

## ORIGINAL ARTICLE

## Plasma protein biomarkers of Alzheimer's disease endophenotypes in asymptomatic older twins: early cognitive decline and regional brain volumes

SJ Kiddle<sup>1,15</sup>, CJ Steves<sup>2,15</sup>, M Mehta<sup>3</sup>, A Simmons<sup>3,4</sup>, X Xu<sup>1,4</sup>, S Newhouse<sup>1,4</sup>, M Sattler<sup>1,4</sup>, NJ Ashton<sup>4,5</sup>, C Bazenet<sup>4,5</sup>, R Killick<sup>5</sup>, J Adnan<sup>5</sup>, E Westman<sup>6</sup>, S Nelson<sup>7</sup>, H Soininen<sup>8,9</sup>, I Kloszewska<sup>10</sup>, P Mecocci<sup>11</sup>, M Tsolaki<sup>12</sup>, B Vellas<sup>13</sup>, C Curtis<sup>1,4</sup>, G Breen<sup>1,4</sup>, SCR Williams<sup>3</sup>, S Lovestone<sup>14</sup>, TD Spector<sup>2,16</sup> and RJB Dobson<sup>1,4,16</sup>

There is great interest in blood-based markers of Alzheimer's disease (AD), especially in its pre-symptomatic stages. Therefore, we aimed to identify plasma proteins whose levels associate with potential markers of pre-symptomatic AD. We also aimed to characterise confounding by genetics and the effect of genetics on blood proteins in general. Panel-based proteomics was performed using SOMAscan on plasma samples from TwinsUK subjects who are asymptomatic for AD, measuring the level of 1129 proteins. Protein levels were compared with 10-year change in CANTAB-paired associates learning (PAL;  $n = 195$ ), and regional brain volumes ( $n = 34$ ). Replication of proteins associated with regional brain volumes was performed in 254 individuals from the AddNeuroMed cohort. Across all the proteins measured, genetic factors were found to explain ~26% of the variability in blood protein levels on average. The plasma level of the mitogen-activated protein kinase (MAPK) MAPKAPK5 protein was found to positively associate with the 10-year change in CANTAB-PAL in both the individual and twin difference context. The plasma level of protein MAP2K4 was found to suggestively associate negatively ( $Q < 0.1$ ) with the volume of the left entorhinal cortex. Future studies will be needed to assess the specificity of MAPKAPK5 and MAP2K4 to eventual conversion to AD.

*Translational Psychiatry* (2015) **5**, e584; doi:10.1038/tp.2015.78; published online 16 June 2015

## INTRODUCTION

Currently, no treatments exist that delay or prevent onset of Alzheimer's disease (AD). Studies using magnetic resonance imaging (MRI) and positron emission tomography brain scans have revealed early signs of AD pathology in subjects up to ~4 and ~17 years before the onset of dementia, respectively.<sup>1</sup> Pathological changes are likely to be least advanced in asymptomatic subjects, and therefore potentially the most remediable in this group. This is the motivation behind prevention trials in asymptomatic subjects with MRI or positron emission tomography evidence of early AD pathology.<sup>2</sup>

However, such brain scans can be relatively expensive and require specialised facilities. Therefore, researchers are looking for surrogate markers that are relatively inexpensive and non-invasive, yet informative enough for enrichment of prevention trials. Preliminary results reviewed in Lista *et al.*,<sup>3</sup> Zürbig and Jahn<sup>4</sup> and Chiam *et al.*<sup>5</sup> suggest that blood may be a useful source of such markers. Despite this, only two discovery studies have been

performed comparing blood proteins to AD-related phenotypes exclusively in asymptomatic individuals.<sup>6,7</sup>

Recently, with the advent of panel-based proteomics, high-throughput studies are becoming more feasible. One such approach, SOMAscan (slow off-rate modified aptamer), allows >1000 proteins to be measured simultaneously and has already been applied in AD research.<sup>8–11</sup> This approach has not yet been applied to discover blood markers of AD-related phenotypes in asymptomatic individuals.

Although most blood protein biomarker studies have used cohorts of individuals, one study of APOA1 blood DNA methylation and protein level has been performed using 24 twin pairs discordant for episodic memory performance. Two APOA1 CpG sites were found to associate with discordant episodic memory. However, the protein level of APOA1 was not found to differ between discordant twins, possibly due to the small sample size available.<sup>12</sup>

An advantage of studying 'omics' markers in twin cohorts is that it allows genetic and environmental influences on markers and

<sup>1</sup>MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; <sup>2</sup>Department of Twin Research & Genetic Epidemiology, King's College London, London, UK; <sup>3</sup>Centre for Neuroimaging Sciences, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; <sup>4</sup>NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK; <sup>5</sup>Department of Old Age Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; <sup>6</sup>Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden; <sup>7</sup>SomaLogic, Boulder, CO, USA; <sup>8</sup>Institute of Clinical Medicine – Neurology, University of Eastern Finland, Kuopio, Finland; <sup>9</sup>NeuroCenter, Kuopio University Hospital, Kuopio, Finland; <sup>10</sup>Department of Old Age Psychiatry and Psychotic disorders, Medical University of Łódź, Łódź, Poland; <sup>11</sup>Institute of Gerontology and Geriatrics, University of Perugia, Perugia, Italy; <sup>12</sup>3rd Department of Neurology, Aristotle University, Thessaloniki, Greece; <sup>13</sup>Department of Internal Medicine and Geriatric Medicine, INSERM University of Toulouse, Toulouse, France and <sup>14</sup>Department of Psychiatry, Oxford University, Warneford Hospital, Oxford, UK. Correspondence: Dr SJ Kiddle or Dr RJB Dobson, Institute of Psychiatry, Psychology and Neuroscience, King's College London, Box P092, SGDP Building, De Crespigny Park, London SE5 8AF, UK. E-mail: [steven.kiddle@kcl.ac.uk](mailto:steven.kiddle@kcl.ac.uk) or [richard.j.dobson@kcl.ac.uk](mailto:richard.j.dobson@kcl.ac.uk)

