



# Discovery and Validation of Key Biomarkers Based on Immune Infiltrates in Alzheimer's Disease

Zhuohang Liu<sup>1,2</sup>, Hang Li<sup>2\*</sup> and Shuyi Pan<sup>1,2\*</sup>

<sup>1</sup> The Fifth Clinical Medical College of Anhui Medical University, Beijing, China, <sup>2</sup> Department of Hyperbaric Oxygen, Sixth Medical Center, Chinese PLA General Hospital, Beijing, China

## OPEN ACCESS

### Edited by:

Tanima Bose,  
Ludwig Maximilian University  
of Munich, Germany

### Reviewed by:

Nora Gray,  
Oregon Health and Science  
University, United States  
Deepika Mathur,  
Icahn School of Medicine at Mount  
Sinai, United States  
Naazneen Khan,  
University of Kentucky, United States

### \*Correspondence:

Hang Li  
363368350@qq.com  
Shuyi Pan  
psy9992011@163.com

### Specialty section:

This article was submitted to  
Neurogenomics,  
a section of the journal  
Frontiers in Genetics

**Received:** 25 January 2021

**Accepted:** 05 May 2021

**Published:** 01 July 2021

### Citation:

Liu Z, Li H and Pan S (2021)  
Discovery and Validation of Key  
Biomarkers Based on Immune  
Infiltrates in Alzheimer's Disease.  
Front. Genet. 12:658323.  
doi: 10.3389/fgene.2021.658323

**Background:** As the most common neurodegenerative disease, Alzheimer's disease (AD) leads to progressive loss of cognition and memory. Presently, the underlying pathogenic genes of AD patients remain elusive, and effective disease-modifying therapy is not available. This study explored novel biomarkers that can affect diagnosis and treatment in AD based on immune infiltration.

**Methods:** The gene expression profiles of 139 AD cases and 134 normal controls were obtained from the NCBI GEO public database. We applied the computational method CIBERSORT to bulk gene expression profiles of AD to quantify 22 subsets of immune cells. Besides, based on the use of the Least Absolute Shrinkage Selection Operator (LASSO), this study also applied SVM-RFE analysis to screen key genes. GO-based semantic similarity and logistic regression model analyses were applied to explore hub genes further.

**Results:** There was a remarkable significance in the infiltration of immune cells between the subgroups. The proportions for monocytes, M0 macrophages, and dendritic cells in the AD group were significantly higher than those in the normal group, while the proportion of some cells was lower than that of the normal group, such as NK cell resting, T-cell CD4 naive, T-cell CD4 memory activation, and eosinophils. Additionally, seven genes (ABCA2, CREBRF, CD72, CETN2, KCNG1, NDUFA2, and RPL36AL) were identified as hub genes. Then we performed the analysis of immune factor correlation, gene set enrichment analysis (GSEA), and GO based on seven hub genes. The AUC of ROC prediction model in test and validation sets were 0.845 and 0.839, respectively. Eventually, the mRNA expression analysis of ABCA2, NDUFA2, CREBRF, and CD72 revealed significant differences among the seven hub genes and then was confirmed by RT-PCR.

**Conclusion:** A model based on immune cell infiltration might be used to forecast AD patients' diagnosis, and it provided a new perspective for AD treatment targets.

**Keywords:** genetic correlation, Alzheimer disease, immune infiltration, bioinformatics analysis, GSEA

## INTRODUCTION

As a neurodegenerative disease, Alzheimer's disease (AD) is usually accompanied by cognitive dysfunction and memory impairment, and even affects speech and motor ability (Ono and Tsuji, 2019). Over 47 million people are estimated to suffer from dementia, and AD as the most common form accounts for 60–80% of them (Barker et al., 2002; Anand et al., 2017). The pathologic changes of AD remain the gold standard for diagnosis characterized by brain atrophy, ventricular enlargement, extracellular amyloid plaques, intracellular neurofibrillary tangles, and so on (Piguat et al., 2009; Serrano-Pozo et al., 2011). However, the preclinical stage of AD patients barely meets the diagnostic criteria until pathology can be found in the postmortem brain. Although amyloid and tau biomarkers can be detected before structural brain changes, there is still a lack of effective diagnosis in the pre-dementia phase.

The risk factors of development for AD diagnosis are important, of which age is dominant (Currais et al., 2019). As the population ages, the prevalence of AD is increasing (Lee et al., 2019). Studies have shown that for people over 60, the incidence doubles per 6.3 years. Besides, family history, Down syndrome, and cerebrovascular disease are also important risk factors (DeTure and Dickson, 2019). These different multifaceted contributors can increase risk for the disease; nevertheless, the genetic factor warrants special emphasis with a higher heritability percentage of about 60–80% (Gratuze et al., 2018).

Familial AD is closely related to mutations in amyloid precursor protein (APP), presenilin 1 (PSEN1), or PSEN2 genes (Ryman et al., 2014). Additionally, the apolipoprotein E4 gene (APOE4) is one of the greatest risks in progressing AD (Ballard et al., 2016; Lane et al., 2018). The genome-wide association studies indicated that TREM2, ADAM10, and PLD3 all affect APP and tau, and regulate directly known functions, including immune response, cholesterol metabolism, and endocytosis (Chouraki and Seshadri, 2014; Karch and Goate, 2015; Jansen et al., 2019). Growing evidence suggests that Alzheimer's disease pathogenesis is strongly associated with immunological mechanism (Shi and Holtzman, 2018; Webers et al., 2020). The binding of AD-related proteins and receptors on microglia can initiate an innate immune response and subsequently cause disease deterioration (Sarlus and Heneka, 2017). Previous studies showed that amyloid plaque burden was negatively correlated with the levels of the immune biomarker, anti-amyloid- $\beta$  (anti-A $\beta$ ) autoantibodies in AD, which links the disease progression and A $\beta$  accumulation (Kellner et al., 2009; Storace et al., 2010). The dysfunction of central and peripheral immune systems was associated with several differentially expressed immune markers including CD3, CD4, CD7, CD28, and CD56 in AD. It has been shown that TREM2 gene mutations greatly increase the risk of AD, and phagocytosis of microglia is mediated with receptors on microglia and astrocytes (Heneka et al., 2015; Yeh et al., 2017). The phagocytic capability for A $\beta$  is weakened, which may be implicated in MGAT-III and Toll-like receptor-3 (Avagyan et al., 2009). Increasing evidence suggests that the rare variants of

TREM2 can promote the development of amyloid and tau pathologies due to the dysfunction of microglia (Gratuze et al., 2018). Also, TREM2 can participate in the interaction with other biomarkers including APOE (Qin et al., 2021). Collectively, Trem2 acts as the immune checkpoint and the scavenger receptor to clear harmful stimuli in the regulation of neurocyte survival (Hickman et al., 2018). Risk genes consolidate a role of the immune system in AD and have been identified relating to complements and various immune cells, including microglia, monocytes, and B and T lymphocytes (Jevtic et al., 2017). Up to now, there is no disease-modifying therapy for Alzheimer's disease. Therefore, finding novel risk factors and targeting these immune mechanisms are needed to develop future therapeutic approaches for Alzheimer's disease.

To explore and identify potential biomarkers of AD, gene chips were obtained from the Gene Expression Omnibus database. We then identified hub genes through LASSO and BORUTA analysis. The combined use of two algorithms for screening biomarkers will make the process more rigorous and standard; finally, identified key biomarkers are strongly associated with AD. Then the CIBERSORT algorithm was applied to calculate the immune infiltration of AD samples. Moreover, the relationship between 22 immune cells and hub genes was analyzed. The logistic regression model was investigated by ROC and further validated by an external dataset. This is of great significance and the first utilization of a combination of LASSO and SVM-RFE method to reveal the key biomarkers in AD.

## MATERIALS AND METHODS

### Gene Expression Data Acquisition

The GSE63061 data file was downloaded from the NCBI Gene Expression Omnibus public database (GEO<sup>1</sup>) annotated by GPL10558 as a Series Matrix File. The file contains data related to 273 groups of patients' expression profiles, including 134 normal groups and 139 AD patients. Meantime, the GSE85426 Series Matrix File was obtained from the GEO public database. The annotation file is GPL14550, which consists of 180 groups of patients' expression profile data involving 90 cases of normal group and 90 cases of AD patients.

### Analysis of Immune Cell Infiltration

The CIBERSORT is widely used in evaluating the type of immune cell in the microenvironment. The tool is based on the principle of linear support vector regression to perform deconvolution analysis on the expression matrix of immune cell subtypes. It contains 547 biomarkers, and 22 phenotypes of human immune cell, covering plasma cells, B cells, T cells, and myeloid cell subsets, have also been defined. Using the CIBERSORT algorithm, this study analyzed the AD patients' data and quantified the relative proportion of 22 infiltrating immune cells. Furthermore, this study performed Spearman correlation analysis on immune cells and gene expression.

<sup>1</sup><http://www.ncbi.nlm.nih.gov/geo/>

## Construction of LASSO Model and SVM-RFE Feature Selection Process

The LASSO logistic regression and the SVM algorithm were used to classify the diagnostic markers of AD. The LASSO analysis was undertaken using the “glmnet” package, the response type was set as binomial, and the alpha was set as 1. Besides, as a surveillant machine learning method to support vectors, the Support Vector Machine (SVM) finds the best variables by deleting the feature vectors generated by the SVM. The SVM classifier from R package e1071 was adopted for classification analysis of the selected biomarkers in the diagnosis of AD;  $k = 5$  was the setting for the  $k$ -fold cross validation, and the parameter of halving above was identified as 100.

