transduction' (GO:0007165; Bonferroni $p = 1.63 \times 10^{-8}$), 'response to stress' (GO:0006950; Bonferroni $p = 1.17 \times 10^{-31}$), 'receptor binding' (GO:0005102; Bonferroni $p = 5.70 \times 10^{-24}$), and 'extracellular space' (GO:0005615; Bonferroni $p = 7.07 \times 10^{-64}$) are particularly enriched. Over-representation analysis using the full GO ontology (Supplementary Table 3) and the KEGG ontology (Table 2) in DAVID revealed specific processes that are enriched. For example, 'activation of plasma proteins involved in the acute inflammatory response' (GO:0002541), a sub-term of 'response to stress', is shown to be over-represented (Bonferroni $p = 3.38 \times 10^{-17}$). Similarly, 'cytokine activity' (GO:0005125; Bonferroni $p = 2.44 \times 10^{-17}$) and 'chemokine activity' (GO:0008009; Bonferroni $p = 1.08 \times 10^{-5}$) are shown to be over-represented, both being sub-terms of 'receptor binding'. Cytokines are also over-represented in the KEGG analysis, along with 'complement and coagulation cascades' (Table 2). It can also be seen that KEGG terms relating

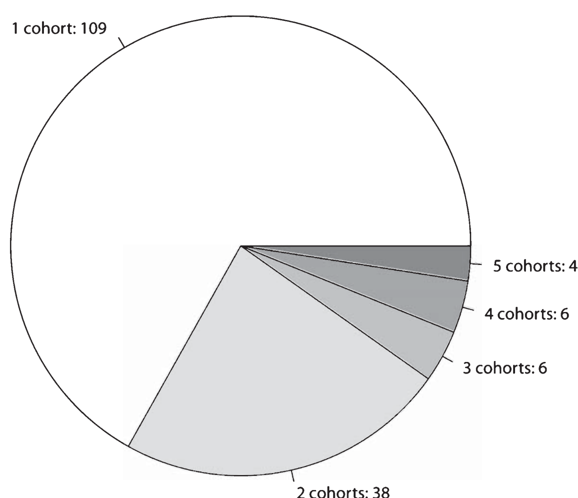


Fig. 2. Pie chart showing the extent to which each candidate biomarker has been replicated in independent research cohorts. The number of independent cohorts each candidate biomarker has been found in is shown. The cohort used by each study is shown in Table 1.

to other diseases are over-represented in the annotations of the 163 candidate proteins, i.e., systemic lupus erythematosus, prion diseases, type I diabetes mellitus, bladder cancer, and allograft rejection.

*Concordance of markers from literature derived studies*

Given that some plasma samples have been used in multiple studies (those using the same cohorts), concordance was assessed at the cohort level, i.e., the number of times each biomarker has been found to associate with an AD-related phenotype in an independent cohort. Different AD-related phenotypes were not distinguished in this analysis, as many studies measured the association of these markers to different AD-related phenotypes making exact comparisons problematic. Results are shown in Supplementary Table 1 and summarized in Fig. 2. It was found that 109 of the 163 candidate biomarkers (∼67%) were only found to associate with an AD-related phenotype in one cohort.

Four candidate biomarkers—apolipoprotein E [31, 39–43], α-2-macroglobulin [5, 8, 20, 31, 34, 37], complement C3 [21, 31, 37, 42, 43], and α-1-antitrypsin [29, 30, 33, 40, 42]—were found to associate with an AD-related phenotype in five independent research cohorts. Additionally, six candidate biomarkers—complement factor H [5, 20, 31, 34, 38], pancreatic prohormone [8, 39–42], plasma protease C1 inhibitor [34, 35, 37, 38], serum amyloid p-component [5, 21, 41, 42], fibrinogen γ chain [21, 29, 41, 42], and serum albumin [5, 21, 40, 43]—were found to associate with an AD-related phenotype in four independent research cohorts. Finally, six candidate biomarkers were found to associate with an AD-related phenotype in three independent research cohorts, and 38 candidate biomarkers with two cohorts, respectively (Supplementary Table 1).

Table 2
DAVID analysis of over-represented KEGG terms in list of candidate plasma protein biomarkers of AD in comparison to all proteins known to be present in plasma

| KEGG term | Count | % | $p$-value | Fold enrichment | Bonferroni corrected $p$-value |
|---|---|---|---|---|---|
| Complement and coagulation cascades | 27 | 17.6 | $9.78 \times 10^{-25}$ | 14.1 | $9.00 \times 10^{-23}$ |
| Cytokine-cytokine receptor interaction | 31 | 20.3 | $2.07 \times 10^{-18}$ | 6.8 | $1.91 \times 10^{-16}$ |
| Systemic lupus erythematosus | 14 | 9.2 | $1.77 \times 10^{-8}$ | 7.4 | $1.63 \times 10^{-6}$ |
| Prion diseases | 10 | 6.5 | $1.05 \times 10^{-7}$ | 11.1 | $9.66 \times 10^{-6}$ |
| Type I diabetes mellitus | 7 | 4.6 | $6.84 \times 10^{-5}$ | 9.3 | $6.27 \times 10^{-3}$ |
| Bladder cancer | 7 | 4.6 | $2.34 \times 10^{-4}$ | 7.5 | 0.0213 |
| Toll-like receptor signaling pathway | 9 | 5.9 | $2.52 \times 10^{-4}$ | 5.2 | 0.023 |
| Allograft rejection | 6 | 3.9 | $3.89 \times 10^{-4}$ | 9 | 0.0352 |

Table 3
Sample characteristics of the AddNeuroMed (ANM) and Alzheimer's Research UK (ARUK)/Maudsley BRC Dementia Case Registry (DCR) at King's Health Partners with available SomaLogic protein data. MCI-MCI, MCI at baseline and a year later; MCI-AD, MCI at baseline and converts to AD within a year

| | Combined ANM+ARUK+DCR cohort (number of subjects) | | | |
|---|---|---|---|---|
| Characteristics | Controls (209) | MCI-MCI (106) | MCI-AD (43) | AD (319) |
| Age (median [IQR]) | 76 [7] | 77 [10] | 76 [9] | 79 [10] |
| Gender (male/female) | 102/107 | 40/66 | 17/26 | 98/221 |
| Number of *APOE* ε4 alleles (0/1/2) | 153/51/5 | 73/29/4 | 17/23/3 | 139/136/44 |
| Baseline MMSE (median [IQR]) | 29.0 [1.00] | 27.0 [2.00] | 26.5 [3.00] | 20.0 [7.00] |
| Number of missing baseline MMSE | 0 | 4 | 3 | 24 |
| Number with sMRI scan | 95 | 62 | 19 | 98 |

