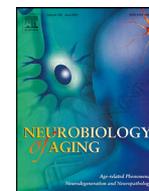




Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging.org

Genetic report

Development and validation of a 13-gene signature associated with immune function for the detection of Alzheimer's disease

Min Zhu^{a,b}, Tingting Hou^{a,b,*}, Longfei Jia^c, Qihua Tan^d, Chengxuan Qiu^{a,e,**}, Yifeng Du^{a,b,***}, Alzheimer's Disease Neuroimaging Initiative[#]^a Department of Neurology, Shandong Provincial Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China^b Department of Neurology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China^c Innovation Center for Neurological Disorders and Department of Neurology, Xuanwu Hospital, Capital Medical University, National Clinical Research Center for Geriatric Diseases, Beijing, China^d Department of Public Health, Epidemiology and Biostatistics, University of Southern Denmark, Odense, Denmark^e Department of Neurobiology, Care Sciences and Society, Aging Research Center, Karolinska Institutet and Stockholm University, Stockholm, Sweden

ARTICLE INFO

Article history:

Received 10 August 2022

Revised 5 December 2022

Accepted 28 December 2022

Available online 31 December 2022

Keywords:

Alzheimer's disease

Detection model

Immune function

Apolipoprotein E

ABSTRACT

Current knowledge of Alzheimer's disease (AD) etiology and effective therapy remains limited. Thus, the identification of biomarkers is crucial to improve the detection and treatment of patients with AD. Using robust rank aggregation method to analyze the microarray data from Gene Expression Omnibus database, we identified 1138 differentially expressed genes in AD. We then explored 13 hub genes by weighted gene co-expression network analysis, least absolute shrinkage, and selection operator, and logistic regression in the training dataset. The detection model, which composed of *CD163*, *CDC42SE1*, *CECR6*, *CSF1R*, *CYP27A1*, *EIF4E3*, *H2AFJ*, *IFIT2*, *IL10RA*, *KIAA1324*, *PSTPIP1*, *SLA*, and *TBC1D2* genes, along with *APOE* gene, showed that the area under the curve for detecting AD was 0.821 (95% confidence interval [CI] = 0.782–0.861) and the model was validated in ADNI dataset (area under the curve = 0.776; 95%CI = 0.686–0.865). Notably, the 13 genes in the model were highly enriched in immune function. These findings have implications for the detection and therapeutic target of AD.

© 2022 Published by Elsevier Inc.

Abbreviations: A β , β -amyloid; AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; APOE, Apolipoprotein E; AUC, Area under the curve; CN, Cognitively normal; GEO, Gene Expression Omnibus; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; LASSO, Least absolute shrinkage and selection operator; ROC, Receiver operating characteristic; RRA, Robust rank aggregation; WGCNA, Weighted gene co-expression network analysis.

* Corresponding author at: Tingting Hou, Department of Neurology, Shandong Provincial Hospital, Cheeloo College of Medicine, Shandong University, No. 324 Jingwuwei Road, Jinan, Shandong 250021, China. Tel.: +86-0531-68773775.

** Corresponding author at: Chengxuan Qiu, Department of Neurobiology, Care Sciences and Society, Aging Research Center, Karolinska Institutet, Tomtebodavägen 18A, SE-171 65 Solna, Sweden. Tel.: +46-852485821.

*** Corresponding author at: Yifeng Du, Department of Neurology, Shandong Provincial Hospital, Cheeloo College of Medicine, Shandong University, No. 324 Jingwuwei Road, Jinan, Shandong 250021, China. Tel.: +86-0531-68776354.

E-mail addresses: houtingting@sdfmu.edu.cn (T. Hou), chengxuan.qiu@ki.se (C. Qiu), duyifeng2013@163.com (Y. Du).

[#] Part of the data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implemen-

1. Introduction

Alzheimer's disease (AD) is the most common type of dementia, manifested by gradually decline in memory and other cognitive domains. The global number of people with AD was projected to reach over 140 million by 2050 (Prince et al., 2015), which would pose a heavy burden on the individuals and society. However, since 1907 when the first patient with AD was described, relatively limited progress has been made in revealing its etiology or in exploring effective therapy to halt its progression (Castellani and Perry, 2012; Hardy, 2006). This is crucial to improve the understanding of molecular mechanisms and to explore biomarkers for the detection, diagnosis, and treatment of AD.

tation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Amyloid plaques and neurofibrillary tangles are neuropathological hallmarks of AD, thus, β amyloid ($A\beta$) and phosphorylated tau (p-tau) in central nervous system become unique biomarkers for AD (Stevenson-Hoare et al., 2022; Zetterberg & Bendlin, 2021). Indeed, measuring the levels of $A\beta$ and p-tau in the cerebrospinal fluid, as well as amyloid or tau protein in the brain by PET scans, would help diagnose AD. However, the invasive nature of lumbar puncture and the high cost of PET scans limit the application of these methods. Therefore, great efforts have been made to identify blood biomarkers for AD diagnosis.

Recently, several blood biomarkers other than $A\beta$ and p-tau have been reported for AD, which provides potential options for early detection and targeted therapies. For example, plasma soluble CD22 (sCD22), a possible biomarker of inflammation and microglial dysfunction, was newly found to be associated with cognitive decline in AD (Bu et al., 2022). Furthermore, plasma concentrations of neurofilament light and glial fibrillary acidic protein (GFAP) were closely related to AD. In addition, a model that includes EEF2 and RPL7 has been reported for predicting the risk of AD, providing possible therapeutic targets (Shigemizu et al., 2020). Hence, it is urgent to discover molecular markers highly associated with AD that could contribute to improving the effect of targeted therapeutic approaches. A recent study tried to develop detection models using blood gene expression data and machine learning methods, which showed different external-validation performances, with the area under the curve (AUC) ranging from 0.619 to 0.859 (Lee & Lee, 2020). However, it is unclear whether the robust differentially expressed genes in AD brain tissues could be sturdy predictors in blood samples for the differentiation of AD patients from cognitively normal (CN) persons.

In the present study, using data from the Gene Expression Omnibus (GEO) dataset, we identified 13 hub genes that were potentially associated with AD risk. Prognostic risk models were built on hub genes levels along with apolipoprotein E (*APOE*) genotypes. Then, the risk models were validated using bootstrap method for internal validation and by Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset for external validation. In addition to *APOE* genotype, the 13 hub genes included in the risk detection model for AD were Cluster of Differentiation 163 (*CD163*), CDC42 small effector 1 (*CDC42SE1*), Cat Eye Syndrome Critical Region Protein 6 (*CECR6*), Colony stimulating factor-1 receptor (*CSF1R*), Cytochrome P450 Family 27 Subfamily A Member 1 (*CYP27A1*), Eukaryotic Translation Initiation Factor 4E Family Member 3 (*EIF4E3*), H2A Histone Family Member J (*H2AFJ*), Interferon Induced Protein With Tetratricopeptide Repeats 2 (*IFIT2*), Interleukin 10 Receptor Subunit Alpha (*IL10RA*), KIAA1324, Proline-Serine-Threonine Phosphatase Interacting Protein 1 (*PSTPIP1*), Src Like Adaptor (*SLA*), and TBC1 Domain Family Member 2 (*TBC1D2*). We found that all the 13 hub genes in the risk detection model were correlated with immune infiltration. Our study presents a new gene-based risk detection model for AD and highlights the role of immune mechanisms in AD, which has potential implications for AD detection and therapeutic targets of the disease.

2. Methods

2.1. Data sources

The overall workflow of the present study is shown in Fig. A.1 (Zhong et al., 2021). We searched the database for the microarray datasets using the keyword "Alzheimer" on GEO dataset (<https://www.ncbi.nlm.nih.gov/geo/>). Datasets were included if they met the following criteria: (1) were from humans; (2) included expression data from cerebral frontal, temporal cortex or the hippocampus of both AD and CN samples; (3) the number of rows in

each platform was $>30,000$; (4) the sample size was ≥ 10 ; and (5) there were no repeated samples among datasets. Finally, 5 datasets from the frontal cortex of AD and CN samples, 5 datasets from temporal cortex, and 3 datasets from hippocampus were selected (Berchtold et al., 2013; Hokama et al., 2014; Liang et al., 2007; McKay et al., 2019; Patel et al., 2019; Piras et al., 2019b). Blood expression data of GSE140829 and ADNI dataset were obtained from GEO database and ADNI database (<https://adni.loni.usc.edu/>). Detailed information for these datasets is shown in Table 1 and Table A.1.

2.2. Identification of differentially expressed genes and robust rank aggregation analysis

GEO series matrix files of the datasets and their corresponding platform files were downloaded for current analysis. Differentially expressed genes (DEGs) were screened using these datasets by the R package "limma" (Ritchie et al., 2015). Those with \log_2 Fold Change (FC) > 0.5 and adjusted p value < 0.05 are defined as DEGs. The genes were ranked according to the adjusted P values. We then used Robust Rank Aggregation (RRA) analysis to integrate the results of those 13 datasets to find the most significant DEGs by the R package "RobustRankAggreg" (Kolde et al., 2012). The RRA method can detect genes that are statistically relevant and rank them after filtering outliers, noise, and errors (Kolde et al., 2012). Genes with adjusted $p < 0.05$ were considered as significant DEGs in the RRA analysis. The R package "OmicCircos" (Hu et al., 2014) was used to show the expression patterns and chromosomal location of the top 200 DEGs (top 100 up-regulated genes and top 100 down-regulated genes according to adjusted p) from the RRA analysis.