## Gene Set Enrichment Analysis (GSEA)

Patients are differentiated into two groups (the high expression group and the low expression group) based on the expression level of seven pivot genes. The GSEA analysis of the two groups was achieved by applying signal pathway differences. Molecular Signature Database v7.0 provided the background gene set data required for this study (MSigDB)<sup>2</sup>. Annotated gene sets were used to distinguish subtypes by the identified differentially expressed genes. We computed the consistency  $P$ -value for each gene set, and  $P$ -values less than 0.05 were considered significantly enriched. Subsequently, significantly enriched gene sets were ranked. The association between disease type and biological processes was analyzed by using GSEA.

## Semantic Similarity GO Annotations

In this paper, we incorporated GO-based semantic similarity and ranked the proteins by the functional similarity of protein–protein interactions. The relationship between semantic similarity in the GO annotation and the correlation of the gene expression profile was verified (Sevilla et al., 2005). It provides a basis for the functional comparison of gene products, so it has been widely used in bioinformatics, such as protein–protein interaction analysis (Jain and Bader, 2010), pathway analysis (Guo et al., 2005), and gene function prediction (Tedder et al., 2010). Here, we measured the functional similarity between proteins. The functional similarity is defined as the geometric mean of the semantic similarity of GO in terms of the molecular function (MF) and the biological process (BP). To quantify the correlations between proteins, the biological function and the pathway were included. The semantic similarity of interacting proteins in MF and BP is performed by using the GO SemSim software package of Wang’s method (Yu et al., 2010). It was carried out in a more accurate method that involved GO topology (Wang et al., 2007). The geometric average of semantic similarity in MF and BP algorithms was used to further estimate the functional similarity.

## Co-expression Analysis of the Hub Genes

Correlation analysis was performed using the corrplot and circlize package in R software. Pearson correlation on the hub

gene expression was carried out using the package corrplot (version 0.84) in R. The circlize plot was generated through the circlize package. The correlation coefficient was depicted by the colors “red” and “green.” Red denotes a positive correlation, whereas green indicates a negative correlation. The darker the color and the thicker the string, the stronger the correlation.

## Differential Expression Analysis

To explore the relationship between hub genes and the development of AD, we performed the differential expression analysis for hub genes in different types of brain tissues using the AlzData. AlzData<sup>3</sup> is a comprehensive AD high-throughput omics database (Xu et al., 2018). AlzData can be used as an in-depth integration system, integrating different levels of data for disease characterization. It is comprised of genomics (GWAS and whole exome sequencing), transcriptome, proteomics, and functional genomics data. In addition, the study also further calculated the CFG (convergent functional genomic) score (Supplementary Table 1) of key genes.

## RT-PCR Validation of the Hub Genes

The serum samples of six patients without AD and six patients with AD were acquired for RT-PCR verification in order to verify the predictive analysis results. The Ethics Committees or institutional review board of the Sixth Medical Center of Chinese PLA General Hospital approved this protocol. Based on the manufacturer’s instructions, this study used a TRIzol reagent to extract total RNA (Invitrogen, 10296028). RNA quality (RNA integrity number, RIN) was determined by Agilent 2100 Bioanalyzer analysis (Agilent Technologies, CA, United States), and  $RIN \geq 8$  for subsequent experiments. RNA samples from total RNA were reverse-transcribed to cDNA, and RT-PCR was performed using the SuperScript III RT (ABI-Invitrogen, 11752050).  $\beta$ -Actin was used as an internal normalization standard. Relative mRNA expression was calculated using the  $\Delta\Delta CT$  method. Primers were as described in the Table 1.

## Statistical Analysis

$T$ -test was used for measurement data (expressed as a mean  $\pm$  SEM), while chi-squared test was used for categorical data (presented as percentages). The logistic regression algorithm was used to build the predictive model. All statistical analyses were performed by R (version 3.6), and GraphPad Prism 8. All experiments were done in triplicate. Significance was defined as  $P < 0.05$  for two-sided tests.

## RESULTS

### Identification of Hub Genes

We downloaded the GSE63061 dataset from the NCBI GEO public database. The total number of patients was 273 (AD group, 139; control group, 134). To explore the biomarkers of AD, we performed feature screening through LASSO regression. The results from the LASSO regression showed that 18 genes were identified as signature genes in AD. On the other hand, we use the

<sup>2</sup><http://www.brAdd.mit.edu/GSEA/msigdb/>

<sup>3</sup><http://www.alzdata.org/>

**TABLE 1** | Primer information.

Target name		Primer
β-Actin	F	GACAGGATGCAGAAGGAGATTACT
	R	TGATCCACATCTGCTGGAAGGT
ABCA2	F	CATGCCATCTTCATCATCC
	R	TTGAGCATGTCCCACTCGAA
CREBRF	F	GGAACAAACTCTGATGC
	R	CATTAGTTGGCTGTTTAC
CD72	F	CATCTCCAGCAGGTTAGGACA
	R	CGGGCACTTGAACATTCTC
CETN2	F	AAAGAATGAGACCTAAGCC
	R	CTTTAACATCTATAGTGCAG
KCNG1	F	CCGAGTTCACCTGCATCCC
	R	CCATAGCCCACCGTCTC
NDUFA2	F	TCTCATCCACTTATGTC
	R	CTTCACATCGGAGCCT
RPL36AL	F	CCTACCTAAAACCCGAAT
	R	TTCTTTGGTCTTAGCGTT

SVM-RFE algorithm to evaluate the characteristic genes in AD. The data illustrated that a total of 7 differentially expressed genes were acquired from the intersection of the top 100 scoring genes and the characteristic genes selected by the LASSO regression

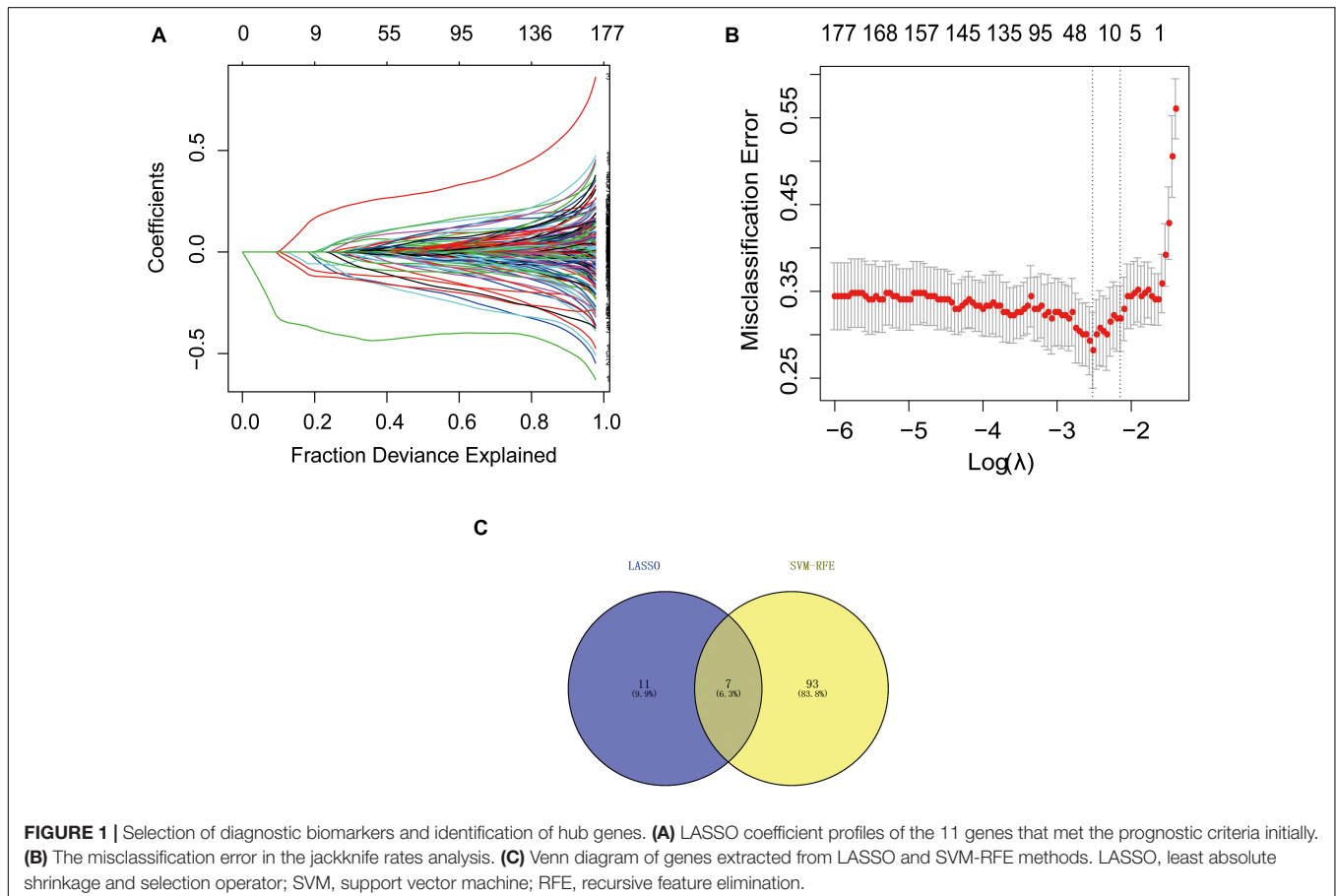
algorithm. The intersection genes were used as the core genes for subsequent research (Figure 1).

