#### Method S1: Recruitment and assessment of participants

This study was approved by the ethics committees of both Universiti Teknologi MARA (UiTM) [reference no: 600-RMI (5/1/6/01)] and the University of Malaya Medical Centre (UMMC) (reference no: PPUM HU-61/12/1-1). The present experimental design was in accordance with principles and guidelines stipulated by the Declaration of Helsinki, World Medical Association (Carlson, Boyd, and Webb 2004). A written informed consent was obtained from either each patient or legal representative prior to blood collection. Recruitment and assessment procedures were conducted as described in detail by Mohd Hasni et al. (2017) and Rehiman et al. (2020). Basically, participants over the age of 65, who had been diagnosed with any one of these conditions; probable AD, VaD or MCI, were recruited from the UMMC Memory Clinic. Those who were chronic, heavy smokers and/ or alcoholic and/ or presented with signiﬁcant unstable systemic illness or organ failure were excluded. Individuals with signiﬁcant neurological or psychiatric illness other than AD, Mini-Mental State Examination (MMSE) score > 26 and functionally independent based on Instrumental Activity of Daily Living (IADL, 34) and Basic Activity of Daily Living (BADL, 35) were also excluded from the study. Dementia severity was assessed using the clinical dementia rating (CDR) scale (Hughes et al. 1982). The Geriatric Depression Scale (GDS) (Yesavage et al. 1982-1983) was used to rule out those with cognitive impairment due to non-organic disorders. In terms of diagnosis, possible or probable AD subjects were diagnosed based on the National Institute of Neurological and Communicative Disease and Stroke and Alzheimer’s disease (NINCDS-ADRDA) (Varma et al. 1999) and Diagnostic and Statistical Manual of Mental Disorder diagnostic (American Psychiatric Association 1994) criteria. MCI patients were required to fulfil the Petersen’s criteria of amnestic MCI (Petersen et al. 1999). All VaD patients were diagnosed in accordance with the National Institute of Neurological Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN) (Román et al. 1993). The NDC group, on the other hand, was mainly made up of spouses of patients with probable AD or MCI. All participants were subjected to physical and psychometric evaluations as well as routine laboratory blood tests. Medical records and clinical data were further reviewed for comorbid conditions, significant medical history, concurrent drugs used and biochemical results. The training cohort (Phase 1) comprised of 184 participants [92 NDC subjects vs 92 probable AD patients] whilst the validation cohort (Phase 2) was made up of a total of 74 participants (25 NDCs vs 25 probable AD vs 12 MCI vs 12 VaD). Figure S1 illustrates the workflow of the present study.

#### Method S2: Blood biochemical profile and RNA extraction

Plasma concentrations of homocysteine (tHcy), holotranscobalamin (holoTC), vitamin B12, folate and serum concentration of ferritin were determined by the ARCHITECT i System (Abbott Laboratories., Wiesbaden, Germany). The serum lipid profile was determined by Gribbles Pathology Sdn. Bhd., Malaysia.

#### Method S3: Microarray

#### Preparation of microarrays, fluorescently labelled cRNA and hybridisation

Total RNA for microarray was extracted from blood samples using the RibopureTM – Blood RNA Isolation Kit (Ambion, USA) and stored at −80 oC before use. Only high integrity RNA with a cut off > 7.0, 260/ 280 and 260/ 230 ratios > 1.8 were used for subsequent analysis. A total of 184 samples were analysed in the microarray study. The gene expression analysis was based on a one colour microarray experiment using commercial oligonucleotide microarray slide. The Oligonucleotide probe (single-stranded RNA fragment) in the slide was 60 bases (mer) in length. The Agilent SurePrint G3 Human GE 8x60K (Agilent Technologies, CA, USA) with 42,405 oligonucleotide probes (60-mer), representing 29,271 annotated genes were used for hybridisation according to the manufacturer’s instructions. Microarray experiments were with one array per sample. All 184 samples were distributed randomly across 23 microarray slides (8 samples/ slide) by taking into consideration biological replicates of each phenotype (4 AD and 4 NDC samples/ slide). The Cy3–CTP-labelled cRNA target (100 ng RNA) was prepared using the One-Color Low Input Quick Amp Kit and the One-Color Spike-In Kit (Agilent Technologies, CA, USA). Labelled cRNA was purified using the RNeasy Mini Kit (Qiagen, Valencia, CA) and quantified using the Nanodrop ND-1000 VIS spectrophotometer. Specific activity of Cy3 was calculated in pmol per µg with a minimum threshold of 0.825. Labelled cRNA was fragmented by incubation with 5 μL 10X blocking reagent and 1 μL 25X fragmentation buffer in a 25 μL reaction volume at 65 °C for 30 min. GE Hybridisation Buffer Hi-RPM (2X; 25 μL) was added to fragmented cRNA, mixed and placed onto the Agilent SurePrint G3 Human GE 8x60K. Hybridisation was performed at 65 °C for 17 h by rotating the mixture at 10 rpm. Microarray slides were washed in GE Wash Buffer 1 for 2 min and pre-warmed GE Wash Buffer 2 at room temperature for 2 min. The slides were then scanned in an Agilent Scanner G2505B (Agilent Technologies, CA, USA). The scanned image was extracted using the Agilent Feature Extraction Software version 9.5.1.1 (Agilent Technologies, CA, USA).