2.3. Weighted gene co-expression network analysis

Top 2000 DEGs (top 1000 up-regulated and top 1000 down-regulated) from RRA analysis were used to conduct Weighted Gene Co-expression Network Analysis (WGCNA). The expression data of top 2000 DEGs were extracted from GSE140829 including blood samples of AD and CN groups. The R package "WGCNA" (Langfelder & Horvath, 2008, 2012) was utilized to find clinical trait-related modules and hub genes among the DEGs. To transform the adjacency matrix to a topological overlap matrix, a soft-threshold power was selected using the code "sft\$powerEstimate." The network and modules were constructed with a cut height of 0.25 and a minimal module size of 20. Modules with significant correlation with AD ($p < 0.05$) were considered as key modules.

2.4. Functional enrichment analyses

Enrichment analyses were performed to further investigate the biological functions of genes. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were conducted using the R package "clusterProfiler" (Yu et al., 2012) among DEGs from RRA analysis and genes in key modules of WGCNA, respectively. GO terms and KEGG pathways with $p < 0.05$ were considered as significant and were exhibited either as chord diagrams using the R package "GOplot" (Walter et al., 2015) or as barplots and dotplots.

2.5. Construction of risk detection models

Genes from key models of WGCNA were screened using Least absolute shrinkage and selection operator (LASSO) regression by the R package "glmnet" (Friedman et al., 2010). After cross-validation, lambda.1se was chosen for gene filtration. Then, genes

Table 1
Characteristics of the included datasets.

Locations	Dataset	Country	Sample size	GPL ID	Number of rows per platform	Usage here
Frontal cortex	GSE118553 (Patel et al., 2019)	UK	23CN; 40AD	GPL10558	48107	RRA
Frontal cortex	GSE122063 (McKay et al., 2019)	USA	22CN; 28AD	GPL16699	62976	
Frontal cortex	GSE36980 (Hokama et al., 2014)	Japan	18CN; 15AD	GPL6244	33297	RRA
Frontal cortex	GSE48350 (Berchtold et al., 2013)	USA	48CN; 21AD	GPL570	54675	
Frontal cortex	GSE5281 (Liang et al., 2007)	USA	11CN; 23AD	GPL570	54675	RRA
Hippocampus	GSE36980 (Hokama et al., 2014)	Japan	10CN; 7AD	GPL6244	33297	
Hippocampus	GSE48350 (Berchtold et al., 2013)	USA	19CN; 43AD	GPL570	54675	RRA
Hippocampus	GSE5281 (Liang et al., 2007)	USA	13CN; 10AD	GPL570	54675	
Temporal cortex	GSE118553 (Patel et al., 2019)	UK	31CN; 52AD	GPL10558	48107	RRA
Temporal cortex	GSE122063 (McKay et al., 2019)	USA	22CN; 28AD	GPL16699	62976	
Temporal cortex	GSE36980 (Hokama et al., 2014)	Japan	19CN; 10AD	GPL6244	33297	RRA
Temporal cortex	GSE132903 (Piras et al., 2019b)	USA	97CN; 98AD	GPL10558	48107	
Temporal cortex	GSE5281 (Liang et al., 2007)	USA	12CN; 16AD	GPL570	54675	RRA
Blood	GSE140829	USA	249CN; 204AD	GPL15988	47322	
Blood	ADNI	USA	138CN; 40AD	GPL13667	49386	WGCNA, model construction model validation

Abbreviations: Detailed information of GSE140829 and ADNI dataset could be referred to the websites at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi> and adni.loni.usc.edu, respectively. AD, Alzheimer disease; CN, cognitively normal; GSE, Gene Expression Omnibus Series; GPL, Gene Expression Omnibus Platform; RRA, robust rank aggregation; WGCNA, weighted correlation network analysis.

filtrated by LASSO, together with age, sex, and *APOE* status, were used to construct the risk detection model by logistic regression. “StepAIC” method with a direction of “both” was utilized to determine the robust variables for the final detection model. A receiver operating characteristic (ROC) analysis was used to estimate the model AUC. The ROC curves were drawn by the R package “pROC” (Robin et al., 2011).

2.6. Internal and external validation of the risk detection models

Internal validation was conducted using the enhanced bootstrap method with $B = 1000$. The adjusted C statistics and brier score were then calculated after bootstrap. Data for external validation were obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessments can be combined to measure the progression of mild cognitive impairment (MCI) and early AD. For up-to-date information on ADNI, see www.adni-info.org. AUC was then used to indicate the model accuracy and bootstrap test was performed for the comparison between 2 ROC curves. Seed was set as 123 and boot.n was 2000. $p < 0.05$ was considered as significant difference in performance.

2.7. Validation of expressions of model genes in datasets

Expression data of model genes extracted from the aforementioned 13 datasets were used to validate the differential expression of these genes in the frontal cortex, temporal cortex, and hippocampus. Then, the GO and KEGG analysis were performed to reveal the possible functions of the model genes.

2.8. Immune infiltration analysis

The Cell-type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT) method was applied to investigating the correlation between these model genes and 22 immune cells. $p < 0.05$ and correlation coefficient > 0.30 were considered a significant correlation between genes and immune cells. Then, the correlations between the expression of the model genes and the immune checkpoints (Table A.2) were examined in AD and CN groups, respectively. $p < 0.05$ was considered a significant correlation between genes and immune checkpoints. The GO and KEGG

analysis were conducted to find the potential functions of immune genes in different groups.

2.9. Statistical analysis

We used LASSO regression for gene infiltration, logistic regression for building the detection models, and enhanced bootstrap method for internal validation. These regression analyses were conducted using R software (version 4.0.2, R Foundation for Statistical Computing, Vienna, Austria). The t-test was applied to assessing the differences between the 2 groups. All the values were presented as means \pm standard deviation (SD). A 2-tailed $p < 0.05$ was considered statistically significant. T-tests were performed using GraphPad Prism software (version 9.1.1, San Diego, CA, www.graphpad.com).

3. Results

3.1. Identification and chromosome locations of DEGs

DEGs were found in 5 datasets from the frontal cortex of AD and CN samples, 5 datasets from temporal cortex, and 3 datasets from hippocampus, respectively (CN: total sample, $N = 345$; AD: total sample, $N = 391$). The volcano plots are shown in Fig. A.2. Using the RRA analysis, 1138 robust DEGs (539 up-regulated and 599 down-regulated) were identified (Table A.3). We described the expression of genes described as “up-regulated” or “downregulated” with respect to AD versus CN groups. The top 200 DEGs that included top 100 up-regulated genes and top 100 down-regulated ones (according to adjusted p values) were chosen to visualize their chromosomal location and expression patterns across these 13 datasets, and to show the logarithm of their adjusted p values in the inner layer (Fig. 1). The top 5 up-regulated genes were *AEBP1*, *GFAP*, *ITPKB*, *PALLD*, and *TOB1*, whereas the top 5 down-regulated genes were *NRN1*, *CACNG3*, *ADCYAP1*, *NRXN3*, and *MAPK9* (Fig. 1).

3.2. Enrichment analysis of DEGs

The chord diagrams show top 5 terms of enrichment analysis of DEGs based on their adjusted p values (Fig. 2). Neurotransmitter transport, vesicle-mediated transport in synapse, synaptic vesicle cycle, regulation of neurotransmitter levels, and protein localization to cell periphery were for the biological process terms for GO analysis (Fig. 2a). Presynapse, exocytic vesicle, synaptic vesicle,

