



OPEN ACCESS

EDITED BY

Can Yang,
Hong Kong University of Science and
Technology, Hong Kong SAR, China

REVIEWED BY

Jiawei Wang,
Yale University, United States
Yiliang Zhang,
Yale University, United States

*CORRESPONDENCE

Xiao-Yu Li,
673426608@qq.com

[†]These authors have contributed equally
to this work

SPECIALTY SECTION

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

RECEIVED 26 September 2022

ACCEPTED 21 November 2022

PUBLISHED 06 December 2022

CITATION

Yao P-A, Sun H-J and Li X-Y (2022),
Identification of key genes in late-onset
major depressive disorder through a co-
expression network module.
Front. Genet. 13:1048761.
doi: 10.3389/fgene.2022.1048761

COPYRIGHT

© 2022 Yao, Sun and Li. This is an open-
access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Identification of key genes in late-onset major depressive disorder through a co-expression network module

Ping-An Yao^{1,2†}, Hai-Ju Sun^{2†} and Xiao-Yu Li^{2,3*}

¹School of Pharmaceutical Sciences, Zhejiang Chinese Medical University, Hangzhou, China, ²Department of Neurobiology and Acupuncture Research, The Third Clinical Medical College, Zhejiang Chinese Medical University, Key Laboratory of Acupuncture and Neurology of Zhejiang Province, Hangzhou, China, ³The Third Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China

Late-onset major depressive disorder (LOD) increases the risk of disability and suicide in elderly patients. However, the complex pathological mechanism of LOD still remains unclear. We selected 10 LOD patients and 12 healthy control samples from the GSE76826 dataset for statistical analysis. Under the screening criteria, 811 differentially expressed genes (DEGs) were screened. We obtained a total of two most clinically significant modules through the weighted gene co-expression network analysis (WGCNA). Functional analysis of the genes in the most clinically significant modules was performed to explore the potential mechanism of LOD, followed by protein–protein interaction (PPI) analysis and hub gene identification in the core area of the PPI network. Furthermore, we identified immune infiltrating cells using the cell-type identification by estimating relative subsets of RNA transcripts (CIBERSORT) algorithm between healthy subjects and LOD patients with the GSE98793 dataset. Next, six hub genes (*CD27*, *IL7R*, *CXCL1*, *CCR7*, *IGLL5*, and *CD79A*) were obtained by intersecting hub genes with DEGs, followed by verifying the diagnostic accuracy with the receiver operating characteristic curve (ROC). In addition, we constructed the least absolute shrinkage and selection operator (LASSO) regression model for hub gene cross-validation. Finally, we found that *CD27* and *IGLL5* were good diagnostic indicators of LOD, and *CD27* may be the key gene of immune function change in LOD. In conclusion, our research shows that the changes in the immune function may be an important mechanism in the development of LOD, which can provide some guidance for the related research of LOD in the future.

KEYWORDS

late-onset major depressive disorder, weighted gene co-expression network analysis, differentially expressed genes, immune infiltration, hub gene, receiver operating characteristic

Introduction

As a common mental disorder with high morbidity, major depressive disorder (MDD) is the main cause of severe disability in patients (Kessler et al., 2003). The age distribution of the MDD onset shows that depression is common during the whole life cycle (Wagner et al., 2020). However, the effective treatment of MDD is currently followed in less than half of the patients (Dowlati et al., 2010). Complex physiological and pathological conditions are the main reason for its refractory state. Late-onset MDD (LOD) usually occurs in patients whose onset age is greater than or equal to 50 years (Miyata et al., 2016). Compared with non-LOD patients, LOD patients have a higher risk of suicide and cognitive and psychosocial disorders (Alexopoulos et al., 1988), which greatly increases the social and family burden. The lack of specificity of diagnostic methods and the decline of body function, such as the decline of the immune system function and the increase of combined diseases, contribute to the intractability of LOD. In most cases, LOD can be accurately identified by the Diagnostic and Statistical Manual of Mental Disorders (DSM). Unfortunately, LOD combined with other psychiatric disorders makes it different (Das-Munshi et al., 2008). Therefore, it is urgent to improve the diagnostic accuracy of LOD patients using physiological and pathological indexes.