### *Independent replication of candidate markers with an orthogonal technology*

To test the literature derived candidate blood-based biomarkers of AD, we examined their levels in plasma samples from the ANM, ARUK, and DCR research cohorts (combined characteristics in Table 3, stratified by cohort in Supplementary Table 5, and subset of subjects with sMRI data in Supplementary Table 5). We used a SOMAmer-based capture array, which has previously been used in studies of aging and cancer [22, 50–54]. SomaLogic SOMAscan aptamer-based assays were available for 94 of the total number of 163 candidate biomarkers (~58%). The proteins complement component C6, apolipoprotein E, complement C3, C-reactive protein, apolipoprotein B-100, and interleukin-11 were found to be affected by the presence of at least one *APOE* ε4 allele at the 0.05 FDR level (Supplementary Table 6).

### *Candidate AD marker association with AD diagnosis*

Logistic regression was used to detect differences in the level of these 94 candidate biomarkers between AD and control subjects in the combined ANM+ARUK+DCR cohort (full results are shown in Supplementary Table 7 and results significant at the 0.1 FDR level are shown in Table 4). The plasma level of one candidate biomarker, pancreatic prohormone (Odds Ratio (OR) 2.4, FDR q $1.6 \times 10^{-4}$), was found to differ at the 0.05 FDR level. At the less stringent 0.1 FDR level, the plasma level of an additional candidate biomarker, insulin-like growth factor binding protein 2 (OR = 1.9, FDR q 0.082), was found to differ. By selection, all of these markers have previously been seen in the literature to associate with AD-related phenotypes. However, in Table 5, it can been seen that we specifically replicate that pancreatic prohormone [8, 39–41] and insulin-like growth factor-binding protein 2 [8, 40] are at a lower level in the blood of AD subjects.

Random forest classification models were used to assess the ability of combinations of co-variates and plasma protein levels to distinguish between AD and control subjects. In the fourth model (co-variates plus candidate proteins selected from the training set), 13 proteins were selected based on their association with AD in the training set: pancreatic prohormone, C-C motif chemokine 18, α-1-antitrypsin, complement C6, insulin-like growth factor-binding protein 2, angiopoietin-2, C-C motif chemokine 15, cystatin C, tumor necrosis factor receptor superfamily member 1B, β-2-microglobulin, prolactin, haptoglobin, and metalloproteinase inhibitor 1. Models were fitted to the training set, and then used to predict AD subjects in the held-out test set. In Table 6, the quality of predictions are assessed by sensitivity, specificity, and accuracy. The most accurate model in the test set was model 2, co-variates plus the most replicated proteins from the 'discovery' literature (accuracy 77%): α-1-antitrypsin, α-2-macroglobulin, apolipoprotein E, and complement C3. All models that included protein data performed better than a model using only co-variates (accuracy 71%). In Supplementary Figure 2, ROC curves for these models applied to the test sets are shown.

### *Candidate AD marker association with MCI to AD conversion*

Logistic regression was used to identify candidate biomarkers that differed between MCI subjects who converted to AD (MCI-AD) within a year of measurement and those who remained MCI during follow-up (MCI-MCI). None of the proteins were found to be associated with conversion, even at the uncorrected 0.05 *p*-value.

### *Candidate AD marker association with rate of cognitive decline*

Partial Spearman's Rank Correlation (SRC) was used to detect associations between plasma levels of

Table 4
Summary of candidate protein markers that associate with AD-related phenotypes in this study. Q-values that pass a FDR threshold of 0.1 are highlighted in bold

| | | FDR corrected q-values from logistic regression or partial SRC | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Protein name | UniProt ID | AD-CTL | Rate of decline | Whole brain | Left hippocampal volume | Right hippocampal volume | Left entorhinal cortex volume | Right entorhinal cortex volume | Number of independent discovery cohorts | Cohort independent from literature discovery cohorts |
| Pancreatic prohormone | P01298 | $\mathbf{1.5 \times 10^{-4}}$ | 0.94 | 0.17 | $\mathbf{8.0 \times 10^{-3}}$ | **0.095** | $\mathbf{7.4 \times 10^{-3}}$ | 0.14 | 4 | Yes |
| Granulocyte colony-stimulating factor | P09919 | 0.56 | 0.94 | 0.79 | 0.23 | 0.2 | $\mathbf{7.4 \times 10^{-3}}$ | **0.022** | 2 | Yes |
| Clusterin | P10909 | 0.26 | $\mathbf{1.2 \times 10^{-3}}$ | 0.72 | 0.72 | 0.85 | 0.96 | 0.94 | 2 | |
| Complement C3 | P01024 | 0.45 | 0.63 | **0.044** | 0.72 | 0.51 | 0.89 | 0.85 | 5 | |
| Complement C6 | P13671 | 0.21 | 0.63 | 0.98 | **0.017** | 0.23 | **0.069** | 0.66 | 1 | Yes |
| Insulin-like growth factor-binding protein 2 | P18065 | **0.083** | 0.22 | 0.96 | 0.72 | 0.67 | 0.52 | 0.86 | 2 | Yes |
| Alpha-1-antitrypsin | P01009 | 0.15 | 0.83 | 0.95 | 0.29 | 0.52 | **0.069** | 0.85 | 5 | Yes |
| Inter-alpha-trypsin inhibitor heavy chain H4 | Q14624 | 0.94 | 0.94 | 0.72 | 0.39 | 0.85 | **0.069** | 0.85 | 2 | |
| C-C motif chemokine 18 | P55774 | 0.25 | 0.87 | 0.99 | 0.29 | 0.32 | **0.069** | 0.55 | 1 | Yes |

*p*-values are from generalized linear models or SRC analysis, with subject age, gender, number/presence of *APOE* $\varepsilon$4 alleles, and study center used as co-variates. FDR correction is performed for each outcome measure separately over *p*-values from comparisons against all 94 candidate protein SOMAmers. For effect sizes and uncorrected *p*-values see Supplementary Tables 7–14. Proteins are sorted first by the number of outcome measures associated at 0.05 FDR and then by the number of outcome measures associated at 0.1 FDR.