<sup>15</sup>These authors contributed equally to this work.

<sup>16</sup>These authors contributed equally to this work.

Received 2 March 2015; accepted 7 May 2015

traits to be studied.<sup>13</sup> If the monozygotic (MZ, that is, 'identical') twin difference in the level of a marker associates with the twin difference in trait value in twin pairs, this suggests that the relationship between marker and trait is not confounded by age, familial environment or genetics. Confounding of AD blood biomarkers by age and/or genetics is a real possibility, as these factors have been shown to affect both AD risk<sup>14,15</sup> and plasma protein levels.<sup>16–18</sup> Despite this, twin differences have not yet been exploited in studies of blood protein biomarkers of AD-related phenotypes.

Genetic markers are generally more practical, robust and inexpensive to measure in comparison with blood proteins. If the association between a blood protein and an AD-related phenotype is confounded by genetics, then it is likely that a genetic test alone, rather than a protein test, would prove more cost effective as a trial enrichment biomarker. One reason to believe that genetic markers alone may not be optimal is that, it has been shown that common variants, considered additively, explain only 24–33% of the total AD variation.<sup>19,20</sup>

In terms of cognitive tests, the mini mental state examination (MMSE) is known to not be very sensitive for detecting very early AD-related changes in cognition. The paired associates learning (PAL) task of the CANTAB cognition battery is thought to be more sensitive at this stage. A 4-year study by de Jager *et al.*<sup>21</sup> following healthy older adults has shown that CANTAB-PAL, along with another test called the Graded Naming Test, associates with significantly higher risk of conversion to mild cognitive impairment or AD. However, the rate of CANTAB-PAL decline, which may be more informative, has not yet been evaluated in this context. It has also not yet been used as an endophenotype in a blood biomarker study. In relation to age-related cognitive decline, Steves *et al.*<sup>22</sup> have shown that a larger proportion of variance in 10-year change in CANTAB-PAL total errors is explained by the non-shared environment than by additive genetics. A plasma protein biomarker measured in a single blood sample could be more practical in some settings than cognitive testing over a 10-year period in healthy older adults.

In this study, we use the TwinsUK and AddNeuroMed (ANM) SOMAscan data sets. First a heritability analysis is performed and then the association between plasma proteins and AD endophenotypes (10-year change in CANTAB-PAL and regional brain volumes) is examined in asymptomatic individuals.

## MATERIALS AND METHODS

### Ethics statement

Informed consent was obtained for all subjects according to the Declaration of Helsinki (1991), and protocols and procedures were approved by the relevant local ethical committees at each site.

### Discovery cohort—TwinsUK

Subjects used in this study were recruited from TwinsUK, a national register of adult twins.<sup>23</sup> Full details of subject selection is given in Supplementary Methods. In brief, a total of 212 subjects (106 twin pairs) were selected with longitudinal CANTAB-PAL data (~1999 and ~2009).<sup>22</sup> These subjects had also been assessed using MMSE at the ~2009 visit. Fasted EDTA plasma samples, taken a median of around 16 months before CANTAB-PAL and MMSE testing (~2009), were used for proteomic analysis, that is, around 8 years after the ~1999 assessment. Given the high level of missingness in the *APOE E4* genotype, and the small number of double-*APOE E4* carriers, we excluded double-*APOE E4* carriers from further analysis. Subjects who were homozygous for the *APOE E4* allele (10 individuals), or whose plasma sample was haemolysed (six individuals, determined by SomaLogic (Boulder, CO, USA)) or failed SOMAscan quality control (1 individual) were excluded leaving 195 subjects (93 twin pairs and 9 individuals).

### Replication cohort—AddNeuroMed

Subjects used in the replication study were individuals from the ANM cohort;<sup>24,25</sup> full details are given in Supplementary Methods and are summarised here. In total, 254 subjects with both plasma SOMAscan and MRI baseline data, excluding subjects homozygous for the *APOE E4* allele (57 individuals), were available.<sup>9,10</sup> Baseline EDTA plasma samples from fasted subjects were used; see Supplementary Table 1 for a comparison of the sample collection, processing and SOMAscan assay between TwinsUK and ANM. Plasma samples were collected within a year and a half of MMSE testing and MRI scanning. Quality control led to the exclusion of seven subjects from ANM (Supplementary Methods). The remaining subjects were diagnosed as either asymptomatic (control,  $N=91$ ), mild cognitive impairment ( $N=81$ ) or AD ( $N=82$ ).