The involvement of immune response in the process of MDD has been confirmed, such as immune cells and cytokines (Chan et al., 2019). Body function decreases with age, including the immune system function. Changes in age-dependent functions contribute to a more harmful effect on adaptive immunity (Schleifer et al., 1989), which is represented by B and T lymphocytes. It has been confirmed that the plasma of MDD patients with non-responsiveness to anti-depressants contains high concentrations of cytokines, such as interleukin 6 (Vogelzangs et al., 2012). Mishra et al. (2018) demonstrated that high concentrations of inflammatory mediators were related to the severity and duration of depressive disorder. Surprisingly, the decreased expression of cytokines and chemokines was found in elderly patients with MDD (Shinko et al., 2020). However, the role of the differential changes of immune response in LOD is poorly understood. The differential changes of immune response may be one of the main factors that make LOD patients different from non-LOD patients.

In recent years, bioinformatics methods have been widely used in different diseases, including LOD (Miyata et al., 2016). A previous study mainly focused on differential genes in the blood of LOD patients, partially revealing the pathological process of LOD (Miyata et al., 2016). However, the ignorance of non-abnormal genes may lead to deviation in the diagnosis of LOD. Different from general MDD, the pathological mechanism of LOD may be more complex. A comprehensive

analysis of all genes in LOD may be the key to clarify the mechanism of LOD. The WGCNA can accurately screen diagnostic genes of diseases through the relationship between genes and clinical characteristics of diseases (Langfelder and Horvath, 2008). As a bioinformatics approach commonly used to analyze the changes in the immune function, the CIBERSORT algorithm can reflect the abundance of 22 kinds of immune cells in different populations through the expression of sample genes (Zhou et al., 2021). Therefore, in our study, we used the WGCNA and CIBERSORT algorithm to identify hub genes and immune infiltrating cells in the plasma of LOD patients and explored the potential mechanism, to provide some guidance for the related research on LOD in the future.