Table 5
Summary of association of blood proteins with AD-related phenotypes in this study and others from the literature. HC, hippocampal; EC, entorhinal cortex

| Protein name | Uniprot ID | This study (i.e., ANM+ARUK+DCR) | Literature discovery studies [phenotype, cohort, reference] | Other literature studies (phenotype, cohort/s, reference/s) | Total number of independent replications |
|---|---|---|---|---|---|
| Alpha-1-antitrypsin | P01009 | - with left EC volume | Increased oxidization in AD, Choi [29] | + with AD, Giometto [55] | 8 |
| | | | + with AD, Yu/Zhang/Liao/AIBL [30, 31, 33, 40] | + with AD, Maes [56] | |
| | | | - with brain amyloid burden, ADNI [42] | | |
| Complement C3 | P01024 | + with whole brain volume | + with AD, Zhang [31] | + with AD, Giometto [55] | 6 |
| | | | + pre-symptomatic AD, RSS [37] | + with whole brain volume, ANM [57] | |
| | | | Hippocampal volume, ARUK [21] | | |
| | | | –/+ Brain amyloid burden, ADNI/BLSA [42, 43] | | |
| Pancreatic prohormone | P01298 | + with AD | + with AD, TARC/AIBL [8, 40] | Not found | 5 |
| | | - with HC volume and left EC volume | + with AD and MCI, ADNI/Hu [39, 41] | | |
| | | | - with brain amyloid burden, ADNI [42] | | |
| Granulocyte colony-stimulating factor | P09919 | - with EC volume | with AD, Ray [7] | - with AD, Laske [58] | 4 |
| | | | - with AD, TARC [8] | | |
| Insulin-like growth factor-binding protein 2 | P18065 | + with AD | + with AD, TARC/AIBL [8, 40] | Not found | 3 |
| Clusterin | P10909 | - with rate of decline | - with AD, Liao [33] | + with AD, Rotterdam Study [59] | 3 |
| | | | Fast AD progression and hippocampal volume, ARUK [21] | | |
| Complement C6 | P13671 | - with left EC volume | + with AD, Henkel [38] | Not found | 2 |
| Inter-alpha-trypsin inhibitor heavy chain H4 | Q14624 | - with left EC volume | - with AD, ARUK/Liao [5, 33] | Not found | 2 |
| C-C motif chemokine 18 | P55774 | - with left EC volume | with AD, Ray [7] | Not found | 2 |

Proteins are ordered by the number of times each protein has been found to associate with an AD-related phenotype in an independent cohort, in either a discovery or a candidate study.

Table 6
Performance of random forest classifier predictions in the training and test sets for a range of models. Subject age, gender, and presence of *APOE* $\varepsilon 4$ alleles are used as co-variates

| | Sensitivity | | Specificity | | Accuracy | |
|---|---|---|---|---|---|---|
| | Train | Test | Train | Test | Train | Test |
| Just co-variates | 74% | 70% | 77% | 74% | 75% | 71% |
| Co-variates + most replicated proteins | 100% | 80% | 100% | 72% | 100% | 77% |
| Co-variates + all candidate proteins | 100% | 85% | 100% | 60% | 100% | 75% |
| Co-variates + selected candidate proteins | 100% | 83% | 100% | 66% | 100% | 76% |

The most replicated proteins, from the literature, are α-1-antitrypsin, α-2-macroglobulin, apolipoprotein E, and complement C3. The candidate proteins selected from the training set are: pancreatic prohormone, C-C motif chemokine 18, α-1-antitrypsin, complement C6, insulin-like growth factor-binding protein 2, angiopoietin-2, C-C motif chemokine 15, cystatin C, tumor necrosis factor receptor superfamily member 1B, β-2-microglobulin, prolactin, haptoglobin,and metalloproteinase inhibitor 1.

the 94 candidate biomarkers and each subject's rate of cognitive decline (rate of change in MMSE scores) in the combined ANM+ARUK+DCR cohort (Supplementary Table 9 and results significant at the 0.1 FDR level are shown in Table 4). The plasma level of clusterin was found to positively associate at the 0.05 FDR level, replicating the finding in a partially over-lapping sample set, that clusterin associates with fast cognitive decline [21].

### *Candidate AD markers association with the volume of AD related brain regions*

To test the utility of the markers in predicting early signs of AD pathology, the plasma level of the 94 proteins were compared to volumes of brain regions known to be related to early AD pathology (whole brain, hippocampi, and entorhinal cortices). Not all of the ANM subjects had MRI data available (Supplementary Table 6), therefore, only the subcohort with this data was included in this analysis (274 ANM subjects: 95 controls, 81 MCI, and 98 AD). Results of partial SRC analysis are shown in Supplementary Tables 10–14, and results significant at the 0.1 FDR level are shown in Table 4. In short, at the 0.05 FDR level, the plasma level of complement C3 was found to positively associate with whole brain volume (FDR q = 0.044); pancreatic prohormone (FDR $q = 8.0 \times 10^{-3}$) and complement C6 (FDR q = 0.017) were found to associate negatively with the volume of the left hippocampus. Additionally, pancreatic prohormone (FDR $q = 7.4 \times 10^{-3}$) and granulocyte colony-stimulating factor (FDR $q = 7.4 \times 10^{-3}$) were found to negatively associate with the volume of the left entorhinal cortex at the 0.05 FDR level. Granulocyte colony-stimulating factor was also found to negatively associate with the volume of the right entorhinal cortex at the 0.05 FDR level (FDR q = 0.022). Complement C3 has previously been shown to positively associate with whole brain volume in a candidate-based study that used a partially over-lapping sample set [57].

At the less stringent 0.1 FDR level, pancreatic prohormone was found to negatively associate with the volume of the right hippocampi (FDR q = 0.095), and four of the candidate markers—complement C6 (FDR q = 0.069), α-1-antitrypsin (FDR q = 0.068), inter-α-trypsin inhibitor heavy chain H4 (FDR q = 0.069), and C-C motif chemokine 18 (FDR q = 0.069)—were found to associate negatively with the volume of the left entorhinal cortex. Scatterplots showing the plasma levels of the nine proteins found to associate with at least one AD-related phenotype at the 0.1 FDR significance threshold are shown in in Supplementary Figure 1.