### Proteomics

SOMAscan methods and data for TwinsUK and ANM have been described previously,<sup>18</sup> but are described again in Supplementary Methods, and are briefly summarised here. A single assay was used per plasma sample. All proteomics data were transformed using the natural logarithm and transformed to zero mean and unit s.d. In addition, protein values  $>2.5$  s.d. from the mean were excluded as outliers.

### Magnetic resonance imaging

Specifics of MRI pre-processing for each cohort are described below. Owing to differences in methods, all MRI data were logged and transformed to zero mean and unit s.d. to increase comparability.

*TwinsUK.* TwinsUK MRI scans were performed ~2 years after plasma sampling (median 805 days, interquartile range (IQR) 151 days). Volumes of the hippocampi and the combined Brodman's areas 28 and 34 (equivalent to the entorhinal cortex) in the TwinsUK cohort were obtained from 38 subjects using Diffeomorphic Anatomical Registration through Exponentiated Lie Algebra,<sup>26</sup> Statistical Parametric Mapping<sup>27</sup> and MarsBar,<sup>28</sup> as fully described in Supplementary Methods.

*AddNeuroMed.* Volumes of the hippocampi and entorhinal cortices in the ANM cohort were obtained using FreeSurfer 5.1.0 (freely available from <http://surfer.nmr.mgh.harvard.edu/>) from 276 ANM subjects who had undergone structural MRI. These regions were selected as they are known to be related to early AD pathology, and were normalised by intracranial volume.<sup>29</sup> Detailed information about data acquisition, pre-processing and quality control assessment have been described for this cohort in detail elsewhere.<sup>29–33</sup>

### APOE genotyping

Full details are given in Supplementary Methods, but in brief single-nucleotide polymorphisms rs429358 and rs7412 were determined from DNA samples using a TaqMan assay.

### Statistical analysis

A detailed account of the statistical methodology is provided in Supplementary Methods, and is summarised here. All statistical analyses were performed in R 3.1.0, except for the transformation and 10-year change calculations for cognitive scores that was performed in STATA 11.<sup>22</sup> All double-*APOE E4* carriers were excluded from analyses. Regressions in the discovery cohort were performed using generalised estimation equations to account for twin dependencies.<sup>34,35</sup> Subject age, gender and recruitment centre were used as co-variables in all relevant regressions. The Benjamini–Hochberg (that is, false discovery rate) multiple testing correction was used to generate  $Q$ -values with thresholds of  $Q < 0.05$  used to indicate association and  $Q < 0.1$  to indicate suggestive association.

Analysis of association between a protein level and 10-year change in CANTAB-PAL in the twin difference context was performed by calculating twin differences in both, and performing a linear regression between the two differences, covarying for twin-pair age. Twin modelling was performed using structural equation modelling to estimate the proportion of variance explained by additive genetics (A), shared environment (C) and non-shared environment (E).<sup>36</sup>

**RESULTS**

Characteristics of the TwinsUK-SOMAscan subcohort

Characteristics of the TwinsUK subjects with SOMAscan data are summarised in Table 1. The MZ and dizygotic twins are matched well, with the exception of *APOE* missingness. Despite not being clinically diagnosed as cognitively impaired, there is variability in the cognitive ability of these subjects (MMSE range 23–30). The subject with MMSE score of 23 was included, as they showed no functional deficit and did not affect the results of the following analyses (data not shown); in all other subjects the MMSE score was >25.

Heritability of plasma protein levels in TwinsUK

Twin modelling was used to estimate the heritability of plasma proteins measured by SOMAscan (Supplementary Table 2). The median proportion of variance of plasma protein levels explained by additive genetics (A) was found to be 26% (IQR 3–46%). The proportion explained by total familial factors (additive genetics (A) or twin-shared environment (C)) was 34% (IQR 3–74%). Finally, a median of 54% (IQR 40–75%) of the variance in plasma protein levels was found to be explained by non-shared environmental factors (E), including technical variability. The results were comparable when a Van der Waerden transformation was used to overcome deviations from the normal distribution (Supplementary Table 2).

Plasma protein markers of cognitive scores in asymptomatic individuals from TwinsUK

