

RESEARCH

Open Access



Blood and brain transcriptome analysis reveals *APOE* genotype-mediated and immune-related pathways involved in Alzheimer disease

Rebecca Panitch¹, Junming Hu¹, Weiming Xia^{2,3}, David A. Bennett⁴, Thor D. Stein^{3,5}, Lindsay A. Farrer^{1,6,7,8,9} and Gyungah R. Jun^{1,7,8*} 

Abstract

Background: While Alzheimer disease (AD) is generally considered as a brain disorder, blood biomarkers may be useful for the diagnosis and prediction of AD brain pathology. The *APOE* $\epsilon 4$ allele has shown cerebrovascular effects including acceleration of blood-brain barrier (BBB) breakdown.

Methods: We evaluated the differential expression of previously established AD genes in brains from 344 pathologically confirmed AD cases and 232 controls and in blood from 112 pathologically confirmed AD cases and 67 controls from the Religious Orders Study and Memory and Aging Project. Differential gene expression between AD cases and controls was analyzed in the blood and brain jointly using a multivariate approach in the total sample and within *APOE* genotype groups. Gene set enrichment analysis was performed within *APOE* genotype groups using the results from the combined blood and brain analyses to identify biologically important pathways. Gene co-expression networks in brain and blood samples were investigated using weighted correlation network analysis. Top-ranked genes from networks and pathways were further evaluated with vascular injury traits.

Results: We observed differentially expressed genes with $P < 0.05$ in both brain and blood for established AD genes *INPP5D* (upregulated) and *HLA-DQA1* (downregulated). *PIGHP1* and *FRAS1* were differentially expressed at the transcriptome-wide level ($P < 3.3 \times 10^{-6}$) within $\epsilon 2/\epsilon 3$ and $\epsilon 3/\epsilon 4$ groups, respectively. Gene set enrichment analysis revealed 21 significant pathways (false discovery rate $P < 0.05$) in at least one *APOE* genotype group. Ten pathways were significantly enriched in the $\epsilon 3/\epsilon 4$ group, and six of these were unique to these subjects. Four pathways (allograft rejection, interferon gamma response, peroxisome, and TNFA signaling via NFKB) were enriched for AD upregulated genes in the $\epsilon 3/\epsilon 4$ group and AD downregulated genes in subjects lacking $\epsilon 4$. We identified a co-expressed gene network in the brain that reproduced in blood and showed higher average expression in $\epsilon 4$ carriers. Twenty-three genes from pathway and network analyses were significantly associated with at least one vascular injury trait.

Conclusion: These results suggest that the *APOE* genotype contributes to unique expression network profiles in both blood and brain. Several genes in these networks are associated with measures of vascular injury and potentially contribute to $\epsilon 4$'s effect on the BBB.

Keywords: Alzheimer's disease, Blood-brain barrier, *APOE*, Differential expression, Co-expression network, Vascular injury

*Correspondence: gyungah@bu.edu

¹ Department of Medicine (Biomedical Genetics), Boston University School of Medicine, 72 East Concord Street, Boston, MA 02118, USA
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Alzheimer disease (AD) is a neurodegenerative disorder characterized by amyloid plaques and neurofibrillary tau tangles in the brain [1]. Because these hallmark proteins are sometimes detectable in blood before clinical symptoms appear, there are on-going efforts to identify blood-based signatures from multi-omics and biomarker data that can facilitate detection of AD preclinically [2, 3]. For example, plasma phosphorylated tau levels are highly correlated with neurodegenerative disorders and AD pathology [4, 5].

Cerebrovascular AD-related pathology that may affect the blood-brain barrier (BBB), such as cerebral amyloid angiopathy (CAA), has been shown to exacerbate neurodegeneration and neuroinflammation [6]. Dysfunction of the BBB, a semi-permeable border separating the extracellular fluid and brain tissue from circulating blood, has been implicated in the accumulation of amyloid- β (A β) and hyperphosphorylation of tau protein [7, 8]. Apolipoprotein E (*APOE*) genotype is the strongest genetic risk factor for late-onset AD and the $\epsilon 4$ allele has been recently associated with BBB dysfunction leading to cognitive decline [9, 10]. Heterozygosity of the *APOE* $\epsilon 4$ allele confers a 3–4-fold increase of AD risk and $\epsilon 4$ homozygotes have a 10–12-fold increased likelihood of a clinical diagnosis of AD among persons of European ancestry [9, 11]. By contrast, among clinically and neuropathologically confirmed AD cases and controls of European ancestry, a single copy of the *APOE* $\epsilon 2$ allele is associated with 0.61-fold decreased risk and $\epsilon 2$ homozygotes have an 0.87-fold reduced risk for AD compared to individuals with the $\epsilon 3/\epsilon 3$ genotype [12]. Cerebrovascular AD-related pathologies have also shown *APOE* genotype-dependent patterns. Both $\epsilon 2$ and $\epsilon 4$ are significantly associated with the risk of CAA [13].

Previous whole transcriptome-wide studies from autopsied brains demonstrate that the classical complement cascade and tau phosphorylation are linked to AD in an *APOE* genotype-specific manner [14, 15]. However, expression profiles associated with AD have not been intensively investigated in the blood and brain from the same individuals, especially separated by the *APOE* genotype. Here, we analyzed gene expression measured in the blood and brain tissue obtained from participants of the Religious Orders Study and Rush Memory and Aging Project (ROSMAP) [16] stratified by the *APOE* genotype in order to discern AD-related differential gene expression, biological pathways, and gene networks shared in the blood and brain.

Methods

Sources of blood transcriptomic and phenotypic data

RNA-sequencing (RNA-seq) data generated from blood donated by 614 ROSMAP participants and phenotypic data collected from those subjects were obtained from the Synapse portal [17]. RNA batches were prepared using a SMART-seq2 protocol (batches 1-2) or a SMART-seq2-like protocol (batch 3). Batch 1 containing 47 samples (2×101 bp) and batch 2 containing 201 samples (2×76 bp) were pooled and sequenced by HiSeq 2500 (Illumina). Batch 3 containing 366 samples (2×50 bp) was pooled and sequenced on Nova Seq 6000 (Illumina) (Supplementary Table 1). A post-mortem diagnosis of AD was established for 112 participants using NIA-Reagan criteria including Braak staging for assessing the severity of neurofibrillary tangles and the Consortium to Establish a Registry for Alzheimer Disease (CERAD) semi-quantitative measure for neuritic plaques (CERAD score). Another 67 participants who were clinically normal showed no pathological evidence of AD and were included in this study as controls (Table 1) [16, 18]. Age, sex, sequencing batch, and library batch information was available for all subjects.

Sources of brain transcriptomic and phenotypic data

Publicly available prefrontal cortex brain RNA-seq and neuropathological data for 639 ROSMAP participants were obtained from the Synapse portal [17] (Supplementary Table 1). Sequencing libraries were prepared using the strand-specific dUTP method with poly-A selection, and all samples were sequenced using an Illumina HiSeq instrument. Of these 639 samples, data from 576 samples with both RNA integrity number (RIN) and post-mortem interval (PMI) were included in subsequent differential expression analyses (Table 1). Samples with RIN < 5 were excluded from further study. Previously reported RNA-seq data were also available which were derived from the frontal cortex tissue region of 208 frontal autopsied brains (64 AD and 129 controls) donated to the Framingham Heart Study and Boston University Alzheimer's Disease Center (FHS/ADRC) [19]. A diagnosis of AD in these brains was established using NIA-Reagan criteria including Braak staging and CERAD score [19].

Quality control, mapping, and quantification of gene expression data and sample

The 614 FASTQ files derived from blood RNA-seq data were processed in batches. Quality control (QC) of the sequence data was performed using FastQC which checked for overabundance of adaptors and overrepresented sequences [20]. Reads passing initial QC were aligned to the human reference genome (GRCh38.95)

Table 1 Number of ROSMAP participants with RNA-seq data by *APOE* genotype and batch

| APOE genotype | Blood batch 1 | | Blood batch 2 | | Combined blood | | Brain | Blood-brain overlap | | | |
|------------------------------|---------------|------|---------------|------|----------------|------|-------|---------------------|----|------|--|
| | AD | CTRL | AD | CTRL | AD | CTRL | AD | CTRL | AD | CTRL | |
| $\epsilon 2/\epsilon 2^{**}$ | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 5 | 0 | 1 | |
| $\epsilon 2/\epsilon 3$ | 5 | 3 | 9 | 9 | 14 | 12 | 32 | 39 | 12 | 8 | |
| $\epsilon 3/\epsilon 3$ | 16 | 11 | 44 | 31 | 60 | 42 | 197 | 158 | 49 | 33 | |
| $\epsilon 2/\epsilon 4^{**}$ | 1 | 1 | 5 | 0 | 6 | 1 | 9 | 4 | 4 | 1 | |
| $\epsilon 3/\epsilon 4^*$ | 8 | 1 | 21 | 10 | 21 | 10 | 101 | 25 | 22 | 8 | |
| $\epsilon 4/\epsilon 4^{**}$ | 1 | 0 | 1 | 0 | 0 | 0 | 5 | 1 | 2 | 0 | |
| NA | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | |
| Total | 31 | 16 | 81 | 51 | 112 | 67 | 344 | 232 | 89 | 51 | |

Blood batch 3 was excluded because it contained controls only

**APOE* $\epsilon 3/\epsilon 4$ subjects from blood batch 1 were excluded from all analyses due to the small sample size

** $\epsilon 2/\epsilon 2$, $\epsilon 2/\epsilon 4$, and $\epsilon 4/\epsilon 4$ subjects were not analyzed separately due to the small sample size

NA, *APOE* genotype not available

using STAR (version 2.6.1c), which implements 2-pass mapping to increase the chances of mapping splice reads from novel junctions [21, 22]. To account for differences in read lengths between batches, we created three genomic alignment index files with read lengths of 50bp, 76bp, and 101bp, respectively, for mapping the study samples to the reference genome.

