Identification of Hub Genes and Immune Cell Infiltration Characteristics in Alzheimer’s Disease

Ming Hu and Jianhua Wang

1Department of Graduate School, Hebei Medical University, Shijiazhuang 050000, Hebei, China
2Department of Neurology, Hebei General Hospital, Shijiazhuang 050051, Hebei, China

Correspondence should be addressed to Jianhua Wang; wangjh6304@sina.com

Received 21 October 2021; Revised 18 November 2021; Accepted 26 November 2021; Published 20 December 2021

The purpose of this study was to identify hub genes closely correlated with Alzheimer’s disease (AD) and their association with immune cell infiltration. In this work, 119 overlapping differentially expressed genes (DEGs) were obtained from GSE5281 and GSE122063 datasets through differential expression analysis. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed on the 119 DEGs, revealing some important biological functions and key pathways. AD immune cell infiltration analysis revealed a significant difference in the proportion of immune cells between the AD group and the control group. Finally, correlation analysis between target hub genes and immune cells indicated that GFAP had a positive or negative correlation with some specific immune cells. Our results provided useful clues, which will help to explain the molecular mechanism of AD and search for precise prognostic markers and potential therapeutic targets.

1. Introduction

Alzheimer’s disease (AD) is a degenerative disease of the central nervous system that occurs in old age and pre-old age and is characterized by progressive cognitive dysfunction and behavioral impairment [1, 2]. It is the most common type of dementia and one of the most common chronic diseases in old age [3], accounting for about 50% to 70% of dementia in old age [4, 5]. While the exact cause of AD has not been elucidated, studies have found that AD is the result of a combination of genes, lifestyle, and environmental factors, caused in part by specific genetic changes [6–9]. A combination of drug therapy, non-drug therapy, and careful nursing can reduce symptoms and delay the progression of the disease [10–12], but there is no specific drug that can cure AD or effectively reverse the progression of the disease. The course of Alzheimer’s disease is about 5–10 years, and a few patients can survive for more than 10 years. Most of them die from complications such as lung infection, urinary tract infection, and pressure ulcers [13–15]. Therefore, it is key to identify the hub genes, explore the pathogenesis, and search for the therapeutic targets of AD.

A new generation of high-throughput sequencing technologies and the development of genomics have produced a wealth of disease gene expression data and clinical information already stored in many public databases [16–18]. This provides a new idea and theoretical basis for in-depth understanding of the pathogenesis and biological characteristics of diseases through bioinformatics analysis.

In this study, we used high-throughput sequencing data for differential gene expression analysis, GO functional and KEGG pathway enrichment analyses, and protein-protein interaction (PPI) network analysis to identify network hub genes and their biological roles. In addition, we also performed immune cell infiltration analysis and correlation analysis between target hub genes and immune cells on all samples, which were main innovative points of this research paper.

2. Materials and Methods

2.1. Downloading AD Transcriptome Data from GEO Database. AD gene expression data were obtained from Gene Expression Omnibus (GEO) database [19]
(https://www.ncbi.nlm.nih.gov/gds). We downloaded the GSE5281 and GSE122063 datasets using the R package GEOquery [20]. A total of 181 AD and 116 normal control samples were collected.

2.2. Data Cleaning and Differential Gene Expression Analysis. Firstly, the gene expression matrices of GSE5281 and GSE122063 datasets were normalized and formatted into input file format of R language. Then, the differentially expressed genes (DEGs) of AD patients were screened by robust rank aggregation [21], and the volcano plots and heatmaps of DEGs were plotted using limma [22] and pheatmap [23] packages of R. P value < 0.05 and | logFC (fold change) | > 1 were considered statistically significant.

2.3. Functional and Pathway Enrichment Analyses. To clarify the biological functions and key pathways of DEGs in AD, we performed Gene Ontology (GO), including biological process (BP), cellular component (CC), and molecular function (MF), and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses [24] using R packages such as clusterProfiler [25], enrichplot, and ggplot2 [26]. P value < 0.05 indicated significant differences.