## DISCUSSION

In this study, the reproducibility of 'discovery' study derived blood protein markers of AD-related phenotypes has been investigated. This was performed across different blood sample types, subject sets, and proteomic assays. A total of 163 candidate proteins were mined from 21 different studies (15 discovery and 6 using panels of over 100 proteins), the majority of which were unique to a single cohort and/or proteomic approach. However, some candidate markers were found to associate with AD-related phenotypes in multiple independent studies, despite considerable methodological differences. Of the 163 candidate blood-based protein markers, 94 are included in the current menu of the SOMAscan assay; these 94 were tested for association with: AD diagnosis, future conversion from MCI to AD, rate of cognitive decline, and the volume of specific brain regions. Additionally, case-control classification using sets of plasma protein markers were explored in combination with age, gender, and the presence of *APOE* $\varepsilon 4$ alleles.

Nine candidates were found to associate with at least one AD-related phenotype at the 0.1 FDR level, taking into account subject age, gender, presence of *APOE* $\varepsilon 4$ alleles, and research center. Three of these candidates had been found to associate with an AD-related phenotype in studies that used an overlapping set of plasma samples, albeit using a different technology, while the remainder were considered independent. Of these six independent markers, two associated with AD clinical diagnosis, and in the same direction previously reported (Table 5): pancreatic prohormone and insulin-like growth factor-binding protein 2. Other markers associate with different but logically consistent AD-related phenotypes, for example we found that α-1-antitrypsin and complement C6 associate negatively with the volume of the left entorhinal cortex. This fits with literature studies showing that these proteins are higher in AD subjects relative to controls [29, 30, 33, 38, 40, 55, 56].

In Table 5, the results of this study are compared with the literature, and while there is general concordance of the direction of association, there are cases where this study finds association of protein markers with AD-related phenotypes in the opposite direction to that we would have expected based on literature findings; for example complement C3 was found to be at a higher plasma level in AD subjects in Zhang et al. [31], Giometto et al. [55], and Maes et al. [56], whereas it was found to associate with whole brain volume in this and another study using ANM samples [57]. As brain volume decreases during the development of AD, one would expect associations with whole brain volume and AD diagnosis to be in the opposite direction. It is noteworthy that complement C3 has been found to associate positively with brain amyloid burden in one study [42] and negatively in another [43]. Taken together, this may indicate that different forms of C3 are being measured in these studies. Another possible explanation for this discrepancy is that this association is being confounded by sample handling; complement C3 is one of the proteins targeted by the SOMAscan assay that is most affected by recruitment center in our study (data not shown). Sample handling is another confounder that should be considered in these studies, but also when considering whether a candidate could be viable in a clinical setting.

Other possible direction of association discrepancies are granulocyte colony-stimulating factor and inter-α-trypsin inhibitor heavy chain H4. Previous studies have found that their plasma level is lowered in AD blood [5, 8, 33, 58], whereas in this study it was found that their plasma level is negatively associated with the volume of the entorhinal cortex; as the entorhinal cortex atrophies during AD; a positive association would fit better with the literature. This could perhaps be explained if these markers are negatively associated with early AD (e.g., pre-clinical brain atrophy) and positively associated with later clinical AD.

In the classification models we tried different sets of markers along with co-variates to predict AD diagnosis. This showed that candidate protein levels measured by SOMAscan consistently add a little to the predictive accuracy, but that co-variates alone contribute most to the classifier. It is promising that the most replicated four protein AD markers from the literature improved the predictive accuracy of the models, although it is disappointing that the improvement in accuracy is so small. This may indicate that either the SOMAscan assay is not the ideal platform to measure the most replicated four proteins, AD signal acting through these proteins is masked by other processes or that a more balanced cohort design with respect to co-variates is necessary to properly assess these biomarker sets.

Although we took into account the most obvious confounding factors—age, gender, *APOE* $\varepsilon 4$ alleles, and recruitment center—it is possible that there are other important confounding factors that we failed to take into account. These could result in misleading correlations between the level of proteins in plasma and AD-related phenotypes in this study and others. This is something that can be explored in future studies. Possible confounding factors include co-morbidities, kidney function, diet, and medications. For example it is interesting that one of the most reproducible protein markers of AD from the literature is α-2-macroglobulin, which has been shown by Gold et al. [22] to be a marker of chronic kidney disease. While we co-varied for the presence of *APOE* $\varepsilon 4$ in all analyses, two replicated candidates, complement component C3/C6, were shown to be associated with it. This could either indicate *APOE* $\varepsilon 4$ independent information for these proteins, or confounding by unknown variables.

The inclusion criteria used to identify blood protein markers from the literature excluded the studies that took a candidate approach, in order to reduce the effect of literature biases. We reported markers against a range of AD-related phenotypes relevant to onset and/or progression of disease, as well as diagnosis. In order to be inclusive of a wide range of studies, the type of blood sample (i.e., plasma/serum/leukocyte) was not considered as exclusion criteria. The overlap in the proteome of these different blood samples is of interest, as has been explored in O'Bryant et al. [60].

Highly replicated candidate biomarkers show promise and are obvious candidates for future studies, especially if they are detected across a range of different proteomic platforms. This is another motivation for applying novel proteomic technologies such as the SOMAscan assay to the challenge of validating blood-based markers of AD. This approach has proved especially fruitful in the replication of pancreatic prohormone plasma levels as a marker of AD clinical diagnosis, which had only previously been observed in studies utilizing Luminex xMAP assays (Table 5). This finding was so robust that it was also a top finding in a 'discovery' analysis we performed using the full SOMAscan panel of 1,001 proteins (Sattlecker et al., unpublished data).

The benefit of the candidate replication approach taken is that it increased the statistical power relative to a univariate analysis using the full SOMAscan panel (Sattlecker et al, unpublished data). This allowed the detection of significant associations (at the 0.1 FDR level) of a greater number of candidate biomarkers with AD-related phenotypes, i.e., complement C6, C-C chemokine 18, α-1-antitrypsin, and inter-α-trypsin inhibitor heavy chain H4.