The 639 binary alignment map (BAM) files containing brain RNA-seq data required additional processing before alignment and thus were converted to FASTQ files using the FastqTosam function in Picard tools [23]. Samples were checked for adaptor overabundance and overrepresented sequences using FastQC [20]. Paired-end reads were aligned to the human reference genome as described above. In order to map brain samples to the reference genome, genomic index files (read length = 101bp) were created.

The resulting BAM files for each brain and blood sample contained mapped paired-end reads and a corresponding alignment report file. Gene and isoform levels were quantified using RSEM (version 1.3.1) [24] and Bowtie2 (version 2.3.4.1) [25] and then annotated using *Homo sapiens* GRCh38.95.gtf annotation files. Files generated by this process for each sample contained several variables for each gene including gene id, gene length, effective gene length, expected count, counts per million (CPM), and fragments per kilobase of exon model per million reads mapped (FPKM) reads.

Gene expression analysis in the blood and brain

Differential expression analysis

Genes with less than two reads on average among 80% or more of the samples were excluded from analyses. Blood and brain samples were corrected for between-sample variability using a trimmed mean of *M*-value

normalization method [26]. Differential gene expression analysis between AD and control subjects in the blood and brain was performed separately using the VOOOM and LIMMA software [27, 28]. For differential gene expression analysis in the brain, the normalized expression of each gene was compared between AD cases and controls using linear regression models adjusting for sex, age at death, RNA integrity number (RIN), post-mortem interval (PMI), and sequencing batch as covariates. Gene expression analysis of the blood samples included only the 179 individuals who were neuropathologically examined and models included covariates for sex, age at exam, and library batch. A total of 140 individuals had genomic data derived from the blood and brain and included in analyses for both tissues. Analyses were performed in the total sample and subgroups defined by the *APOE* genotype ($\epsilon 2/\epsilon 3$, $\epsilon 3/\epsilon 3$, and $\epsilon 3/\epsilon 4$). Subjects with genotypes $\epsilon 2/\epsilon 2$, $\epsilon 2/\epsilon 4$, and $\epsilon 4/\epsilon 4$ were excluded from analyses due to small sample sizes (Supplementary Table 1, Table 1). Analyses of gene expression in blood were further stratified by RNA batch due to differences in read length and sample substructure (Supplementary Fig. 1a), while we did not observe batch differences in the brain (Supplementary Fig. 1b). For the $\epsilon 3/\epsilon 4$ subgroup, data from batch 2 in blood were only analyzed because the batch 1 sample size was too small (Table 1). Analyses were not stratified by batch in the brain because there was no obvious batch effect and sample sizes in each of the nine batches were too small, especially within the $\epsilon 2/\epsilon 3$ and $\epsilon 3/\epsilon 4$ subgroups (Supplementary Table 2). Results from analyses of each batch and *APOE* genotype group were combined by meta-analysis weighting for the number of AD cases and accounting for effect direction using the METAL program [29].

To evaluate differential gene expression patterns in the joint blood and brain datasets, we combined univariate results from the blood and brain using the R package CUMP which incorporates O'Brien's method [30]. In this method, a combined Z -score was calculated using t -value estimates derived from the LIMMA linear regression analyses and/or from Z -score estimates from the meta-analysis of the blood batches. All analyses were weighted by the number of samples within *APOE* genotype groups or in the total sample.

Single-cell gene expression analysis

A normalized single-cell RNA-sequencing expression matrix from ~2400 cells collected from the blood of healthy individuals ages 25 to 40 and proportions of dendritic cells and monocytes for each sample were obtained from the Single Cell Portal [31]. Additional details of these subjects and single-cell RNA sequencing are reported elsewhere [32]. FASTQ single nuclei RNA-sequencing data from the prefrontal cortex of 48 brains from ROSMAP participants (24 AD cases, 24 controls) were obtained from the Synapse portal [17] and processed as previously described [19]. Among the 48 individuals in this dataset, 26 are included in the ROSMAP bulk brain RNA-seq dataset and four of these 48 individuals overlap with the ROSMAP bulk blood RNA-seq dataset. The average expression for each cell type in the blood and brain RNA-seq datasets was calculated for each gene.

Gene set enrichment analysis

Differentially expressed genes in the total sample or within *APOE* genotype groups were ranked by a combined Z -score from the blood and brain using the O'Brien method. Gene set enrichment analysis was performed using this ranked list and hallmark gene set pathway information obtained from the Molecular Signatures Database (MSigDB) as previously described [33, 34]. The hallmark gene set is focused on biological processes obtained by aggregated MsigDB signatures. Pathway enrichment scores were determined based on the degree to which a set of genes was overrepresented by the largest positive and smallest negative Z -scores. Genes that contributed the most to the enrichment score of each pathway were designated as leading-edge genes.

Co-expressed gene network analysis in the blood and brain

Co-expressed genes in networks were identified using 14,456 coding genes in the brain and 11,379 coding genes in the blood in the ROSMAP RNA-seq dataset using the weighted gene correlation network analysis (WGCNA) algorithm [35]. Analyses of data from blood included

only 141 batch 2 samples with and without post-mortem examination to avoid batch effects, and analyses of data from the brain comprised 636 samples excluding lack of RIN or batch information (Supplementary Table 3). We used gene expression levels calculated as log-transformed fragments per kilobase of transcript per million (FPKM). Soft-power parameters of 12.0 and 12.5 were selected for analyses of brain and blood data, respectively, as previously described [19]. Expression data were clustered hierarchically by implementing a dissimilarity topological overlap matrix (TOM). Initial modules with a minimal network size of 100 genes were identified and labeled using dynamic tree cutting. Eigengenes were derived from the first principle component for each module and served as representative values of gene expression in a given module [36]. Networks with high eigengene similarity and a height of 0 were merged using the mergeCloseModules function in WGCNA. Fuzzy module membership was assigned using the signedKME function.

Network modules identified in the brain were examined for preservation in blood using the modulePreservation function in WGCNA. Brain networks with a Z summary score > 5 were considered preserved in blood networks [37]. Relevance of the networks to AD pathology was established based on enrichment of AD-related genes that was determined using the userListEnrichment function in WGCNA. For the purpose of this analysis, we defined AD-related genes which included those within 20kb of single nucleotide polymorphisms (SNPs) showing at least modest evidence ($P < 0.001$) for association with AD risk [38] or AD-related neuropathological measures of Tau and A β proteins [39]. We used EnrichR to identify KEGG pathways enriched for AD-related genes in the preserved networks [40]. Next, genes in networks contributing to significant pathways were further evaluated using Ingenuity Pathway Analysis software (QIAGEN Inc.) to identify biological subnetworks.

Measurements and association with vascular injury-related proteins

Intercellular adhesion molecule 1 (ICAM-1), vascular cell adhesion molecule 1 (VCAM-1), and serum amyloid α (SAA) were detected and measured in fresh tissue lysate from the dorsolateral prefrontal cortex area using the Mesoscale Discovery V-PLEX Plus Vascular Injury Panel Kit (Mesoscale Discovery, K15198G, Rockville MD). Gray matter was separated from frozen brain tissue on dry ice and weighed. Ice-cold RIPA buffer (ThermoScientific, #89901) was added to the gray matter at 5mL RIPA: 1g brain wet weight, and homogenized with Qiagen Tissue Lyser LT at 50Hz for 5 min (Qiagen, Germany) (ThermoScientific, Waltham MA). The homogenate was

centrifuged at 17,000g at 4°C for 15 min, then the supernatant was aliquoted and stored at -80°C until use. Buffers and immunoassay plates were prepared according to the manufactory instructions and the brain homogenate was further diluted 5-fold. The immunoassay plates were read using the multi-detection SPECTOR 6000 Imager to quantitate protein levels (Mesoscale Discovery).

Additional analyses of 107 top-ranked genes emerging from pathways identified by differential gene expression and network analyses were performed using log-transformed FPKM values obtained previously from these FHS/ADRC donor brains [19]. Levels of ICAM-1, VCAM-1, and SAA proteins were rank-transformed after adjusting for age and sex. We performed association analyses using the expression levels of the selected genes with the levels of vascular injury-related proteins as quantitative outcomes in linear regression models further adjusting for RIN.