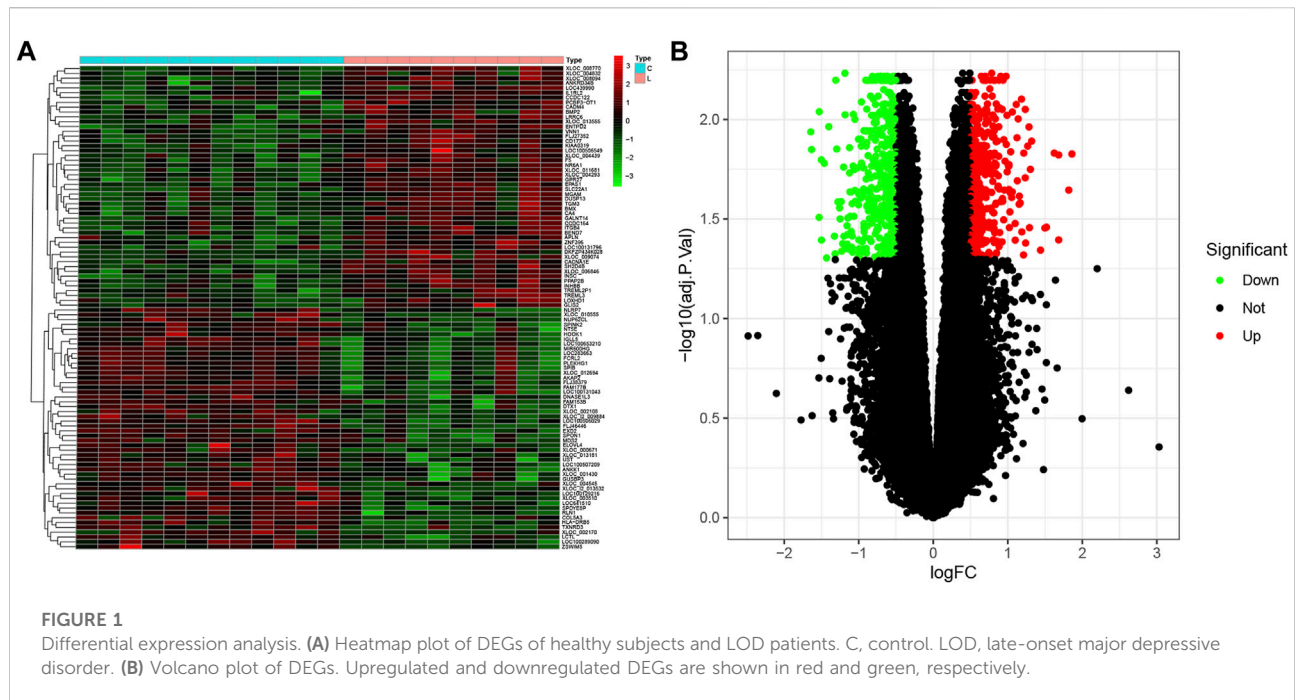
Materials and methods

Dataset collection

All datasets (GSE76826 and GSE98793) used in this study were downloaded from the Gene Expression Omnibus database. The GSE76826 dataset contains 12 healthy control individuals, 10 patients with LOD (onset age ≥ 50 years), and 10 patients with LOD syndromal remission. Also, 12 healthy control individuals and 10 patients with LOD were selected for the WGCNA. In addition, 12 healthy control individuals and 10 patients with LOD syndromal remission were selected for the cross-validation of hub genes. The GSE98793 dataset contains 64 patients with MDD, 64 healthy controls, and 64 patients with MDD and generalized anxiety disorder. From the GSE98793 dataset, we selected healthy controls and MDD patients without anxiety disorder for the immune infiltration analysis, whose age ≥ 50 years. We used the Perl language (version 5.32.1) to convert the probe ID extracted from the downloaded series matrix into a gene symbol. After that, R software (version 3.6.1) was used on the gene expression datasets for normalization and statistics (Ren et al., 2020). We followed the methods of dataset sorting, which were used in a previous literature report (Xu et al., 2022).

Differentially expressed genes (DEGs) between healthy control individuals and patients with LOD

Compared with the healthy control individuals, the genes with fold changes < -1.4 or > 1.4 and an adjusted p -value < 0.05 were defined as DEGs. According to the screening criteria for DEGs, we used R software with the “limma” package to complete the screening of DEGs. A heatmap and volcano map were obtained to visualize the top 100 DEGs (50 upregulations and 50 downregulations) through the “pheatmap” package in R software.



Co-expression network

We used R software with the “WGCNA” package to identify and visualize the co-expressed modules of the genes in the GSE76826 dataset (Langfelder and Horvath, 2008; Xu et al., 2022). The construction of a co-expression network was in accordance with the steps in the previous literature (Xu et al., 2022). First, the sample information on healthy control individuals and 10 patients with LOD in the GSE76826 dataset was matched with the clinical characteristics, followed by clustering to identify outliers. Then, 1–20 were set as the candidate power values. According to the similarity of co-expression, we used the pickSoftThreshold function to calculate the corresponding power values and drew the scatter diagram corresponding to the power values, including the analysis of network topology for various soft thresholds and mean connectivity. Among them, when the R-square of the scale-free topological network fitting model is about 0.9, the corresponding soft threshold is the best power value to establish the adjacency matrix. Next, the TOM matrix was calculated, followed by detecting the modules through the dynamic tree cutting function and merging similar modules according to MEDissThres = 0.25. Finally, all modules had calculated the clinical correlation values and recorded the genes in each module. Only the modules with a correlation index value greater than 0.5 and a *p*-value less than 0.05 were considered the most clinically significant modules for LOD. The gene information in the most clinically significant modules was selected.