3. Results

3.1. Identification of DEGs. Datasets GSE5281 and GSE122063 were downloaded from GEO database. The former included 87 AD brain tissue and 74 normal tissue samples, while the latter included 92 AD brain tissue and 44 normal tissue samples. After data preprocessing and gene differential expression analysis, 119 differentially expressed genes (AD/normal control tissue) were obtained using robust rank aggregation, of which 30 genes were significantly upregulated and 89 genes were downregulated in AD patients, as shown in Figures 1(a) and 1(b). The heatmap showed the top 50 DEGs with most significant upregulation and downregulation, as shown in Figure 1(c). The P values < 0.05 and | logFC | ≥ 1 were the cutoff criteria.

3.2. GO and KEGG Enrichment Analyses of the 119 DEGs. We also ran GO function and KEGG pathway enrichment analyses for the 119 overlapping DEGs by R package clusterProfiler. Figure 2 shows the result of GO enrichment analysis. The biological processes (BPs) of the 119 DEGs focused predominantly on chemical synaptic transmission, nervous system development, ion transport, and positive regulation of neuron projection development, as shown in Figure 2(a). With regard to the cellular components (CCs), it was found that these DEGs were strongly associated with Golgi membrane, cell junction, and neuronal cell body, as shown in Figure 2(b). Furthermore, in terms of molecular function (MF), those 119 DEGs were associated with calmodulin binding, extracellular ligand-gated channel activity, and GABA, as shown in Figure 2(c). Searching the KEGG database revealed that the DEGs mainly matched to retrograde endocannabinoid signaling, morphine addiction, and GABAergic synapse, as shown in Figure 2(d).

3.3. Identification of Hub Genes by PPI Network Analysis. We constructed the PPI network among these overlapping DEGs by using the STRING database and visualized them using Cytoscape software, as shown in Figure 3(a). Cytoscape was used to screen out two key modules from PPI network by MCODE algorithm, as shown in Figure 3(b). Network hub genes were identified by Degree algorithm, as shown in Figure 3(c). The top 10 network hub genes were SLC32A1, STMN2, GFAP, GABRA1, SST, GABRG2, SYN2, GNG3, PVALB, and SH3GL2, as shown in Figure 3(d).

3.4. Composition and Differential Expression of the Infiltrating Immune Cells. We performed CIBERSORT immune cell infiltration analysis on the GSE12206 dataset to compare the
composition and differential expression of immune cells between the AD group and the normal control group. Figure 4(a) summarizes the infiltration of 22 types of immune cells in each sample. Figure 4(b) shows the overall composition of immune cells in AD group and control group. Figure 4(c) shows the co-expression correlation between 22 immune cell proportions. As shown in Figure 4(d), compared with normal control group, higher proportions of T cells CD4 memory activated, macrophages M2, and neutrophils could be detected in AD group, along with lower proportions of T cells follicular helper, T cells regulatory (Tregs), NK cells activated, and mast cells resting ($P < 0.05$).

3.5. The Relationship between Target Hub Genes and Immune Cells. Through the PPI network analysis, we obtained 10 hub genes, among which GFAP was the upregulated gene in the AD group, so we conducted correlation analysis between GFAP and various immune-infiltrating cells. Figures 5 and 6 show the strong correlation between GFAP and immune-infiltrating cells. GFAP had a positive correlation with T cell CD4 memory activated, macrophages M2, neutrophils, plasma cells, and macrophages M1. GFAP has a negative correlation with T cells regulatory (Tregs), Mast cells resting, NK cells activated, and T cells follicular helper (Correlation Coefficient < 0 and $P$ value < 0.05).
Figure 2: Bubble plots of GO and KEGG enrichment analyses of the 119 overlapping DEGs. (a) GO-BP enrichment analysis. (b) GO-CC enrichment analysis. (c) GO-MF enrichment analysis. (d) KEGG enrichment analysis. Circle size indicates the number of DEGs, and circle color indicates the P values.

