Supplementary Methods

1.1 Patient recruitment and blood sample collection.

1.1.1 Study subjects.

Subjects were recruited from the Alzheimer's disease (AD) Center at the University of California at Davis. Demographic variables included age, gender, race, years of education, history of hypertension, hyperlipidemia, heart disease, diabetes and WMH. The institutional review board at the University of California at Davis approved this study.

1.1.2 Clinical assessment and Brain imaging.

The medical history was obtained from medical records and current medication based on formal questionnaire. Cognitive evaluations included a Mini-Mental State Examination (MMSE) and the Clinical Dementia Rating Scale (CDR). The diagnosis of AD was made according to the NINCDS and Communication Disorders and Stroke/AD and Related Disorders Association (NINDS-ADRDA) criteria. AD with and without WMH was included in the present study as well as cognitively normal subjects with and without WMH.

Brain imaging was performed using a 1.5 T GE Signa Horizon LX Echospeed MRI scanner. A T1-weighted coronal three-dimensional spoiled gradient-recalled echo acquisition and a FLAIR sequence designed to enhance WMH segmentation were employed. WMH segmentation and WMH volume measurements were performed as described previously [1]. All the WMH+ (WMH positive) cases had head size adjusted WMH volume above the 75th percentile of a normal population (n=20, 11 non-demented control subjects and 9 AD patients) and all the WMH- (WMH negative) subjects had

head size adjusted WMH volume below the 25th percentile of a normal population (n=18, 10 non-demented control subjects and 8 AD patients).

1.1.3 Blood collection and processing.

Whole blood (15 mL) was collected from each subject into six PAXgene tubes via antecubital fossa venipuncture. After 2h at room temperature, PAXgene tubes were frozen at -80^{0}C. Total RNA was isolated according to the manufacturer's protocol (PAXgene blood RNA kit; Pre-AnalytiX) on an automated workstation Qiacube (Qiagen, Valencia, CA). The RNA is from polymorphonuclear cells (neutrophils, basophils, and eosinophils), mononuclear cells (lymphocytes and monocytes/macrophages), and platelet precursors and red blood cell precursors.

1.2 Microarray experiment and data analysis.

1.2.1 cDNA Synthesis

Biotin-labeled cDNA was synthesized from 50 ng of total RNA using the Ovation™ Whole Blood Solution system according to the manufacturer's protocol (NuGEN Technologies Inc., San Carlos, CA, USA). Ovation® Whole Blood Solution system consists of optimized protocols for global gene expression profiling of whole blood RNA without the need for globin reduction procedures. Details of the procedures have been published previously [1].

1.2.2 Microarray hybridization.

Subsequent biotin-labeled cDNA hybridization to Affymetrix Human U133 Plus 2.0 Arrays (Affymetrix, Santa Clara, CA, USA) was performed according to standard Affymetrix Protocols (Affymetrix Expression Analysis Technical Manual). Affymetrix Human U133 Plus 2.0 Arrays contain more than 54,000 probe sets, which analyze the

expression level of over 47,000 transcripts and variants, including >20,000 well-characterized human genes.

1.2.3 Analysis of Demographic Data and Microarray Analysis in Partek Genomics Suite.

Demographic data were analyzed with Student's t-test or Fisher's exact tests. Microarray probe-level data were summarized with Robust Multi-array Average (RMA) algorithm. Quality control analysis and Analysis of Covariance (ANCOVA) were conducted in Partek Genomics Suite (http://www.partek.com, Partek Inc., St. Louis, MI, USA). A multivariate ANCOVA was performed on AD status with adjustments for potential confounders, including sample batch, sex, age, heart disease, hyperlipidemia, years of education and presence of WMH. Principal Component Analysis (PCA) and unsupervised hierarchical clustering analysis were then performed in Partek Genomics Suite based on genes identified as differentially regulated to visualize the differential expression pattern between subjects with and without AD and with and without WMH.

1.2.4 Prediction Analysis of Microarrays (PAM) Analysis.

Prediction Analysis of Microarrays (PAM) analysis was performed to illustrate the feasibility of using the identified genomic expression profiles to predict the presence of AD, independent of whether WMH were present or not [1]. PAM performs sample classification using the nearest shrunken centroid method. After training with PAM, self cross-validation was conducted with leave-one-sample-out-at-a-time method.