## Immune Infiltration Analyses

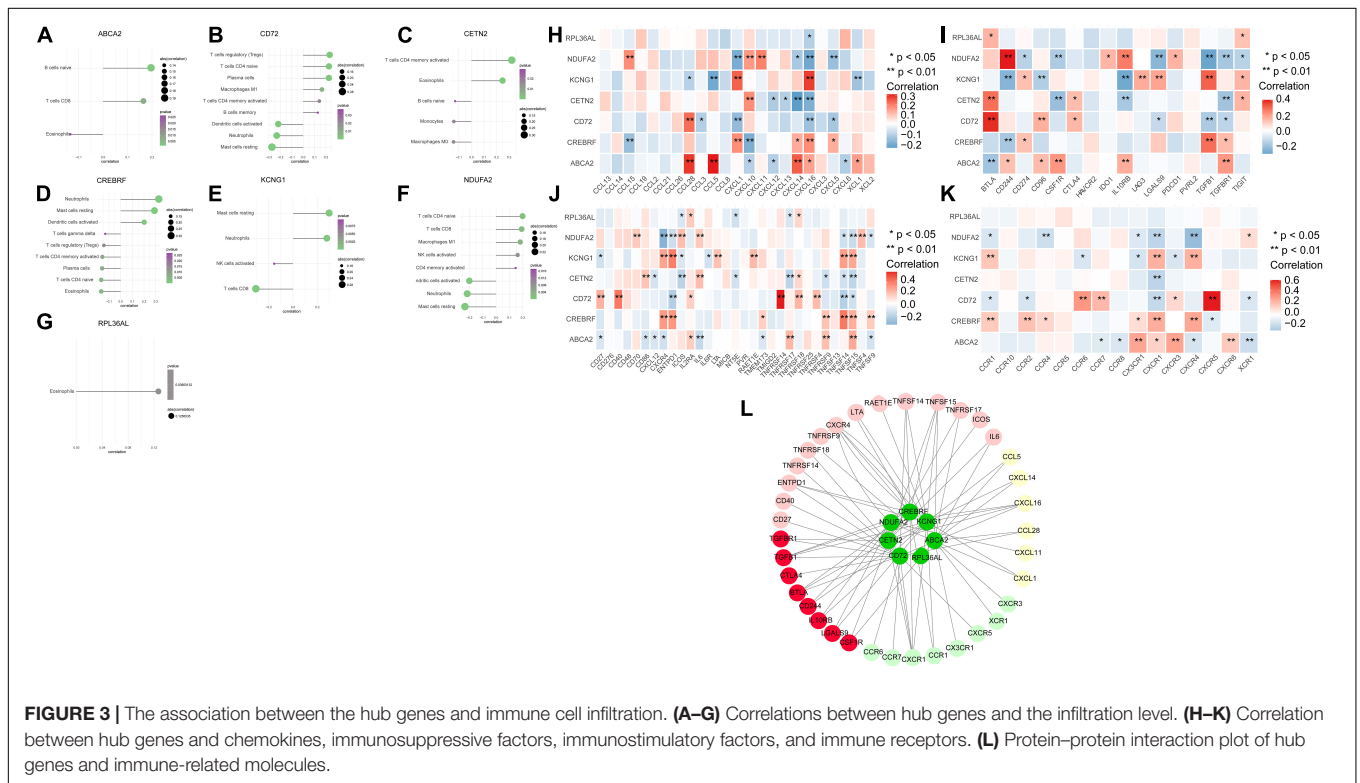
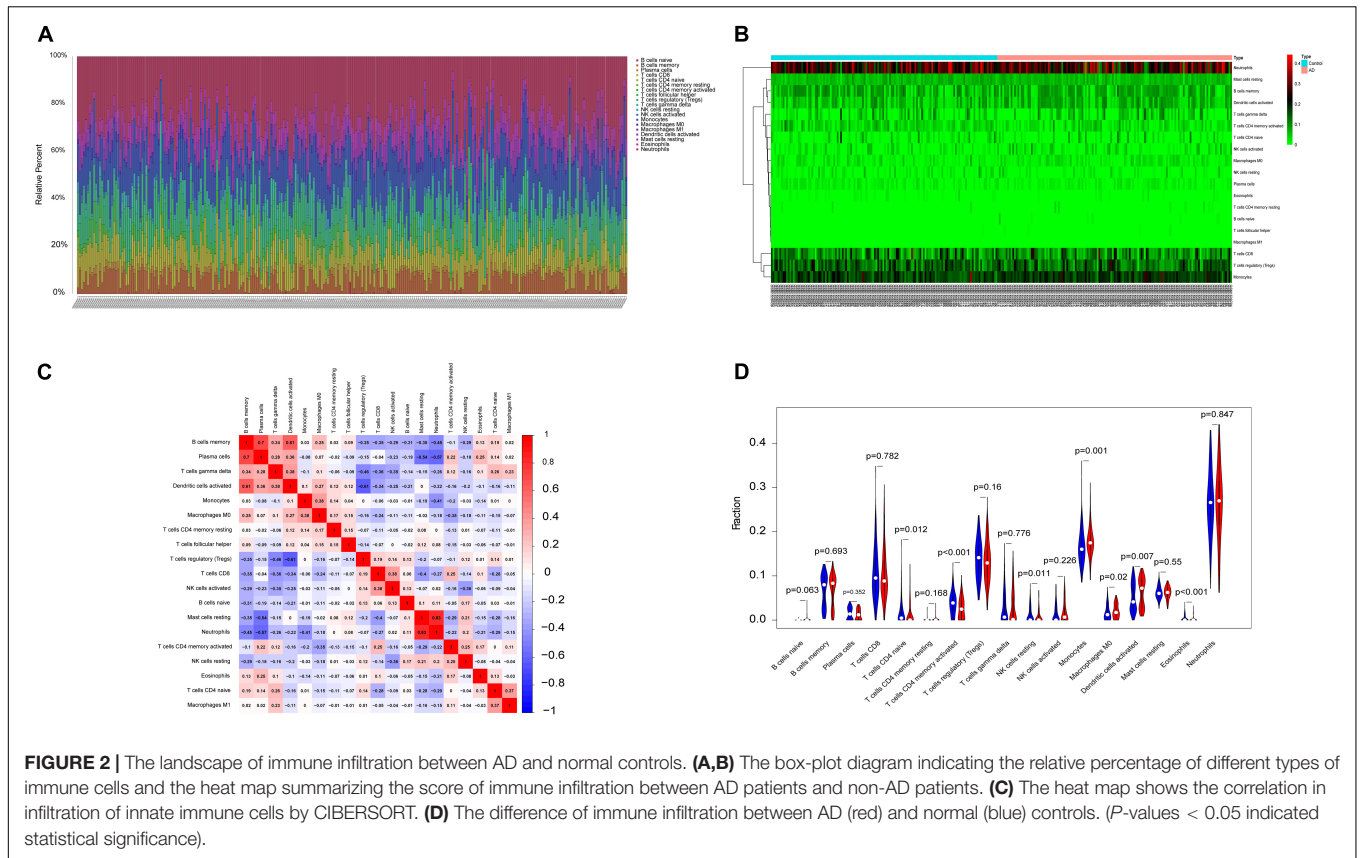
The microenvironment is composed of immune cells, inflammatory factors, extracellular matrix, and various growths, and it has an important impact on clinical treatment sensitivity and disease diagnosis. Through studying the relationship between immune infiltration and the hub gene in the AD dataset, the potential molecular mechanisms of core genes influencing the AD progression have been further explored. The results indicated that the fractions for monocytes, M0 macrophages, and dendritic cells in the AD group were remarkably higher compared with those of the normal patients, while the fractions of many cells are lower than those of the normal patients, such as T-cell CD4 memory activation, T-cell CD4 naive, eosinophils, and NK cell resting. The interaction between the immune cells is shown in Figure 2.