Functional enrichment

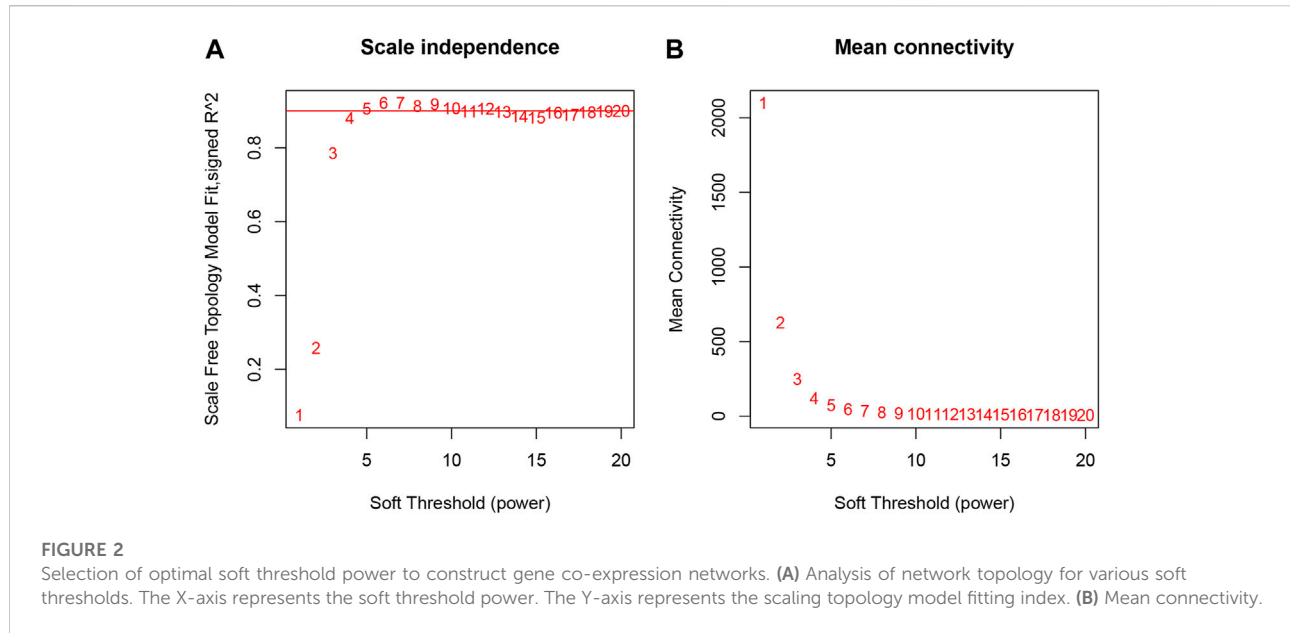
We performed functional enrichment analysis on genes in the first two most clinically significant modules, including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG). The functional enrichment results were visualized through corresponding R packages. The criteria for functional enrichment were the *p* threshold <0.05.

Immune infiltration analysis

To explore the difference in the proportion of immune infiltrating cells (22 kinds in total) in healthy control individuals and patients with LOD, we used the CIBERSORT algorithm to analyze selected samples from the GSE98793 dataset. The results of the immune infiltrating cells were visualized by R software with corresponding R packages.

Protein–protein interaction (PPI) analysis

Genes in the first two most clinically significant modules were selected to construct a PPI network through the Search Tool for the Retrieval of Interacting Genes (STRING) (Szklarczyk et al., 2011; Szklarczyk et al., 2019). The setting standards of the PPI network are as follows: interaction score >0.7 and isolated nodes removed. We exported the results of the PPI networks as TSV format files, followed by importing into Cytoscape software (version: 3.7.1). The top 10 hub genes in the PPI networks of the selected two modules were calculated



by the cytoHubba tool. Finally, the hub genes of selected modules and DEGs were intersected to screen the hub genes, followed by GO, and KEGG analyses.

Receiver operating characteristic curve (ROC)

We applied the ROC to verify the diagnostic accuracy of the six hub genes in LOD. The value of the area under curve (AUC) of the ROC represented the diagnostic accuracy. We used R software with the “pROC” package to perform ROC analysis and visualize the results. In the GSE76826 dataset, healthy control individuals and 10 patients with LOD were defined as the control group and case group, respectively. According to the respective real expression values of hub genes in the two groups, the ROC curves were drawn using the “pROC” package with R software. Meanwhile, the ROC curves were smoothed using the “smooth” function, and the AUCs of the ROC curves were calculated for each hub gene. AUC values >0.7000 were defined as accurate. For the cross-validation, we re-drew the ROC curve of each hub gene, according to the actual expression values of hub genes in healthy control individuals and patients with LOD syndrome remission to approach the real testing power.

Least absolute shrinkage and selection operator (LASSO) regression model and correlation between genes and immune cells

According to the expressions of hub genes, we used R software with the “glmnet” package to construct the LASSO

regression model. The parameters of the LASSO regression model were set as follows: family = “binomial” and alpha = 1. Tenfold cross-validation was utilized to determine the optimal value of the penalty parameter λ . Next, the genes screened by the LASSO regression model and the genes verified by the ROC were intersected to obtain the final genes. We used “ggpubr” and “limma” packages to visualize the expression level of the final genes. Finally, the correlations between final genes and immune cells were analyzed by Spearman’s method.