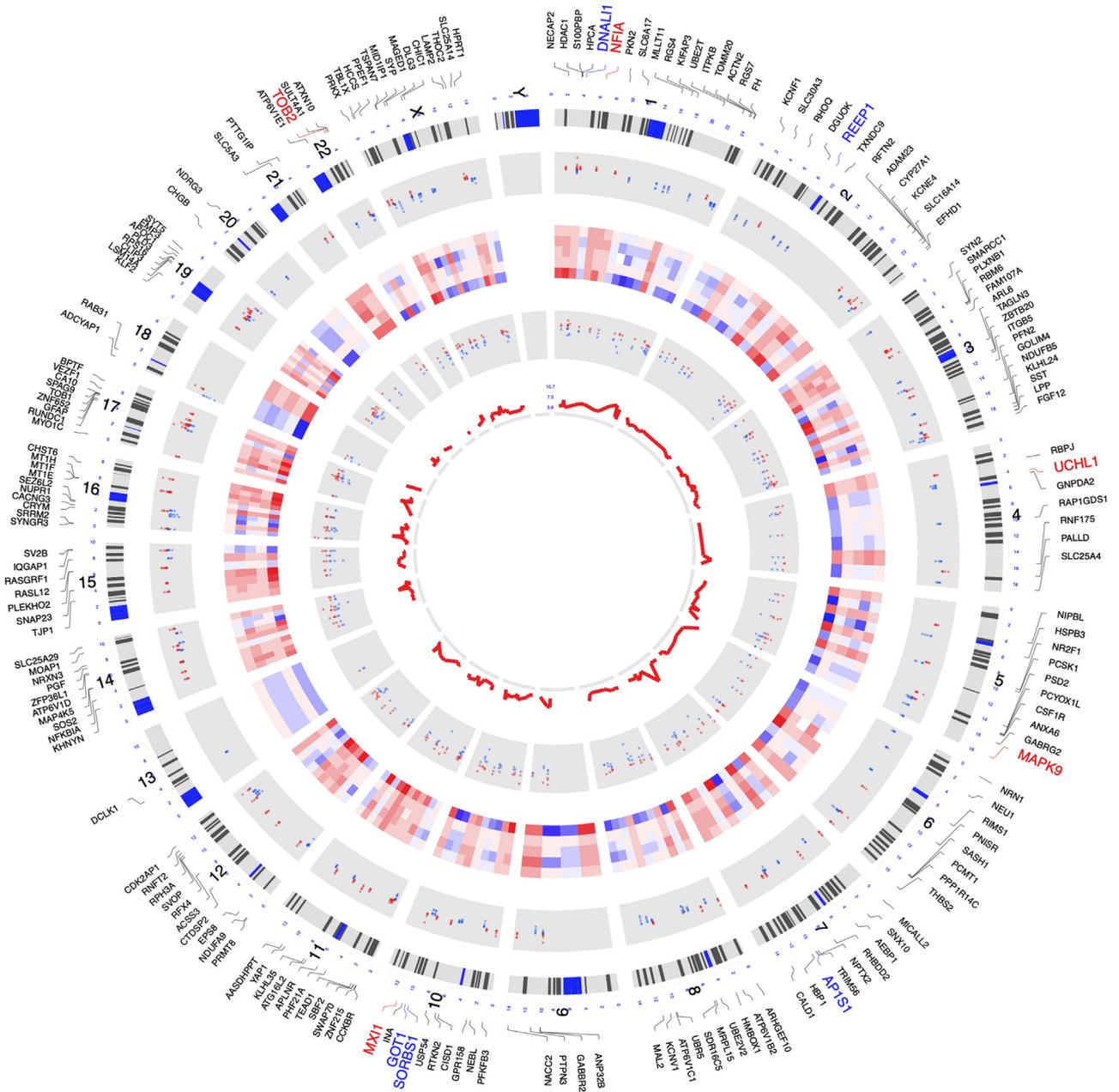


Fig. 1. Circos plot of expression patterns and chromosomal positions of top 200 differentially expressed genes (DEGs). The outer circle represents chromosomes, and lines coming from each gene point to their specific chromosomal locations. The Alzheimer disease (AD) microarray datasets from Gene Expression Omnibus (GEO) used for robust rank aggregation (RRA) analysis are represented in the inner circular multiple points and heatmaps. Datasets from frontal cortex, temporal cortex, and hippocampus are presented from the outside to the inside. According to the value of \log_2 fold change, red indicates up-regulation in AD samples while blue indicates down-regulation. The red lines in the inner layer indicate $-\log_{10}$ (adjusted p -value for RRA) of each gene. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

transport vesicle, and synaptic vesicle membrane were for the cellular component terms for GO analysis (Fig. 2b). The top 5 molecular function terms for GO analysis were ATPase activity, coupled to transmembrane movement of ions, rotational mechanism; proton-transporting ATPase activity, rotational mechanism; ATPase-coupled cation transmembrane transporter activity; pyrophosphate hydrolysis-driven proton transmembrane transporter activity; and ATPase-coupled ion transmembrane transporter activity (Fig. 2c). For KEGG analysis, DEGs were mostly related to Epithelial cell signaling in Helicobacter pylori infection, Synaptic vesicle cycle, Citrate cycle (TCA cycle), Vibrio cholerae infection, and Collecting duct acid secretion (Fig. 2d).

3.3. WGCNA and identification of key modules

Expression data of top 2000 DEGs from RRA analysis (top 1000 up-regulated and top 1000 down-regulated) were extracted from GSE140829 (36 genes not available) and used to conduct WGCNA (Fig. 3, Fig. A.3). By setting the soft-threshold power as 3 (scale-free $R^2 = 0.85$, slope = -1.36 ; Fig. A.3) and cut height as 0.25, we acquired 14 modules, including the grey module that contained non-clustering DEGs. Genes in each module are shown in Table A.3. From the heatmap of module-trait correlations, we found that the blue, purple, red, salmon, turquoise, and yellow modules were significantly correlated with AD ($p < 0.05$ or < 0.001 ; Fig. 3c). The

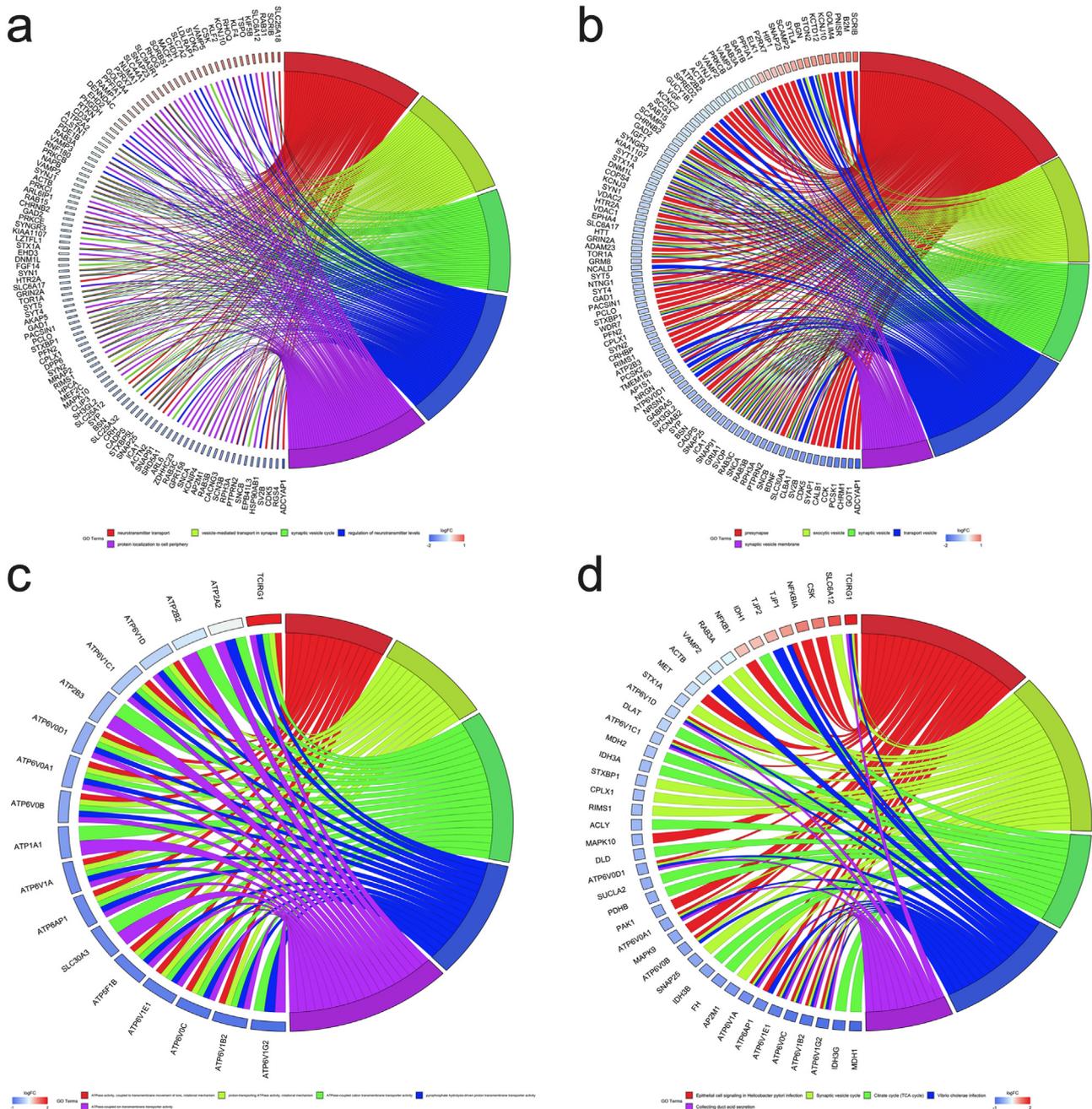


Fig. 2. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses of all DEGs from RRA analysis. Chord plots indicate enrichment analysis of genes. (a) Biological process of GO analysis. (b) Cellular component of GO analysis. (c) Molecular function of GO analysis. (d) KEGG pathways. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

relationships between gene significance and module membership of the 6 key modules are also shown in Fig. A.4 (correlation coefficient = 0.48 ($p < 0.0001$), 0.41 ($p = 0.027$), 0.41 ($p < 0.0001$), 0.34 ($p < 0.0001$), 0.57 ($p < 0.001$), and 0.24 ($p < 0.0001$), respectively).

3.4. Functional enrichment analysis of the key modules

All the 6 key modules contained 941 genes (Table A.3). The enrichment analysis showed that the blue module was mostly enriched in protein localization to cell periphery and membrane depolarization; the purple module enriched in myoblast differentiation and RNA splicing; the red module in regulation of protein de-

phosphorylation and regulation of cell cycle G2/M phase transition; the salmon module in positive regulation of neuron death and positive regulation of neuron apoptotic process; the turquoise module in regulation of ubiquitin-dependent protein catabolic process and regulation of protein modification by small protein conjugation or removal; and the yellow module enriched in neutrophil degranulation and transferrin transport (Fig. A.5).

