Gene Set Index Based on Different Modules May Help Differentiate the Mechanisms of Alzheimer’s Disease and Vascular Dementia

Fengkun Zhou1,2,*
Deyao Chen1,2,*
Guoying Chen1,2,*
Peilin Liao1,2
Rongjie Li1,2
Qingfang Nong1,2
Youshi Meng1,2
Donghua Zou1
Xianfeng Li1,2

1Department of Neurology, The Fifth Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, 530022, People’s Republic of China; 2Department of Neurology, The First People’s Hospital of Nanning, Nanning, Guangxi, 530022, People’s Republic of China

*These authors contributed equally to this work

Introduction

The risk of Alzheimer’s disease (AD), one of the most common neurodegenerative diseases, increases with increasing inflammation, autoimmune disease, cholinergic synaptic receptor activity and cortisol secretion.1,2 AD involves the accumulation of extracellular amyloid-β1 (Aβ1) and deposits of intracellular tau neurofibrillary tangles.3–5 The main symptoms of AD are difficulty in concentration, memory loss and the gradual loss of specific cognitive function, which eventually leads to loss of independence and death.6 The loss of independence places a heavy burden on patients and their families. So far, AD treatments are only palliative and delay the development of symptoms, without halting or reversing the disease.7

Like AD, vascular dementia manifests as dementia, but it is distinct from AD. Vascular dementia is caused by ischemia, hemorrhage, hypoxia or hypoxic brain damage. Its incidence is increasing, together with the incidence of cerebrovascular...
diseases in general. Although vascular dementia and AD are distinct diseases, they can co-exist in patients, who are then diagnosed as having “mixed dementia.” In fact, VD is the second most common cause of dementia, second only to AD. The two diseases share several features, and both may involve impairment of cholinergic function. Nevertheless, the two diseases are distinct and can be difficult to distinguish clinically. Research that clarifies similarities and differences between their underlying mechanisms may facilitate differential diagnosis and treatment.

Given the association of AD with cholinergic dysfunction, current treatments focus on acetylcholinesterase inhibitors that act on the central nervous system, which can restore cholinergic input and partially improve memory loss. AD also involves activation of astrocytes and microglia, which can be found in neuronal plaques, as well as an increase in levels of pro-inflammatory molecules. These findings implicate systemic inflammation in AD. Consistent with this idea, levels of chemokine receptors, such as CCL2, also known as monocyte chemoattractant protein-1 (MCP-1), can influence the risk of AD and accelerate its development. CCL2, secreted mainly by astrocytes and microglia, promotes inflammation as well as the differentiation and migration of macrophages.

After binding to its cognate receptor, CCL2 activates the kinases PKC, PI2K, Akt and ERK. Higher CCL2 levels activate microglia near the inflammatory site, and such microglia contribute to amyloid plaques, plaque aggregation and cognitive impairment. Elevated CCL2 levels are risk factors not only for AD but also for neuroinflammatory disease and other neurodegenerative diseases, such as multiple sclerosis, cerebral ischemia, and traumatic brain injury. At the same time, CCL2 down-regulation in AD mice also accelerates amyloid pathology. These results indicate that too low or too high CCL2 levels can harm the brain’s innate immunity and cognitive function.

In order to further explore the mechanisms of AD and how they may be different from those of vascular dementia, we conducted a comprehensive module exploration based on publicly available data. We found that both diseases involve immune-related biological processes, including the regulation of inflammatory responses and T cell activation. These processes are significantly involved in signaling based on chemokines, cytokines and Toll-like receptors. While AD may be associated with cholinergic synaptic disorder and decreased neurotransmitter secretion, vascular dementia is more likely to be associated with abnormal lipid metabolism and cortisol production.

Materials and Methods
Data Collection and Processing
Gene expression datasets of patients with AD in the Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo) were searched. GSE122063, which was based on the GPL16699 platform, included brain samples from 56 samples with AD, 36 samples with vascular dementia, and 44 healthy samples. The samples were from cases matched for age, postmortem interval, and sex.