Results

Differentially expressed genes in the blood and brain

Gene expression levels in 179 blood and 576 brain samples from the ROSMAP dataset were compared between AD cases and controls (Fig. 1). In the total sample, no genes in the combined data from the blood and brain were differentially expressed at the transcriptome-wide significance level ($P < 3.3 \times 10^{-6}$). Of 78 genes containing or nearest to SNPs associated with AD at a genome-wide significance level in a recent large genome-wide association study (GWAS) [41], 64 passed QC and were expressed in both brain and blood. The expression of five of these 64 genes (*HLA-DQA1*, *INPP5D*, *SPDYE3*, *TSPOAPI*, and *SIGLEC11*) were nominally significant ($P < 0.05$) in the analysis of the combined blood and brain data (Table 2, Supplemental Table 4, Supplementary Fig. 2). Differential expression of *HLA-DQA1* and *INPP5D* was nominally significant at $P < 0.05$ in both blood and brain with the same direction of effect. Differentially expressed genes (DEGs) after multiple testing correction at $P < 6.4 \times 10^{-4}$ were evident only in the brain for *BCKDK* ($P = 5.1 \times 10^{-4}$, $P_{\text{adj}} = 0.04$), *TSPOAPI* ($P = 2.6 \times 10^{-4}$, $P_{\text{adj}} = 0.02$), and *SIGLEC11* ($P = 1.6 \times 10^{-4}$, $P_{\text{adj}} = 0.01$).

Two genes were differentially expressed between AD cases and controls at the transcriptome-wide level ($P <$

3.3×10^{-6}) within a particular *APOE* genotype group. *PIGHP1* was significantly upregulated in AD cases in the combined brain and blood samples in the $\epsilon 2/\epsilon 3$ group ($Z = 4.67$, $P = 3.1 \times 10^{-6}$, $P_{\text{adj}} = 0.05$), a pattern predominated by the evidence in the brain but also apparent in the blood (Table 3, Supplementary Figs. 2, 3a). Among $\epsilon 3/\epsilon 4$ subjects, the expression of *FRAS1* was significantly downregulated in AD cases in blood only ($Z = -4.66$, $P = 3.2 \times 10^{-6}$, $P_{\text{adj}} = 0.05$) (Supplementary Fig. 3b). No transcriptome-wide significant DEGs were identified in the brain from any *APOE* genotype groups. Among genes previously associated with AD among $\epsilon 2/\epsilon 3$ subjects [19], *C4A*, *C4B*, and *HSPA2* were moderately ($P < 10^{-3}$) upregulated in the brain but not blood from AD ROSMAP Study participants in the $\epsilon 2/\epsilon 3$ subgroup and total sample (Supplemental Table 5). Notably, *C4B* expression trended in the opposite direction (i.e., downregulated in AD cases) in blood from $\epsilon 2/\epsilon 3$ subjects ($P = 0.08$, $P_{\text{adj}} = 1.0$).

Examination of cell-level expression profiles of the DEGs in Tables 1 and 2 revealed that in blood cell types *HLA-DQA1* and *INPP5D* were more highly expressed in dendritic cells and monocytes compared to other genes in this group (Supplementary Fig. 4a). *INPP5D* was the only gene in this group expressed in brain cell types and specifically in microglia (Supplementary Fig. 4b).

APOE genotype-dependent pathways in combined blood and brain expression profiles

We identified 21 pathways that were significantly enriched for upregulated or downregulated genes in the combined blood and brain expression levels in at least one *APOE* genotype group (Fig. 2a and Supplementary Table 6). Enrichment scores from significant pathways identified in the $\epsilon 3/\epsilon 4$ group were generally downregulated and had the opposite effect direction compared to those for the other *APOE* genotype groups (Fig. 2a, Table 4). Six pathways were significantly and uniquely enriched in the $\epsilon 3/\epsilon 4$ subgroup including apoptosis, estrogen response late, hypoxia, il6/jak/stat3 signaling, inflammatory response, and p53 pathway. Pathways for allograft rejection, interferon gamma response, peroxisome, and TNFA signaling via NFkB were enriched for upregulated AD genes in the $\epsilon 3/\epsilon 4$ group but for downregulated AD genes in the $\epsilon 2/\epsilon 3$ and $\epsilon 3/\epsilon 3$ groups.

(See figure on next page.)

Fig. 1 Analysis design and workflow. RNA-sequencing data were obtained from the blood and post-mortem frozen brain of neuropathologically verified AD cases and controls. Data were analyzed in two ways. First, gene co-expression analysis identified networks in the brain that reproduced in the blood. A second analysis identified genes differentially expressed between AD cases in controls in the total sample as well as within *APOE* genotype groups in both blood and brain. The expression of genes in the co-expression networks that were previously associated with AD by GWAS was tested for associated with AD-related traits measured in the brain. Next, genes in significant co-expression networks and differentially expressed genes in the blood and brain were incorporated as seeds in pathway analysis. Finally, the expression of genes from the most significant pathways was tested for association with levels of several vascular damage proteins. Figure created with BioRender.com

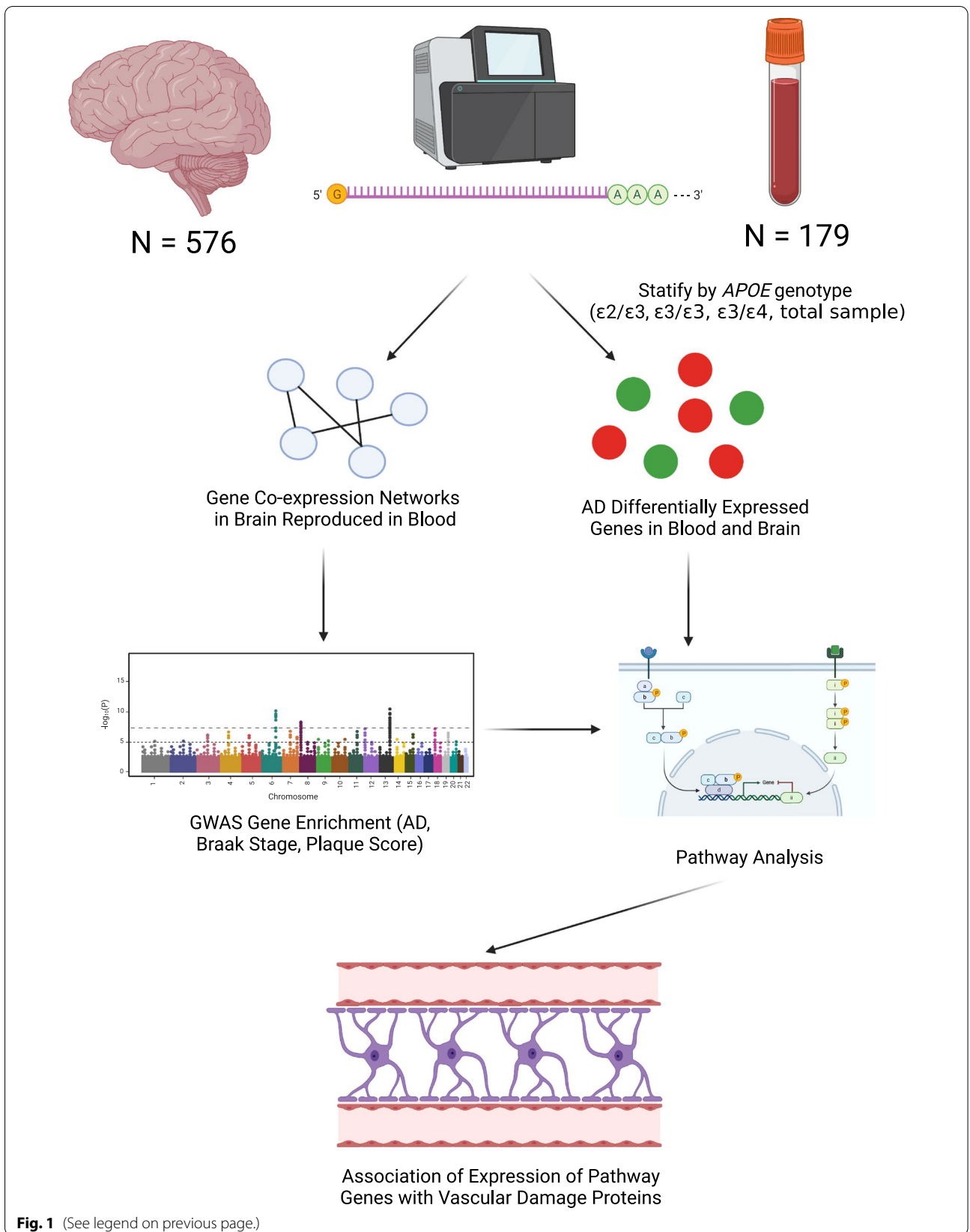


Fig. 1 (See legend on previous page.)