4. Discussion

AD is a central neurodegenerative disease occurring in the early and old age. It is mainly characterized by progressive cognitive dysfunction and behavioral impairment. The etiology is not clear, and there is no cure at present [32]. Therefore, it is particularly urgent to find precise prognostic biomarkers and therapeutic targets for AD. In this paper, 119 overlapping DEGs were first identified between GSE5281 and GSE12206 datasets by differential gene expression analysis. Second, GO and KEGG enrichment analyses were performed on the 119 DEGs, revealing some important biological functions and key pathways, such as chemical synaptic transmission, Golgi membrane, calmodulin binding, retrograde endocannabinoid signaling, morphine addiction, and GABAergic synapse. Also, we used the STRING database to build a PPI network among these overlapping DEGs, screened two key modules from the PPI network, and identified 10 network hub genes. These were GFAP (Glial fibrillary acidic protein) is one of the groups of protein components that make up intermediate silk.

GFAP (Glia fibrillary acidic protein) is one of a group of protein components that make up intermediate silk. Intermediate filaments are found in astrocytes and help maintain normal structure and function of the brain and spinal cord. When GFAP is defective, the protein products it expresses become abnormal, which can lead to what is known as Alzheimer’s disease the rapid development of the automobile industry, automobile practitioners have proposed several n, a rare condition in which brain tissue is gradually destroyed. In recent years, many studies have reported the close relationship between GFAP and AD. Chatterjee et al [33], used Simoa assay to measure plasma proteins in cognitively unimpaired older adults (CU) and found that GFAP and p-tau181 were upregulated in the CU group with cerebral amyloidosis, which indicated the clinical potential of GFAP and p-tau for the diagnosis and longitudinal monitoring of preclinical AD. Cicognola et al [34] conducted a follow-up study of 160 patients with mild cognitive impairment (MCI) for an average of 4.7 years to detect the associated amyloid proteins in the cerebrospinal fluid. The result showed that plasma GFAP can detect the pathology of AD and predict the transformation to AD dementia in patients with MCI. Teitsdottir et al [35] quantitatively measured novel biomarkers, including GFAP, in cerebrospinal fluid of 52 subjects using enzyme-linked immunosorbent assay (ELISA) and bioinformatics analysis.
These results suggested that GFAP may be a marker of cognitive decline in predementia and early AD. AD is a disease of the nervous system, but it also presents with systemic inflammation, with higher levels of inflammatory cytokines and chemokines in the patient’s peripheral and central nerves [36, 37]. Goldeck et al. [38] studied the phenotype of circulating immune cells in AD patients by flow cytometry and confirmed that the proportion of cells expressing CD25 (a T cell CD4 memory activated) in AD patients was significantly higher than that in the control group. The proportion of CCR6+ cells was also increased, and this chemokine receptor was mainly expressed in pro-inflammatory memory cells and Th17 cells. AD patients also had a greater proportion of cells expressing CCR4 (expressed on Th2 cells) and CCR5 (Th1 cells and dendritic cells). Kasus-Jacobi et al. [39] used mass spectrometry and in vitro aggregation methods to detect the activity of neutrophil elastase (NE) and cathepsin G (CG) against amyloid-beta peptide $\alpha\beta_{1-42}$ and found that the peptide derived from CAP37 mimics the quenching and inhibitory aggregation effects of $\alpha\beta_{1-42}$ full-length protein. In addition, the peptide inhibited the neurotoxicity of the most toxic $\alpha\beta_{1-42}$ aggregates. These results provide possible strategies for the development of novel AD-modifying drugs. By constructing a neuropathic AD transgenic mouse model, St-Amour et al. [40] analyzed the important characteristics of the adaptive immune system in the serum, bone marrow, and spleen of the mice by flow cytometry and ELISPOT. The results showed that the proportion of hematopoietic stem cells decreased in the bone marrow of 12-month-old triple transgenic mouse model (3xTg-AD), and the number of lymphocytes, granulocytes, and monocytes remained unchanged. These results suggest that the 3xTg-AD model validates the adaptive immune response observed in patients with AD and confirms the activation of valuable immune pathways in AD.