Gene expression in frontal and temporal cortex of patients with AD, patients with vascular dementia or non-demented controls was profiled using data from the University of Michigan Brain Bank. Controls and AD patients had no infarcts in the autopsied hemisphere. The “normalize between arrays” function in the limma package was used to normalize the gene expression profiles. If a gene was detected by multiple probes, the expression values of all probes were averaged to obtain the value for the gene.

The workflow of the study is shown in Figure 1.

Analysis of Differentially Expressed Genes (DEGs)
In GSE122063, genes differentially expressed between patients with AD or VD and controls were identified using the limma package in R. DEGs were included in further analysis if \(|\log_2(\text{fold change})| > 1\) and the P-value after adjustment by the false discovery rate was below 0.05.

Weighted Gene Co-Expression Network Analysis (WGCNA) and Identification of Hub Genes
For WGCNA, the sample phenotypes were converted into a number matrix with three columns (AD, VD, and control) and samples in rows. For example, an AD sample was marked 1 in AD column, 0 in VD and control column. The gene expression profile was rotated into a new matrix with gene in column and observers in rows. We extracted the gene expression profile of DEGs from GSE122063 in order to run the WGCNA package in R. First, hierarchical clustering analysis was performed using the “hclust”
function, then the soft thresholding power value was screened from 1 to 30 during module construction using the “pickSoftThreshold” function. The average connectivity among different modules and their independence were assessed at the different power values. The power was considered appropriate if the degree of independence exceeded 0.9. Then, coexpression networks (modules) were constructed in the WGCNA package with a minimum module size of 30. Each module was assigned a unique color.

In WGCNA, the module eigengene was the first principal component of the expression matrix for a module. The module eigengene was considered an average gene expression value for genes in a module. Phenotype was converted to a numerical value, and regression was performed between the module eigengene values and the phenotype. Module membership (MM) was defined as the association between a gene and its module, and gene significance (GS) was defined as the correlation of a gene with a phenotype. Genes with GS > 0.2 and MM > 0.8 were considered to be hub genes in the module.

**Functional Enrichment**

Enrichment for gene ontology (GO) functions and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways was analyzed using the clusterProfiler package in R. Enrichment was considered significant if P < 0.05. The GSEA program (version 2.2.4 in Java) and default parameters were used to screen for GO terms and KEGG pathways that may be related to AD or vascular dementia. The reference gene sets were c5.bp.v6.2. symbols.gmt and c2.cp.kegg.v6.2. symbols.gmt in the MsigDB V6.2 database.

**Construction of the Protein–Protein Interaction (PPI) Network**

Based on the interactions of human transcription factors (TFs) with their target genes in the TRRUST v2 database, the hypergeometric test was used to identify TFs that may regulate functional modules. In order to predict complex cellular functions, a PPI was constructed using Cytoscape (http://cytoscape.org/) and the STRING database.

**Prediction of Structural Domain Combination**

TFs were docked onto potential target genes using hex software (version 8.0.0). Results were visualized using PyMOL.

**Definition of an Index to Differentiate AD and Vascular Dementia**

The least absolute shrinkage and selection operator (LASSO) is effective at selecting features from high-dimensional data that can be used to make classifications. Firstly, GSE122063 data set was randomly assigned to the training set (70%) and test set (30%). We used the cv.glmnet function from glmnet package (https://CRAN.R-project.org/package=glmnet) to select the optimal genes with non-zero coefficients, which were entered into a LASSO model to differentiate among patients with AD, patients with vascular dementia and non-demented controls. The two main parameters are “nfolds = 10” representing 10-fold cross-validation, and “family = ‘binomial’” representing two-class logistic regression analysis. A model index for
each sample was created by weighting the expression values of the selected genes by the regression coefficients from LASSO:

\[ \text{Index} = (\text{expression value} \times \text{coefficient})_{\text{gene 1}} + (\text{expression value} \times \text{coefficient})_{\text{gene 2}} + (\text{expression value} \times \text{coefficient})_{\text{gene 3}} \ldots \]