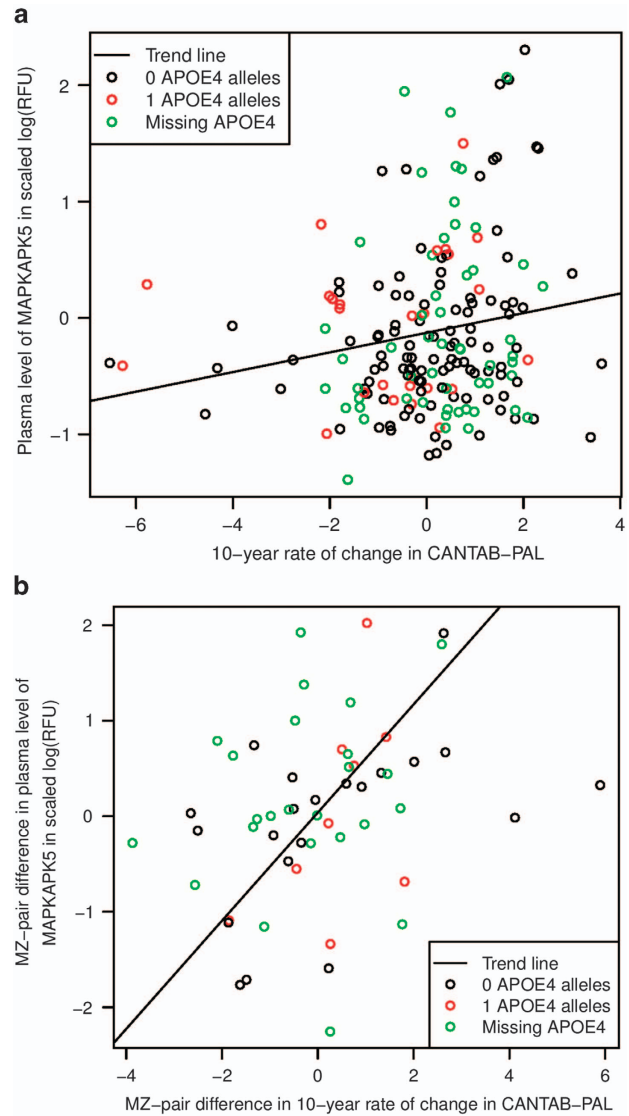
Levels of plasma proteins in TwinsUK subjects were first compared with MMSE scores, both as a continuous and dichotomised variable. No protein was found to be associated in either analysis (Supplementary Tables 4–5).

No protein passed multiple testing corrections for association with CANTAB-PAL total errors in 1999 or 2009; however, MAPKAPK5 was found to be associated with the 10-year change in CANTAB-PAL total errors (Figure 1a;  $\beta=0.48$ ,  $Q=0.0059$ ; Supplementary Tables 5–7).

In total, 54 MZ twin pairs had complete data on plasma MAPKAPK5 levels and 10-year change in CANTAB-PAL total errors. Within these subjects, MAPKAPK5 was found to associate with 10-year change in CANTAB-PAL total errors in the context of MZ-pair differences (Figure 1b;  $\beta=0.55$ ,  $P=0.030$ ). In a twin model, non-shared environmental factors (E) were found to explain 90% (95% confidence interval (CI): 64–100%) of the variance in the plasma level of MAPKAPK5 when a Van der Waerden transformation was applied (A = 10% (CI: 0–36%), C = 0% (CI: 0–23%) and E = 90% (CI: 64–100%); Supplementary Table 1).

Plasma protein markers of MRI measures in asymptomatic individuals from TwinsUK

Plasma protein levels were then compared with selected regional brain volumes derived from structural MRI scans, taking into



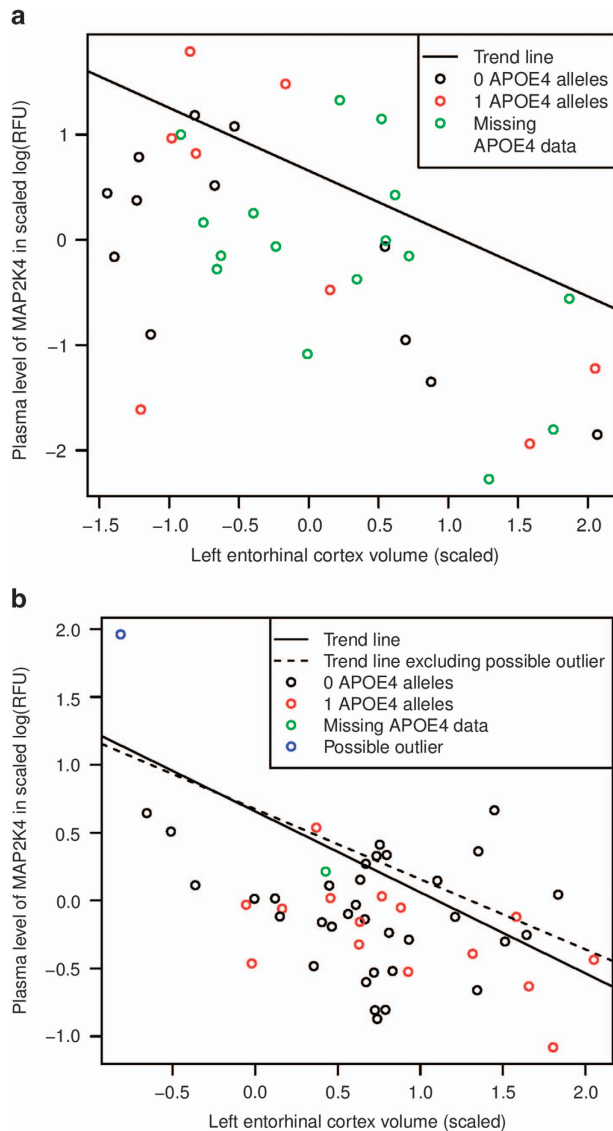
**Figure 1.** The relationship between plasma MAPKAPK5 levels and 10-year change in CANTAB-PAL in TwinsUK. Scatterplot comparing (a) plasma MAPKAPK5 levels and 10-year change in CANTAB-PAL in TwinsUK, or (b) MZ twin differences in both. Points are coloured by *APOE E4* data. (a) GEEs covarying for age and taking into account twin relatedness show an association between plasma MAPKAPK5 and the 10-year change in CANTAB-PAL ( $\beta=0.48$ ,  $Q=0.0059$ ). (b) Linear regression covarying for age shows an association between plasma MAPKAPK5 and the 10-year change in CANTAB-PAL in the MZ twin difference context ( $\beta=0.55$ ,  $P=0.030$ ). CANTAB-PAL, CANTAB-paired associates learning; GEE, generalised estimation equation; MZ, monozygotic; RFU, relative fluorescence unit.