Table 2 Differentially expressed known AD genes in the combined blood and brain datasets in the total sample

| Gene | Dataset | APOE ε2/ε3 | | | APOE ε3/ε3 | | | APOE ε3/ε4 | | | Total | | |
|----------|----------|------------|---------|---------|------------|---------|----------------------|------------|---------|----------------------|-------|---------|----------------------|
| | | N | Z-score | P-value | N | Z-score | P-value | N | Z-score | P-value | N | Z-score | P-value |
| INPP5D | Blood | 26 | -0.62 | 0.53 | 102 | 1.67 | 0.10 | 31 | 2.12 | 0.03 | 179 | 2.16 | 0.03 |
| | Brain | 71 | 1.80 | 0.07 | 355 | 0.92 | 0.36 | 126 | 0.66 | 0.51 | 576 | 2.30 | 0.02 |
| | Combined | 97 | 1.51 | 0.13 | 457 | 1.29 | 0.20 | 157 | 1.31 | 0.19 | 755 | 2.42 | 0.02 |
| HLA-DQA1 | Blood | 26 | -1.49 | 0.14 | 102 | -1.80 | 0.07 | 31 | 0.96 | 0.34 | 179 | -2.26 | 0.03 |
| | Brain | 71 | 0.30 | 0.76 | 355 | -3.13 | 1.7×10^{-3} | 126 | -1.01 | 0.31 | 576 | -2.23 | 0.03 |
| | Combined | 97 | -0.16 | 0.87 | 457 | -2.84 | 4.6×10^{-3} | 157 | -0.66 | 0.51 | 755 | -2.42 | 0.02 |
| SPDYE3 | Blood | 26 | -0.08 | 0.94 | 102 | -0.76 | 0.45 | 31 | -0.96 | 0.34 | 179 | -1.47 | 0.14 |
| | Brain | 71 | -0.79 | 0.43 | 355 | -0.97 | 0.33 | 126 | -2.16 | 0.03 | 576 | -2.73 | 6.4×10^{-3} |
| | Combined | 97 | -0.77 | 0.44 | 457 | -0.96 | 0.33 | 157 | -2.37 | 0.02 | 755 | -2.26 | 0.02 |
| TSPOAP1 | Blood | 26 | -1.38 | 0.17 | 102 | 0.30 | 0.77 | 31 | -0.16 | 0.87 | 179 | -0.55 | 0.58 |
| | Brain | 71 | -1.04 | 0.30 | 355 | -2.17 | 0.03 | 126 | -1.85 | 0.06 | 576 | -3.65 | 2.6×10^{-4} |
| | Combined | 97 | -1.39 | 0.16 | 457 | -1.35 | 0.18 | 157 | -1.82 | 0.07 | 755 | -2.24 | 0.03 |
| SIGLEC11 | Blood | 26 | 1.43 | 0.15 | 102 | 0.47 | 0.64 | 31 | 0.24 | 0.81 | 179 | 0.33 | 0.74 |
| | Brain | 71 | 1.05 | 0.29 | 355 | 2.46 | 0.01 | 126 | 2.58 | 9.9×10^{-3} | 576 | 3.77 | 1.6×10^{-4} |
| | Combined | 97 | 1.42 | 0.16 | 457 | 1.86 | 0.06 | 157 | 2.54 | 0.01 | 755 | 2.17 | 0.03 |

The combined column reflects the meta-analysis results or single batch results in the case of one batch including low gene filtering

Table 3 Novel differentially expressed genes in the blood or brain within APOE genotype groups

| Gene | Dataset | APOE ε2/ε3 | | | APOE ε3/ε3 | | | APOE ε3/ε4 | | | Total | | |
|--------|----------|------------|---------|----------------------|------------|---------|---------|------------|---------|----------------------|-------|---------|---------|
| | | N | Z-score | P-value | N | Z-score | P-value | N | Z-score | P-value | N | Z-score | P-value |
| FRAS1 | Blood | NA | NA | NA | 75 | -0.24 | 0.81 | 31 | -4.66 | 3.2×10^{-6} | 132 | -2.26 | 0.02 |
| | Brain | NA | NA | NA | 355 | -0.97 | 0.33 | 126 | 0.28 | 0.78 | 576 | -0.49 | 0.62 |
| | Combined | NA | NA | NA | 430 | -0.75 | 0.45 | 157 | -1.23 | 0.22 | 708 | -1.27 | 0.20 |
| PIGHP1 | Blood | 26 | 1.67 | 0.09 | 102 | 0.43 | 0.67 | 31 | 1.30 | 0.19 | 179 | 1.14 | 0.25 |
| | Brain | 71 | 4.42 | 9.8×10^{-6} | 355 | 0.53 | 0.59 | 126 | 0.47 | 0.64 | 576 | 1.98 | 0.05 |
| | Combined | 97 | 4.67 | 3.1×10^{-6} | 457 | 0.53 | 0.59 | 157 | 0.87 | 0.39 | 755 | 1.68 | 0.09 |

NA not available due to low expression. Bolded P-values pass transcriptome-wide multiple testing threshold (3.3×10^{-6})

There was little overlap of the leading-edge genes for these four pathways across APOE genotype groups with at most one overlapping gene per pathway (Table 4). Leading-edge genes in the same pathway shared by different APOE genotypes included *HLA-DRA* in allograft rejection, *CD74* in interferon gamma response, and *KYNU* in TNFA signaling via NFKB. Among 92 leading-edge genes from the four pathways that had enrichment scores in opposite directions among subjects with and without ε4 (Table 4), the expression of 21 genes in FHS/ADRC brains was significantly associated ($P < 4.67 \times 10^{-4}$) with at least one of the three vascular injury-related proteins (Fig. 2b, Table 5). The expression of 11 of these 21 genes (52%) was significantly associated with the SAA level, and the expression of 13 genes (62%) was significantly associated with the VCAM-1 level. The expression of only two genes, *TRIP10* and

FOSL1, which are both involved in signaling via NFKB, was significantly associated with ICAM-1 levels.

Co-expression networks common to the brain and blood

Four co-expression networks identified in the brain were preserved in the blood (Supplementary Table 7). The eigengene value (i.e., first principle component of gene expression across the network) in the light green network was significantly higher among ε4 carriers than non-carriers ($P = 4.7 \times 10^{-3}$) (Fig. 3a). The light green network is significantly enriched for genes previously associated with AD risk [38] and plaque score [39] (Supplementary Table 7). The AD-related genes in this network were significantly enriched in nine KEGG pathways and four hallmark pathways (Fig. 3b, Supplementary Tables 8, 9). Seventeen genes contributing to the significant KEGG pathways form a biological subnetwork (Fig. 3c). One of these genes, *NFKBIA*, is a leading-edge gene from the