3.5. Establishment of the Risk Detection Models

LASSO regression was performed for the screening of all the 941 genes from six key modules. GSE140829 was used as the training

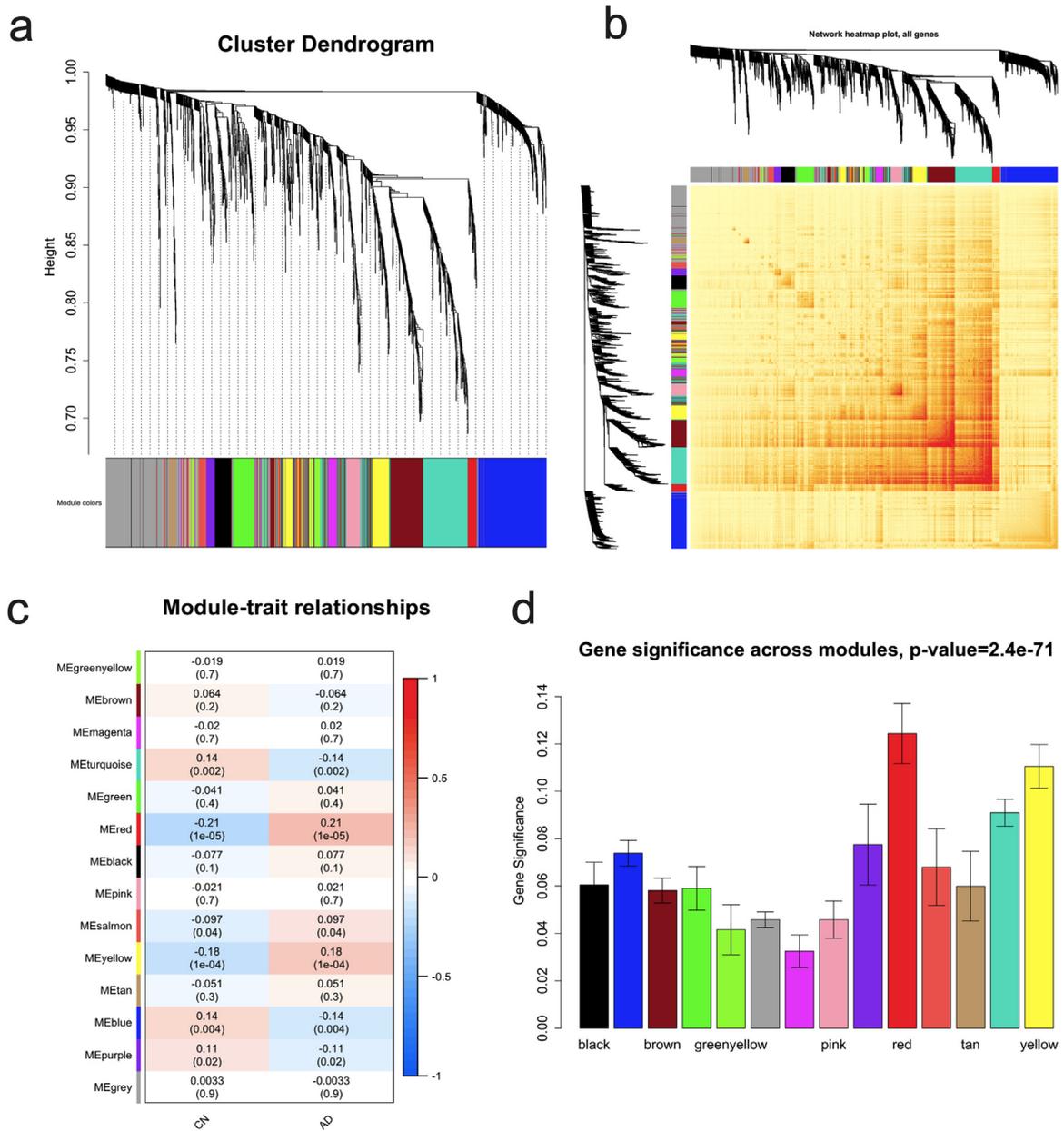


Fig. 3. Key modules correlated with Alzheimer disease identified by weighted gene co-expression network analysis (WGCNA). (a) Cluster dendrogram of genes. (b) Network heatmap plot of all genes. (c) Heatmap shows the relationships between different modules and clinical traits. (d) Gene significance in different modules associated with AD. AD, Alzheimer's disease; CN, cognitively normal. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

dataset (CN: N = 249; AD: N = 204). Using $\lambda = 0.01076384$, 27 genes were remained in the cross-validation of the LASSO regression (Table A.4). Together with age, sex, and *APOE* genotype, all 30 variables were put into the logistic regression model to create the detection model. After stepwise regression, 14 gene variables, including *APOE* genotype, *CD163*, *CDC42SE1*, *CECR6*, *CSF1R*, *CYP27A1*, *EIF4E3*, *H2AFJ*, *IFIT2*, *IL10RA*, *KIAA1324*, *PSTPIP1*, *SLA*, and *TBC1D2*, were included in the final risk detection model (Table 2; Table A.5). The AUC of the training model with *APOE* genotype was 0.821 [95% confidence interval (CI): 0.782–0.861]. To investigate the influence of the *APOE* genotype, we also built model only with *APOE* genotype using the training dataset (Table A.5), and the AUC was 0.672 [95%CI: 0.628–0.716]. Then, we compared the ROC curves of the 2 models and found significant differences between the detection

model including 13 genes and *APOE* genotype, and the model with only *APOE* genotype ($p < 0.001$; Fig. 4a).

3.6. Validation of the risk detection models

The enhanced bootstrap method was used for internal validation of the detection models. In the enhanced bootstrap analysis, the AUC of the detection model including 13 genes and *APOE* genotype changed from 0.821 to 0.771, and the brier score changed from 0.170 to 0.197 (Fig. A.6a and b).

The ADNI dataset was used for the external validation (CN: N = 138; AD: N = 40). The *APOE* genotype and expression of 13 genes were extracted from the baseline data of ADNI. The AUC of the detection model with *APOE* genotype in the external valida-

Table 2Associations of 13 hub genes and *APOE* in the final detection model with Alzheimer's disease: Logistic regression analysis.

Variables	β coefficient	SE	Exp(β)	95% CI	<i>p</i> value
<i>CD163</i>	0.329	0.171	1.389	0.996-1.949	0.054
<i>CDC42SE1</i>	-0.410	0.131	0.664	0.511-0.854	0.002
<i>CECR6</i>	0.243	0.150	1.275	0.953-1.715	0.104
<i>CSF1R</i>	-0.683	0.175	0.505	0.356-0.708	<0.001
<i>CYP27A1</i>	-0.243	0.132	0.784	0.604-1.014	0.066
<i>EIF4E3</i>	0.338	0.145	1.402	1.060-1.872	0.020
<i>H2AFJ</i>	0.239	0.128	1.270	0.991-1.641	0.062
<i>IFIT2</i>	-0.336	0.138	0.715	0.543-0.934	0.015
<i>IL10RA</i>	0.398	0.158	1.489	1.104-2.042	0.012
<i>KIAA1324</i>	-0.295	0.135	0.745	0.570-0.968	0.029
<i>PSTPIP1</i>	0.354	0.156	1.425	1.052-1.941	0.023
<i>SLA</i>	0.254	0.144	1.289	0.974-1.713	0.077
<i>TBC1D2</i>	0.483	0.186	1.620	1.120-2.348	0.010
With one <i>APOE</i> $\epsilon 4$ allele	1.230	0.250	3.422	2.107-5.627	<0.001
With 2 <i>APOE</i> $\epsilon 4$ alleles	2.388	0.589	10.891	3.785-40.071	<0.001

Abbreviations: CI, confidence interval; SE, standard error.

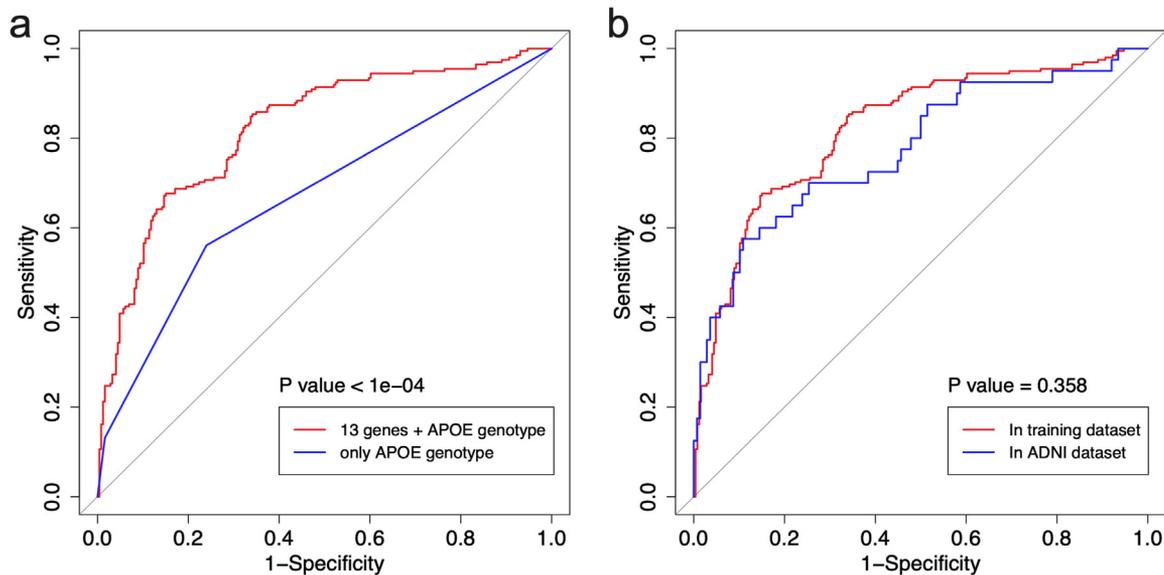


Fig. 4. Receiver operating characteristic (ROC) curve of the immune-related variables for differentiating Alzheimer's disease and cognitively normal. (a) Comparison between the model with *APOE* genotype and the model only including *APOE* genotype in the training dataset. (b) Comparison of models with *APOE* genotype in the training dataset and the ADNI dataset. *p* value was for the tests of comparison of ROC curves using bootstrap method. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

tion was 0.776 [95%CI: 0.686–0.865]. We compared the ROC curves of the training dataset and ADNI dataset, and found no significant difference in the detection models with *APOE* genotype ($p > 0.05$; Fig. 4b). In addition, the *APOE*-genotype hierarchical analysis showed no significant difference of ROC curves between the training dataset and ADNI dataset either in *APOE* $\epsilon 4$ allele carriers or in non-carriers of the *APOE* $\epsilon 4$ allele ($p > 0.05$; Fig. A.7).