**Table 1.** Characteristics of the discovery cohort (TwinsUK) with SOMAscan data.

	<i>Monozygotic subjects</i>	<i>Dizygotic subjects</i>	<i>Subjects without twin-pair data</i>
Number of subjects	110 (55 pairs)	76 (38 pairs)	9
Number of MRI scan subjects	34	0	2
Age (median (IQR))	65 (62–71)	63 (59–68)	67 (63–74)
% Female	100	100	100
<i>APOE E4</i> (# alleles 0/1 {NA})	44/18 {48}	61/9 {6}	8/0 {1}
MMSE (median (IQR) {# NA})	29 (29–30) {7}	29 (29–30) {2}	28 (27–30)

Abbreviations: IQR, interquartile range; MMSE, mini mental state examination; MRI, magnetic resonance imaging; NA, individuals with missing data for this measure. Double-*APOE E4* carriers were excluded from this study.





**Figure 2.** The association between plasma MAP2K4 levels and (a) left parahippocampal volume in TwinsUK or (b) left entorhinal cortex volume in ANM. Scatterplot comparing plasma MAP2K4 levels with the volume of the left entorhinal cortex are given for the two cohorts. Points are coloured by *APOE E4* data and whether they appear to be outliers. (a) GEEs covarying for age and taking into account twin relatedness show a suggestive association ( $Q < 0.1$ ) between plasma MAP2K4 and the volume of the left entorhinal cortex in TwinsUK. (b) Linear regression covarying for age and centre shows a suggestive association ( $Q < 0.1$ ) between plasma MAP2K4 and the volume of the left entorhinal cortex in ANM, when a possible outlier (in blue) is included ( $\beta = -0.64$ ,  $P = 0.0025$  and  $Q = 0.088$ ), or nominal association ( $P < 0.05$ ) when it is excluded ( $\beta = -0.54$  and  $P = 0.014$ ). ANM, AddNeuroMed; GEE, generalised estimation equation.

account age and twin relatedness (Supplementary Tables 8–11). These selected regions of interest—the left and right hippocampus and entorhinal cortices (LE and RE, respectively)—are known to show early pathology and greater atrophy in AD.<sup>1,37</sup>

Full results are given in Supplementary Tables 8–11. The plasma level of one protein—FAM107B—was found to associate with the volume of both the LE ( $\beta = 1.05$ ,  $Q = 0.034$ ) and RE ( $\beta = 0.77$ ,  $Q = 0.015$ ) cortices. Two proteins associated with the volume of one region ( $Q < 0.05$ ) were also associated with the volume of three out of four regions at a suggestive threshold of  $Q < 0.1$ ,

NSF1C (positively with left hippocampus, LE and RE) and MAP2K4 (negatively with right hippocampus, LE and RE; Supplementary Tables 8–11). The association between MAP2K4 and LE cortex volume in TwinsUK is shown in Figure 2a.

Of the three proteins most consistently associated with multiple brain regions—FAM107B, NSF1C and MAP2K4—only MAP2K4 was also found to be nominally associated with the 10-year change in CANTAB-PAL total errors, (rank = 16th,  $\beta = -0.24$ ,  $P$ -value = 0.026,  $Q \sim 1.0$ ; Supplementary Table 7). Similarly, MAPKAP5 was nominally associated with the volume of the right hippocampus (rank = 93rd,  $\beta = -0.28$ ,  $P$ -value = 0.039,  $Q = 0.48$ ; Supplementary Table 6).

#### Replication of discovery results in ANM

We wished to replicate the association of plasma levels of the proteins FAM107B, NSF1C and MP2K4 with MRI regional volumes in the ANM SOMAScan data. However, as TwinsUK consists of asymptomatic females, whereas ANM also includes males and cognitively impaired individuals, we wished to see whether this had any impact on replication. We therefore performed the replication analysis in ANM in all subjects, in just the asymptomatic (control) subjects and in just the asymptomatic females separately. The characteristics of the relevant ANM subcohorts are given in Table 2.

The plasma level of the protein MAP2K4 in asymptomatic female ANM subjects was found to suggestively associate with the volume of LE ( $\beta = -0.64$ ,  $P = 0.0025$ ,  $Q = 0.088$ ), in a direction consistent with the discovery cohort (Figure 1). Although this trend is visible in Figure 2b, there is a possible outlier remaining in the data set (with the lowest volume and highest protein level). Removing this possible outlier reduces the significance but does not remove the association of MAP2K4 with the volume of LE in asymptomatic females ( $\beta = -0.54$ ,  $P = 0.014$ ).

No other combination of proteins, subjects or regions showed an association (Supplementary Table 12). However, the directions of association were consistent for all the proteins between asymptomatic females from ANM and TwinsUK, and the models including these proteins appeared to explain a higher proportion of the variance ( $R^2$ ) in asymptomatic females versus the other subject groups.

Twin models suggest that plasma levels of MAP2K4 may be more affected by additive genetics than by environmental influences ( $A = 72\%$  (CI: 15–83%),  $C = 0\%$  (CI: 0–55%) and  $E = 28\%$  (CI: 17–39%); Supplementary Table 1).