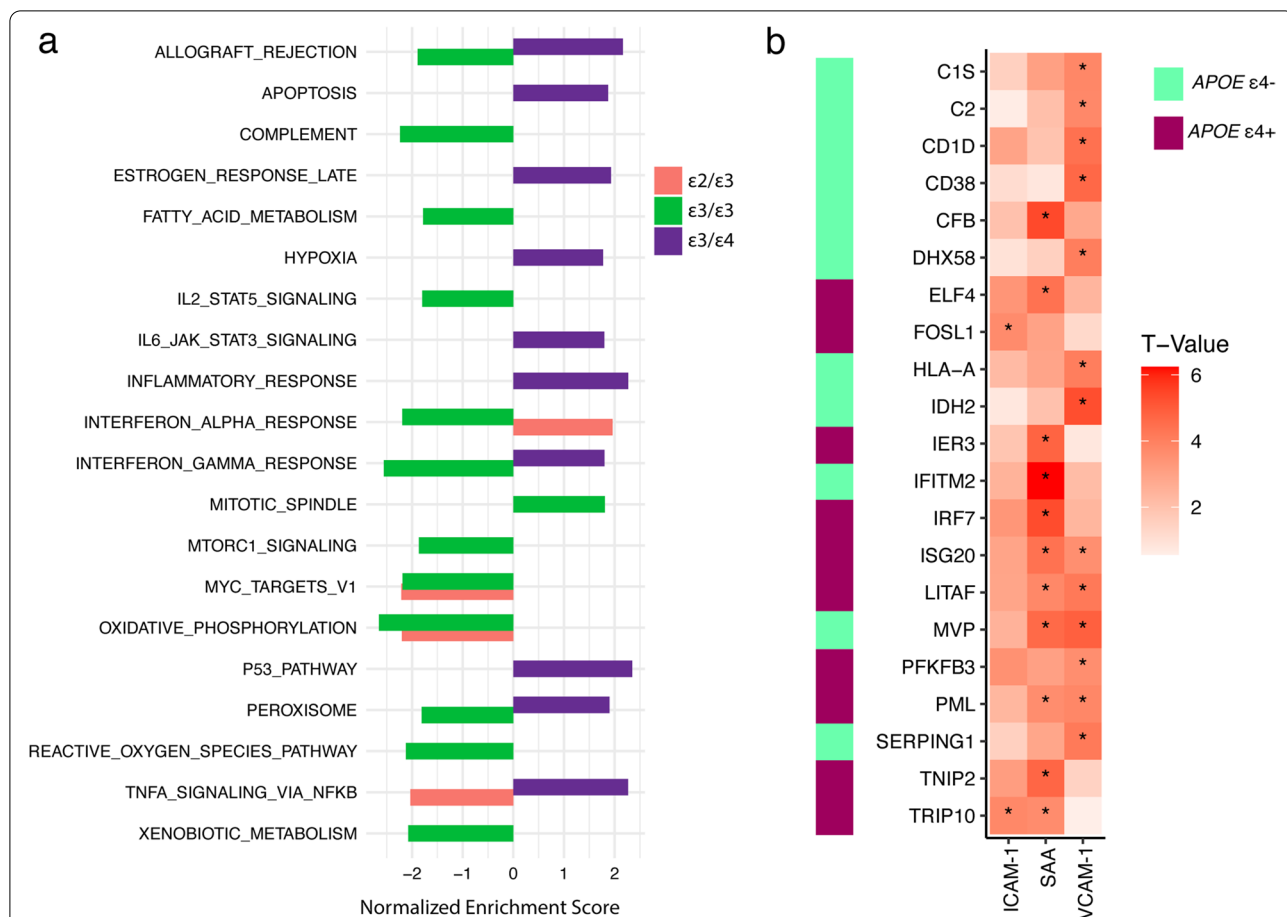


Fig. 2 Significant pathways in the blood and brain by the *APOE* genotype. **a** Significant pathways (adjusted $P < 0.05$) within *APOE* genotype groups that are enriched for differentially expressed genes in the blood and brain combined are plotted according to the normalized enrichment score. Enrichment score indicates whether the genes in the pathway are upregulated (positive) or downregulated (negative) in AD. **b** Heatmap shows strength and direction of association of levels of proteins involved in vascular damage with the expression of leading-edge genes from significant pathways where enrichment scores are in opposite directions between *APOE* $\epsilon 4$ carriers (+) and non-carriers (-). Genes whose expression was significantly ($P < 4.67 \times 10^{-4}$) associated with the level of at least one protein (indicated by an asterisk) are shown

Table 4 Significant co-expressed gene pathways in the combined blood and brain datasets

| Hallmark pathway | <i>APOE</i> genotype | NES | Adjusted <i>P</i> -value | Leading-edge genes |
|---------------------------|-------------------------|-------|--------------------------|--|
| Allograft rejection | $\epsilon 3/\epsilon 3$ | -1.89 | 0.02 | C2, HLA-DQA1, FAS, HLA-A, UBE2N, HLA-DOB, LTB, F2R, HLA-DRA, TAP2, B2M, CD1D, CD74, MAP3K7 |
| | $\epsilon 3/\epsilon 4$ | 2.16 | 7.4×10^{-3} | IRF4, CCL22, IRF7, CD74, HLA-DRA, ELF4, IL16, IFNGR2, IL27RA, IL1B |
| Interferon gamma response | $\epsilon 3/\epsilon 3$ | -2.56 | 1.3×10^{-5} | HLA-DQA1, FAS, CFB, BPGM, C1S, HLA-A, LAP3, MVP, PSME2, PSMA2, UBE2L6, SERPING1, DHX58, IFITM2, CD38, B2M, CD74 |
| | $\epsilon 3/\epsilon 4$ | 1.80 | 0.04 | CD274, CD69, BTG1, ISG20, PML, IRF4, NFKBIA, IRF7, CD74, IL10RA, IRF9 |
| Peroxisome | $\epsilon 3/\epsilon 3$ | -1.81 | 0.03 | IDH2, EHHADH, MVP, ALDH1A1, SCP2, SOD1, ABCD2, MSH2 |
| | $\epsilon 3/\epsilon 4$ | 1.90 | 0.03 | RDH11, ELOVL5, SLC25A19, CTPS1, SLC23A2, SEMA3C |
| Tnfa signaling via nfkb | $\epsilon 2/\epsilon 3$ | -2.03 | 0.01 | DUSP4, NR4A1, NR4A3, MARCKS, NFAT5, PHLDA1, DUSP2, KYNU, GOS2, ETS2, PTGS2, GCH1, MSC, SOD2, EGR2 |
| | $\epsilon 3/\epsilon 4$ | 2.27 | 2.5×10^{-3} | TRIP10, CD69, BTG1, DENND5A, PFKFB3, FOS, NFKBIA, LDLR, IER2, JUN, IL1A, PANX1, PNRC1, DUSP1, IFNGR2, OLR1, MAFF, IL1B, TNIP2, CCL20, BIRC2, IER3, GADD45B, KYNU, LITAF, CCRL2, SPHK1, FOSL1 |

NES normalized enrichment score. Only pathways with adjusted *P*-value < 0.05 were included

Table 5 Association of expression of leading-edge genes from co-expressed gene networks with vascular damage protein levels

| Gene | APOE genotype | Pathway(s) | I-CAM1 | | SAA | | V-CAM1 | |
|----------|-------------------------|------------|---------|----------------------|---------|----------------------|---------|----------------------|
| | | | β | P-value | β | P-value | β | P-value |
| C1S | $\epsilon 3/\epsilon 3$ | IGR | 0.11 | 0.13 | 0.24 | 2.5×10^{-3} | 0.30 | 2.0×10^{-4} |
| C2 | $\epsilon 3/\epsilon 3$ | AR | 0.05 | 0.53 | 0.17 | 0.04 | 0.30 | 2.2×10^{-4} |
| CD1D | $\epsilon 3/\epsilon 3$ | AR | 0.28 | 3.4×10^{-3} | 0.21 | 0.05 | 0.46 | 1.7×10^{-5} |
| CD38 | $\epsilon 3/\epsilon 4$ | IGR | 0.09 | 0.25 | 0.07 | 0.41 | 0.39 | 6.1×10^{-6} |
| CFB | $\epsilon 3/\epsilon 3$ | IGR | 0.16 | 0.05 | 0.46 | 2.2×10^{-7} | 0.26 | 5.4×10^{-3} |
| DHX58 | $\epsilon 3/\epsilon 3$ | IGR | 0.11 | 0.34 | 0.20 | 0.13 | 0.51 | 7.1×10^{-5} |
| ELF4 | $\epsilon 3/\epsilon 4$ | AR | 0.32 | 8.7×10^{-4} | 0.46 | 1.9×10^{-5} | 0.26 | 0.02 |
| FOSL1 | $\epsilon 3/\epsilon 4$ | TSN | 0.20 | 3.1×10^{-4} | 0.18 | 3.1×10^{-3} | 0.08 | 0.21 |
| HLA-A | $\epsilon 3/\epsilon 3$ | AR; IGR | 0.27 | 0.03 | 0.38 | 3.9×10^{-3} | 0.53 | 7.0×10^{-5} |
| IDH2 | $\epsilon 3/\epsilon 3$ | P | 0.09 | 0.43 | 0.24 | 0.05 | 0.61 | 3.6×10^{-7} |
| IER3 | $\epsilon 3/\epsilon 4$ | TSN | 0.14 | 0.06 | 0.37 | 3.5×10^{-6} | 0.07 | 0.43 |
| IFITM2 | $\epsilon 3/\epsilon 3$ | IGR | 0.19 | 0.01 | 0.48 | 2.9×10^{-9} | 0.18 | 0.03 |
| IRF7 | $\epsilon 3/\epsilon 4$ | AR; IGR | 0.24 | 9.5×10^{-4} | 0.41 | 2.8×10^{-7} | 0.20 | 0.02 |
| ISG20 | $\epsilon 3/\epsilon 4$ | IGR | 0.21 | 3.6×10^{-3} | 0.34 | 1.9×10^{-5} | 0.29 | 4.4×10^{-4} |
| LITAF | $\epsilon 3/\epsilon 4$ | TSN | 0.27 | 4.1×10^{-3} | 0.39 | 2.1×10^{-4} | 0.43 | 3.9×10^{-5} |
| MVP | $\epsilon 3/\epsilon 3$ | IGR; P | 0.21 | 0.01 | 0.42 | 7.0×10^{-6} | 0.45 | 2.4×10^{-6} |
| PFKFB3 | $\epsilon 3/\epsilon 4$ | TSN | 0.30 | 5.6×10^{-4} | 0.30 | 2.6×10^{-3} | 0.35 | 4.0×10^{-4} |
| PML | $\epsilon 3/\epsilon 4$ | IGR | 0.28 | 0.02 | 0.47 | 3.8×10^{-6} | 0.50 | 2.0×10^{-4} |
| SERPING1 | $\epsilon 3/\epsilon 3$ | IGR | 0.17 | 0.13 | 0.35 | 4.4×10^{-3} | 0.50 | 4.7×10^{-5} |
| TNIP2 | $\epsilon 3/\epsilon 4$ | TSN | 0.36 | 1.9×10^{-3} | 0.59 | 4.7×10^{-6} | 0.19 | 0.15 |
| TRIP10 | $\epsilon 3/\epsilon 4$ | TSN | 0.32 | 2.1×10^{-4} | 0.35 | 3.4×10^{-4} | 0.06 | 0.58 |

I-CAM intercellular adhesion molecule 1, SAA serum amyloid A, V-CAM1 vascular cell adhesion molecule 1, AR allograft rejection, IGR interferon gamma response, P peroxisome, TSN Tnfa signaling via nfkb

Results in bold surpass the multiple testing threshold ($P < 4.67 \times 10^{-4}$)