It seems that a larger number of proteins associate with the left hippocampus and entorhinal cortex than to their right equivalent, which may relate to a difference in the onset and severity of atrophy of left versus right brain regions; this has preliminarily been reported for the hippocampi previously [61, 62].

We have used longitudinal MMSE scores to calculate subjects' rate of cognitive decline. MMSE was chosen as it was the most widely available cognitive measure in these cohorts, but it should be noted that the MMSE is less sensitive to cognitive decline than other cognitive assessments like the Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-cog). More sensitive cognitive tests could be used to identify plasma markers that are more sensitive to cognitive decline in a future study. However, it is interesting that given the lack of sensitivity of MMSE to monitor the rate of cognitive decline, we are still able to find a protein marker, clusterin, that associates with its rate of change, replicating previous findings [21].

Pathway analysis has revealed that inflammatory proteins are over-represented in the literature-derived candidate markers. It is known that neuroinflammation is a feature of AD, but it is also believed to be a feature of many other psychiatric disorders [63]. It would be interesting to know if these proteins represent a distinctive AD signal in blood or represent general inflammation/neuroinflammation. Proteins involved in the complement pathway are also over-represented in the list of candidate markers. This is encouraging since complement proteins are known to have a more direct involvement in AD. They have been identified as genetic risk factors for AD [64] and are activated in amyloid plaques in very early AD [65].

A key factor in the utility of markers of disease is their specificity; it should be noted that some of the candidates in this study were found to be relevant to other diseases in the KEGG analysis. Similarly, candidate proteins that were found to associate with AD-related phenotypes in this study have been found to be potential markers of other diseases, for example the level of matrix metalloproteinase-9 has been found to be altered in serum samples of subjects with schizophrenia [66]. Similarly, the level of apolipoprotein A1 and complement C9 in serum have been considered as markers of malignant pleural mesothelioma in a previous study utilizing the SomaLogic technology [52]. This suggests that some single protein markers of AD-related disease in blood may have limited disease specificity, possibly due to their involvement in multiple disease processes. Further studies will be needed to test this and to evaluate the specific potential utility of these proteins as prognostic, diagnostic, and pharmacodynamics markers of AD.

## CONCLUSION

We reviewed published discovery or panel-based blood proteomics studies of AD. Out of the total of 164 potential markers and 18 independent cohorts described in these studies, ∼66% of the potential biomarkers identified were only found to associate with AD-related measures in a single cohort. However, some proteins do appear in multiple studies; for example, four candidate biomarkers are found to associate with AD-related pathologies in five independent research cohorts in these 21 studies. Using SomaLogic's SOMAscan proteomics technology, we replicated nine candidate biomarkers at the 0.1 FDR level. Five of these associate with AD-related phenotypes in a direction compatible with literature findings. This was performed on plasma samples from the AddNeuroMed and ARUK/Maudsley BRC Dementia Case Registry at King's Health Partners research cohorts. Overall, our findings replicate some previously reported markers, and suggest that AD pathology affects the blood proteome with some consistency.

Given the unmet need for blood-based biomarkers of AD, it is imperative that the most well replicated

candidates are tested in larger and more extensively phenotyped cohorts (for example followed longitudinally). In parallel with this, a greater focus must be applied to specificity and potential confounders (including genotype). Finally, it is hoped that novel proteomics approaches may reveal stronger candidate markers of AD. Overall, a replicated blood protein signature of AD has been found, across this and many other studies, and is worthy of greater investigation.

## ACKNOWLEDGMENTS

This work was supported by InnoMed, (Innovative Medicines in Europe), an Integrated Project funded by the European Union of the Sixth Framework program priority [FP6-2004-LIFESCIHEALTH-5]; Alzheimer's Research UK (formerly Alzheimer's Research Trust); The John and Lucille van Geest Foundation and the NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at the South London and Maudsley NHS Foundation Trust and Kings College London and a joint infrastructure grant from Guy's and St Thomas' Charity and the Maudsley Charity; Kuopio University Hospital (HS) and funding from UEFBRAIN (HS). SOMAscan™ and SOMAmer™ are trademarks of SomaLogic, Inc. We would like to acknowledge Abhishek Dixit for his help with data management and access. We would also like to thank peer reviewers for their advice, which has improved the manuscript. This paper presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

Authors' disclosures available online (http://www.j-alz.com/disclosures/view.php?id=1878).

## SUPPLEMENTARY MATERIAL

Supplementary figures and tables are available in the electronic version of this article: http://dx.doi.org/10.3233/JAD-130380.

## REFERENCES

[1] Frankfort SV, Turner LR, van Campen JP, Verbeek MM, Jansen RW, Beijnen JH (2008) Amyloid beta protein and tau in cerebrospinal fluid and plasma as biomarkers for dementia: A review of recent literature. *Curr Clin Pharamacol* **3**, 123-131.

[2] Buchhave P, Minthon L, Zetterberg H, Wallin AK, Blennow K, Hansson O (2012) Cerebrospinal fluid levels of beta-amyloid 1-42, but not of tau, are fully changed already 5 to 10 years before the onset of Alzheimer dementia. *Arch Gen Psychiatry* **69**, 98-106.

[3] Bazenet C, Lovestone S (2012) Plasma biomarkers for Alzheimer's disease: Much needed but tough to find. *Biomarkers Med* **6**, 441-454.

[4] Rissman RA, Trojanowski JQ, Shaw LM, Aisen PS (2012) Longitudinal plasma amyloid beta as a biomarker of Alzheimer's disease. *J Neural Transm* **119**, 843-850.

[5] Hye A, Lynham S, Thambisetty M, Causevic M, Campbell J, Byers HL, Hooper C, Rijsdijk F, Tabrizi SJ, Banner S, Shaw CE, Foy C, Poppe M, Archer N, Hamilton G, Powell J, Brown RG, Sham P, Ward M, Lovestone S (2006) Proteome-based plasma biomarkers for Alzheimer's disease. *Brain* **129**, 3042-3050.

[6] Güntert A, Campbell J, Saleem M, O'Brien DP, Thompson AJ, Byers HL, Ward MA, Lovestone S (2010) Plasma gelsolin is decreased and correlates with rate of decline in Alzheimer's disease. *J Alzheimers Dis* **21**, 585-596.