Results

Research diagram

As shown in [Supplementary Figure S1](#), GSE76826 and GSE98793 datasets were downloaded for this study. We constructed a co-expression network of selected samples from the GSE76826 dataset and identified the hub genes in LOD. To determine the types of immune infiltrating cells of LOD, we used the CIBERSORT algorithm to analyze the selected samples in the GSE98793 dataset.

DEGs between healthy control individuals and patients with LOD

According to the defined criteria (fold changes < -1.4 or >1.4 and adjusted p -value < 0.05.), we applied R software with the “limma” package to identify the DEGs between healthy control individuals and patients with LOD. A total of 811 DEGs, which consisted of 383 upregulated

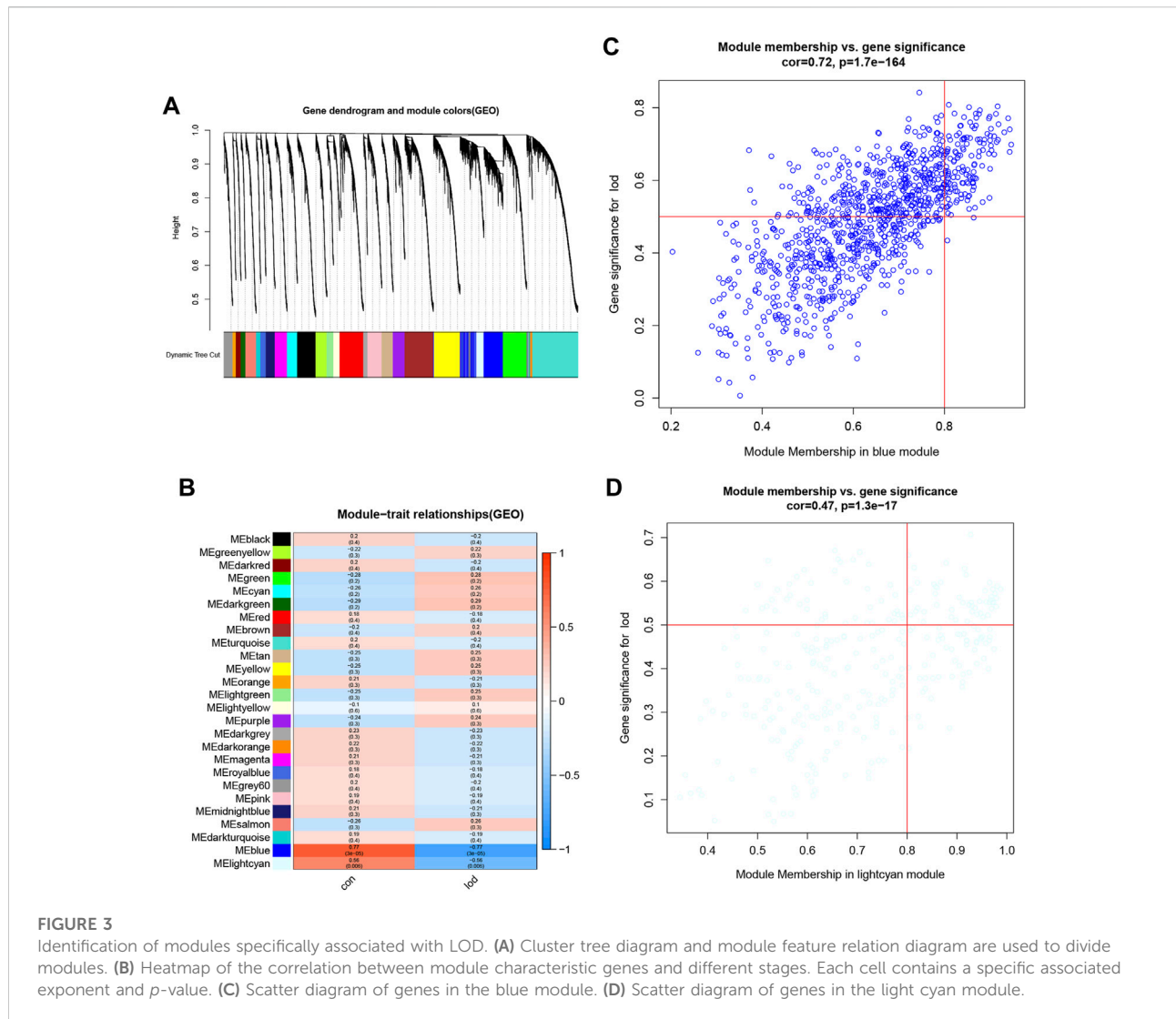


TABLE 1 Gene number information of modules by the WGCNA.