3.7. Differential expression of the model genes in datasets

The expression data of the 13 genes (*CD163*, *CDC42SE1*, *CECR6*, *CSF1R*, *CYP27A1*, *EIF4E3*, *H2AFJ*, *IFIT2*, *IL10RA*, *KIAA1324*, *PSTPIP1*, *SLA*, and *TBC1D2*) were extracted from the 13 datasets used for RRA analysis. Except for *H2AFJ* in frontal cortex and *PSTPIP1*, *SLA*, and *TBC1D2* in hippocampus, all 13 genes were differentially expressed in the brain tissues (Fig. A.8). In addition, GO and KEGG enrichment analyses were conducted to explore the potential functions of these genes. The results suggested that these genes were most

associated with positive regulation of receptor signaling pathway via JAK-STAT, which is critical for immune cell development and inflammatory responses (Fig. A.9).

3.8. Results of immune infiltration analysis

To further validate the potential immune-regulation functions of the 13 genes, immune infiltration analysis was performed. The infiltration levels of activated Dendritic cells, M0 macrophages, M1 macrophages, M2 macrophages, and follicular helper T cells were found to be correlated with model genes in AD (Fig. A.10). The model genes were correlated with different immune checkpoints between AD and CN groups (Fig. 5). The distinct immune checkpoints in AD group and CN group were significantly associated in different biological functions (Table A.6). The immune functions of T cell costimulation, lymphocyte costimulation, and T cell proliferation were enriched in AD group, while the functions of T cell

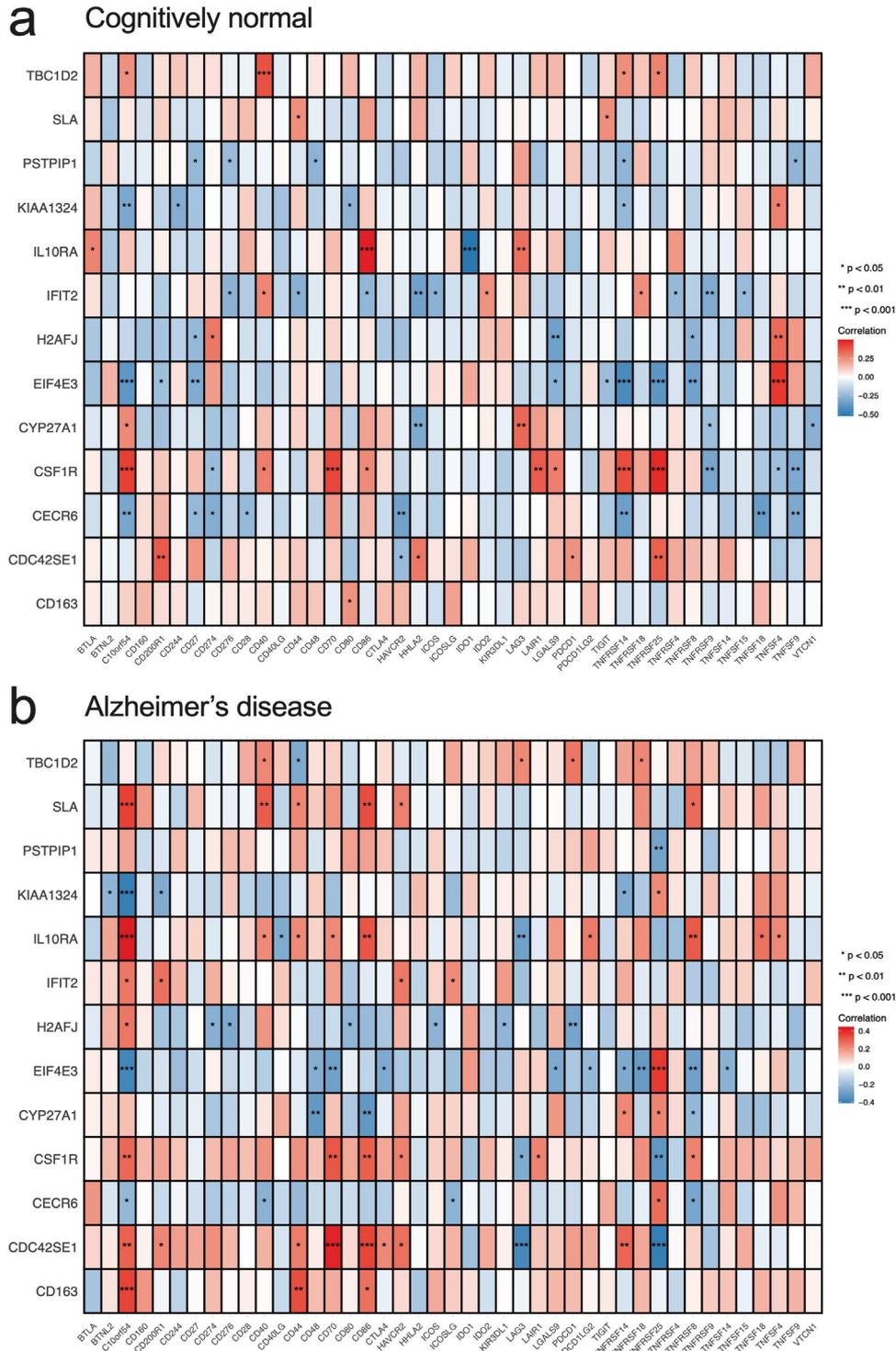


Fig. 5. The correlation between immune-related variables and immune checkpoints. (a) Correlation in samples of cognitively normal. (b) Correlation in samples of Alzheimer's disease. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

activation, regulation of lymphocyte activation, and regulation of leukocyte cell-cell adhesion were enriched in CN group.

4. Discussion

In the present study, we first identified 1138 robust DEGs between AD and CN samples. Then, a WGCNA was built, and 13

hub genes were explored by LASSO and multivariable regression analysis. Moreover, the risk detection model, composed of *CD163*, *CDC42SE1*, *CECR6*, *CSF1R*, *CYP27A1*, *EIF4E3*, *H2AFJ*, *IFIT2*, *IL10RA*, *KIAA1324*, *PSTPIP1*, *SLA*, and *TBC1D2*, along with *APOE* genotype, served as a novel tool that was validated to be valuable for the detection of AD risk. Notably, the model genes were highly correlated with immune cells as well as immune checkpoints, indicating

a potential role of these genes in AD via immune-related biological pathways.

We first analyzed the expression profiles from 13 GEO datasets to explore key genes related to AD. The most up-regulated gene *AEBP1* has been associated with $A\beta$ plaques and neurofibrillary tangle density (Piras et al., 2019a; Piras et al., 2019b; Shijo et al., 2018). Meanwhile, the most down-regulated gene *NRN1* was also found to be highly correlated with Braak stages and may represent as a novel target for the maintenance of cognition in old age (Piras et al., 2019a; Yu et al., 2020). Using the GO and KEGG enrichment analysis, we found that DEGs might correlate with neurotransmitter transport. The WGCNA helped identify the co-expression modules associated with the disease. In the key modules, genes were enriched in protein localization, RNA splicing, protein dephosphorylation, positive regulation of neuron death, ubiquitin-dependent protein catabolic process, and neutrophil degranulation, respectively. These findings revealed the potential roles of DEGs in neuron development and degradation.

Using LASSO and logistic regression analyses, 13 hub genes (*CD163*, *CDC42SE1*, *CECR6*, *CSF1R*, *CYP27A1*, *EIF4E3*, *H2AFJ*, *IFIT2*, *IL10RA*, *KIAA1324*, *PSTPIP1*, *SLA*, and *TBC1D2*) were identified to be valuable for the detection of AD. *CD163* has been implicated in macrophage activation and could increase $A\beta$ phagocytosis in AD (Koronyo-Hamaoui et al., 2020). Consistent with our findings, the level of *CD163* was increased in the frontal cortices of AD patients. Moreover, *CD163*-positive microglia were also enhanced and most of them were associated with $A\beta$ plaques (Pey et al., 2014). The role of *CDC42SE1* encoding CDC42 small effector 1 in AD has been poorly understood. The meta-analysis of genome-wide association studies showed that *CDC42SE1* was significantly associated with sodium excretion (Kho et al., 2020). *CSF1R* was considered as a potential therapeutic target for neurodegenerative diseases, as its inhibitors had beneficial effects in preclinical models (Han et al., 2022). *CSF1R* was also decreased in schizophrenia (Snijders et al., 2021). In patients with AD, *CYP27A1* was decreased in neurons but increased in oligodendrocytes (Brown et al., 2004), which might partly explain inconsistent differences in brain tissues between AD and CN groups in the present study. In addition, increased expression of *CYP27A1* in the brain could contribute to learning and memory impairment in rats (Zhang et al., 2018). *TBC1D2* was reported to be associated with ventricular enlargement in non-demented elders and might be involved in Facial Neuralgia (Li et al., 2019), indicating its possible roles in neurological diseases. Taken together, these studies suggest that the model genes are promising biomarkers in the development and progression of AD.

However, little is known about *CECR6*, *EIF4E3*, *H2AFJ*, *IFIT2*, *IL10RA*, *KIAA1324*, *PSTPIP1*, and *SLA* in neurological diseases or aging-related disorders. As indicated in GeneCards (www.genecards.org), *CECR6*, known as Transmembrane Protein 121B (TMEM121B), is associated with Cat Eye Syndrome and Postmenopausal Atrophic Vaginitis. *EIF4E3* is related to Eyelid Disease. *H2AFJ* encodes a replication-independent histone that is a variant H2A histone. *IFIT2* has been associated with Japanese Encephalitis and Hermansky-Pudlak Syndrome 1, while *IL10RA* is related to Inflammatory Bowel Disease 28 and Autosomal Recessive and Inflammatory Bowel Disease 28. *KIAA1324*, which is also called Endosome-Lysosome Associated Apoptosis And Autophagy Regulator 1 (*ELAPOR1*), may protect cells from cell death by upregulating the autophagy pathway (Deng et al., 2010). *PSTPIP1* encoded proteins were involved in cytoskeletal organization and inflammatory processes. *SLA* may be relevant to cell differentiation and innate immune response. Thus, these seldom-studied genes were highly correlated with immune functions and future research is required

both *in vitro* and *in vivo* to further understand their roles in the neurological diseases.