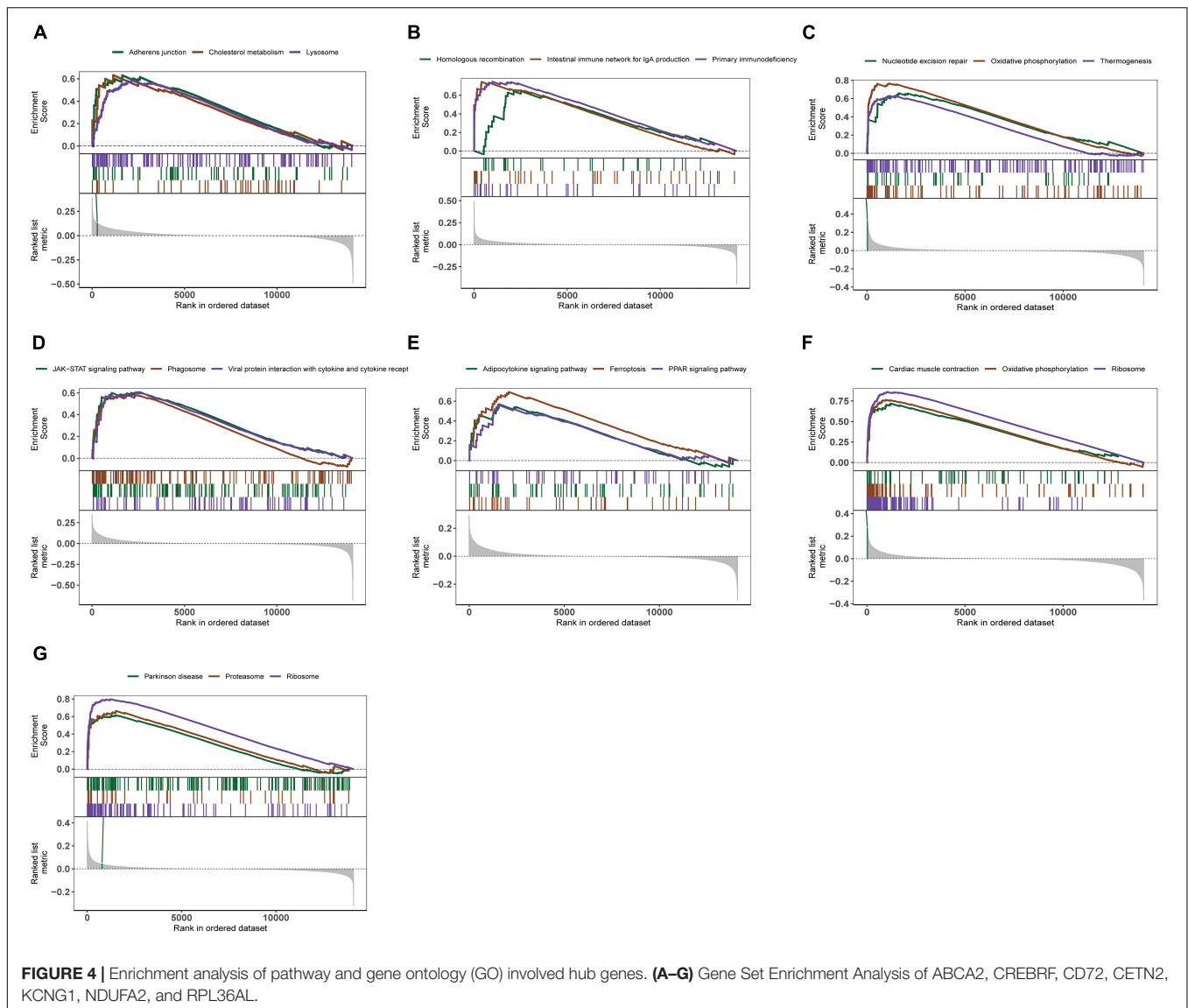
## Analysis of Hub Genes and Immune Infiltration

To explore further the relationship between core genes and immune cells, the results indicated that ABCA2 was positively correlated with B cells naive. Besides, CREBRF was positively correlated with neutrophils, dendritic cells activated, and mast







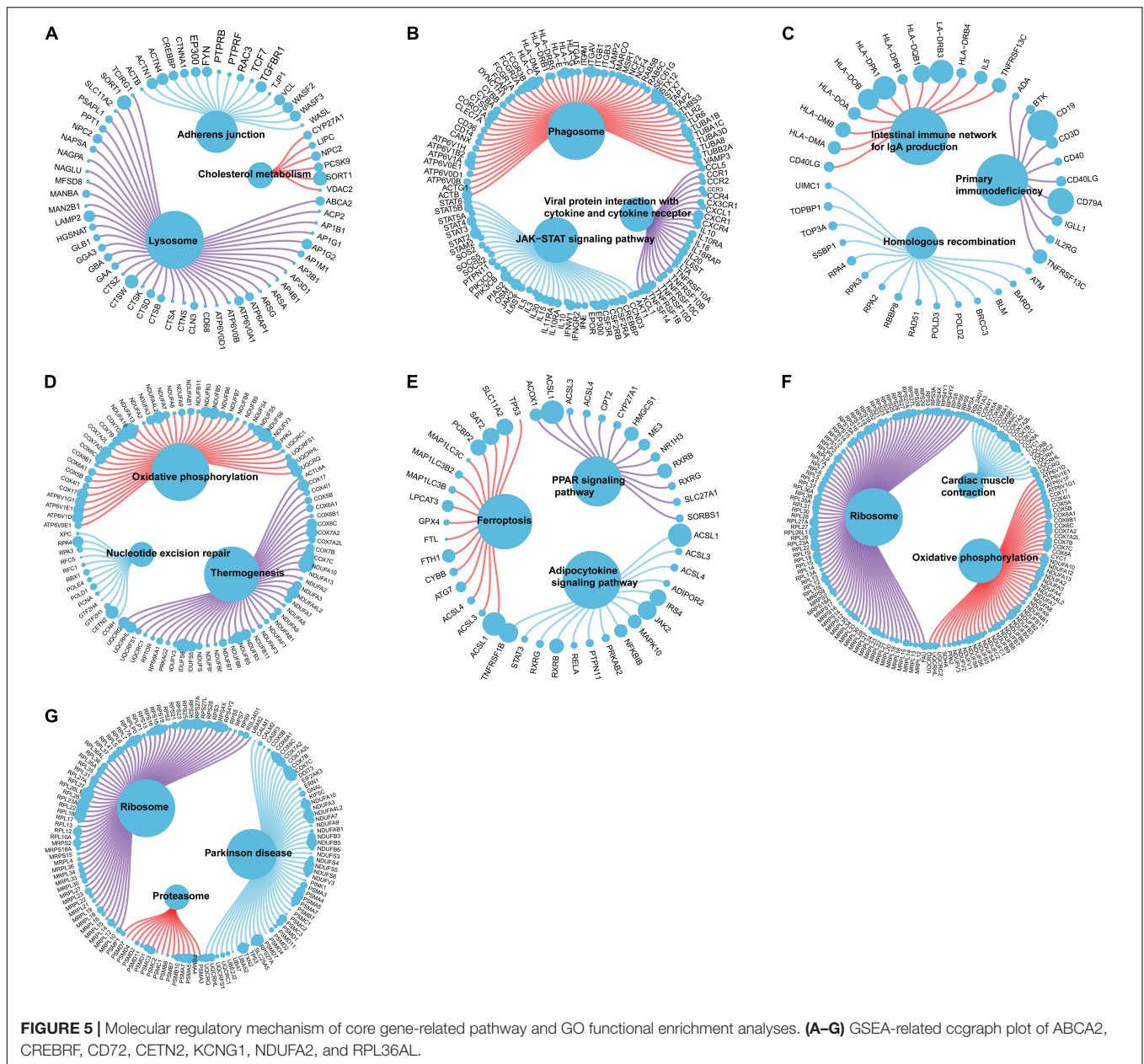


cells resting. The CD72 was positively correlated with T regulatory cells (Tregs), T-cell CD4 naïve, and plasma cells but negatively correlated with dendritic cells activated, neutrophils, and mast cells resting. CETN2 had a positive relationship with T cells, eosinophils, and CD4 memory activation. KCNG1 was positively correlated with mast cells resting and neutrophils, while it was negatively correlated with T-cell CD8. There was a positive correlation between NDUFA2 and T-cell CD4 naïve, and it was negatively associated with dendritic cells activated, neutrophils, and mast cells resting (**Figures 3A–G**). We next obtained the correlations between these seven key genes and different immune factors including immunosuppressive factors, immunostimulatory factors, chemokines, and receptors from the TISIDB database. The graphs of the correlational relationships between immune factors and AD core genes were constructed (**Figures 3H–K**). We selected immune factors related to the hub gene (the average correlation coefficient greater than 0.2) and

used STRING and Cytoscape to construct an interaction network (**Figure 3L**). These analyses confirmed that key genes were closely related to the level of immune cell infiltration and played a crucial role in the immune microenvironment.

## Analyses of Biological Processes and Core Pathways Enriched for Genes in AD

We further explored the specific signaling pathways involved in the seven core genes, and explore the potential molecular mechanisms of them that affect the progression of AD. GSEA results showed that the pathway involved in the high expression of ABCA2 is cholesterol metabolism. Meanwhile, the pathway involved in the high expression of KCNG1 is the adipocytokine signaling pathway. The pathways involved in the high expression of CETN2 are oxidative phosphorylation and thermogenesis, all related to energy metabolism. It suggested that these three core