1.2.5 Differentially expressed genes filtering and functional annotation.

We chose 0.005 and 1.2 as the detection p value and absolute fold change value threshold respectively to obtained a list of differentially expressed genes (DEGs) between AD and controls, and performed gene sets enrichment analysis (GSEA) on the DEG list using the DAVID Functional Annotation web tool online. (http://david.abcc.ncifcrf. gov/tools.jsp). The results of GSEA analysis were filtered for p value < 0.01 and FDR < 0.05.

1.3 Other Microarray Datasets for Data analysis.

1.3.1 Microarray Datasets for AD.

Microarray datasets for AD blood transcriptome studies were downloaded from the online databases -- Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/). The major datasets re-analyzed in this study included GSE6613 [2] and GSE4229 [3] from GEO. In addition, DEG lists were retrieved directly from published papers by Fehlbaum et al., Rye et al. and Lunnon et al. [4-6]. We also conducted a re-analysis of our independent blood microarray study in a Chinese cohort [7].

1.3.2 Microarray Data Analysis

Microarray datasets were pre-processed and filtered and the results were mapped from probes to Entrez gene IDs. Differential expression genes were calculated using the method of RankProd [8]. Functional enrichment analysis was based on one-tailed Fisher's exact test and multiple comparison corrections were performed using the False Discovery Rate approach (FDR $p<0.05$), and functional categories were those

used in the GO database (http://www.geneontology.org/). The detailed procedures have been published previously [9, 10].

REFERENCES

1. Xu H, Stamova B, Jickling G, Tian Y, Zhan X, Ander BP, Liu D, Turner R, Rosand J, Goldstein LB *et al*: **Distinctive RNA expression profiles in blood associated with white matter hyperintensities in brain**. *Stroke* 2010, **41**(12):2744-2749.
2. Scherzer CR, Eklund AC, Morse LJ, Liao Z, Locascio JJ, Fefer D, Schwarzschild MA, Schlossmacher MG, Hauser MA, Vance JM *et al*: **Molecular markers of early Parkinson's disease based on gene expression in blood**. *Proc Natl Acad Sci U S A* 2007, **104**(3):955-960.
3. Maes OC, Schipper HM, Chertkow HM, Wang E: **Methodology for discovery of Alzheimer's disease blood-based biomarkers**. *J Gerontol A Biol Sci Med Sci* 2009, **64**(6):636-645.
4. Fehlbaum-Beurdeley P, Jarrige-Le Prado AC, Pallares D, Carriere J, Guihal C, Soucaille C, Rouet F, Drouin D, Sol O, Jordan H *et al*: **Toward an Alzheimer's disease diagnosis via high-resolution blood gene expression**. *Alzheimer's & dementia : the journal of the Alzheimer's Association* 2010, **6**(1):25-38.
5. Rye PD, Booij BB, Grave G, Lindahl T, Kristiansen L, Andersen HM, Horndalsveen PO, Nygaard HA, Naik M, Hoprekstad D *et al*: **A novel blood test for the early detection of Alzheimer's disease**. *Journal of Alzheimer's disease : JAD* 2011, **23**(1):121-129.
6. Lunnon K, Ibrahim Z, Proitsi P, Lourdusamy A, Newhouse S, Sattlecker M, Furney S, Saleem M, Soininen H, Kloszewska I *et al*: **Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood**. *Journal of Alzheimer's disease : JAD* 2012, **30**(3):685-710.
7. Han G, Wang J, Zeng F, Feng X, Yu J, Cao HY, Yi X, Zhou H, Jin LW, Duan Y *et al*: **Characteristic Transformation of Blood Transcriptome in Alzheimer's Disease**. *Journal of Alzheimer's disease : JAD* 2013, **35**:373-386.
8. Hong F, Breitling R, McEntee CW, Wittner BS, Nemhauser JL, Chory J: **RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis**. *Bioinformatics* 2006, **22**(22):2825-2827.
9. Liang D, Han G, Feng X, Sun J, Duan Y, Lei H: **Concerted perturbation observed in a hub network in Alzheimer's disease**. *PLoS One* 2012, **7**(7):e40498.
10. Sun J, Feng X, Liang D, Duan Y, Lei H: **Down-regulation of energy metabolism in Alzheimer's disease is a protective response of neurons to the microenvironment**. *Journal of Alzheimer's disease : JAD* 2012, **28**(2):389-402.