[7] Ray S, Britschgi M, Herbert C, Takeda-Uchimura Y, Boxer A, Blennow K, Friedman LF, Galasko DR, Jutel M, Karydas A, Kaye JA, Leszek J, Miller BL, Minthon L, Quinn JF, Rabinovici GD, Robinson WH, Sabbagh MN, So YT, Sparks DL, Tabaton M, Tinklenberg J, Yesavage JA, Tibshirani R, Wyss-Coray T (2007) Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. *Nat Med* **13**, 1359-1362.

[8] O'Bryant SE, Xiao G, Barber R, Reisch J, Doody R, Fairchild T, Adams P, Waring S, Diaz-Arrastia R, Texas Alzheimer's Research Consortium (2010) A serum protein-based algorithm for the detection of Alzheimer disease. *Arch Neurol* **67**, 1077.

[9] Lista S, Faltraco F, Prvulovic D, Hampel H (2013) Blood and plasma-based proteomic biomarker research in Alzheimer's disease. *Prog Neurobiol* **101-102**, 1-17.

[10] Zürbig P, Jahn H (2012) Use of proteomic methods in the analysis of human body fluids in Alzheimer research. *Electrophoresis* **33**, 3617-3630.

[11] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-29.

[12] Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* **27**, 29-34.

[13] Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57.

[14] Huang DW, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1-13.

[15] Maere S, Heymans K, Kuiper M (2005) BiNGO: A Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics* **21**, 3448-3449.

[16] Muthusamy B, Hanumanthu G, Suresh S, Rekha B, Srinivas D, Karthick L, Vrushabendra BM, Sharma S, Mishra G, Chatterjee P, Mangala KS, Shivashankar HN, Chandrika KN, Deshpande N, Suresh M, Kannabiran N, Niranjan V, Nalli A, Prasad TS, Arun KS, Reddy R, Chandran S, Jadhav T, Julie D, Mahesh M, John SL, Palvankar K, Sudhir D, Bala P, Rashmi NS, Vishnupriya G, Dhar K, Reshma S, Chaerkady

R, Gandhi TK, Harsha HC, Mohan SS, Deshpande KS, Sarker M, Pandey A (2005) Plasma Proteome Database as a resource for proteomics research. *Proteomics* **5**, 3531-3536.

[17] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504.

[18] Lovestone S, Francis P, Strandgaard K (2007) Biomarkers for disease modification trials-The innovative medicines initiative and AddNeuroMed. *J Nutr Health Aging* **11**, 359-361.

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

[20] Thambisetty M, Hye A, Foy C, Daly E, Glover A, Cooper A, Simmons A, Murphy D, Lovestone S (2008) Proteome-based identification of plasma proteins associated with hippocampal metabolism in early Alzheimer's disease. *J Neurol* **255**, 1712-1720.

[21] Thambisetty M, Simmons A, Velayudhan L, Hye A, Campbell J, Zhang Y, Wahlund LO, Westman E, Kinsey A, Güntert A, Proitsi P, Powell J, Causevic M, Killick R, Lunnon K, Lynham S, Broadstock M, Choudhry F, Howlett DR, Williams RJ, Sharp SI, Mitchelmore C, Tunnard C, Leung R, Foy C, O'Brien D, Breen G, Furney SJ, Ward M, Kloszewska I, Mecocci P, Soininen H, Tsolaki M, Vellas B, Hodges A, Murphy DG, Parkins S, Richardson JC, Resnick SM, Ferrucci L, Wong DF, Zhou Y, Muehlboeck S, Evans A, Francis PT, Spenger C, Lovestone S (2010) Association of plasma clusterin concentration with severity, pathology, and progression in Alzheimer disease. *Arch Gen Psychiatry* **67**, 739-748.

[22] 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.

[23] Westman E, Aguilar C, Muehlboeck JS, Simmons A (2013) Regional magnetic resonance imaging measures for multivariate analysis in Alzheimer's disease and mild cognitive impairment. *Brain Topogr* **26**, 9-23.

[24] Simmons A, Westman E, Muehlboeck S, Mecocci P, Vellas B, Tsolaki M, Kłoszewska I, Wahlund LO, Soininen H, Lovestone S, Evans A, Spenger C, AddNeuroMed consortium (2009) MRI measures of Alzheimer's disease and the AddNeuroMed study. *Ann N Y Acad Sci* **1180**, 47-55.

[25] Simmons A, Westman E, Muehlboeck S, Mecocci P, Vellas B, Tsolaki M, Kłoszewska I, Wahlund LO, Soininen H, Lovestone S, Evans A, Spenger C (2011) The AddNeuroMed framework for multi-centre MRI assessment of Alzheimer's disease: Experience from the first 24 months. *Int J Geriatr Psychiatry* **26**, 75-82.

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

[27] Westman E, Simmons A, Zhang Y, Muehlboeck JS, Tunnard C, Liu Y, Collins L, Evans A, Mecocci P, Vellas B, Tsolaki M, Kłoszewska I, Soininen H, Lovestone S, Spenger C, Wahlund LO, AddNeuroMed consortium (2011) Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls. *Neuroimage* **54**, 1178-1187.

[28] Sing T, Sander O, Berenwinkel N, Lengauer T (2005) ROCR: Visualizing classifier performance in R. *Bioinformatics* **21**, 3940-3941.

[29] Choi J, Malakowsky CA, Talent JM, Conrad CC, Gracy RW (2002) Identification of oxidized plasma proteins in Alzheimer's disease. *Biochem Biophys Res Commun* **293**, 1566-1570.

[30] Yu HL, Chertkow HM, Bergman H, Schipper HM (2003) Aberrant profiles of native and oxidized glycoproteins in Alzheimer plasma. *Proteomics* **3**, 2240-2248.

[31] Zhang R, Barker L, Pinchev D, Marshall J, Rasamoelisolo M, Smith C, Kupchak P, Kireeva I, Ingratta L, Jackowski G (2004) Mining biomarkers in human sera using proteomic tools. *Proteomics* **4**, 244-256.

[32] Liu HC, Hu CJ, Chang JG, Sung SM, Lee LS, Yuan RY, Leu SJ (2006) Proteomic identification of lower apolipoprotein A-I in Alzheimer's disease. *Dement Geriatr Cogn Disord* **21**, 155-161.