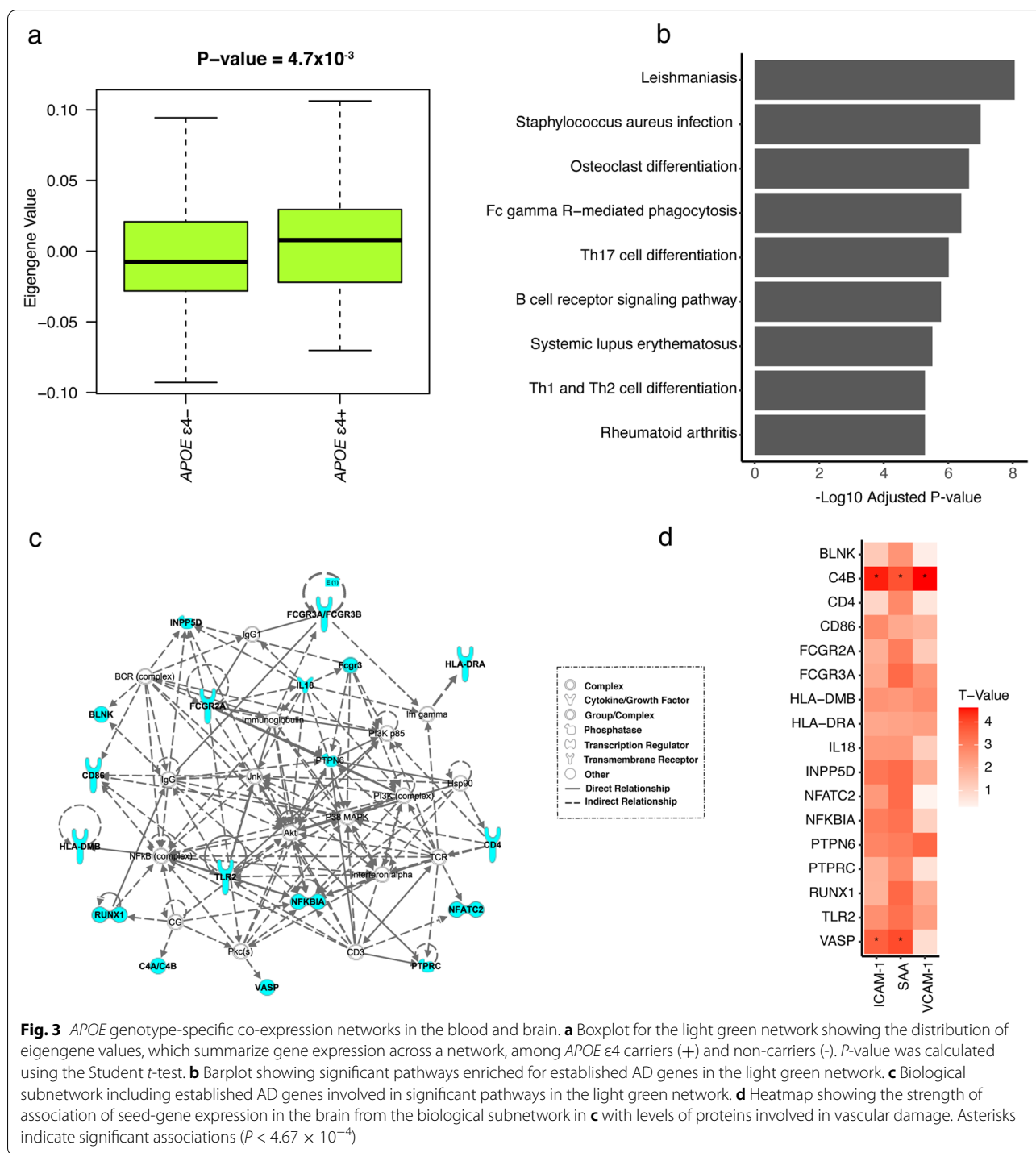
signaling via NFKB pathway and was involved in five out of nine significant KEGG pathways and two out of four hallmark pathways in the light green network (Table 4, Supplementary Table 8). HLA-DRA is involved in six of the nine significant KEGG pathways and the allograft rejection hallmark pathway in the light green network and is a leading-edge gene in the allograft rejection pathway identified in the $\epsilon 3/\epsilon 3$ and $\epsilon 3/\epsilon 4$ groups. INPP5D, which is differentially expressed in both blood and brain (Table 2, Supplementary Fig. 2), is involved in two significant KEGG pathways (*Fc gamma R-mediated phagocytosis* and *B cell receptor signaling*) in the light green network. C4B, which is upregulated in the brain from AD cases compared to controls in the $\epsilon 2/\epsilon 3$ group (Supplementary Table 5), was included in the light green network pathways involved in *Staphylococcus aureus infection* and *systemic lupus erythematosus*.

Two of the 17 subnetwork genes in the significant pathways enriched for AD genes in the light green network (Supplementary Table 10) were significantly associated with the level of at least one of the three vascular injury proteins after multiple testing correction (Fig. 3d). Specifically, VASP expression was significantly associated

with levels of ICAM-1 ($P = 3.7 \times 10^{-4}$, $P_{\text{adj}} = 0.04$) and SAA ($P = 1.0 \times 10^{-4}$, $P_{\text{adj}} = 0.01$) and C4B expression was significantly associated with levels of ICAM-1 ($P = 1.3 \times 10^{-5}$, $P_{\text{adj}} = 1.4 \times 10^{-3}$), SAA ($P = 1.6 \times 10^{-4}$, $P_{\text{adj}} = 0.02$), and VCAM-1 ($P = 7.0 \times 10^{-6}$, $P_{\text{adj}} = 7.6 \times 10^{-4}$).

Discussion

The primary purpose of this study was to identify genes previously associated with AD and in biological pathways enriched for AD genes whose expression differs between AD cases and controls in both blood and brain, especially in an APOE genotype-specific manner. We observed that two established AD genes, INPP5D and HLA-DQA1, were differentially expressed in both blood and brain. Among the 21 top-ranked pathways in the combined blood and brain expression profiles, 10 pathways were specific to persons having the APOE $\epsilon 3/\epsilon 4$ genotype. Additionally, we identified a co-expression network enriched for AD genes in the brain that was preserved in the blood and showed significantly higher average expression in $\epsilon 4$ carriers than non-carriers. Lastly, several genes from the top-ranked pathways and co-expression networks were significantly associated



with levels of vascular injury proteins. These findings suggest that AD genes that are differentially expressed in both blood and brain and associated with vascular markers, and their effects are dependent on *APOE* genotypes.

The BBB is a semi-permeable endothelial cell membrane regulating transport between cerebral blood vessels and the central nervous system [42]. The dysregulation of the BBB has been implicated in early cognitive decline and exacerbation of neuroinflammation and neurodegeneration [43]. A recent study showed that *APOE*

$\epsilon 4$ carriers exhibit BBB dysfunction and cognitive decline independent of AD pathology [10]. Our analyses identified six pathways uniquely enriched for DEGs among $\epsilon 3/\epsilon 4$ carriers in combined blood and brain expression data. Expression of *INPP5D* and *HLA-DQA-1* was significantly greater in both blood and brain from individuals with AD compared to controls. Increased expression of *INPP5D* in blood has been previously linked with an increased risk of hemorrhagic transformation [44], which is associated also with BBB permeability [45]. *INPP5D* is highly expressed in microglia and encodes the protein SHIP1 which has been implicated in many neuroinflammatory processes [46]. Additionally, *HLA-DQA1* and *INPP5D* are expressed in dendritic cells and monocytes and involved in immune processes, and the migration of monocytes across an inflamed BBB can cause differentiation into dendritic cells [47]. *FRAS1* was significantly downregulated in AD compared to controls in blood from $\epsilon 3/\epsilon 4$ AD individuals and a recent study showed that *FRAS1* knockdown mice were impaired in memory and learning behaviors [48].

We identified four pathways (allograft rejection, interferon gamma response, peroxisome, and TNFA signaling via NFKB) containing gene sets that, with respect to AD, were significantly upregulated in the blood and brain from $\epsilon 4$ carriers and other gene sets from the same pathways that were downregulated in individuals without $\epsilon 4$. The inflammatory cytokine interferon gamma has been shown to impact directly brain endothelium to cause BBB breakdown [49] and can inhibit ApoE production in macrophages [50]. Peroxisomes synthesize fatty acids which have been implicated in the development of AD [51]. The TNFA via NFKB signaling pathway has been implicated in BBB dysfunction [52], and the TNFA and NFKB pathways have been independently associated with increased neuroinflammation related to *APOE* $\epsilon 4$ [53, 54].

Multiple genes from networks we observed to be preserved in the brain and blood transcriptome data and enriched in pathways from combined blood and brain expression profiles showed significant association with the vascular injury proteins ICAM-1, SAA, and VCAM-1. The SAA level increases in the presence of BBB dysfunction [55]. ICAM-1 is a cytokine involved in the regulation of the BBB [56], and increased ICAM-1 level has been associated with BBB damage and neuroinflammation [57]. Under inflammatory conditions, the VCAM-1 level is upregulated and the BBB can release soluble VCAM-1 which in turn can compromise BBB function [58]. Our study showed that *FOSL1* and *TRIP10* were among the genes enriched in the TNFA via the NFKB pathway, and their expression was associated with ICAM-1. *TRIP10* was previously included in an AD network derived from multi-omic integration [59] and *FOSL-1* was identified

in conjunction with PIAS1, a protein associated with AD and inflammatory response [60]. We identified *VASP* and *C4B* in an *APOE* genotype-specific co-expressed gene network in the brain that was reproduced in the blood, and the expression of these genes was significantly associated with levels of multiple vascular damage proteins. *VASP* encodes vasodilator-stimulated phosphoprotein which regulates BBB function [61]. Additionally, *VASP* has been recently implicated in a microglial network in AD [62]. The pattern of *C4B* expression in the brain is dependent on the *APOE* genotype [19], and dysregulation of the complement system can cause or exacerbate BBB dysfunction [63]. *C4B*-binding-protein levels in cerebral spinal fluid have been shown to correlate with BBB integrity [64]. These genes require further investigation in their role with AD specifically related to BBB function and *APOE*.