#### DISCUSSION

In this study we have applied the following methodologies to the study of blood protein biomarkers of AD endophenotypes for the first time: (1) the use of panel-based proteomics in an asymptomatic cohort, (2) the use of CANTAB-PAL as an endophenotype and (3) the use of twin samples. We also believe this is the third discovery study of blood protein markers of AD endophenotypes in asymptomatic individuals, and the largest twin high-throughput proteomics study performed to date.

We are only aware of one other study estimating the heritability of blood protein levels in a high-throughput manner.<sup>17</sup> The estimates from our study suggest a heritability that is at least twice as high on average, further demonstrating the potential for genetic confounding of blood protein biomarkers.

The suggestive/nominal replication of MAP2K4 in ANM, again in asymptomatic individuals, is consistent with the hypothesis that it shows the greatest variability in the asymptomatic phase of AD, which is what is required of an early biomarker. This may be why it has not been reported in previous AD blood biomarker studies, which have mostly focused on later disease stages. As well as being negatively associated with entorhinal cortex volume,

**Table 2.** Characteristics of the replication cohort (ANM) with complete SOMAscan and MRI data.

	All	Controls only	Female controls only
Number of subjects	254	91	51
% Female	100%	100%	100%
Median age (years)	74	72	72
IQR age	70–78	68–76	67–76
APOE E4 (# alleles 0/1 {NA})	158/90 {6}	66/24 {1}	35/15 {1}
Diagnostic groups (control/MCI/AD)	91/81/82	91/--	51/--
Median MMSE by diagnostic group (control/MCI/AD)	29/28/22	29/--	30/--
IQR MMSE (control/MCI/AD)	29–30 / 26–29 / 17 - 25	29–30 /--	29–30 /--

Abbreviations: AD, Alzheimer's disease; ANM, AddNeuroMed; IQR, interquartile range; MCI, mild cognitive impairment; MMSE, mini mental state examination; MRI, magnetic resonance imaging; NA, individuals with missing data for this measure. Double-APOE E4 carriers were excluded from this study.

MAP2K4 has been linked to phosphorylation of tau<sup>38</sup> and of amyloid precursor protein through c-Jun N-terminal kinase.<sup>39</sup> In both cases, these phosphorylation events are believed to lead to increased AD pathology.

MAPKAPK5 has mostly been studied in the context of cancer<sup>40</sup> and rheumatoid arthritis,<sup>41</sup> and, to the best of our knowledge, has not previously been implicated in AD. However, it is interesting that we find two different mitogen-activated protein kinases (MAPKs) whose levels in plasma associate with AD endophenotypes.

The use of longitudinal change in CANTAB-PAL, a cognitive test that is sensitive to early cognitive decline, was critical in the discovery of MAPKAPK5 as a potential biomarker relevant to AD. This demonstrates the importance of studying decline in cognitive abilities using sensitive tests in asymptomatic older adults.

By showing an association between MAPKAPK5 and 10-year change in CANTAB-PAL total errors in the MZ twin difference context, we have demonstrated that the association of MAPKAPK5 and change in cognitive ability in older adults is independent of genetic and twin-shared environmental factors. This is consistent with twin modelling, which suggests that plasma MAPKAPK5 levels are mostly affected by non-shared environmental factors. This has also been demonstrated for 10-year change in CANTAB-PAL total errors in a previous study.<sup>22</sup> Genetic risk, and many aspects of twin-shared environment (such as maternal and family factors) have been implicated in cognitive ageing and dementia, but are currently not modifiable. Given our findings, we hypothesise that MAPKAPK5 may be a biomarker of modifiable cognitive ageing. Selection of individuals for intervention studies whose risk is modifiable may lead to improved outcomes.

The failed replication of findings for proteins FAM107B and NSF11C may indicate that the original findings are false positives, or they may indicate failed replication due to the technical differences or small sample sizes present in this study. However, in asymptomatic females of the replication cohort, the directions of association of all of the proteins were consistent with discovery findings. In addition, the proportion of variance explained by the models was largest in the asymptomatic females. This is consistent with a failure of replication due to small sample size and therefore insufficient statistical power. This suggests that the sample size of this replication cohort may have been a limitation.

There is great demand for inexpensive and relatively non-invasive markers that could be used to identify asymptomatic subjects at risk of AD to recruit into prevention trials. In this study, we started by demonstrating heritability of plasma protein levels, which raise the potential of genetic confounding of plasma protein biomarkers of disease. We then looked for plasma protein biomarkers of AD endophenotypes in asymptomatic twins. The plasma protein level of two MAPKs—MAP2K4 and MAPKAPK5—were identified as possible biomarkers of early AD. MAPKAPK5 was shown to be associated with decline in cognitive ability in the context of MZ twin differences. This suggests that it conveys

information on cognitive ability that is complementary to genetic markers, and may be a biomarker of modifiable cognitive ageing. Future studies will need to assess the specificity of MAPKAPK5 and MAP2K4 to eventual AD, and their potential utility as enrichment biomarkers for clinical trials.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGMENTS

SOMAscan and SOMAmer are trademarks of SomaLogic.