#### Data processing of microarray

The microarray scanned images were normalised and analysed using the GeneSpring GX 12.0 Software (Agilent Technologies, CA, USA). Signals were normalised against the quantile with threshold raw signal to 1.0 and the median of all samples. Baseline transformation was conducted using the median of all samples as a baseline for batch effect removal. The quality and reproducibility of the microarray data were subjected to the Principal Component Analysis (PCA). This was subsequently followed by a four-step filtering process [by expression, flags, data set and finally 20% coefficient of variation (CV)]. The data has been made available in the National Centre for Biotechnology Information’s (NCBI) Gene Expression Omnibus and are accessible through GEO series accession number GSE85426 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85426).

#### Diagnostic classifier development

The Benjamini Hochberg false discovery rate (FDR) multiple testing correction with *p* < 0.05 was applied on genes with regulation in expression levels of fold change (FC) ≥ 2.0 (Irizarry et al. 2005; Dunckley et al. 2006). Probes that did not have official gene symbols were excluded. The mean of the selected gene expression levels was determined for both the probable AD and NDC groups. After calculation of the z-score for each gene, gene clustering was performed by unsupervised clustering using a hierarchical algorithm (HCA) based on an average linkage method and Euclidean distance created by the GeneSpring GX12. HCA was used to group similar gene expression patterns across probable AD patients and NDCs. Next, multivariate of the partial least square discriminant analysis (PLSDA) was performed using the GeneSpring GX 12.0 (Agilent Technologies, CA, USA) for determination of accuracy, sensitivity and specificity of selected genes. The area under receiver operating curve characteristic (ROC AUC) analysis was then performed to validate the significance of genes (based on AUC) and to determine both positive predictive value (PPV) and negative predictive value (NPV). A rough guide for measuring the utility of a biomarker based on its AUC is as follows: 0.90 – 1.00 = excellent; 0.80 – 0.90 = good; 0.70 – 0.80 = fair; 0.60 – 0.70 = poor; 0.50 – 0.60 = fail. The Pearson’s correlation test was utilised to verify possible confounders in any bivariate correlation (Brazma 2009).

#### Gene ontology and pathway analysis

In order to systematically elucidate the biological functions of genes discovered in AD, gene ontology (GO) was performed using the Database for Annotation, Visualisation and Integrated Discovery (DAVID) v 6.7.14. GO with unadjusted *p* < 0.01. Significant and differentially expressed genes were also subjected to the Ingenuity Pathway Analysis (IPA, Ingenuity Systems, www.ingenuity.com, USA) to highlight deregulated canonical pathway. In the Ingenuity Pathway Knowledge Base, the respective gene identifier was mapped to its analogous gene entities. Consequently, genes from the dataset were put into a global molecular network derived from information contained in the knowledge base to create a gene network. The Fisher’s exact test was used to calculate *p*-value for the likelihood by chance alone of the association between the canonical pathway and the genes in the dataset. Gene networks were then algorithmically generated based on their connectivity. The DAVID v6.7 (http://david.abcc. ncirf.gov) Gene Functional Classification Tool was adopted for functional analysis of the genes.

#### Method S4: Real time quantitative-PCR (RT-qPCR) of blood samples from the training and validation cohorts

Only six genes with FC ≥ 2.0 (namely *DDIT4*, *CNOT8*, *SESN1*, *MAP2K1*, *ABCA9* and *UCP3*) were selected based on their significant up or downregulation (*p* < 0.05) for verification of the microarray results using the RT-qPCR technique. 30 AD subjects and 30 NDCs were randomly selected from the training cohort. Total RNA (1 µg) was reverse transcribed (20 µL) to cDNA using the Tetro cDNA Synthesis Kit (Bioline, USA). The ensemble database was referred for primer design (www.ensembl.org) and synthesised by Integrated DNA Technologies (Table S1). RT-qPCR was performed in triplicates using the Corbett 3000 RotorGene (Corbett Research, NSW, Australia) with a final volume of 20 µL [Sensifast SYBR No ROX Kit (Bioline, USA)]. The conditions used were as follows: initial polymerase enzyme activation step (95⁰C for 2 min), followed by 40 cycles of denaturation (95⁰C for 5 sec) and finally annealing/ extension (60⁰C for 15 sec). Subsequent melt curve analysis was carried out using the following conditions: 95 °C for 15 min, 60 °C for 1 min and 95 °C for 15 min. The housekeeping genes, β-Actin and glyceraldehyde-3-phosphate dehydrogenase (GAPDH), were used as references for normalisation of target cDNAs. The FC in AD was determined by the Pfaffl method (Pfaffl 2001). The panel of six genes was finally evaluated in the validation cohort (independent of microarray study, Figure S1) which comprised of 74 subjects, (25 AD, 25 NDC, 12 MCI and 12 VaD).