Module name	Number of genes	Module name	Number of genes
Black	597	Light yellow	202
Green-yellow	366	Purple	389
Dark red	160	Dark gray	137
Green	768	Dark orange	65
Cyan	342	Magenta	392
Dark green	154	Royal blue	171
Red	765	Gray	284
Brown	923	Pink	460
Turquoise	1,512	Midnight blue	303
Tan	363	Salmon	346
Yellow	857	Dark turquoise	147
Orange	108	Blue	1,025
Light green	224	Light cyan	295

genes and 428 downregulated genes, were obtained. As shown in Figure 1, we selected the top 50 genes, upregulated or downregulated, respectively, to visualize through a heatmap and volcano plot.

Co-expression network

The samples selected from the GSE76826 dataset were used for the WGCNA to explore the hub genes in LOD. The co-expression network was constructed under the soft threshold optimized as 4 (scale-free $R^2 = 0.88$), which meets the approximate scale-free topology criterion (Figure 2). A total of 26 modules (Figure 3B) were generated after the combination of similar modules under the criterion that MEDissThres was set as 0.25. In brief, there were 1,025 genes in the blue module, 923 genes in the brown module, 296 genes in the light cyan module, and 1,512 genes in the turquoise module. The specific module information is shown in Table 1.

The first two most clinically significant modules

As shown in Figure 3, the correlation values and p -values between each module and clinical feature are given in detail and visualized. Among them, the blue module had the most significant correlation with LOD ($r = -0.77$ and $P = 3e-05$). The light cyan module had the second significant correlation with LOD ($r = -0.56$ and $p = 0.006$) (Figure 3). Therefore, the blue module and the light cyan module were defined as the first two most clinically significant modules. We selected the genes in the first two most clinically significant modules for detailed analysis (Figure 3).

Functional enrichment analysis of the genes in the first two most clinically significant modules

Genes in the blue module ($n = 1,025$) and light cyan module ($n = 296$) were selected to perform GO and KEGG enrichment analyses. The results of the GO analysis were divided into the biological process (BP), cellular constituent (CC), and molecular function (MF). As shown in the GO results, in the blue module, the BP of LOD was mainly manifested in T-cell activation, regulation of cell–cell adhesion, and leukocyte cell–cell adhesion. The CC of LOD was mainly manifested in the external side of the plasma membrane, collagen-containing extracellular matrix, and the apical part of cells. The MF of LOD was mainly manifested in carbohydrate binding, immune receptor activity, and cytokine receptor activity. Meanwhile, the results of the KEGG analysis showed that the cytokine–cytokine receptor interaction, hematopoietic cell lineage,