Limitations

Our study has several limitations. First, the sample sizes of the *APOE* genotype groups in the blood dataset were relatively small which limited statistical power. Additionally, the ROSMAP blood dataset exhibited significant batch effects. However, we were able to account for these batch effects by running each batch separately and meta-analyzing our results. Second, the software WGCNA creates networks based on strong computational correlations but does not account for underlying biological implications. We evaluated biological connections using the IPA software by rebuilding subnetworks of the leading-edge genes. Third, publicly available single-cell data were available only for dendritic cells and monocytes in the blood, and therefore, we could not analyze a wider array of blood cell types. Fourth, we were unable to account for RIN in the ROSMAP blood dataset because this information was unavailable. Fifth, because expression profiles may differ between tissues, lack of overlap between brain and blood does not necessarily exclude the relevance of some of our discordant findings across tissues to AD and BBB dysfunction. Finally, although there was little overlap of leading-edge genes in GSEA pathways across *APOE* genotype groups, several particular pathways containing a different complement of genes were significant among individuals with different *APOE* genotypes. Experimental studies are needed to confirm the mechanisms involving these genes.

Conclusions

Our study provides evidence of the importance of evaluating brain and blood transcriptome data together with genetic information derived from the same subjects to identify meaningful correlations of blood biomarkers with AD-related proteins in the brain. Future

studies are required to investigate further, how the genes and biological pathways identified in this study in the context of the *APOE* genotype influence the BBB and contribute to and/or exacerbate AD-related pathology.

Abbreviations

A β : Amyloid- β ; AD: Alzheimer's disease; BBB: Blood-brain barrier; CAA: Cerebral amyloid angiopathy; CERAD: Consortium to Establish a Registry for Alzheimer Disease; CPM: Counts per million; DE: Differential expression; DGE: Differential gene expression; FHS/ADRC: Framingham Heart Study/Boston University Alzheimer's Disease Research Center; FPKM: Fragments per kilobase of transcript per million; GWAS: Genome-wide association study; ICAM-1: Intercellular adhesion molecule 1; PMI: Post-mortem interval; RIN: RNA integrity number; ROSMAP: Religious Orders Study and Rush Memory and Aging Project; RNA-seq: RNA sequencing; SAA: Serum amyloid α ; VCAM-1: Vascular cell adhesion molecule 1; WGCNA: Weighted gene correlation network analysis.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13195-022-00975-z>.

Additional file 1. Supplementary Figs. 1–4 and Supplementary Tables 1–10.

Acknowledgements

Not applicable.

Authors' contributions

R.P., L.A.F., and G.R.J. conceived the overall study design. R.P. and G.R.J. perceived the statistical analysis. R.P. and J.H. performed data analyses. R.P., J.H., T.D.S., L.A.F., and G.R.J. wrote the manuscript. W.X. and T.D.S. conducted immunoassay experiments using autopsied brains. D.A.B. and T.D.S. provided neuropathological data. R.P., D.A.B., T.D.S., W.X., L.A.F., and G.R.J. reviewed and edited the manuscript. G.R.J. and L.A.F. supervised and L.A.F. and G.R.J. obtained funding for the project. The authors read and approved the final manuscript.

Funding

This study was supported by the National Institute on Aging (NIA) grants RF1-AG057519, R01-AG069453, U01-AG068057, U19-AG068753, R01-AG048927, U01-AG062602, U01-AG058654, RF1-AG057768, RF1-AG054156, P30-AG072978, U01-AG032984, and RF1AG063913. Collection of study data provided by the Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, was supported through funding by NIA grants P30-AG10161, R01-AG15819, R01-AG17917, R01-AG30146, R01-AG36836, U01-AG32984, U01-AG46152, and U01-AG61358, and funding from the Illinois Department of Public Health and the Translational Genomics Research Institute.

Availability of data and materials

The datasets supporting the conclusions of this article are available in <https://singlecell.broadinstitute.org/> and <http://www.synapse.org>, as well as from the corresponding author upon request.

Declarations

Ethics approval and consent to participate

The study protocol, design, and performance of the current study were approved by the Boston University Institutional Review Board.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Medicine (Biomedical Genetics), Boston University School of Medicine, 72 East Concord Street, Boston, MA 02118, USA. ²Department of Pharmacology & Experimental Therapeutics, Boston University School of Medicine, 72 East Concord Street, Boston, MA 02118, USA. ³Department of Veterans Affairs Medical Center, Bedford, MA 01730, USA. ⁴Rush Alzheimer's Disease Center, Rush University Medical Center, 1750 W. Harrison Street, Suite 1000, Chicago, IL 60612, USA. ⁵Department of Pathology & Laboratory Medicine, Boston University School of Medicine, 72 East Concord Street, Boston, MA 02118, USA. ⁶Department of Neurology, Boston University School of Medicine, 72 East Concord Street, Boston, MA 02118, USA. ⁷Department of Ophthalmology, Boston University School of Medicine, 72 East Concord Street, Boston, MA 02118, USA. ⁸Department of Biostatistics, Boston University School of Public Health, 715 Albany Street, Boston, MA 02118, USA. ⁹Department of Epidemiology, Boston University School of Public Health, 715 Albany Street, Boston, MA 02118, USA.

Received: 10 September 2021 Accepted: 2 February 2022

Published online: 09 February 2022

References

- DeTure MA, Dickson DW. The neuropathological diagnosis of Alzheimer's disease. *Mol Neurodegeneration*. 2019;14:32.
- Fehlbaum-Beurdeley P, Jarrige-Le Prado AC, Pallares D, Carrière J, Guihal C, Soucaille C, et al. Toward an Alzheimer's disease diagnosis via high-resolution blood gene expression. *Alzheimers Dement*. 2010;6:25–38.
- Zetterberg H, Burnham SC. Blood-based molecular biomarkers for Alzheimer's disease. *Mol Brain*. 2019;12:26.
- Palmqvist S, Janelidze S, Quiroz YT, Zetterberg H, Lopera F, Stomrud E, et al. Discriminative accuracy of plasma phospho-tau217 for Alzheimer disease vs other neurodegenerative disorders. *JAMA*. 2020;324:1–11.
- Ashton NJ, Pascoal TA, Karikari TK, Benedet AL, Lantero-Rodriguez J, Brinkmalm G, et al. Plasma p-tau231: a new biomarker for incipient Alzheimer's disease pathology. *Acta Neuropathol*. 2021;141:709–24.
- Bell RD, Zlokovic BV. Neurovascular mechanisms and blood–brain barrier disorder in Alzheimer's disease. *Acta Neuropathol*. 2009;118:103–13.
- Cai Z, Qiao P-F, Wan C-Q, Cai M, Zhou N-K, Li Q. Role of blood-brain barrier in Alzheimer's disease. *J Alzheimers Dis*. IOS Press. 2018;63:1223–34.
- Ramos-Cejudo J, Wisniewski T, Marmar C, Zetterberg H, Blennow K, de Leon MJ, et al. Traumatic brain injury and Alzheimer's disease: the cerebrovascular link. *EBioMedicine*. 2018;28:21–30.
- Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. *APOE and Alzheimer Disease Meta Analysis Consortium*. *JAMA*. 1997;278:1349–56.
- Montagne A, Nation DA, Sagare AP, Barisano G, Sweeney MD, Chakhoyan A, et al. *APOE4* leads to blood–brain barrier dysfunction predicting cognitive decline. *Nature*. Nature Publishing Group. 2020;581:71–6.
- Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*. 1993;261:921–3.
- Reiman EM, Arboleda-Velasquez JF, Quiroz YT, Huentelman MJ, Beach TG, Caselli RJ, et al. Exceptionally low likelihood of Alzheimer's dementia in *APOE2* homozygotes from a 5,000-person neuropathological study. *Nat Commun*. 2020;11:1–11.
- Yu L, Boyle PA, Nag S, Leurgans S, Buchman AS, Wilson RS, et al. *APOE* and cerebral amyloid angiopathy in community-dwelling older persons. *Neurobiol Aging*. 2015;36:2946–53.
- Panitch R, Hu J, Chung J, Zhu C, Meng G, Xia W, et al. Integrative brain transcriptome analysis links complement component 4 and *HSPA2* to the *APOE* ϵ 2 protective effect in Alzheimer disease. *Mol Psychiatry*. 2021;26:6054–64.
- Jun G, You Y, Zhu C, Meng G, Chung J, Panitch R, et al. Protein phosphatase 2A, complement component 4, and *APOE* genotype linked to Alzheimer disease using a systems biology approach. *medRxiv*. Cold Spring Harbor Laboratory Press. 2020;2020(11):20.20235051.