*TwinsUK*: this work was supported by: the EU Framework Programme 7 small-scale focused research collaborative project EurHEALTHAging 277849; *TwinsUK* cognitive phenotyping and imaging and was funded by the Wellcome Trust in a Clinical Research Training Fellowship to Claire Steves (086904 /Z/08/Z). The study also receives support from the European Community's Seventh Framework Programme (FP7/2007–2013), the National Institute for Health Research (NIHR) Clinical Research Facility at Guy's and St Thomas' NHS Foundation Trust and NIHR Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London. TS is an NIHR senior Investigator and is a holder of an ERC Advanced Principal Investigator award.

*ANM*: this work was supported by InnoMed (Innovative Medicines in Europe), an Integrated Project funded by the European Union of the Sixth Framework programme 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; a joint infrastructure grant from Guy's and St Thomas' Charity and the Maudsley Charity; Kuopio University Hospital (HS); and funding from UEFBRAIN (HS).

We would like to acknowledge Abhishek Dixit, Helena Zavos, Maciej Trzaskowski and Fruhling Rijdsdijk for support. SJK is supported by an MRC Career Development Award in Biostatistics (MR/L011859/1). The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no 115372, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007–2013) and in-kind contribution from EFPIA companies. This paper presents an independent research part funded by the NIHR Mental Health Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

## REFERENCES

- Villemagne VL, Burnham S, Bourgeat P, Brown B, Ellis KA, Salvado O *et al*. Amyloid  $\beta$  deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol* 2013; **12**: 357–367.
- Aisen PS, Vellas B, Hampel H. Moving towards early clinical trials for amyloid-targeted therapy in Alzheimer's disease. *Nat Rev Drug Discov* 2013; **12**: 324.
- Lista S, Faltraco F, Hampel H. Blood and plasma-based proteomic biomarker research in Alzheimer's disease. *Prog Neurobiol* 2013; **101–102**: 1–17.
- Zürbig P, Jahn H. Use of proteomic methods in the analysis of human body fluids in Alzheimer research. *Electrophoresis* 2012; **33**: 3617–3630.