Through comprehensive bioinformatics analysis, we identified the hub genes closely related to the molecular mechanism of AD, verified the biological functions and key pathways of the hub genes, and conducted immune cell infiltration analysis and correlation analysis for the target genes.
Figure 4: Differences and correlation of immune cell infiltration between the AD group and control group. (a) Summary of immune infiltration in 22 immune cell subpopulations from 100 samples. (b) Heatmap of 22 immune cell infiltrations between the two groups. (c) Correlation matrix of 22 immune cell infiltration in 100 samples. (d) Violin plot of differential expression of 22 infiltrating immune cells. Abscissa: immune cell types; ordinate: relative immune cell content. Blue: AD group; red: control group.
core genes. Our work will help clarify the pathogenesis of AD and provide new candidate biomarkers and potential therapeutic targets for clinical application in the future. The limitation of this study is the lack of attention to different subtypes of AD, and the results still need to be verified in vivo and in vitro.

**Figure 5:** Scatter plots of correlation in GFAP expression and different immune cell contents (a–i). (a) T cells CD4 memory activated, (b) Macrophages M2, (c) Neutrophils, (d) Plasma cells, (e) Macrophages M1, (f) T cells regulatory (Tregs), (g) Mast cells resting, (h) NK cells activated, (i) T cells follicular helper. (R) correlation coefficient; $P < 0.05$ indicates a significant correlation.
5. Conclusions

In this study, we identified 10 network hub genes (SLC32A1, STMN2, GFAP, GABRA1, SST, GABRG2, SYN2, GNG3, PVALB, and SH3GL2). GFAP had a positive or negative correlation with some specific immune cells. These genes could be candidate precise prognostic markers and potential therapeutic targets.

Data Availability

The simulation experiment data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] C. Reitz, C. Brayne, and R. Mayeux, “Epidemiology of alzheimer disease,” Nature Reviews Neurology, vol. 7, no. 3, pp. 137–152, 2011.

[2] M. S. Jilani, D. Tagwireyi, L. L. Gadaga, C. C. Maponga, and C. Mutsimhu, “Cognitive-enhancing effect of a hydroethanolic extract of against memory impairment induced by aluminum chloride in BALB/c mice,” Behavioural Neurology, vol. 2018, Article ID 2057219, 2018.

[3] J. Pitt, K. C. Wilcox, V. Tortelli et al., “Neuroprotective astrocyte-derived insulin/insulin-like growth factor 1 stimulates endocytic processing and extracellular release of neuron-bound Aβ oligomers,” Molecular Biology of the Cell, vol. 28, no. 2, pp. 2623–2636, 2017.

[4] T. Ernest James Phillips and E. Maguire, “Phosphoinositides: roles in the development of microglial-mediated neuroinflammation and neurodegeneration,” Frontiers in Cellular Neuroscience, vol. 15, Article ID 652593, 2021.

[5] D. H. Cribbs, A. Ghochikyan, V. Vasilevko et al., “Adjuvant-dependent modulation of Th1 and Th2 responses to immunization with beta-amyloid,” International Immunology, vol. 15, no. 4, pp. 505–514, 2003.

[6] G. Barthet, T. Jordà-Siquier, J. Rumi-Masante, F. Bernadou, U. Müller, and C. Mulle, “Presenilin-mediated cleavage of APP regulates synaptotagmin-7 and presynaptic plasticity,” Nature Communications, vol. 9, no. 1, p. 4780, 2018.

[7] A. McKeever and M. Agius, “Dementia risk assessment and risk reduction using cardiovascular risk factors,” Psychiatria Danubina, vol. 30, pp. 469–474, 2018.

[8] J. Campdelacreu, “Parkinson’s disease and Alzheimer disease: environmental risk factors,” Neurologia, vol. 29, no. 9, pp. 541–549, 2014.

[9] M. A. Williams, G. J. McKay, R. Carson, D. Craig, G. Silvestri, and P. Passmore, “Age-related macular degeneration-associated genes in alzheimer disease,” American Journal of Geriatric Psychiatry, vol. 23, no. 12, pp. 1290–1296, 2015.