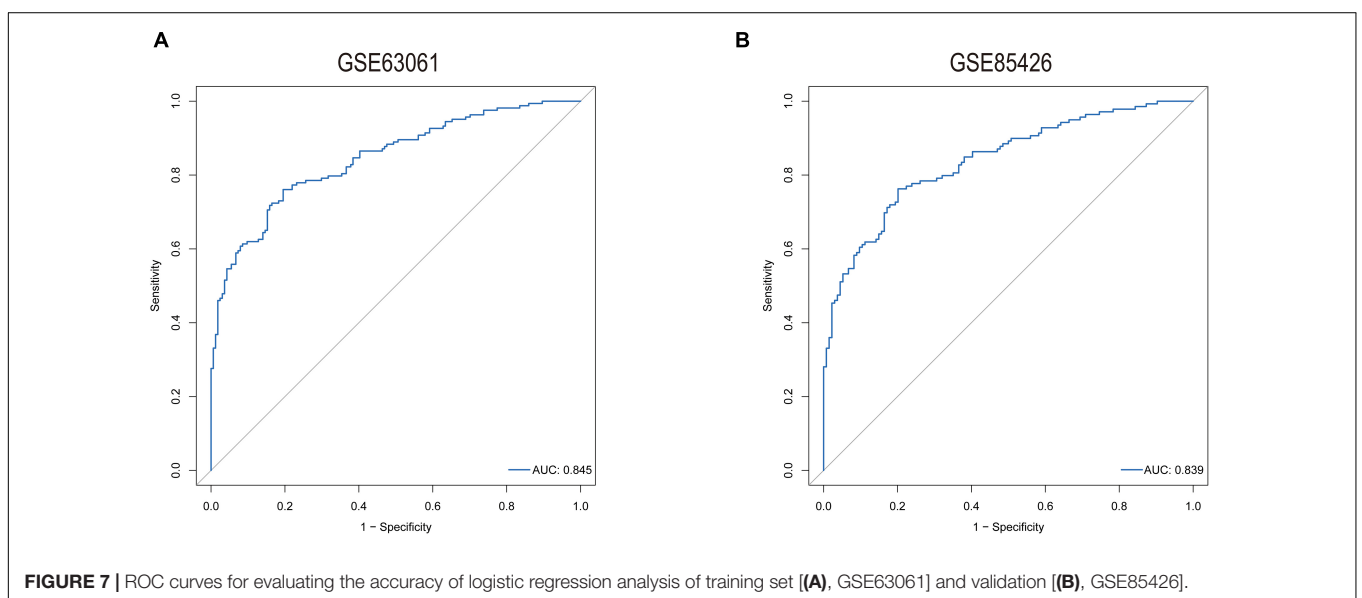
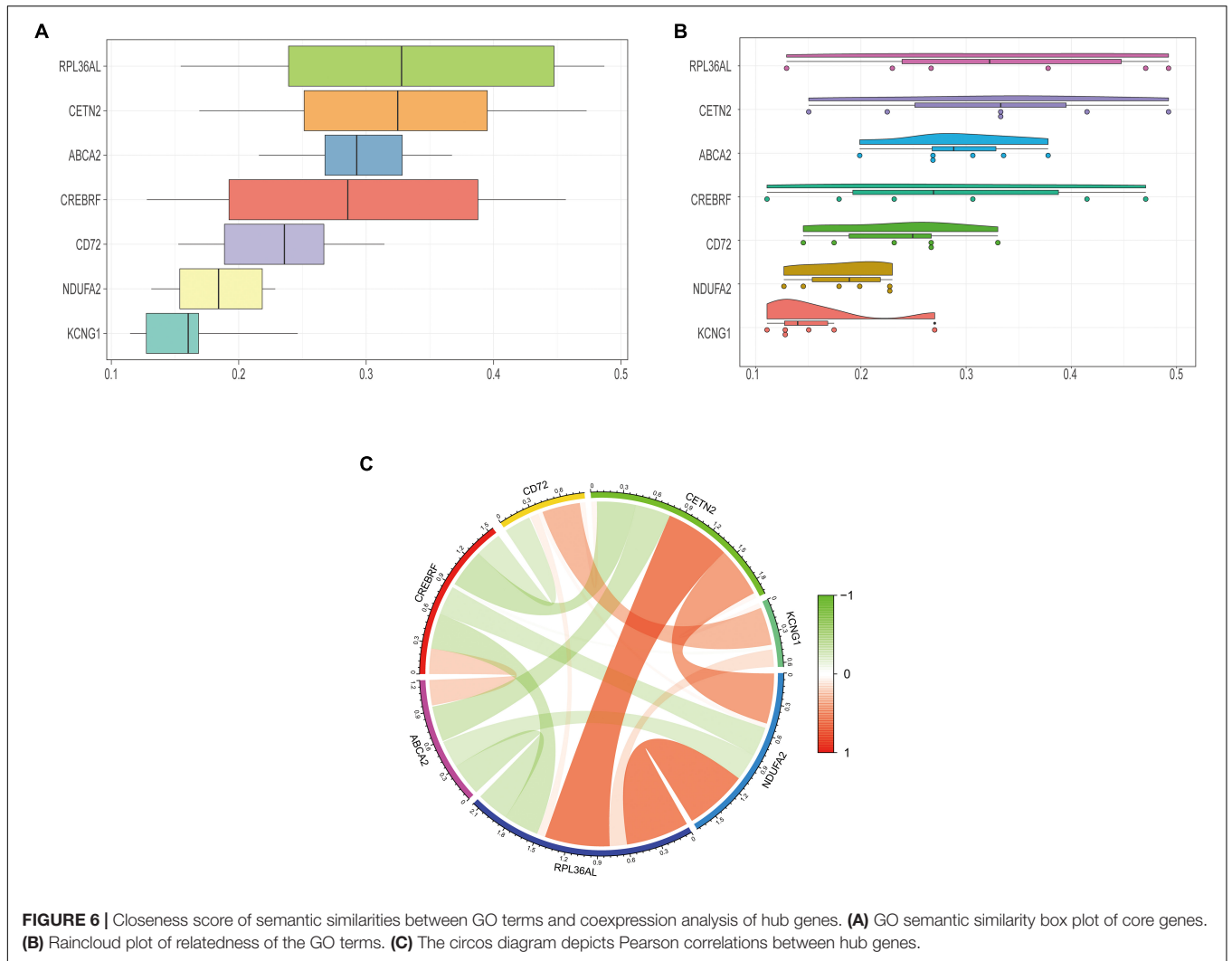


genes may participate in the progression of AD by influencing the body’s metabolism. Concurrently, high expressions of NDUFA2 and RPL36AL are related to the ribosome pathway. It suggested that both of them can affect the progression of AD by regulating the function of the ribosome (Figure 4). The molecular interaction network between each pathway is shown in Figure 5. Also, to further identify the core genes that play a key role in AD, we ranked the core genes according to the average functional similarity relationship between the proteins. The results demonstrated that RPL36AL, CETN2, and ABCA2 are the top 3 molecules based on the GO similarity score in AD; their median scores were around 0.35, 0.35, and 0.30, respectively. Besides, the rest of them were under 0.3 remarkably (Figures 6A,B). Correlation analysis of seven hub genes was indicated using the circle diagram in Figure 6C. Furthermore,

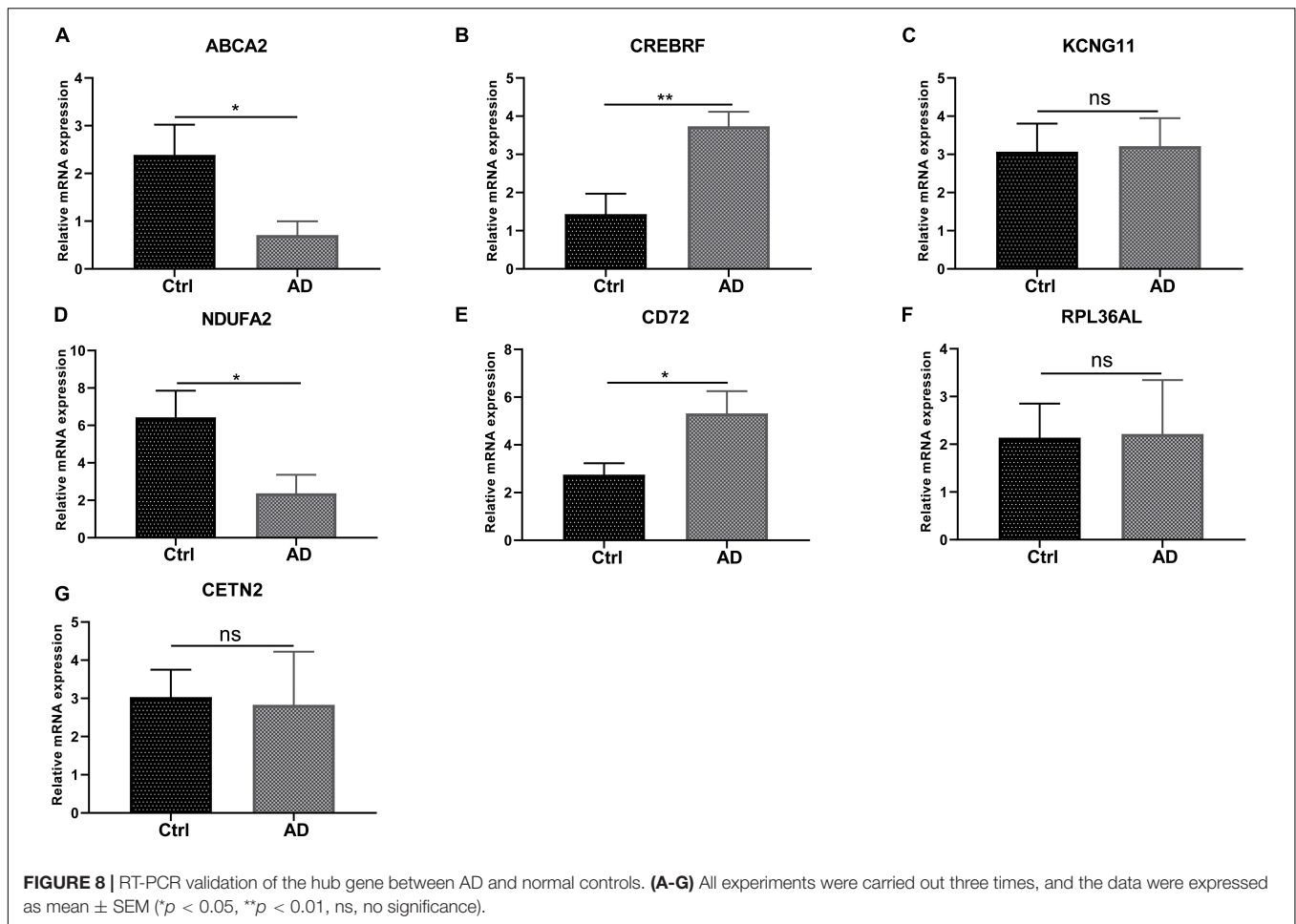
RPL36AL, CETN2, KCNG1, and NDUFA2 displayed closely positive relationships, but the remaining three genes mainly revealed negative associations with each other. The mechanism remains to be further explored.