The ability of the LASSO model to identify AD or vascular dementia was assessed by receiver operating characteristic curve analysis using the pROC package in R.\textsuperscript{35}

**Results**

**DEGs and Modules Associated with AD or Vascular Dementia**

In the data set GSE122063, 8537 DEGs between AD patients and controls were identified, of which 2827 were up-regulated and 3800 were down-regulated.

![Figure 2](https://example.com/figure2.png)

**Figure 2** Differentially expressed gene (DEG) analysis and Weighted Gene Co-Expression Network Analysis. (A) Manhattan map of differential gene expression. (B) Definition of power related to modules. (C) Recognition module. (D) The brown module positively correlated with both AD and vascular dementia, the blue module negatively correlated with vascular dementia, and the turquoise module negatively correlated with AD. The grey module did not correlate with either type of dementia. (E) Module membership and gene significance strongly correlated with each other within each module.
in AD. There were 5833 DEGs between patients with vascular dementia and controls, of which 1849 were up-regulated and 2131 were down-regulated in the disease. There were 3027 DEGs between patients with AD and vascular dementia, of which 521 were up-regulated and 1003 were down-regulated in AD. There were 7330 DEGs between AD patients, on one hand, and vascular dementia patients and controls, on the other, of which 2182 were up-regulated and 3097 were down-regulated in AD (Figure 2A).

To identify the key module most associated with AD, WGCNA was performed based on the DEG expression profile. The lowest power that gave independence larger than 0.90 was 0.885 (Figure 2B). Four modules were identified (Figure 2C). The brown module positively correlated with AD (correlation coefficient = 0.22, P = 0.01) and vascular dementia (correlation coefficient = 0.24, P = 0.006). The blue module negatively correlated with vascular dementia (correlation coefficient = −0.24, P = 0.005), while the turquoise module negatively correlated with AD (correlation coefficient = −0.44, P = 1 x 10−7; Figure 2D).

In the brown module, the following 11 genes were identified as up-regulated hub genes in AD and vascular dementia: SLAMF8, C1QB, MS4A6A, C1QC, MS4A4A, SAMS1, CCR5, STAB1, SLC7A7, EBI3, and ITGB2. In the blue module, six genes were identified as down-regulated hub genes in vascular dementia: DOCK3, WDTCL1, IDUA, FBX1L19, ARMC6, and CREB3L1. In the turquoise module, 129 genes were identified as down-regulated hub genes in AD (Table S1). Correlation analysis confirmed that hub genes in the brown module were strongly associated with AD and vascular dementia;

Figure 3 Results of functional enrichment. (A) Biological processes involving genes of the blue module. (B) KEGG pathways involving genes of the blue module. (C) Biological processes involving genes of the brown module. (D) KEGG pathways involving genes of the brown module. (E) Biological processes involving genes of the turquoise module. (F) KEGG pathways involving genes of the turquoise module.
hub genes in the blue module, with vascular dementia; and hub genes in the turquoise module, with AD (Figure 2E).