[10] S. Franzen, J. Smith, E. V. D. Berg et al., “Diversity in Alzheimer’s disease drug trials: the importance of eligibility criteria,” Alzheimer’s and Dementia, 2021.
cognitive function in patients with Alzheimer’s disease: A Chinese consensus on the comprehensive management of Alzheimer’s disease,” *Neuroscience bulletin*, vol. 37, no. 7, pp. 1023–1038, 2021.

[12] E. Scherder and F. van Manen, “Pain in Alzheimer’s disease: nursing assistants’ and patients’ evaluations,” *Journal of Advanced Nursing*, vol. 52, no. 2, pp. 151–158, 2005.

[13] R. Balczon, M. T. Lin, J. Y. Lee et al., “Pneumonia initiates a tauopathy,” *The FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, vol. 35, no. 9, Article ID e21807, 2021.

[14] A. P. Pan, J. Meeks, T. Potter et al., “SARS-CoV-2 susceptibility and COVID-19 mortality among older adults with cognitive impairment: cross-sectional analysis from hospital records in a diverse US metropolitan area,” *Frontiers in Neurology*, vol. 12, Article ID 692662, 2021.

[15] S. Mitchell, E. D’Agata, L. Hanson et al., “The trial to reduce antimicrobial use in nursing home residents with alzheimer disease and other dementias (TRAIN-AD): a cluster randomized clinical trial,” *JAMA internal medicine*, vol. 181, no. 9, p. 117, 2021.

[16] Y. Hamaguchi, C. Zeng, and M. Hamada, “Impact of human gene annotations on RNA-seq differential expression analysis,” *BMC Genomics*, vol. 22, no. 1, p. 730, 2021.

[17] T. Barrett, S. E. Willhite, P. Ledoux et al., “NCBI GEO: archive for functional genomics data sets—update,” *Nucleic Acids Research*, vol. 41, no. D1, pp. D991–D995, 2013.

[18] C. Hutter and J. C. Zenklusen, “The cancer genome atlas: creating lasting value beyond its data,” *Cell*, vol. 173, no. 2, pp. 283–285, 2018.

[19] T. Barrett and R. Edgar, “Gene expression Omnibus: microarray data storage, submission, retrieval, and analysis,” *Methods in Enzymology*, vol. 411, pp. 352–369, 2006.

[20] S. Davis and P. S. Meltzer, “GEQquery: GEOquery: a bridge between the gene expression Omnibus (GEO) and BioConductor,” *Bioinformatics*, vol. 23, no. 14, pp. 1846-1847, 2007.

[21] Z. Cao, Y. Guo, Y. Ao, and S. Zhou, “Dysregulated microRNAs in laryngeal cancer: a comprehensive meta-analysis using a robust rank aggregation approach,” *Future Oncology*, vol. 16, no. 33, pp. 2723–2734, 2020.

[22] M. E. Ritchie, B. Phipson, D. Wu et al., “Limma powers differential expression analysis for RNA-sequencing and microarray studies,” *Nucleic Acids Research*, vol. 43, no. 7, p. e47, 2015.

[23] Y. Pan, L. Wu, S. He, J. Wu, T. Wang, and H. Zang, “Identification of hub genes in thyroid carcinoma to predict prognosis by integrated bioinformatics analysis,” *Bioengineered*, vol. 12, no. 1, pp. 2928–2940, 2021.

[24] Y.-H. Zhang, T. Zeng, L. Chen, T. Huang, and Y.-D. Cai, “Determining protein-protein functional associations by functional rules based on gene ontology and KEGG pathway,” *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics*, vol. 1869, no. 6, p. 140621, 2021.

[25] G. Yu, L.-G. Wang, Y. Han, and Q.-Y. He, “ClusterProfiler: an R package for comparing biological themes among gene clusters,” *OMICS: A Journal of Integrative Biology*, vol. 16, no. 5, pp. 284–287, 2012.