### Logistic Regression Is Potential to Build a Core Gene Correlation Model

The predictive model was constructed by using the logistic regression algorithm. The results showed that the predictive model constructed by seven genes had better diagnostic performance, and the area value under the receiver operating characteristic curve (AUC) was 0.845. We further downloaded the dataset GSE85426, of which 90 people had AD and 90 were normal patients. Fivefold cross-validation was performed







to further verify the diagnosis model, as external validation data. The results showed that the model had strong stability, and the AUC value was 0.839 (Figure 7).

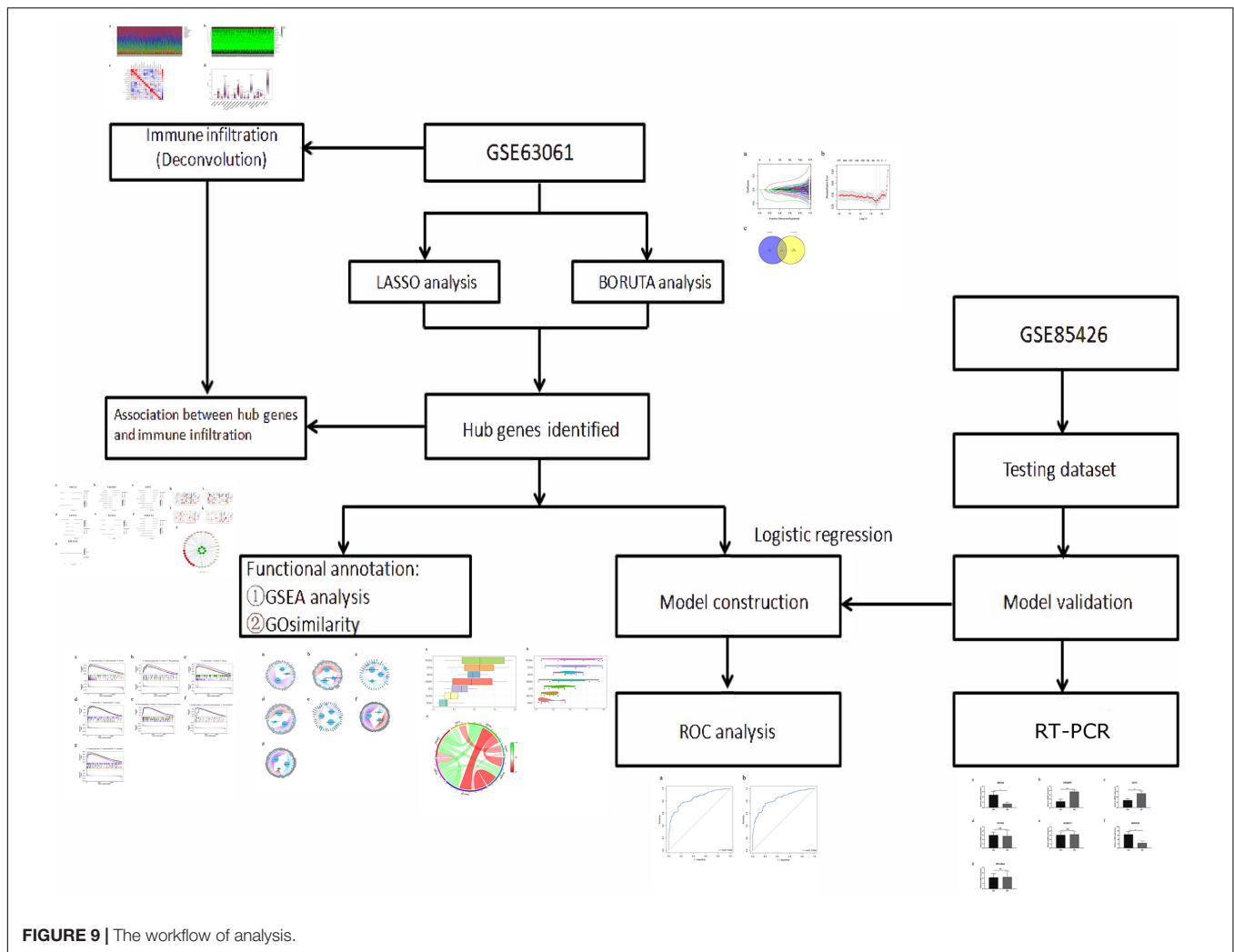
### The Experiment of Seven Hub Genes

We next conducted RT-PCR experiments to detect the hub genes' relative expression level in AD and normal control groups. The research data demonstrated that the mRNA expression level of ABCA2 and NDUFA2 in AD decreases compared with that of the controls ( $p < 0.05$ ). Conversely, the opposite results were observed for CREBRF and CD72 ( $p < 0.05$ , CREBRF;  $p < 0.01$ , CD72). Besides, there was no significant difference in the levels of CETN2, KCNG1, and RPL36AL between the AD group and the control group (Figure 8). These identified seven genes might function as the potential diagnostic and prognostic biomarkers.

## DISCUSSION

The most common cause of degenerative dementia is Alzheimer's disease (AD). So far, amyloid plaques and neurofibrillary tangles are its main pathological diagnostic features (DeTure and Dickson, 2019). Nevertheless, detection of AD brain lesions using

histopathology is suboptimal, and treatments were currently in a limited pharmacotherapy phase. A growing number of researchers realize that immune infiltration is related to the diagnosis in cancer and other diseases (Barnes and Amir, 2018; Varn et al., 2018; Cai et al., 2020). In this context, the immune system plays an essential role in the etiology of AD. Additionally, immune infiltrates were linked to the target and clearance of amyloid beta-peptide plaques in the brain (Monsonogo et al., 2006). The central immune system compositions such as microglia and complements, as well as peripheral immune-related components including monocytes and lymphocytes, can influence the pathology of AD (Jevtic et al., 2017). The previous research demonstrated that C1q is highly correlated with A $\beta$  deposition and activates microglia phagocytosis, which, in turn, leads to synapse loss (Hong et al., 2016). In addition to C1q, elevated A $\beta$  burden can be mediated through C3, also associated with NF- $\kappa$ B in AD (Ascolani et al., 2012; Shi et al., 2017). Besides, CD33 as pivotal genes of microglia reduces phagocytosis, but TREM2 has a reverse effect (Jiang et al., 2008; Griciuc et al., 2013). Equally, the peripheral immune cells such as monocytes, neutrophils, and lymphocytes can promote neuroinflammation in the progress of AD (Erny et al., 2015). When the monocytes were deleted in a mice model, it will worsen the AD pathology (El Khoury et al., 2007). Except for monocyte, neutrophil was proven



**FIGURE 9** | The workflow of analysis.

that its neutrophil extravasation ability and amyloid pathology were ameliorated by the inhibition of LFA-1 (Zenaro et al., 2015). In absence of certain lymphocytes (functional B and T cells), Rag2 knockout can decrease A $\beta$  levels in PSAPP AD mice brain (Späni et al., 2015). Yet, few studies have systematically screened the biomarkers related to AD immune infiltration and their value in assessing AD immune infiltration.

In the present study, we used LASSO regression analysis and the SVM-RFE method to screen out seven potential genes (ABCA2, CREBRF, CD72, CETN2, KCNG1, NDUFA2, and RPL36AL). Subsequently, the CIBERSORT algorithm performed deconvolution analysis on the immune microenvironment to assess the proportion of the immune cells in AD. What is more, previous research confirmed the effectiveness of the CIBERSORT technique (Dai et al., 2018; Mao et al., 2018; Chang et al., 2020). Besides, the immune cell infiltration scoring model was built by LASSO regression based on 22 immune cells. GO semantic similarity and GSEA analysis annotated the profiles of the seven hub genes. GSEA analysis suggested that there were considerably abundant genes in diverse immune-related biological processes and pathways. The RNA characteristics

as indicators of immune cell infiltration in AD patients were identified through adopting this calculation framework to AD patient datasets and immune cell lines. Our research combined with the results of bioinformatics analysis and RT-PCR and showed that there were significant differences in the samples of AD patients containing four genes. We found that ABCA2 and NDUFA2 in the AD group were significantly lower than those in the control group; CREBRF and CD72 mRNA in the AD patients had significantly higher mRNA expression level compared to the control group (**Figure 9**). Likewise, differential expression analysis of hub genes in various brain cell types was revealed (**Supplementary Figure 1** and **Supplementary Table 1**). ABCA2 is rich in the frontal and temporal areas of the AD brain, involving resistance to reactive oxygen species in the ABCA2 transfected cell line (Chen et al., 2004). In agreement, it is ranked in the top 3 for both GO similarity analyses. Previous reports have pointed out that ABCA2 is a key regulator of endogenous APP expression and AD truncation. In research, CREBRF was proposed as a novel biomarker using weighted gene co-expression network analysis (WGCNA) based on a total of 329 samples (Soleimani Zakeri et al., 2020). Moreover, it has been documented to block

autophagy through the CREB3/ATG5 pathway in brain tumor (Xue et al., 2016). Studies showed that CETN2 modulates male mice infertility or dysosmia and affects embryonic development, which is regulated by the FGF/FGFR gene expression (Shi et al., 2015; Ying et al., 2019). CD72 plays a negative role in the regulation of B-cell activation (Davis, 2010). It is reported that the combined effects between Sema4D and CD72 induce immune activation in the central nervous system (Tsubata, 2019). However, the study for CD72 in AD remains poorly understood. Therefore, we call for further research to better elucidate this aspect.