16. Bennett DA, Buchman AS, Boyle PA, Barnes LL, Wilson RS, Schneider JA. Religious Orders Study and Rush Memory and Aging Project. *J Alzheimers Dis*. 2018;64:S161–89.
17. De Jager PL, Ma Y, McCabe C, Xu J, Vardarajan BN, Felsky D, et al. A multi-omic atlas of the human frontal cortex for aging and Alzheimer's disease research. *Sci Data*. 2018;5:180142.
18. Hyman BT, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Carrillo MC, et al. National Institute on Aging–Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement*. 2012;8:1–13.
19. Panitch R, Hu J, Chung J, Zhu C, Meng G, Xia W, et al. Integrative brain transcriptome analysis links complement component 4 and HSPA2 to the APOE ε2 protective effect in Alzheimer disease. *medRxiv*. Cold Spring Harbor Laboratory Press. 2020;2020(11):23.20235762.
20. Babraham Bioinformatics - FastQC: a quality control tool for high throughput sequence data [Internet]. [cited 2019 Sep 6]. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
21. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. Oxford University Press. 2013;29:15.
22. Dobin A, Gingeras TR. Mapping RNA-seq reads with STAR. *Curr Protoc Bioinformatics*. 2015;51:11.14.1–11.14.19.
23. Picard Tools - By Broad Institute [Internet]. [cited 2020 Dec 17]. Available from: <http://broadinstitute.github.io/picard/>.
24. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323.
25. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–9.
26. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11:R25.
27. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15:R29.
28. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
29. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26:2190–1.
30. Yang Q, Wang Y. Methods for analyzing multivariate phenotypes in genetic association studies. Shao Y, editor. *J Probability Stat*. Hindawi Publishing Corporation. 2012;2012:652569.
31. Single Cell Portal [Internet]. [cited 2021 Dec 9]. Available from: https://singlecell.broadinstitute.org/single_cell.
32. Villani A-C, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* [Internet], vol. 356: American Association for the Advancement of Science; 2017. [cited 2021 Jan 4]; Available from: <https://science.sciencemag.org/content/356/6335/eaah4573>
33. Korotkevich G, Sukhov V, Sergushichev A. Fast gene set enrichment analysis. *bioRxiv*. Cold Spring Harbor Laboratory. 2019:060012.
34. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *PNAS*. National Academy of Sciences. 2005;102:15545–50.
35. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
36. Langfelder P, Horvath S. Eigengene networks for studying the relationships between co-expression modules. *BMC Syst Biol*. 2007;1:54.
37. Bakhtiarzadeh MR, Hosseinpour B, Shahhoseini M, Korte A, Gifani P. Weighted gene co-expression network analysis of endometriosis and identification of functional modules associated with its main hallmarks. *Front Genet*. 2018;9:453.
38. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat Genet*. 2019;51:414–30.
39. Beecham GW, Hamilton K, Naj AC, Martin ER, Huentelman M, Myers AJ, et al. Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. *PLOS Genetics*. Public Library of Science. 2014;10:e1004606.
40. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013;14:128.
41. Bellenguez C, Küçükali F, Jansen I, Andrade V, Moreno-Grau S, Amin N, et al. New insights on the genetic etiology of Alzheimer's and related dementia. *medRxiv*. Cold Spring Harbor Laboratory Press. 2020;2020.10.01.20200659.
42. Abbott NJ, Patabendige AAK, Dolman DEM, Yusof SR, Begley DJ. Structure and function of the blood-brain barrier. *Neurobiol Dis*. 2010;37:13–25.
43. Zenaro E, Piacentino G, Constantin G. The blood-brain barrier in Alzheimer's disease. *Neurobiol Dis*. 2017;107:41–56.
44. Jickling GC, Ander BP, Stamova B, Zhan X, Liu D, Rothstein L, et al. RNA in blood is altered prior to hemorrhagic transformation in ischemic stroke. *Ann Neurol*. 2013;74:232–40.
45. Rosenthal SL, Kamboh MI. Late-onset Alzheimer's disease genes and the potentially implicated pathways. *Curr Genet Med Rep*. 2014;2:85–101.
46. Malik M, Parikh I, Vasquez JB, Smith C, Tai L, Bu G, et al. Genetics ignite focus on microglial inflammation in Alzheimer's disease. *Mol Neurodegen*. 2015;10 [cited 2021 May 3]; Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4595327/>.
47. Ifergan I, Kébir H, Bernard M, Wosik K, Dodelet-Devillers A, Cayrol R, et al. The blood–brain barrier induces differentiation of migrating monocytes into Th17-polarizing dendritic cells. *Brain*. 2008;131:785–99.
48. Kalpachidou T, Makrygiannis AK, Pavlakis E, Stylianopoulou F, Chalepakis G, Stamatakis A. Behavioural effects of extracellular matrix protein Fras1 depletion in the mouse. *Eur J Neurosci*. 2020;53(12):3905–19.
49. Bonney S, Seitz S, Ryan CA, Jones KL, Clarke P, Tyler KL, et al. Gamma interferon alters junctional integrity via rho kinase, resulting in blood-brain barrier leakage in experimental viral encephalitis. *mBio*. 2019;10(4):e01675–19.
50. Brand K, Mackman N, Curtiss LK. Interferon-gamma inhibits macrophage apolipoprotein E production by posttranslational mechanisms. *J Clin Invest*. 1993;91:2031–9.
51. Lizard G, Rouaud O, Demarquoy J, Cherkaoui-Malki M, Iuliano L. Potential roles of peroxisomes in Alzheimer's disease and in dementia of the Alzheimer's type. *J Alzheimers Dis*. 2012;29:241–54.
52. Coelho-Santos V, Leitão RA, Cardoso FL, Palmela I, Rito M, Barbosa M, et al. The TNF-α/NF-κB signaling pathway has a key role in methamphetamine-induced blood-brain barrier dysfunction. *J Cereb Blood Flow Metab*. 2015;35:1260–71.
53. Ophir G, Amariglio N, Jacob-Hirsch J, Elkon R, Rechavi G, Michaelson DM. Apolipoprotein E4 enhances brain inflammation by modulation of the NF-κappaB signaling cascade. *Neurobiol Dis*. 2005;20:709–18.
54. Fan Y-Y, Cai Q-L, Gao Z-Y, Lin X, Huang Q, Tang W, et al. APOE ε4 allele elevates the expressions of inflammatory factors and promotes Alzheimer's disease progression: a comparative study based on Han and She populations in the Wenzhou area. *Brain Res Bull*. 2017;132:39–43.
55. Bowman GL, Dayon L, Kirkland R, Wojcik J, Peyratout G, Severin IC, et al. Blood-brain barrier breakdown, neuroinflammation, and cognitive decline in older adults. *Alzheimers Dement*. 2018;14:1640–50.
56. Dietrich J-B. The adhesion molecule ICAM-1 and its regulation in relation with the blood-brain barrier. *J Neuroimmunol*. 2002;128:58–68.
57. Huber JD, Campos CR, Mark KS, Davis TP. Alterations in blood-brain barrier ICAM-1 expression and brain microglial activation after λ-carrageenan-induced inflammatory pain. *Am J Physiol Heart Circ Physiol*. 2006;290:H732–40.
58. Haarmann A, Nowak E, Deiß A, van der Pol S, Monoranu C-M, Kooij G, et al. Soluble VCAM-1 impairs human brain endothelial barrier integrity via integrin α-4-transduced outside-in signalling. *Acta Neuropathol*. 2015;129:639–52.
59. Klein H-U, Schäfer M, Bennett DA, Schwender H, De Jager PL. Bayesian integrative analysis of epigenomic and transcriptomic data identifies Alzheimer's disease candidate genes and networks. *PLoS Comput Biol*. 2020;16 [cited 2021 Apr 9]; Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7138305/>.
60. He K, Zhang J, Liu J, Cui Y, Liu LG, Ye S, et al. Functional genomics study of protein inhibitor of activated STAT1 in mouse hippocampal neuronal cells revealed by RNA sequencing. *Aging (Albany NY)*. 2021;13:9011–27.
61. Kraft P, Benz PM, Austinat M, Brede ME, Schuh K, Walter U, et al. Deficiency of vasodilator-stimulated phosphoprotein (VASP) increases

blood-brain-barrier damage and edema formation after ischemic stroke in mice. *PLOS ONE*. Public Library of Science. 2010;5:e15106.

62. Patrick E, Olah M, Taga M, Klein H-U, Xu J, White CC, et al. A cortical immune network map identifies distinct microglial transcriptional programs associated with β -amyloid and Tau pathologies. *Transl Psychiatry*. 2021;11:1–17.
63. Alexander JJ. Blood-brain barrier (BBB) and the complement landscape. *Mol Immunol*. 2018;102:26–31.
64. Trouw Leendert A, Nielsen HM, Minthon L, Londos E, Landberg G, Veerhuis R, et al. C4b-binding protein in Alzheimer's disease: binding to A β 1–42 and to dead cells. *Mol Immunol*. 2008;45:3649–60.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