To further explore utility of the identified genes, the risk detection models for AD were built using the expression of key genes, with or without *APOE* genotype. The internal and external validation analyses showed that the model that included 13 genes and *APOE* genotype performed well for detection of AD risk, indicating the model as a useful tool for AD detection and the potential involvement of these genes in AD. Intriguingly, the model genes were found to be enriched in immune-related pathways and were associated with immune infiltration. In the past decades, increasing attention has been paid to immune dysregulations in AD. Several studies have demonstrated that innate immunity and inflammation are critical to AD onset and progression, which may be a therapeutic strategy to treat AD (Ennerfelt & Lukens, 2020; Pons & Rivest, 2022). For example, rare coding variants in *PLCG2*, *ABI3*, and *TREM2* were identified in patients with AD, implicating the dysregulation of microglial-mediated innate immunity in the disease pathogenesis (Sims et al., 2017). It was reported that targeted immunotherapy of the central nervous system in animals could reduce the activation of innate immunity and neurodegeneration, suggesting the possible clinical benefit for AD therapy (Ryu et al., 2018). Furthermore, we also found that several genes in the AD detection models were involved in neuroinflammation and immune activation that were associated with development and progression of AD. For example, the expression of *CD163* could be enhanced by the over-expression of ACE, which is an amyloid- β protein degrading enzyme (Koronyo-Hamaoui et al., 2020). *CD163* can act as a surface scavenger receptor and may mediate internalization of oligomeric $A\beta$ (Huang et al., 2013). In addition, previous studies indicated that targeting the activity of *CSF1R* could inhibit neuroinflammation and slow neuronal damage and disease progression, especially in mouse models of AD (Gomez-Nicola et al., 2013; Mancuso et al., 2019). These findings suggested the potential involvement of the model genes in immune activation and AD progression, which deserves further investigation.

In addition, our study suggests that *APOE* is one of the robust genes in the detection models. It has been well established that *APOE* $\epsilon 4$ allele is a risk gene for AD onset, especially in sporadic AD (Jia et al., 2020; Martens et al., 2022). The AD risk is dose dependent, as those carrying 1 *APOE* $\epsilon 4$ allele have a 2-3-fold increased risk for AD, while those carrying 2 $\epsilon 4$ alleles have a 10-15-fold increased AD risk (Troutwine et al., 2022). Recent research also found that the *APOE* genotype-mediated and immune-related pathways were involved in AD (Panitch et al., 2022), indicating the effect of *APOE* $\epsilon 4$ allele and immunity in the development and progression of the disease.

Nevertheless, this study has several limitations. First, even though our cross-sectional study showed that the detection model worked well both in the internal and external datasets, future large-scale longitudinal studies among different ethnic populations are warranted to validate our results and further develop the risk prediction models for AD. Second, the interactions between the 13 hub genes and the potential mechanisms in AD development and progression have not been explored clearly and should be examined in the future.

5. Conclusions

In summary, by analyzing the microarray data from GEO database, we identified 13 key genes that were associated with AD risk. A detection model consisting of the key genes and *APOE* genotype was developed and validated in internal and external datasets. Furthermore, all the model genes were found to be related to immune infiltration, providing a promising strategy for AD diagnosis

and treatment. Future experimental research may help clarify the biological function and mechanisms of these genes in AD pathogenesis and progression.

CRedit authorship contribution statement

Min Zhu: Conceptualization, Resources, Formal analysis, Validation, Visualization, Writing – original draft, Writing – review & editing. **Tingting Hou:** Writing – review & editing, Resources, Funding acquisition, Supervision. **Longfei Jia:** Writing – review & editing, Resources. **Qihua Tan:** Writing – review & editing, Resources. **Chengxuan Qiu:** Writing – review & editing, Resources, Funding acquisition, Supervision. **Yifeng Du:** Writing – review & editing, Resources, Funding acquisition, Supervision.

Disclosure statement

The authors declare no competing interests.

Acknowledgements

The present study was supported in part by grants from the Brain Science and Brain-Like Intelligence Technology Research Projects of China (grant no.: 2021ZD0201801 and 2021ZD0201808), the National Key R&D Program of China (grant no.: 2017YFC1310100 and 2022YFC3501404); the National Nature Science Foundation of China (grant no.: 81861138008, 81772448, 82011530139, 82171175, and 82200980); the Academic Promotion Program of Shandong First Medical University (grant no.: 2019QL020); the Integrated Traditional Chinese and Western Medicine Program in Shandong Province (grant no.: YXH2019ZXY008); the Natural Science Foundation of Shandong Province (grant no.: ZR2021MH392 and ZR2021QH240), the Post-doctoral Innovation Project of Shandong Province, the Technology Development Plan Project of Jinan City (grant no.: 202134028); the Shandong Provincial Key Research and Development Program (grant no.: 2021LCZX03); and the Taishan Scholar Program of Shandong Province, China. C Qiu received grants from the Swedish Research Council (grant no.: 2017-05819 for the China-Sweden Joint Research Projects and 2020-01574 for the research project), and the Swedish Foundation for International Cooperation in Research and Higher Education (STINT) (grant no.: CH2019-8320) for the Joint China-Sweden Mobility program, and Karolinska Institutet, Stockholm, Sweden.

Thanks to the researchers of the public datasets for sharing the data used in this study. Data collection and sharing for this project from ADNI dataset was funded by the ADNI (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by

the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neurobiolaging.2022.12.014](https://doi.org/10.1016/j.neurobiolaging.2022.12.014).

References

- Berchtold, N.C., Coleman, P.D., Cribbs, D.H., Rogers, J., Gillen, D.L., Cotman, C.W., 2013. Synaptic genes are extensively downregulated across multiple brain regions in normal human aging and Alzheimer's disease. *Neurobiol Aging* 34 (6), 1653–1661.
- Brown 3rd, J., Theisler, C., Silberman, S., Magnuson, D., Gottardi-Littell, N., Lee, J.M., Yager, D., Crowley, J., Sambamurti, K., Rahman, M.M., Reiss, A.B., Eckman, C.B., Wolozin, B., 2004. Differential expression of cholesterol hydroxylases in Alzheimer's disease. *J Biol Chem* 279 (33), 34674–34681.
- Bu, X.L., Sun, P.Y., Fan, D.Y., Wang, J., Sun, H.L., Cheng, Y., Zeng, G.H., Chen, D.W., Li, H.Y., Yi, X., Shen, Y.Y., Miles, L.A., Maruff, P., Gu, B.J., Fowler, C.J., Masters, C.L., Wang, Y.J., 2022. Associations of plasma soluble CD22 levels with brain amyloid burden and cognitive decline in Alzheimer's disease. *Sci Adv* 8 (13), eabm5667.
- Castellani, R.J., Perry, G., 2012. Pathogenesis and disease-modifying therapy in Alzheimer's disease: the flat line of progress. *Arch Med Res* 43 (8), 694–698.
- Deng, L., Feng, J., Broadus, R.R., 2010. The novel estrogen-induced gene EIG121 regulates autophagy and promotes cell survival under stress. *Cell Death Dis* 1, e32.
- Ennerfelt, H.E., Lukens, J.R., 2020. The role of innate immunity in Alzheimer's disease. *Immunol Rev* 297 (1), 225–246.
- Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 33 (1), 1–22.
- Gomez-Nicola, D., Franssen, N.L., Suzzi, S., Perry, V.H., 2013. Regulation of microglial proliferation during chronic neurodegeneration. *J Neurosci* 33 (6), 2481–2493.
- Han, J., Chitu, V., Stanley, E.R., Wszolek, Z.K., Karrenbauer, V.D., Harris, R.A., 2022. Inhibition of colony stimulating factor-1 receptor (CSF-1R) as a potential therapeutic strategy for neurodegenerative diseases: opportunities and challenges. *Cell Mol Life Sci* 79 (4), 219.
- Hardy, J., 2006. A hundred years of Alzheimer's disease research. *Neuron* 52 (1), 3–13.
- Hokama, M., Oka, S., Leon, J., Ninomiya, T., Honda, H., Sasaki, K., Iwaki, T., Ohara, T., Sasaki, T., LaFerla, F.M., Kiyohara, Y., Nakabeppu, Y., 2014. Altered expression of diabetes-related genes in Alzheimer's disease brains: the Hisayama study. *Cereb Cortex* 24 (9), 2476–2488.
- Hu, Y., Yan, C., Hsu, C.H., Chen, Q.R., Niu, K., Komatsoulis, G.A., Meerzaman, D., 2014. OmicCircos: a simple-to-use R Package for the circular visualization of multidimensional omics data. *Cancer Inform* 13, 13–20.
- Huang, F.L., Shiao, Y.J., Hou, S.J., Yang, C.N., Chen, Y.J., Lin, C.H., Shie, F.S., Tsay, H.J., 2013. Cysteine-rich domain of scavenger receptor AI modulates the efficacy of surface targeting and mediates oligomeric Aβ internalization. *J Biomed Sci* 20, 54.
- Jia, L., Xu, H., Chen, S., Wang, X., Yang, J., Gong, M., Wei, C., Tang, Y., Qu, Q., Chu, L., Shen, L., Zhou, C., Wang, Q., Zhao, T., Zhou, A., Li, Y., Li, F., Li, Y., Jin, H., Qin, Q., Jiao, H., Li, Y., Zhang, H., Lyu, D., Shi, Y., Song, Y., Jia, J., 2020. The APOE ε4 exerts differential effects on familial and other subtypes of Alzheimer's disease. *Alzheimers Dement* 16 (12), 1613–1623.
- Kho, M., Smith, J.A., Verweij, N., Shang, L., Ryan, K.A., Zhao, W., Ware, E.B., Gansevoort, R.T., Irvin, M.R., Lee, J.E., Turner, S.T., Sung, J., van der Harst, P., Arnett, D.K., Baylín, A., Park, S.K., Seo, Y.A., Kelly, K.M., Chang, Y.P.C., Zhou, X., Lieske, J.C., Kardina, S.L.R., 2020. Genome-wide association meta-analysis of individuals of European ancestry identifies suggestive loci for sodium intake, potassium intake, and their ratio measured from 24-hour or half-day urine samples. *J Nutr* 150 (10), 2635–2645.
- Kolde, R., Laur, S., Adler, P., Vilo, J., 2012. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* 28 (4), 573–580.
- Koronyo-Hamaoui, M., Sheyn, J., Hayden, E.Y., Li, S., Fuchs, D.T., Regis, G.C., Lopes, D.H.J., Black, K.L., Bernstein, K.E., Teplow, D.B., Fuchs, S., Koronyo, Y., Rentsendorj, A., 2020. Peripherally derived angiotensin converting enzyme-enhanced macrophages alleviate Alzheimer-related disease. *Brain* 143 (1), 336–358.
- Langfelder, P., Horvath, S., 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559.
- Langfelder, P., Horvath, S., 2012. Fast R functions for robust correlations and hierarchical clustering. *J Stat Softw* 46 (11), i11.