**Reference:**

American Psychiatric Association. 1994. *Diagnostic and statistical manual of mental disorders*. 4th edition (DSM-IV) ed. Washington, D C.

Brazma, Alvis. 2009. "Minimum Information About a Microarray Experiment (MIAME)-successes, failures, challenges " *ScientificWorldJournal* 9: 420-423.

Carlson, Robert V, Kenneth M Boyd, and David J Webb. 2004. "The revision of the Declaration of Helsinki: past, present and future." *Br J Clin Pharmacol* 57 (6): 695–713. <https://doi.org/https://doi.org/10.1111/j.1365-2125.2004.02103.x>.

Dunckley, Travis, Thomas G Beach, Keri E Ramsey, Andrew Grover, Diego Mastroeni, Douglas G Walker, Bonnie J LaFleur, Keith D Coon, Kevin M Brown, Richard Caselli, Walter Kukull, Roger Higdon, Daniel McKeel, John C Morris, Christine Hulette, Donald Schmechel, Eric M Reiman, Joseph Rogers, and Dietrich A Stephan. 2006. "Gene expression correlates of neurofibrillary tangles in Alzheimer's disease " *Neurobiol Aging* 27 (10): 1359-1371.

Hughes, C P, L Berg, W L Danziger, L A Coben, and R L Martin. 1982. "A new clinical scale for the staging of dementia " *Br J Psychiatry* 140: 566-572.

Irizarry, Rafael A, Daniel Warren, Forrest Spencer, Irene F Kim, Shyam Biswal, Bryan C Frank, Edward Gabrielson, Joe G N Garcia , Joel Geoghegan, Gregory Germino, Constance Griffin, Sara C Hilmer, Eric Hoffman, Anne E Jedlicka, Ernest Kawasaki, Francisco Martínez-Murillo, Laura Morsberger, Hannah Lee, David Petersen, John Quackenbush, Alan Scott, Michael Wilson, Yanqin Yang, Shui Qing Ye, and Wayne Yu. 2005. "Multiple-laboratory comparison of microarray platforms." *Nat Methods* 2 (5): 345-350.

Mohd Hasni, Dayana Sazereen, Siong Meng Lim, Ai Vyrn Chin, Maw Pin Tan, Philip Jun Hua Poi, Shahrul Bahyah Kamaruzzaman, Abu Bakar Abdul Majeed, and Kalavathy Ramasamy. 2017. "Peripheral cytokines, C-X-C motif ligand10 and interleukin-13, are associated with Malaysian Alzheimer's disease." *Geriatr Gerontol Int* 17 (5): 839-846. <https://doi.org/https://doi.org/10.1111/ggi.12783>.

Petersen, R C, G E Smith, S C Waring, R J Ivnik, E G Tangalos, and E Kokmen. 1999. "Mild cognitive impairment: clinical characterization and outcome." *Arch Neurol* 56 (3): 303-308.

Pfaffl, M W 2001. "A new mathematical model for relative quantification in real-time RT-PCR." *Nucleic Acids Res* 29 (9): e45. <https://doi.org/10.1093/nar/29.9.e45>.

Rehiman, Siti Hajar, Siong Meng Lim, Fei Tieng Lim, Ai-Vyrn Chin, Maw Pin Tan, Shahrul Bahyah Kamaruzzaman, Kalavathy Ramasamy, and Abu Bakar Abdul Majeed. 2020. "Fibrinogen isoforms as potential blood-based biomarkers of Alzheimer's disease using a proteomics approach." *Int J Neurosci*: <https://doi.org/10.1080/00207454.2020.1860038>.

Román, G C, T K Tatemichi, T Erkinjuntti, J L Cummings, J C Masdeu, J H Garcia, L Amaducci, J-M Orgogozo, A Brun, A Hofman, D M Moody, M D O'Brien, T Yamaguchi, J Grafman, B P Drayer, D A Bennett, M Fisher, J Ogata, E Kokmen, F Bermejo, P A Wolf, P B Gorelick, K L Bick, A K Pajeau, M A Bell, C DeCarli, A Culebras, A D Korczyn, J Bogousslavsky, A Hartmann, and P Scheinberg. 1993. "Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop." *Neurology* 43 (2): 250-260.

Varma, A R, J S Snowden, J J Lloyd, P R Talbot, D M Mann, and D Neary. 1999. "Evaluation of the NINCDS-ADRDA criteria in the differentiation of Alzheimer's disease and frontotemporal dementia " *J Neurol Neurosurg Psychiatry* 66 (2): 184-188.

Yesavage, J A, T L Brink, T L Rose, O Lum, V Huang, M Adey, and V O Leirer. 1982-1983. "Development and validation of a geriatric depression screening scale: a preliminary report." *J Psychiatr Res* 17 (1): 37-49.