Due to the limited number of samples, there is still a need to confirm these preclinic observations in future clinic studies for novel biomarkers. Moreover, most AD cases have other neurodegenerative comorbidities. Similarly, this phenomenon was also observed in the profile of 1,153 AD pathologically diagnosed patients acquired from Mayo Clinic Brain Bank data (DeTure and Dickson, 2019). Compared to other neurodegenerative diseases, many pathological changes are highly supportive but not specific to AD. Besides, there are similar pathological features between various diseases. For example, Down's syndrome patients with chromosome 21 trisomy where APP is located can have Alzheimer-like pathology (Kang et al., 1987). APOE4, as a common risk factor of AD, also carries certain morbidity risk for vascular dementia, Down's syndrome, and brain injury (Verghese et al., 2011). These will need further study.

## CONCLUSION

In this study, the overlap of the LASSO model and the SVM-RFE algorithm was obtained, and eventually, seven hub genes were recognized. RT-PCR detected these hub genes' expression levels. To understand AD development, the GSEA and GO analysis of the selected genes supplied a more specific molecular mechanism. Through carrying out the deconvolution algorithms on the patient data in the GEO database, this study revealed that there was distinctness in immune infiltration between the two groups. So far, the relationship between key genes and immune infiltration has been little reported. The mechanism of key genes and immune infiltration-related factors in the diagnosis of AD remains to be explored.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## REFERENCES

- Anand, A., Patience, A. A., Sharma, N., and Khurana, N. (2017). The present and future of pharmacotherapy of Alzheimer's disease: a comprehensive review. *Eur. J. Pharmacol.* 815, 364–375. doi: 10.1016/j.ejphar.2017.09.043
- Ascolani, A., Balestrieri, E., Minutolo, A., Mosti, S., Spalletta, G., Bramanti, P., et al. (2012). Dysregulated NF- $\kappa$ B pathway in peripheral mononuclear cells of Alzheimer's disease patients. *Curr. Alzheimer Res.* 9, 128–137. doi: 10.2174/156720512799015091

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Sixth Medical Center of Chinese PLA General Hospital approved this protocol. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

ZL: conceptualization, methodology, data curation, writing—original draft preparation, visualization, investigation, software, and performed the experiments. HL and SP: conceptualization, methodology, supervision, writing—reviewing and editing. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by the Beijing Nova Program (Z171100001117092), Research Found of Chinese PLA General Hospital—Sixth Medical Center (CXPY201816), the National Nature Science Foundation of China (81500669), and the National Nature Science Foundation of Beijing (7154241).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.658323/full#supplementary-material>

**Supplementary Figure 1** | Differential expression analysis of hub genes in the AlzData database. **(A–F)** Differential expression analysis of ABCA2, CD72, CETN2, KCNG1, NDUFA2, and RPL36AL for various cell types in the brain tissue. No data available about CREBRF in AlzData.

**Supplementary Table 1** | Convergent functional genomic (CFG) ranking for target genes. 1. eQTL: expression of target gene is regulated by AD genetic variants (genetic variants: IGAP GWAS P-val < 1E-3; eQTL: P-val < 1E-3); 2. GWAS: IGAP P-val < 1E-3; 3. Early\_DEG: target gene is differentially expressed in AD mouse models before AD pathology emergence; 4. CFG: total CFG score of target gene, 1 CFG point is added if any of above evidence is significant, CFG point ranges from 0 to 5.

- Avagyan, H., Goldenson, B., Tse, E., Masoumi, A., Porter, V., Wiedau-Pazos, M., et al. (2009). Immune blood biomarkers of Alzheimer disease patients. *J. Neuroimmunol.* 210, 67–72.
- Ballard, C., Mobley, W., Hardy, J., Williams, G., and Corbett, A. (2016). Dementia in down's syndrome. *Lancet Neurol.* 15, 622–636.
- Barker, W. W., Luis, C. A., Kashuba, A., Luis, M., Harwood, D. G., and Loewenstein, D. (2002). Relative frequencies of Alzheimer disease, Lewy body, vascular and frontotemporal dementia, and hippocampal sclerosis in the State of Florida