- Lee, T., Lee, H., 2020. Prediction of Alzheimer's disease using blood gene expression data. *Sci Rep* 10 (1), 3485.
- Li, X., Chu, S.G., Shen, X.N., Hou, X.H., Xu, W., Ou, Y.N., Dong, Q., Tan, L., Yu, J.T., 2019. Genome-wide association study identifies SIAH3 locus influencing the rate of ventricular enlargement in non-demented elders. *Aging (Albany NY)* 11 (21), 9862–9874.
- Liang, W.S., Dunckley, T., Beach, T.G., Grover, A., Mastroeni, D., Walker, D.G., Caselli, R.J., Kukull, W.A., McKeel, D., Morris, J.C., Hulette, C., Schmechel, D., Alexander, G.E., Reiman, E.M., Rogers, J., Stephan, D.A., 2007. Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. *Physiol Genomics* 28 (3), 311–322.
- Mancuso, R., Fryatt, G., Cleal, M., Obst, J., Pipi, E., Monzon-Sandoval, J., Ribe, E., Winchester, L., Webber, C., Nevado, A., Jacobs, T., Austin, N., Theunis, C., Grauwen, K., Daniela Ruiz, E., Mudher, A., Vicente-Rodriguez, M., Parker, C.A., Simmons, C., Cash, D., Richardson, J., Consortium, N., Jones, D.N.C., Lovestone, S., Gomez-Nicola, D., Perry, V.H., 2019. CSF1R inhibitor JNJ-40346527 attenuates microglial proliferation and neurodegeneration in P301S mice. *Brain* 142 (10), 3243–3264.
- Martens, Y.A., Zhao, N., Liu, C.C., Kanekiyo, T., Yang, A.J., Goate, A.M., Holtzman, D.M., Bu, G., 2022. ApoE cascade hypothesis in the pathogenesis of alzheimer's disease and related dementias. *Neuron* 110 (8), 1304–1317.
- McKay, E.C., Beck, J.S., Khoo, S.K., Dykema, K.J., Cottingham, S.L., Winn, M.E., Paulson, H.L., Lieberman, A.P., Counts, S.E., 2019. Peri-infarct upregulation of the oxytocin receptor in vascular dementia. *J Neuropathol Exp Neurol* 78 (5), 436–452.
- Panitch, R., Hu, J., Xia, W., Bennett, D.A., Stein, T.D., Farrer, L.A., Jun, G.R., 2022. Blood and brain transcriptome analysis reveals APOE genotype-mediated and immune-related pathways involved in Alzheimer disease. *Alzheimers Res Ther* 14 (1), 30.
- Patel, H., Hodges, A.K., Curtis, C., Lee, S.H., Troakes, C., Dobson, R.J.B., Newhouse, S.J., 2019. Transcriptomic analysis of probable asymptomatic and symptomatic alzheimer brains. *Brain Behav Immun* 80, 644–656.
- Pey, P., Pearce, R.K., Kalaitzakis, M.E., Griffin, W.S., Gentleman, S.M., 2014. Phenotypic profile of alternative activation marker CD163 is different in Alzheimer's and Parkinson's disease. *Acta Neuropathol Commun* 2 (2014), 21.
- Piras, I.S., Krate, J., Delvaux, E., Nolz, J., De Both, M.D., Mastroeni, D.F., Serrano, G.E., Sue, L.L., Beach, T.G., Coleman, P.D., Huentelman, M.J., 2019a. Association of AEBP1 and NRN1 RNA expression with Alzheimer's disease and neurofibrillary tangle density in middle temporal gyrus. *Brain Res* 1719, 217–224.
- Piras, I.S., Krate, J., Delvaux, E., Nolz, J., Mastroeni, D.F., Persico, A.M., Jepsen, W.M., Beach, T.G., Huentelman, M.J., Coleman, P.D., 2019b. Transcriptome changes in the alzheimer's disease middle temporal gyrus: importance of RNA metabolism and mitochondria-associated membrane genes. *J Alzheimers Dis* 70 (3), 691–713.
- Pons, V., Rivest, S., 2022. Targeting systemic innate immune cells as a therapeutic avenue for alzheimer disease. *Pharmacol Rev* 74 (1), 1–17.
- Prince, M.J., Wimo, A., Guerchet, M.M., Ali, G.C., Wu, Y.-T., Prina, M., 2015. World Alzheimer Report 2015 - The Global Impact of Dementia. Alzheimer's Disease International, London.
- Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., Smyth, G.K., 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43 (7) e47.
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., Muller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12, 77.
- Ryu, J.K., Rafalski, V.A., Meyer-Franke, A., Adams, R.A., Poda, S.B., Rios Coronado, P.E., Pedersen, L.O., Menon, V., Baeten, K.M., Sikorski, S.L., Bedard, C., Hanspers, K., Bardehle, S., Mendiola, A.S., Davalos, D., Machado, M.R., Chan, J.P., Plastira, I., Petersen, M.A., Pfaff, S.J., Ang, K.K., Hallenbeck, K.K., Syme, C., Hakozaki, H., Ellisman, M.H., Swanson, R.A., Zamvil, S.S., Arkin, M.R., Zorn, S.H., Pico, A.R., Mucke, L., Freedman, S.B., Stavenhagen, J.B., Nelson, R.B., Akassoglou, K., 2018. Fibrin-targeting immunotherapy protects against neuroinflammation and neurodegeneration. *Nat Immunol* 19 (11), 1212–1223.
- Shigemizu, D., Mori, T., Akiyama, S., Higaki, S., Watanabe, H., Sakurai, T., Niida, S., Ozaki, K., 2020. Identification of potential blood biomarkers for early diagnosis of Alzheimer's disease through RNA sequencing analysis. *Alzheimers Res Ther* 12 (1), 87.
- Shijo, M., Honda, H., Suzuki, S.O., Hamasaki, H., Hokama, M., Abolhassani, N., Nakabeppu, Y., Ninomiya, T., Kitazono, T., Iwaki, T., 2018. Association of adipocyte enhancer-binding protein 1 with Alzheimer's disease pathology in human hippocampi. *Brain Pathol* 28 (1), 58–71.
- Sims, R., van der Lee, S.J., Naj, A.C., Bellenguez, C., Badarinarayan, N., Jakobsdottir, J., Kunkle, B.W., Boland, A., Raybould, R., Bis, J.C., Martin, E.R., Grenier-Boley, B., Heilmann-Heimbach, S., Chouraki, V., Kuzma, A.B., Slegers, K., Vronskaya, M., Ruiz, A., Graham, R.R., O'Laso, R., Hoffmann, P., Grove, M.L., Vardarajan, B.N., Hiltunen, M., Nothen, M.M., White, C.C., Hamilton-Nelson, K.L., Epelbaum, J., Maier, W., Choi, S.H., Beecham, G.W., Dulary, C., Herms, S., Smith, A.V., Funk, C.C., Derbois, C., Forstner, A.J., Ahmad, S., Li, H., Bacq, D., Harold, D., Satizabal, C.L., Valladares, O., Squassina, A., Thomas, R., Brody, J.A., Qu, L., Sanchez-Juan, P., Morgan, T., Wolters, F.J., Zhao, Y., Garcia, F.S., Denning, N., Fornage, M., Malamon, J., Naranjo, M.C.D., Majounie, E., Mosley, T.H., Dombroski, B., Wallon, D., Lupton, M.K., Dupuis, J., Whitehead, P., Fratiglioni, L., Medway, C., Jian, X., Mukherjee, S., Keller, L., Brown, K., Lin, H., Cantwell, L.B., Panza, F., McGuinness, B., Moreno-Grau, S., Burgess, J.D., Solfrizzi, V., Proitsis, P., Adams, H.H., Allen, M., Seripa, D., Pastor, P., Cupples, L.A., Price, N.G., Hannequin, D., Frank-Garcia, A., Levy, D., Chakrabarty, P., Caffarra, P., Nieding, I., Beiser, A.S., Giedraitis, V., Hampel, H., Garcia, M.E., Wang, X., Lannfelt, L., Mecocci, P., Eiriksdottir, G., Crane, P.K., Pasquier, F., Boccardi, V., Henandez, I., Barber, R.C., Scherer, M., Tarraga, L., Adams, P.M., Leber, M., Chen, Y., Albert, M.S., Riedel-Heller, S., Emilsson, V., Beekly, D., Braae, A., Schmidt, R., Blacker, D., Mاضullo, C., Schmidt, H., Doody, R.S., Spalletta, G., Longstreth