[33] Liao PC, Yu L, Kuo CC, Kuo YM (2007) Proteomics analysis of plasma for potential biomarkers in the diagnosis of Alzheimer's disease. *Proteomics Clin Appl* **1**, 506-512.

[34] Akuffo EL, Davis JB, Fox SM, Gloger IS, Hosford D, Kinsey EE, Jones NA, Nock CM, Roses AD, Saunders AM, Skehel JM, Smith MA, Cutler P (2008) The discovery and early validation of novel plasma biomarkers in mild-to-moderate Alzheimer's disease patients responding to treatment with rosiglitazone. *Biomarkers* **13**, 618-636.

[35] Cutler P, Akuffo EL, Bodnar WM, Briggs DM, Davis JB, Debouck CM, Fox SM, Gibson RA, Gormley DA, Holbrook JD, Hunter AJ, Kinsey EE, Prinjha R, Richardson JC, Roses AD, Smith MA, Tsokanas N, Willé DR, Wu W, Yates JW, Gloger IS (2008) Proteomic identification and early validation of complement 1 inhibitor and pigment epithelium-derived factor: Two novel biomarkers of Alzheimer's disease in human plasma. *Prot Clin Appl* **2**, 467-477.

[36] Mhyre TR, Loy R, Tariot PN, Profenno LA, Maguire-Zeiss KA, Zhang D, Coleman PD, Federoff HJ (2008) Proteomic analysis of peripheral leukocytes in Alzheimer's disease patients treated with divalproex sodium. *Neurobiol Aging* **29**, 1631-1643.

[37] Ijsselstijn L, Dekker LJ, Stingl C, van der Weiden MM, Hofman A, Kros JM, Koudstaal PJ, Sillevis Smitt PA, Ikram MA, Breteler MM, Luider TM (2011) Serum levels of pregnancy zone protein are elevated in presymptomatic Alzheimer's disease. *J Proteome Res* **10**, 4902-4910.

[38] Henkel AW, Müller K, Lewczuk P, Müller T, Marcus K, Kornhuber J, Wiltfang J (2012) Multidimensional plasma protein separation technique for identification of potential Alzheimer's disease plasma biomarkers: A pilot study. *J Neural Transm* **119**, 779-788.

[39] Soares HD, Potter WZ, Pickering E, Kuhn M, Immermann FW, Shera DM, Ferm M, Dean RA, Simon AJ, Swen-

son F, Siuciak JA, Kaplow J, Thambisetty M, Zagouras P, Koroshetz WJ, Wan HI, Trojanowski JQ, Shaw LM, Biomarkers Consortium Alzheimer's Disease Plasma Proteomics Project (2012) Plasma biomarkers associated with the apolipoprotein E genotype and Alzheimer. *Arch Neurol* **69**, 1310-1317.

[40] Doecke JD, Laws SM, Faux NG, Wilson W, Burnham SC, Lam CP, Mondal A, Bedo J, Bush AI, Brown B, De Ruyck K, Ellis KA, Fowler C, Gupta VB, Head R, Macaulay SL, Pertile K, Rowe CC, Rembach A, Rodrigues M, Rumble R, Szoeke C, Taddei K, Taddei T, Trounson B, Ames D, Masters CL, Martins RN, Alzheimer's Disease Neuroimaging Initiative, Australian Imaging Biomarker and Lifestyle Research Group (2012) Blood-based protein biomarkers for diagnosis of Alzheimer disease. *Arch Neurol* **69**, 1318-1325.

[41] Hu WT, Holtzman DM, Fagan AM, Shaw LM, Perrin R, Arnold SE, Grossman M, Xiong C, Craig-Schapiro R, Clark CM, Pickering E, Kuhn M, Chen Y, Van Deerlin VM, McCluskey L, Elman L, Karlawish J, Chen-Plotkin A, Hurtig HI, Siderowf A, Swenson F, Lee VM, Morris JC, Trojanowski JQ, Soares H, Alzheimer's Disease Neuroimaging, Initiative (2012) Plasma multianalyte profiling in mild cognitive impairment and Alzheimer disease. *Neurology* **79**, 897-905.

[42] Kiddle SJ, Thambisetty M, Simmons A, Riddoch-Contreras J, Hye A, Westman E, Pike I, Ward M, Johnston C, Lupton MK, Lunnon K, Soininen H, Kloszewska I, Tsolaki M, Vellas B, Mecocci P, Lovestone S, Newhouse S, Dobson R, Alzheimers Disease Neuroimaging Initiative (2012) Plasma based markers of [11C] PiB-PET brain amyloid burden. *PLoS One* **7**, e44260.

[43] Thambisetty M, Tripaldi R, Riddoch-Contreras J, Hye A, An Y, Campbell J, Sojkova J, Kinsey A, Lynham S, Zhou Y, Ferrucci L, Wong DF, Lovestone S, Resnick SM (2010) Proteome-based plasma markers of brain amyloid-β deposition in non-demented older individuals. *J Alzheimers Dis* **22**, 1099-1109.

[44] Farrah T, Deutsch EW, Omenn GS, Campbell DS, Sun Z, Bletz JA, Mallick P, Katz JE, Malmström J, Ossola R, Watts JD, Lin B, Zhang H, Moritz RL, Aebersold R (2011) A high-confidence human plasma proteome reference set with estimated concentrations in PeptideAtlas. *Mol Cell Proteomics* **10**, M110.006353.

[45] Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack CR, Jagust W, Trojanowski JQ, Toga AW, Beckett L (2005) Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement* **1**, 55-66.

[46] Ellis KA, Bush AI, Darby D, De Fazio D, Foster J, Hudson P, Lautenschlager NT, Lenzo N, Martins RN, Maruff P, Masters C, Milner A, Pike K, Rowe C, Savage G, Szoeke C, Taddei K, Villemagne V, Woodward M, Ames D, AIBL Research Group (2009) The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: Methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *Int Psychogeriatr* **21**, 672-687.

[47] Resnick SM, Goldszal AF, Davatzikos C, Golski S, Kraut MA, Metter EJ, Bryan RN, Zonderman AB (2000) One-year age changes in MRI brain volumes in older adults. *Cereb Cortex* **10**, 464-472.