[26] C. Zhang, Y. Zheng, X. Li, X. Hu, F. Qi, and J. Luo, “Genome-wide mutation profiling and related risk signature for prognosis of papillary renal cell carcinoma,” *Annals of Translational Medicine*, vol. 7, no. 18, p. 427, 2019.

[27] F. Sibai, A. A. El-Moursy, A. Asaduzzaman, and S. Majzoub, “Hardware acceleration of the STRIKE string kernel algorithm for estimating protein to protein interactions,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, p. 1, 2021.

[28] H. Cui, L. Xu, Z. Li et al., “Integrated bioinformatics analysis for the identification of potential key genes affecting the pathogenesis of clear cell renal cell carcinoma,” *Oncology letters*, vol. 20, no. 2, pp. 1573–1584, 2020.

[29] C.-H. Chin, S.-H. Chen, H.-H. Wu, C.-W. Ho, M.-T. Ko, and C.-Y. Lin, “cytoHubba: identifying hub objects and subnetworks from complex interactome,” *BMC Systems Biology*, vol. 8, no. S4, p. S11, 2014.

[30] Y. Zhang, B. Shen, L. Zhuge, and Y. Xie, “Identification of differentially expressed genes between the colon and ileum of patients with inflammatory bowel disease by gene co-expression analysis,” *Journal of International Medical Research*, vol. 48, no. 5, Article ID 300060519887268, 2020.

[31] M. Yan, M. Song, R. Bai, S. Cheng, and W. Yan, “Identification of potential therapeutic targets for colorectal cancer by bioinformatics analysis,” *Oncology letters*, vol. 12, no. 2, pp. 5092–5098, 2016.

[32] X. Dong, J. Nao, J. Shi, and D. Zheng, “Predictive value of routine peripheral blood markers in Alzheimer’s disease,” *Frontiers in Aging Neuroscience*, vol. 11, p. 332, 2019.

[33] P. Chatterjee, S. Pedrini, N. Ashton et al., “Diagnostic and prognostic plasma biomarkers for preclinical Alzheimer’s disease,” *Alzheimer’s and Dementia*, 2021.

[34] C. Cigognola, S. Janelidze, J. Hertz et al., “Plasma glial fibrillary acidic protein detects Alzheimer pathology and predicts future conversion to Alzheimer dementia in patients with mild cognitive impairment,” *Alzheimer’s Research & Therapy*, vol. 13, no. 1, p. 68, 2021.

[35] U. D. Teitsdottir, M. K. Jónsdóttir, S. H. Lund, T. Darreh-Shori, J. Snaedal, and P. H. Petersen, “Association of glial and neuronal degeneration markers with Alzheimer’s disease cerebrospinal fluid profile and cognitive functions,” *Alzheimer’s Research & Therapy*, vol. 12, no. 1, p. 92, 2020.

[36] S. Westfall, N. Lomis, I. Kahouli, S. Y. Dia, S. P. Singh, and S. Prakash, “Microbiome, probiotics and neurodegenerative diseases: deciphering the gut brain axis,” *Cellular and Molecular Life Sciences*, vol. 74, no. 20, pp. 3769–3787, 2017.

[37] P. McGeer and E. G. McGeer, “Inflammation, autoimmunity and Alzheimer disease,” *Neurobiology of Aging*, vol. 22, no. 6, pp. 799–809, 2001.

[38] D. Goldeck, A. Larbi, M. Pellicano et al., “Enhanced chemokine receptor expression on leukocytes of patients with Alzheimer’s disease,” *PloS one*, vol. 8, no. 6, p. e66664, 2013.

[39] A. Kasus-Jacobi, J. L. Washburn, C. A. Land et al., “Association of glial and neuronal degeneration markers with Alzheimer’s disease,” *PloS one*, vol. 16, no. 3, 2019.

[40] I. St-Amour, C. Bosoi, I. Paré et al., “Peripheral adaptive immunity of the triple transgenic mouse model of Alzheimer’s disease,” *Journal of Neuroinflammation*, vol. 16, no. 3, 2019.