Jr., W.T., Fairchild, T.J., Bossu, P., Lopez, O.L., Frosch, M.P., Sacchinelli, E., Ghetti, B., Yang, Q., Huebinger, R.M., Jessen, F., Li, S., Kamboh, M.I., Morris, J., Sotolongo-Grau, O., Katz, M.J., Corcoran, C., Dunstan, M., Braddel, A., Thomas, C., Meggy, A., Marshall, R., Gerrish, A., Chapman, J., Aguilar, M., Taylor, S., Hill, M., Fairen, M.D., Hodges, A., Vellas, B., Soininen, H., Kloszewska, I., Daniilidou, M., Uphill, J., Patel, Y., Hughes, J.T., Lord, J., Turton, J., Hartmann, A.M., Cecchetti, R., Fenoglio, C., Serpente, M., Arcaro, M., Caltagirone, C., Orfei, M.D., Ciaramella, A., Pichler, S., Mayhants, M., Gu, W., Lleo, A., Fortea, J., Blesa, R., Barber, I.S., Brookes, K., Cupidi, C., Maletta, R.G., Carrell, D., Sorbi, S., Moebus, S., Urbano, M., Pilotto, A., Kornhuber, J., Bosco, P., Todd, S., Craig, D., Johnston, J., Gill, M., Lawlor, B., Lynch, A., Fox, N.C., Hardy, J., Consortium, A., Albin, R.L., Apostolova, L.G., Arnold, S.E., Asthana, S., Atwood, C.S., Baldwin, C.T., Barnes, L.L., Barral, S., Beach, T.G., Becker, J.T., Bigio, E.H., Bird, T.D., Boeve, B.F., Bowen, J.D., Boxer, A., Burke, J.R., Burns, J.M., Buxbaum, J.D., Cairns, N.J., Cao, C., Carlson, C.S., Carlson, C.M., Carney, R.M., Carrasquillo, M.M., Carroll, S.L., Diaz, C.C., Chui, H.C., Clark, D.G., Cribbs, D.H., Crocco, E.A., DeCarli, C., Dick, M., Duara, R., Evans, D.A., Faber, K.M., Fallon, K.B., Fardo, D.W., Farlow, M.R., Ferris, S., Foroud, T.M., Galasko, D.R., Gearing, M., Geschwind, D.H., Gilbert, J.R., Graff-Radford, N.R., Green, R.C., Growdon, J.H., Hamilton, R.L., Harrell, L.E., Honig, L.S., Huentelman, M.J., Hulette, C.M., Hyman, B.T., Jarvik, G.P., Abner, E., Jin, L.W., Jun, G., Karydas, A., Kaye, J.A., Kim, R., Kowall, N.W., Kramer, J.H., LaFerla, F.M., Lah, J.J., Leverenz, J.B., Levey, A.I., Li, G., Lieberman, A.P., Lunetta, K.L., Lyketsos, C.G., Marson, D.C., Martiniuk, F., Mash, D.C., Masliah, E., McCormick, W.C., McCurry, S.M., McDavid, A.N., McKee, A.C., Mesulam, M., Miller, B.L., Miller, C.A., Miller, J.W., Morris, J.C., Murrell, J.R., Myers, A.J., O'Bryant, S., Olichney, J.M., Pankratz, V.S., Parisi, J.E., Paulson, H.L., Perry, W., Peskind, E., Pierce, A., Poon, W.W., Potter, H., Quinn, J.F., Raj, A., Raskind, M., Reisberg, B., Reitz, C., Ringman, J.M., Roberson, E.D., Rogaeve, E., Rosen, H.J., Rosenberg, R.N., Sager, M.A., Saykin, A.J., Schneider, J.A., Schneider, L.S., Seeley, W.W., Smith, A.G., Sonnen, J.A., Spina, S., Stern, R.A., Swerdlow, R.H., Tanzi, R.E., Thornton-Wells, T.A., Trojanowski, J.Q., Troncoso, J.C., Van Deerlin, V.M., Van Eldik, L.J., Vinters, H.V., Vonsattel, J.P., Weintraub, S., Welsh-Bohmer, K.A., Wilhelmson, K.C., Williamson, J., Wingo, T.S., Woltjer, R.L., Wright, C.B., Yu, C.E., Yu, L., Garzia, F., Golamauilly, F., Septier, G., Engelborghs, S., Vandenberghe, R., De Deyn, P.P., Fernandez, C.M., Benito, Y.A., Thonberg, H., Forsell, C., Lilius, L., Kinhult-Stahlbom, A., Kilander, L., Brundin, R., Concari, L., Helisalmi, S., Koivisto, A.M., Haapasalo, A., Derme-court, V., Fievet, N., Hanon, O., Dufouil, C., Brice, A., Ritchie, K., Dubois, B., Hilmali, J.J., Keene, C.D., Tschanz, J., Fitzpatrick, A.L., Kukull, W.A., Norton, M., Aspelund, T., Larson, E.B., Munger, R., Rotter, J.I., Lipton, R.B., Bullido, M.J., Hofman, A., Montine, T.J., Coto, E., Boerwinkle, E., Petersen, R.C., Alvarez, V., Rivadeneira, F., Reiman, E.M., Gallo, M., O'Donnell, C.J., Reich, J.S., Bruni, A.C., Royall, D.R., Dichgans, M., Sano, M., Galimberti, D., St George-Hyslop, P., Scarpini, E., Tsuang, D.W., Mancuso, M., Bonuccelli, U., Winslow, A.R., Daniele, A., Wu, C.K., Gerard/Perades, C.A.E., Peters, O., Nacmias, B., Riemenschneider, M., Heun, R., Brayne, C., Rubinsztein, D.C., Bras, J., Guerreiro, R., Al-Chalabi, A., Shaw, C.E., Collinge, J., Mann, D., Tsolaki, M., Clarimon, J., Sussams, R., Lovestone, S., O'Donovan, M.C., Owen, M.J., Behrens, T.W., Mead, S., Goate, A.M., Uitterlinden, A.G., Holmes, C., Cruchaga, C., Ingelsson, M., Bennett, D.A., Powell, J., Golde, T.E., Graff, C., De Jager, P.L., Morgan, K., Ertekin-Taner, N., Combarros, O., Psaty, B.M., Passmore, P., Younkin, S.G., Berr, C., Gudnason, V., Rujescu, D., Dickson, D.W., Dartigues, J.F., DeStefano, A.L., Ortega-Cubero, S., Hakonarson, H., Campion, D., Boada, M., Kawe, J.K., Farrer, L.A., Van Broeckhoven, C., Ikram, M.A., Jones, L., Haines, J.L., Tzourio, C., Launer, L.J., Escott-Price, V., Mayeux, R., Deleuze, J.F., Amin, N., Holmans, P.A., Pericak-Vance, M.A., Amouyel, P., van Duijn, C.M., Ramirez, A., Wang, L.S., Lambert, J.C., Seshadri, S., Williams, J., Schellenberg, G.D., 2017. Rare coding variants in PLCG2, Aβ3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. *Nat Genet* 49 (9), 1373–1384.
- Snijders, G., van Zuiden, W., Sneeuwer, M.A.M., Berdenis van Berlekom, A., van der Geest, A.T., Schnieder, T., MacIntyre, D.J., Hol, E.M., Kahn, R.S., de Witte, L.D., 2021. A loss of mature microglial markers without immune activation in schizophrenia. *Glia* 69 (5), 1251–1267.
- Stevenson-Hoare, J., Heslegrave, A., Leonenko, G., Fathalla, D., Bellou, E., Luckcuck, L., Marshall, R., Sims, R., Morgan, B.P., Hardy, J., de Strooper, B., Williams, J., Zetterberg, H., Escott-Price, V., 2022. Plasma biomarkers and genetics in the diagnosis and prediction of Alzheimer's disease. *Brain* 146 (2), 690–699.
- Troutwine, B.R., Hamid, L., Lysaker, C.R., Strope, T.A., Wilkins, H.M., 2022. Apolipoprotein E and Alzheimer's disease. *Acta Pharm Sin B* 12 (2), 496–510.
- Walter, W., Sanchez-Cabo, F., Ricote, M., 2015. GOpot: an R package for visually combining expression data with functional analysis. *Bioinformatics* 31 (17), 2912–2914.
- Yu, G., Wang, L.G., Han, Y., He, Q.Y., 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16 (5), 284–287.
- Yu, L., Tasaki, S., Schneider, J.A., Arfanakis, K., Duong, D.M., Wingo, A.P., Wingo, T.S., Kearns, N., Thatcher, G.R.J., Seyfried, N.T., Levey, A.I., De Jager, P.L., Bennett, D.A., 2020. Cortical proteins associated with cognitive resilience in community-dwelling older persons. *JAMA Psychiatry* 77 (11), 1172–1180.
- Zetterberg, H., Bendlin, B.B., 2021. Biomarkers for Alzheimer's disease-preparing for a new era of disease-modifying therapies. *Mol Psychiatry* 26 (1), 296–308.

Zhang, X., Lv, C., An, Y., Liu, Q., Rong, H., Tao, L., Wang, Y., Wang, Y., Xiao, R., 2018. Increased levels of 27-hydroxycholesterol induced by dietary cholesterol in brain contribute to learning and memory impairment in rats. *Mol Nutr Food Res* 62 (3). doi:[10.1002/mnfr.201700531](https://doi.org/10.1002/mnfr.201700531).

Zhong, H., Wang, J., Zhu, Y., Shen, Y., 2021. Comprehensive analysis of a nine-gene signature related to tumor microenvironment in lung adenocarcinoma. *Front Cell Dev Biol* 9, 700607.