- Brain Bank. *Alzheimer Dis. Assoc. Disord.* 16, 203–212. doi: 10.1097/00002093-200210000-00001
- Barnes, T. A., and Amir, E. (2018). HYPE or HOPE: the prognostic value of infiltrating immune cells in cancer. *Br. J. Cancer* 118, 451–460. doi: 10.1038/bjc.2017.220
- Cai, W., Li, H., Zhang, Y., and Han, G. (2020). Identification of key biomarkers and immune infiltration in the synovial tissue of osteoarthritis by bioinformatics analysis. *PeerJ* 8:e8390. doi: 10.7717/peerj.8390
- Chang, Z., Huang, R., Fu, W., Li, J., Ji, G., Huang, J., et al. (2020). The construction and analysis of ceRNA network and patterns of immune infiltration in colon adenocarcinoma metastasis. *Front. Cell Dev. Biol.* 8:688.
- Chen, Z. J., Vulevic, B., Ile, K. E., Soulika, A., Davis, W., Reiner, P. B., et al. (2004). Association of ABCA2 expression with determinants of Alzheimer's disease. *FASEB J.* 18, 1129–1131. doi: 10.1096/fj.03-1490fje
- Chouraki, V., and Seshadri, S. (2014). Genetics of Alzheimer's disease. *Adv. Genet.* 87, 245–294.
- Currais, A., Huang, L., Goldberg, J., Petrascheck, M., Ates, G., Pinto-Duarte, A., et al. (2019). Elevating acetyl-CoA levels reduces aspects of brain aging. *Elife* 8:e47866.
- Dai, W., Feng, Y., Mo, S., Xiang, W., Li, Q., Wang, R., et al. (2018). Transcriptome profiling reveals an integrated mRNA-lncRNA signature with predictive value of early relapse in colon cancer. *Carcinogenesis* 39, 1235–1244. doi: 10.1093/carcin/bgy087
- Davis, W. (2010). The ATP-binding cassette transporter-2 (ABCA2) increases endogenous amyloid precursor protein expression and A $\beta$  fragment generation. *Curr. Alzheimer Res.* 7, 566–577. doi: 10.2174/156720510793499002
- DeTure, M. A., and Dickson, D. W. (2019). The neuropathological diagnosis of Alzheimer's disease. *Mol. Neurodegener.* 14:32.
- El Khoury, J., Toft, M., Hickman, S. E., Means, T. K., Terada, K., Geula, C., et al. (2007). Ccr2 deficiency impairs microglial accumulation and accelerates progression of Alzheimer-like disease. *Nat. Med.* 13, 432–438. doi: 10.1038/nm1555
- Erny, D., Hrabě de Angelis, A. L., Jaitin, D., Wieghofer, P., Staszewski, O., and David, E. (2015). Host microbiota constantly control maturation and function of microglia in the CNS. *Nat. Neurosci.* 18, 965–977. doi: 10.1038/nn.4030
- Gratuzze, M., Leyns, C. E. G., and Holtzman, D. M. (2018). New insights into the role of TREM2 in Alzheimer's disease. *Mol. Neurodegener.* 13:66.
- Griciuc, A., Serrano-Pozo, A., Parrado, A. R., Lesinski, A. N., Asselin, C. N., and Mullin, K. (2013). Alzheimer's disease risk gene CD33 inhibits microglial uptake of amyloid beta. *Neuron* 78, 631–643. doi: 10.1016/j.neuron.2013.04.014
- Guo, X., Shriver, C. D., Hu, H., and Liebman, M. N. (2005). Analysis of metabolic and regulatory pathways through gene ontology-derived semantic similarity measures. *Annu. Symp. Proc.* 2005:972.
- Heneka, M. T., Carson, M. J., El Khoury, J., Landreth, G. E., Brosseron, F., and Feinstein, D. L. (2015). Neuroinflammation in Alzheimer's disease. *Lancet Neurol.* 14, 388–405.
- Hickman, S., Izzy, S., Sen, P., Morsett, L., and El Khoury, J. (2018). Microglia in neurodegeneration. *Nat. Neurosci.* 21, 1359–1369.
- Hong, S., Beja-Glasser, V. F., Nfonoyim, B. M., Frouin, A., Li, S., Ramakrishnan, S., et al. (2016). Complement and microglia mediate early synapse loss in Alzheimer mouse models. *Science* 352, 712–716. doi: 10.1126/science.aad8373
- Jain, S., and Bader, G. D. (2010). An improved method for scoring protein-protein interactions using semantic similarity within the gene ontology. *BMC Bioinformatics* 11:562.
- Jansen, I. E., Savage, J. E., Watanabe, K., Bryois, J., Williams, D. M., and Steinberg, S. (2019). Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* 51, 404–413.
- Jevtic, S., Sengar, A. S., Salter, M. W., and McLaurin, J. (2017). The role of the immune system in Alzheimer disease: etiology and treatment. *Ageing Res. Rev.* 40, 84–94. doi: 10.1016/j.arr.2017.08.005
- Jiang, Q., Lee, C. Y., Mandrekar, S., Wilkinson, B., Cramer, P., and Zelcer, N. (2008). ApoE promotes the proteolytic degradation of Abeta. *Neuron* 58, 681–693. doi: 10.1016/j.neuron.2008.04.010
- Kang, J., Lemaire, H. G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K. H., et al. (1987). The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature* 325, 733–736. doi: 10.1038/325733a0
- Karch, C. M., and Goate, A. M. (2015). Alzheimer's disease risk genes and mechanisms of disease pathogenesis. *Biol. Psychiatry* 77, 43–51. doi: 10.1016/j.biopsych.2014.05.006
- Kellner, A., Matschke, J., Bernreuther, C., Moch, H., Ferrer, I., and Glatzel, M. (2009). Autoantibodies against beta-amyloid are common in Alzheimer's disease and help control plaque burden. *Ann. Neurol.* 65, 24–31. doi: 10.1002/ana.21475
- Lane, C. A., Hardy, J., and Schott, J. M. (2018). Alzheimer's disease. *Eur. J. Neurol.* 25, 59–70.
- Lee, S. Y., Song, M. Y., Kim, D., Park, C., Park, D. K., Kim, D. G., et al. (2019). A proteotranscriptomic-based computational drug-repositioning method for Alzheimer's disease. *Front. Pharmacol.* 10:1653.
- Mao, Y., Feng, Q., Zheng, P., Yang, L., Zhu, D., Chang, W., et al. (2018). Low tumor infiltrating mast cell density confers prognostic benefit and reflects immunoactivation in colorectal cancer. *Int. J. Cancer* 143, 2271–2280. doi: 10.1002/ijc.31613
- Monsonogo, A., Imitola, J., Petrovic, S., Zota, V., Nemirovsky, A., Baron, R., et al. (2006). Abeta-induced meningoencephalitis is IFN-gamma-dependent and is associated with T cell-dependent clearance of Abeta in a mouse model of Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.* 103, 5048–5053. doi: 10.1073/pnas.0506209103
- Ono, K., and Tsuji, M. (2019). Pharmacological potential of cilostazol for Alzheimer's disease. *Front. Pharmacol.* 10:559.
- Piguet, O., Double, K. L., Kril, J. J., Harasty, J., Macdonald, V., McRitchie, D. A., et al. (2009). White matter loss in healthy aging: a postmortem analysis. *Neurobiol. Aging* 30, 1288–1295. doi: 10.1016/j.neurobiolaging.2007.10.015
- Qin, Q., Teng, Z., Liu, C., Li, Q., Yin, Y., and Tang, Y. (2021). TREM2, microglia, and Alzheimer's disease. *Mech. Ageing Dev.* 195:111438.
- Ryman, D. C., Acosta-Baena, N., Aisen, P. S., Bird, T., Danek, A., and Fox, N. C. (2014). Symptom onset in autosomal dominant Alzheimer disease: a systematic review and meta-analysis. *Neurology* 83, 253–260.
- Sarlus, H., and Heneka, M. T. (2017). Microglia in Alzheimer's disease. *J. Clin. Invest.* 127, 3240–3249.
- Serrano-Pozo, A., Froesch, M. P., Masliah, E., and Hyman, B. T. (2011). Neuropathological alterations in Alzheimer disease. *Cold Spring Harb. Perspect. Med.* 1:a006189. doi: 10.1101/cshperspect.a006189
- Sevilla, J. L., Segura, V., Podhorski, A., Guruceaga, E., Mato, J. M., Martínez-Cruz, L. A., et al. (2005). Correlation between gene expression and GO semantic similarity. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 2, 330–338. doi: 10.1109/tcbb.2005.50
- Shi, J., Zhao, Y., Vonderfecht, T., Winey, M., and Klymkowsky, M. W. (2015). Centrin-2 (Cetn2) mediated regulation of FGF/FGFR gene expression in *Xenopus*. *Sci. Rep.* 5:10283.
- Shi, Q., Chowdhury, S., Ma, R., Le, K. X., Hong, S., Caldaroni, B. J., et al. (2017). Complement C3 deficiency protects against neurodegeneration in aged plaque-rich APP/PS1 mice. *Sci. Transl. Med.* 9:eaf6295. doi: 10.1126/scitranslmed.aaf6295
- Shi, Y., and Holtzman, D. M. (2018). Interplay between innate immunity and Alzheimer disease: APOE and TREM2 in the spotlight. *Nat. Rev. Immunol.* 18, 759–772. doi: 10.1038/s41577-018-0051-1
- Soleimani Zakeri, N. S., Pashazadeh, S., and MotieGhader, H. (2020). Gene biomarker discovery at different stages of Alzheimer using gene co-expression network approach. *Sci. Rep.* 10:12210.
- Späni, C., Suter, T., Derungs, R., Ferretti, M. T., Welt, T., Wirth, F., et al. (2015). Reduced  $\beta$ -amyloid pathology in an APP transgenic mouse model of Alzheimer's disease lacking functional B and T cells. *Acta Neuropathol. Commun.* 3:71.
- Storace, D., Cammarata, S., Borghi, R., Sanguineti, R., Giliberto, L., Piccini, A., et al. (2010). Elevation of  $\beta$ -amyloid 1-42 autoantibodies in the blood of amnesic patients with mild cognitive impairment. *Arch. Neurol.* 67, 867–872.
- Tedder, P. M., Bradford, J. R., Needham, C. J., McConkey, G. A., Bulpitt, A. J., and Westhead, D. R. (2010). Gene function prediction using semantic similarity clustering and enrichment analysis in the malaria parasite *Plasmodium falciparum*. *Bioinformatics* 26, 2431–2437. doi: 10.1093/bioinformatics/btq450
- Tsubata, T. (2019). CD72 is a negative regulator of B cell responses to nuclear lupus self-antigens and development of systemic lupus erythematosus. *Immune Netw.* 19:e1.



- Varn, F. S., Tafe, L. J., Amos, C. I., and Cheng, C. (2018). Computational immune profiling in lung adenocarcinoma reveals reproducible prognostic associations with implications for immunotherapy. *Oncoimmunology* 7:e1431084. doi: 10.1080/2162402x.2018.1431084
- Verghese, P. B., Castellano, J. M., and Holtzman, D. M. (2011). Apolipoprotein E in Alzheimer's disease and other neurological disorders. *Lancet Neurol.* 10, 241–252.
- Wang, J. Z., Du, Z., Payattakool, R., Yu, P. S., and Chen, C. F. (2007). A new method to measure the semantic similarity of GO terms. *Bioinformatics* 23, 1274–1281. doi: 10.1093/bioinformatics/btm087
- Webers, A., Heneka, M. T., and Gleeson, P. A. (2020). The role of innate immune responses and neuroinflammation in amyloid accumulation and progression of Alzheimer's disease. *Immunol. Cell Biol.* 98, 28–41. doi: 10.1111/imcb.12301
- Xu, M., Zhang, D.-F., Luo, R., Wu, Y., Zhou, H., Kong, L.-L., et al. (2018). A systematic integrated analysis of brain expression profiles reveals YAP1 and other prioritized hub genes as important upstream regulators in Alzheimer's disease. *Alzheimer's Dement.* 14, 215–229. doi: 10.1016/j.jalz.2017.08.012
- Xue, H., Zhang, J., Guo, X., Wang, J., Li, J., Gao, X., et al. (2016). CREBRF is a potent tumor suppressor of glioblastoma by blocking hypoxia-induced autophagy via the CREB3/ATG5 pathway. *Int. J. Oncol.* 49, 519–528. doi: 10.3892/ijo.2016.3576
- Yeh, F. L., Hansen, D. V., and Sheng, M. (2017). TREM2, microglia, and neurodegenerative diseases. *Trends Mol. Med.* 23, 512–533. doi: 10.1016/j.molmed.2017.03.008
- Ying, G., Frederick, J. M., and Baehr, W. (2019). Deletion of both centrin 2 (CETN2) and CETN3 destabilizes the distal connecting cilium of mouse photoreceptors. *J. Biol. Chem.* 294, 3957–3973. doi: 10.1074/jbc.ra118.006371
- Yu, G., Li, F., Qin, Y., Bo, X., Wu, Y., and Wang, S. (2010). GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* 26, 976–978. doi: 10.1093/bioinformatics/btq064
- Zenaro, E., Pietronigro, E., Della Bianca, V., Piacentino, G., Marongiu, L., and Budui, S. (2015). Neutrophils promote Alzheimer's disease-like pathology and cognitive decline via LFA-1 integrin. *Nat. Med.* 21, 880–886. doi: 10.1038/nm.3913

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

Copyright © 2021 Liu, Li and Pan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.