[48] Li H, Wetten S, Li L, St Jean PL, Upmanyu R, Surh L, Hosford D, Barnes MR, Briley JD, Borrie M, Coletta N, Delisle R, Dhalla D, Ehm MG, Feldman HH, Fornazzari L, Gauthier S, Goodgame N, Guzman D, Hammond S, Hollingworth P, Hsiung GY, Johnson J, Kelly DD, Keren R, Kertesz A, King KS, Lovestone S, Loy-English I, Matthews PM, Owen MJ, Plumpton M, Pryse-Phillips W, Prinjha RK, Richardson JC, Saunders A, Slater AJ, St George-Hyslop PH, Stinnett SW, Swartz JE, Taylor RL, Wherrett J, Williams J, Yarnall DP, Gibson RA, Irizarry MC, Middleton LT, Roses AD (2008) Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease. *Arch Neurol* **65**, 45-53.

[49] Waring S, O'Bryant SE, Reisch JS, Diaz-Arrastia R, Knebl J, Doody R, Texas Alzheimer's Research Consortium (2008) The Texas Alzheimer's Research Consortium longitudinal research cohort: Study design and baseline characteristics. *Texas Public Health J* **60**, 10-13.

[50] Robinson W, Nikrad MP, Robinson S, Williams S, Ostroff R (2010) New SOMAmer-based assay to discover biomarkers relevant to malignant melanoma. *AACR Meeting Abstracts*, A10.

[51] Ostroff RM, Bigbee WL, Franklin W, Gold L, Mehan M, Miller YE, Pass HI, Rom WN, Siegfried JM, Stewart A, Walker JJ, Weissfeld JL, Williams S, Zichi D, Brody EN (2010) Unlocking biomarker discovery: Large scale application of aptamer proteomic technology for early detection of lung cancer. *PLoS One* **5**, e15003.

[52] Ostroff RM, Mehan MR, Stewart A, Ayers D, Brody EN, Williams SA, Levin S, Black B, Harbut M, Carbone M, Goparaju C, Pass HI (2012) Early detection of malignant pleural mesothelioma in asbestos-exposed individuals with a noninvasive proteomics-based surveillance tool. *PLoS One* **7**, e46091.

[53] Baird GS, Nelson SK, Keeney TR, Stewart A, Williams S, Kraemer S, Peskind ER, Montine TJ (2012) Age-dependent changes in the cerebrospinal fluid proteome by slow off-rate modified aptamer array. *Am J Pathol* **180**, 446-456.

[54] Lourdusamy A, Newhouse S, Lunnon K, Proitsi P, Powell J, Hodges A, Nelson SK, Stewart A, Williams S, Kloszewska I, Mecocci P, Soininen H, Tsolaki M, Vellas B, Lovestone S; AddNeuroMed Consortium, Dobson R, Alzheimer's Disease Neuroimaging Initiative (2012) Identification of cis-regulatory variation influencing protein abundance levels in human plasma. *Hum Mol Genet* **21**, 3719-3726.

[55] Giometto B, Argentiero V, Sanson F, Ongaro G, Tavolato B (1988) Acute-phase proteins in Alzheimer's disease. *Eur Neurol* **28**, 30-33.

[56] Maes OC, Kravitz S, Mawal Y, Su H, Liberman A, Mehindate K, Berlin D, Sahlas DJ, Chertkow HM, Bergman H, Melmed C, Schipper HM (2006) Characterization of α1-antitrypsin as a heme oxygenase-1 suppressor in Alzheimer plasma. *Neurobiol Dis* **24**, 89-100.

[57] Thambisetty M, Simmons A, Hye A, Campbell J, Westman E, Zhang Y, Wahlund LO, Kinsey A, Causevic M, Killick R, Kloszewska I, Mecocci P, Soininen H, Tsolaki M, Vellas B, Spenger C, Lovestone S, AddNeuroMed Consortium (2011) Plasma biomarkers of brain atrophy in Alzheimer's disease. *PLoS One* **6**, e28527.

[58] Laske C, Stellos K, Stransky E, Leyhe T, Gawaz M (2009) Decreased plasma levels of granulocyte-colony stimulating factor (G-CSF) in patients with early Alzheimer's disease. *J Alzheimers Dis* **17**, 115-123.

[59] Schrijvers EMC, Koudstaal PJ, Hofman A, Breteler MMB (2011) Plasma clusterin and the risk of Alzheimer disease. *JAMA* **305**, 1322-1326.

[60] O'Bryant SE, Xiao G, Barber R, Huebinger R, Wilhelmsen K, Edwards M, Graff-Radford N, Doody R, Diaz-Arrastia R, Texas Alzheimer's Research & Care Consor-

tium, Alzheimer's Disease Neuroimaging Initiative (2011) A blood-based screening tool for Alzheimer's disease that spans serum and plasma: Findings from TARC and ADNI. *PLoS One* **6**, e28092.

[61] Wolf H, Grunwald M, Kruggel F, Riedel-Heller SG, Angerhöfer S, Hojjatoleslami A, Hensel A, Arendt T, Gertz H (2001) Hippocampal volume discriminates between normal cognition; questionable and mild dementia in the elderly. *Neurobiol Aging* **22**, 177-186.

[62] Barnes J, Scahill RI, Schott JM, Frost C, Rossor MN, Fox NC (2005) Does Alzheimer's disease affect hippocampal asymmetry? Evidence from a cross-sectional and longitudinal volumetric MRI study. *Dement Geriatr Cogn Disord* **19**, 338-344.

[63] Najjar S, Pearlman DM, Alper K, Najjar A, Devinsky O (2013) Neuroinflammation and psychiatric illness. *J Neuroinflammation* **10**, 1-24.

[64] Aiyaz M, Lupton MK, Proitsi P, Powell JF, Lovestone S (2012) Complement activation as a biomarker for Alzheimer's disease. *Immunobiology* **217**, 204-215.

[65] Zanjani H, Finch CE, Kemper C, Atkinson J, McKeel D, Morris JC, Price JL (2005) Complement activation in very early Alzheimer disease. *Alzheimer Dis Assoc Disord* **19**, 55-66.

[66] Schwarz E, Guest PC, Steiner J, Bogerts B, Bahn S (2012) Identification of blood-based molecular signatures for prediction of response and relapse in schizophrenia patients. *Transl Psychiatry* **2**, e82.