The Global Regulatory Landscape of AD and Vascular Dementia

Analysis of function enrichment showed that blue module genes were significantly involved in biological processes related to cell proliferation, such as cell proliferation in the midbrain and negative regulation of apoptosis in development, and in biological processes related to abnormal cortisol and lipid metabolism, such as cortisol synthesis and secretion as well as Cushing syndrome (Figure 3A and B). Brown module genes were involved in immune-related biological processes, including regulation of innate immune response, T cell activation, regulation of inflammatory response, cellular responses to interferon-g, chemokine signaling, interaction of cytokines with their receptors, and signaling involving Toll-like receptors (Figure 3C and D). Turquoise module genes were significantly enriched in biological processes regulating neurotransmitter transport, synapse organization, neurotransmitter secretion, neuropeptide signaling, insulin secretion, MAPK signaling, and cholinergic synapse transmission (Figure 3E and F). ClueGo enrichment analysis confirmed that the main biological processes enriched in the blue module were sensory perception of chemical stimuli and keratinocyte differentiation. The enriched processes in the brown module were cell activation involved in immune responses, cellular responses to interferon-g, neuroinflammatory response, and inflammatory response. The enriched processes in the turquoise module were mainly receptor regulator activity, synapse organization, regulation of neurotransmitter transport, and neuropeptide signaling (Figure 4A–C).

Validation of Biological Processes and Key Pathways in AD and Vascular Dementia

Gene Set Enrichment Analysis (GSEA) showed that biological processes such as astrocyte differentiation glial cell development, negative regulation of adaptive immune response, etc. are significantly enriched in AD (Figure 5A). Biological processes such as calcium ion regulate exocytosis, cognition, glutamate receptor signaling pathway, positive regulation of synaptic transmission were significantly enriched in vascular dementia (Figure 5B). Key pathways in AD were complement and coagulation cascades, interaction between extracellular matrix and receptors, and hematopoietic cell lineage (Figure 5C); key pathways in vascular dementia involved calcium signaling and gap junctions (Figure 5D).

![Figure 4 Enrichment analysis using ClueGo. Biological processes significantly enriched in the (A) blue, (B) brown or (C) turquoise modules.](image)

Altered Expression of Hub Genes in Both AD and Vascular Dementia

Using the STRING database, we extracted the genes from modules and assembled them into a global regulatory network based on the pathways in which those genes were significantly involved (Figure 6A). TFs regulating the
pathways were extracted, and correlations between TFs and pathway genes were further analyzed (Figure 6B–D). Docking studies of TFs with target genes suggested that the TF SPI1 and target ITGB2, the TF YBX1 and EGFR as well as the TF PAX4 and INS have negative docking energies (Figure 7A–C). Finally, we explored how different modules may contribute to AD or vascular dementia (Figure 7D–F).

LASSO Can Predict AD and Vascular Dementia

We extracted the expression profile of hub genes to construct a LASSO model (Figure 8A–C), in which 146 genes in three modules were found to have non-zero regression coefficients. The model was able to differentiate AD patients from controls with an area under the receiver operating characteristic curve (AUC) of 0.961 in the training set and 1.000 in the test set. Similarly, the index was able to distinguish between vascular patients and controls with an AUC of 1.000 in the training set and 1.000 in the test set. Finally, the index was able to distinguish between patients with AD or vascular dementia with an AUC of 0.857 in the training set and 0.997 in the test set (Figure 8D–F). These results suggest the index may be a biomarker of AD and vascular dementia and may even differentiate the two diseases from each other.

Discussion

AD and vascular dementia are the two most common forms of dementia and place a heavy burden on many families. Despite decades of research, no specific treatments for dementia have been described. In this study, WGCNA identified nine gene modules, among which the brown module positively correlated with the development of both forms of dementia. The brown module
mainly involves immune-related biological processes and pathways, such as the innate immune response, T cell activation, inflammatory response regulation, chemokine signaling, interaction between cytokines and their receptors, and Toll-like receptor signaling. These indicated the dysfunction of immune progress may be the similar molecular mechanisms underlying AD and vascular dementia. The blue module negatively correlated with the development of vascular dementia, and the turquoise module negatively correlated with the development of AD. These results suggest that AD and vascular dementia result from different dysfunctional gene modules, and the underlying mechanisms of AD and vascular dementia are different. For instance, the blue module involves in NF-kappa B signaling pathway, toll-like receptor signaling pathway, and NOD-like receptor signaling pathway. NF-KB and activator protein 1 (AP1) are modulated by activated toll-like receptor (TLR)-mediated recruiting of adaptor proteins. The turquoise module involved in MAPK signaling pathway, insulin signaling, and HIF-1 signaling pathway. Unsurprisingly, all these pathways were reported associated with dementia. Our results indicated that these pathways show different activities in AD and vascular dementia.