- 5 Chiam J, Dobson R, Kiddle SJ, Sattler M. Are blood-based protein biomarkers for Alzheimer's disease also involved in other brain disorders? A systematic review. *J Alzheimers Dis* **43**: 303–314.
- 6 Thambisetty M, Tripaldi R, Riddoch-Contreras J, Hye A, An Y, Campbell J et al. Proteome-based plasma markers of brain amyloid- $\beta$  deposition in non-demented older individuals. *J Alzheimers Dis* **2010**; **22**: 1099–1109.
- 7 Ijsselstijn L, Dekker LJM, Stingl C, van der Weiden MM, Hofman A, Kros JM et al. Serum levels of pregnancy zone protein are elevated in presymptomatic Alzheimer's disease. *J Proteome Res* **2011**; **10**: 4902–4910.
- 8 Kiddle SJ, Sattler M, Proitsi P, Simmons A, Westman E, Bazenec C et al. Candidate blood proteome markers of Alzheimer's disease onset and progression: a systematic review and replication study. *J Alzheimers Dis* **2014**; **38**: 515–531.
- 9 Sattler M, Kiddle SJ, Newhouse S, Proitsi P, Nelson S, Williams S et al. Alzheimer's disease biomarker discovery using SOMAscan multiplexed protein technology. *Alzheimers Dement* **2014**; **10**: 724–734.
- 10 Zhao X, Lejnine S, Spond J, Zhang C, Ramaraj TC, Holder DJ et al. A Candidate plasma protein classifier to identify Alzheimer's disease. *J Alzheimers Dis* **2015**; **43**: 549–563.
- 11 Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One* **2010**; **5**: e15004.
- 12 Lazarus J, Mather KA, Armstrong NJ, Song F, Poljak A, Thalamuthu A et al. DNA methylation in the Apolipoprotein-A1 gene is associated with episodic memory performance in healthy older individuals. *J Alzheimers Dis* **44**: 175–182.
- 13 Jenny van Dongen P, Draisma HH, Martin NG, Boomsma DI. The continuing value of twin studies in the omics era. *Nat Rev Genet* **2012**; **13**: 640–653.
- 14 Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C et al. Variance decomposition of protein profiles identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **2013**; **45**: 1452–1458.
- 15 Jorm AF, Jolley D. The incidence of dementia: a meta-analysis. *Neurology* **1998**; **51**: 728–733.
- 16 Lourdasamy A, Newhouse S, Lunnon K, Proitsi P, Powell J, Hodges A et al. Identification of cis-regulatory variation influencing protein abundance levels in human plasma. *Hum Mol Genet* **2012**; **21**: 3719–3726.
- 17 Kato BS, Nicholson G, Neiman M, Rantalainen M, Holmes CC, Barrett A et al. Variance decomposition of protein profiles from antibody arrays using a longitudinal twin model. *Proteome Sci* **2011**; **9**: 73.
- 18 Menni C, Kiddle SJ, Mangino M, Viñuela A, Psatha M, Steves C et al. Circulating proteomic signatures of chronological age. *J Gerontol A Biol Sci Med Sci* **2014**. e-pub ahead of print 14 August 2014.10.1093/gerona/glu121.
- 19 Ridge PG, Mukherjee S, Crane PK, Kauwe JS. Alzheimer's Disease Genetics Consortium Alzheimer's disease: analyzing the missing heritability. *PLoS One* **2013**; **8**: e79771.
- 20 Lee SH, Harold D, Nyholt DR, ANZGene Consortium, International Endogene Consortium, Genetic and Environmental Risk for Alzheimer's disease Consortium et al. Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. *Hum Mol Genet* **2013**; **22**: 832–841.
- 21 De Jager C, Blackwell AD, Budge MM, Sahakian BJ. Predicting cognitive decline in healthy older adults. *Am J Geriatr Psychiatry* **2005**; **13**: 735–740.
- 22 Steves CJ, Jackson SH, Spector TD. Cognitive change in older women using a computerised battery: a Longitudinal Quantitative Genetic Twin Study. *Behav Genet* **2013**; **43**: 468–479.
- 23 Moayyeri A, Hammond CJ, Hart DJ, Spector TD. The UK Adult Twin Registry (TwinsUK Resource). *Twin Res Hum Genet* **2012**; **16**: 144–149.
- 24 Lovestone S, Francis P, Kloszewska I, Mecocci P, Simmons A, Soininen H et al. AddNeuroMed—the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. *Ann NY Acad Sci* **2009**; **1180**: 36–46.
- 25 Lovestone S, Francis P, Strandgaard K. Biomarkers for disease modification trials—the innovative medicines initiative and AddNeuroMed. *J Nutr Health Aging* **2007**; **11**: 359–361.
- 26 Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage* **2007**; **38**: 95–113.
- 27 Ashburner J, Friston KJ. Unified segmentation. *Neuroimage* **2005**; **26**: 839–851.
- 28 Brett M, Anton JL, Valabregue R, Poline JB. Region of interest analysis using the MarsBar toolbox for SPM 99. *Neuroimage* **2002**; **16**: S497.
- 29 Westman E, Aguilar C, Muehlboeck J-S, Simmons A. Regional magnetic resonance imaging measures for multivariate analysis in Alzheimer's disease and mild cognitive impairment. *Brain Topogr* **2013**; **26**: 9–23.
- 30 Simmons A, Westman E, Muehlboeck S, Mecocci P, Vellas B, Tsolaki M et al. The AddNeuroMed framework for multi-centre MRI assessment of Alzheimer's disease: experience from the first 24 months. *Int J Geriatr Psych* **2011**; **26**: 75–82.
- 31 Simmons A, Westman E, Muehlboeck S, Mecocci P, Vellas B, Tsolaki M et al. MRI measures of Alzheimer's disease and the AddNeuroMed study. *Ann NY Acad Sci* **2009**; **1180**: 47–55.
- 32 Westman E, Simmons A, Muehlboeck J-S, Mecocci P, Vellas B, Tsolaki M et al. AddNeuroMed and ADNI: similar patterns of Alzheimer's atrophy and automated MRI classification accuracy in Europe and North America. *Neuroimage* **2011**; **58**: 818–828.
- 33 Westman E, Simmons A, Zhang Y, Muehlboeck J-S, Tunnard C, Liu Y et al. Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls. *Neuroimage* **2011**; **54**: 1178–1187.
- 34 Carlin JB, Gurrin LC, Sterne JA, Morley R, Dwyer T. Regression models for twin studies: a critical review. *Int J Epidemiol* **2005**; **34**: 1089–1099.
- 35 Hanley JA, Negassa A, Edwards MB, Forrester JE. Statistical analysis of correlated data using generalized estimating equations: an orientation. *Am J Epidemiol* **2003**; **157**: 364–375.
- 36 Boker S, Neale M, Maes H, Wilde M, Spiegel M, Brick T et al. OpenMx: An open source extended structural equation modeling framework. *Psychometrika* **2011**; **76**: 306–317.
- 37 De Toledo-Morrell L, Goncharova I, Dickerson B, Wilson RS, Bennett DA. From healthy aging to early Alzheimer's Disease: in vivo detection of entorhinal cortex atrophy. *Ann NY Acad Sci* **2006**; **911**: 240–253.
- 38 Grueninger F, Bohrmann B, Christensen K, Graf M, Roth D, Czech C. Novel screening cascade identifies MKK4 as key kinase regulating Tau phosphorylation at Ser422. *Mol Cell Biochem* **2011**; **357**: 199–207.
- 39 Mazzitelli S, Xu P, Ferrer I, Davis RJ, Tournier C. The loss of c-Jun N-terminal protein kinase activity prevents the amyloidogenic cleavage of amyloid precursor protein and the formation of amyloid plaques in vivo. *J Neurosci* **2011**; **31**: 16969–16976.
- 40 Sun P, Yoshizuka N, New L, Moser BA, Li Y, Liao R et al. PRAK is essential for ras-induced senescence and tumor suppression. *Cell* **2007**; **128**: 295–308.
- 41 Westhovens R, De Keyser F, Rekalov D, Nasonov EL, Beetens J, Van der Aa A et al. Oral administration of GLPG0259, an inhibitor of MAPKAPK5, a new target for the treatment of rheumatoid arthritis: a phase II, randomised, double-blind, placebo-controlled, multicentre trial. *Ann Rheum Dis* **2013**; **72**: 741–744.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

Supplementary Information accompanies the paper on the Translational Psychiatry website (<http://www.nature.com/tp>)