In this study, 146 genes were identified as hub genes. We extracted TFs from each module and constructed a global regulatory network of TF-hub genes-pathways. Through this network, we analyzed how TFs may regulate pathway genes directly or indirectly. Some pathways identified here have already been associated with AD or vascular dementia. In previously, a study identified blood-cell-derived molecular signatures and potential therapeutic targets through integrated analysis in two AD-related data set. Another study identified molecular signatures and therapeutic agents in Alzheimer’s disease using a network-based approach. A previous study revealed biomarkers and pathways shared by AD and ischemic Stroke. The dysfunction of cell adhesion molecules, purine metabolism, intracellular transmembrane protein transport, immune system, and cholesterol metabolism contributes to AD according to the previous genome-wide association study. Compared with these studies, some related pathways were also found in our present study. We also proposed the shared pathways and potentially different pathways
between AD and VaD and constructed a potential transcriptional regulatory network.

Mutations in chemokine receptors can increase the risk of AD and accelerate its progression, and the chemokine CCL2. Consistent with this literature, we found here that the chemokine pathway positively correlated with the development of AD and vascular dementia. We also found an association between AD and changes in cholinergic synapse transmission. This is consistent with studies that attributed inflammation-related cognitive decline mainly to the degeneration of central cholinergic neurons, altering expression of genes such as those encoding muscarinic and nicotinic receptors. This leads to reduced synthesis and release of acetylcholine and can trigger intracellular signal transduction disorders. Cholinergic dysfunction is often accompanied by the formation and deposition of Aβ, which is a hallmark of AD.

Analysis of our blue module revealed an association of cortisol synthesis and secretion with vascular dementia. This
Assessment of models for identifying AD and vascular dementia. (A–C) LASSO model. (D–F) Receiver operating characteristic curves for differentiating (D) AD patients and non-demented controls, (E) vascular dementia patients and non-demented controls, or (F) patients with AD or vascular dementia. AUC, area under the curve.
is consistent with previous reports that suggest that neurosteroids regulate cognitive function. Both AD and vascular dementia can dysregulate the adrenocortical axis. A LASSO model based on 146 genes with non-zero regression coefficients was able to distinguish AD or vascular dementia from non-dementia controls with extremely high AUCs. It may be an effective biological index to identify the types of dementia.

Our study shows that bioinformatics analysis can make predictions about the molecular pathways behind AD and vascular dementia and thereby generate testable hypotheses for future experiments in vitro and in preclinical models.

Conclusion
The pathogenesis of AD may involve down-regulation of genes involved in cholinergic synapase transmission, while that of vascular dementia may be more closely related to down-regulation of cortisol metabolism and secretion. The chemokine pathway may be quite dysregulated in both types of dementia. A LASSO model based on hub genes of different modules may be useful for diagnosing the two diseases.

Data Sharing Statement
The raw analyses from this study and codes can be obtained from the corresponding author upon reasonable request.

Author Contributions
FKZ, DYC, GYC, DHZ and XFL conceived and designed the study. All authors contributed to data analysis, drafting or revising the article, have agreed on the journal to which the article will be submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Funding
This study was supported by the Project of Nanning Scientific Research and Technology Development Plan (20193092 and 20193109), the Project of Qingxiu District of Nanning Scientific Research and Technology Development Plan (2019028, 2019039 and 2020058), the Scientific Research Project of Guangxi Zhuang Autonomous Region Health Commission (Z20200201), the High-Level Medical Expert Training Program of Guangxi “139” Plan Funding (G201903049). This study was also sponsored by the Nanning Excellent Young Scientist Program and the Guangxi Beibu Gulf Economic Zone Major Talent Program (RC20190103).

Disclosure
The authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References
1. Calsolaro V, Edison P. Neuroinflammation in Alzheimer’s disease: current evidence and future directions. Alzheimers Dement. 2016;12(6):719–732. doi:10.1016/j.jalz.2016.02.010
2. Zou D, Li R, Huang X, et al. Identification of molecular correlations of RBM8A with autophagy in Alzheimer’s disease. Aging. 2019;11(23):11673–11685. doi:10.18632/aging.102571
3. Bose S, Cho J. Role of chemokine CCL2 and its receptor CCR2 in neurodegenerative diseases. Arch Pharm Res. 2013;36(9):1039–1050. doi:10.1007/s12272-013-0161-x
4. Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer’s disease: progress and problems on the road to therapeutics. Science. 2002;297(5580):353–356. doi:10.1126/science.1072994
5. Jian C, Lu M, Zhang Z, et al. miR-34a knockout attenuates cognitive deficits in APP/PS1 mice through inhibition of the amyloidogenic processing of APP. Life Sci. 2017;182:104–111.
6. Zhu JB, Tan CC, Tan L, Yu JT. State of Play in Alzheimer’s Disease. Genet. 2019;58(3):631–659.
7. Zhou C, Wang J, Huang X, Jian C, Zou D, Li X. Analysis of transcription factor- and ncRNA-mediated potential pathogenic gene modules in Alzheimer’s disease. Aging. 2019;11(16):6109–6119.
8. Zou D, Liu C, Zhang Q, et al. Association between polymorphisms in microRNAs and ischemic stroke in an Asian population: evidence based on 6033 cases and 7248 controls. Clin Interv Aging. 2018;13:1709–1726.
9. Zekry D, Duyckaerts C, Hauw JJ. Défences vasculaires et défences mixtes: vascular dementia and mixed dementia. [Vascular dementia and mixed dementia]. Presse Med. 2007;36(10 Pt 2):1469–1476. French.
10. Bennett D. Public health importance of vascular dementia and Alzheimer’s disease with cerebrovascular disease. Int J Clin Pract Suppl. 2001;1(120):41–48.
11. Roman G. Diagnosis of vascular dementia and Alzheimer’s disease. Int J Clin Pract Suppl. 2001;1(120):9–13.
12. Otsuka M, Yamaguchi K, Ueki A. Similarities and differences between Alzheimer’s disease and vascular dementia from the viewpoint of nutrition. Ann N Y Acad Sci. 2002;977(1):155–161. doi:10.1111/j.1749-6632.2002.tb04811.x
13. Schliebs R, Arendt T. The cholinergic system in aging and neuronal degeneration. Behav Brain Res. 2011;221(2):555–563. doi:10.1016/j.bbr.2010.11.058
14. Heneka MT, Carson MJ, El Khoury J, et al. Neuroinflammation in Alzheimer’s disease. Lancet Neurol. 2015;14(4):388–405.
15. Kauwe JS, Bailey MH, Ridge PG, et al. Genome-wide association study of CSF levels of 59 Alzheimer’s disease candidate proteins: significant associations with proteins involved in amyloid processing and inflammation. PLoS Genet. 2014;10(10):e1004758. doi:10.1371/journal.pgen.1004758
16. Semple BD, Kossmann T, Morganti-Kossman MC. Role of chemokines in CNS health and pathology: a focus on the CCL2/CCR2 and CXCL8/CXCR2 networks. J Cereb Blood Flow Metab. 2010;30(3):459–473. doi:10.1038/jcbfm.2009.240
17. Sokolova A, Hill MD, Rahimi F, Warden LA, Halliday GM, Shepherd CE. Monocyte chemotactant protein-1 plays a dominant role in the chronic inflammation observed in Alzheimer’s disease. Brain Pathol. 2009;19(3):392–398. doi:10.1111/j.1750-3639.2008.00188.x