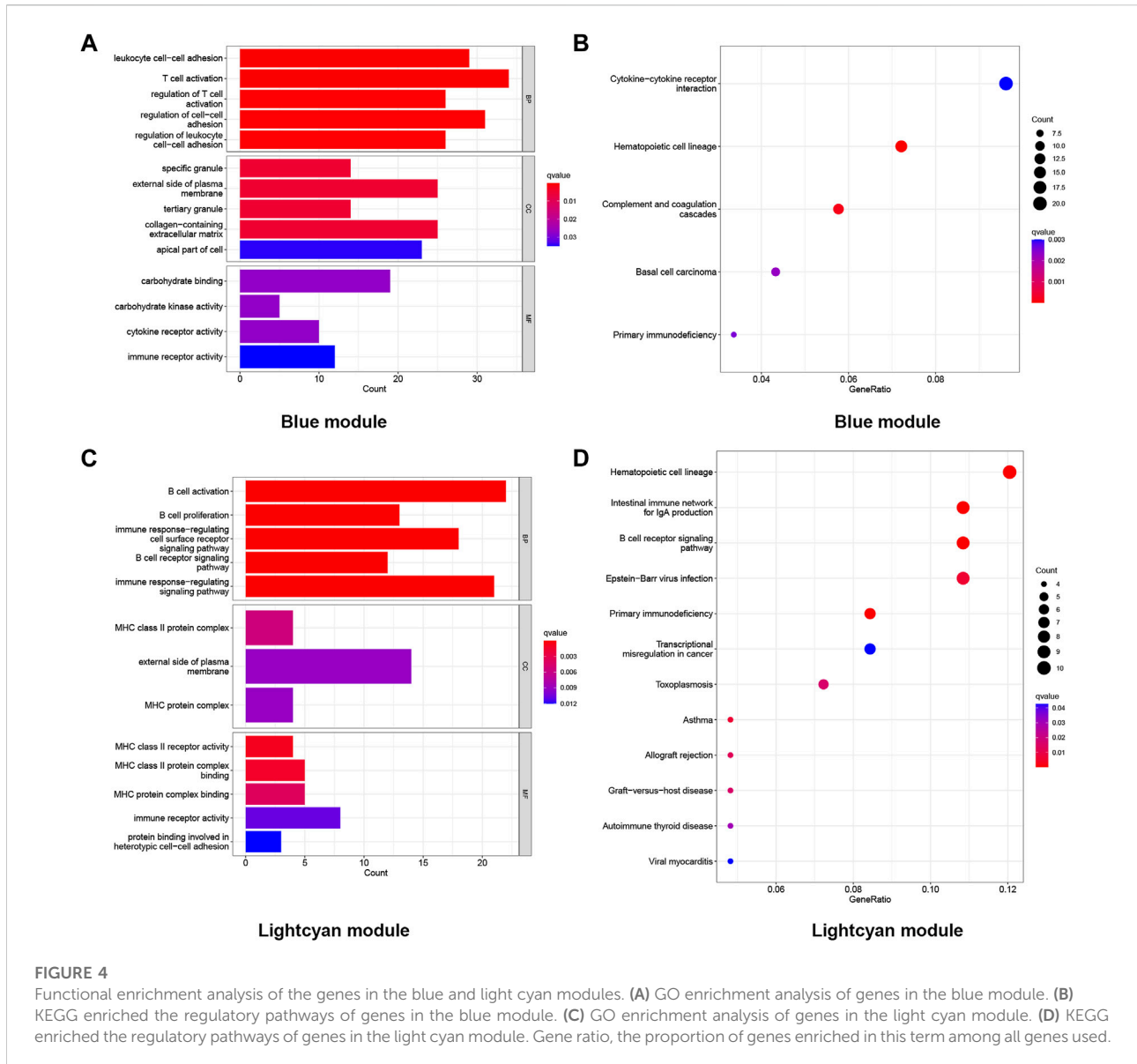
and complement and coagulation cascades were the main manifestations. Similarly, as shown in the GO results, in the light cyan module, the BP of LOD was mainly manifested in B-cell activation, the immune response-regulating signaling pathway, and immune response-regulating cell surface receptor signaling pathway. The CC of LOD was mainly manifested in the external side of the plasma membrane, MHC class II protein complex, and MHC protein complex. The MF of LOD was mainly manifested in immune receptor activity, MHC class II protein complex binding, and MHC protein complex binding. In the KEGG analysis of the light cyan module, the hematopoietic cell lineage, intestinal immune network for IgA production, and B-cell receptor signaling pathway were the main manifestations. All functional enrichment results are shown in Figure 4.

Immune infiltrating cells of LOD

A previous study confirmed that immune function changes play an important role in MDD (Berk et al., 2013). However, the change in the immune function in LOD patients remains unclear. In this study, the results of the functional enrichment analysis of genes in the first two most clinically significant modules showed that the pathological mechanism of LOD involved immune cell response and the regulation of cytokines and chemokines, suggesting immune-related mechanisms may play a role in LOD. To determine the immune infiltrating cells in LOD, we used the CIBERSORT algorithm to identify the immune infiltrating cells of the samples selected from the GSE98793 dataset. The proportion and percentage of 22 subpopulations of immune cells were depicted. It was observed that neutrophils, monocytes, CD8 T cells, and naïve CD4 T cells were the main infiltrating cells. Meanwhile, the results showed that CD4 memory T cells positively correlated with B-cell memory, while CD8 T-cell activation negatively correlated with neutrophils. Compared with healthy controls, LOD patients generally contained a higher proportion of plasma cells ($p = 0.003$), regulatory T cells (Tregs) ($p = 0.034$), and neutrophils ($p = 0.049$), whereas the gamma delta T-cell ($p = 0.003$) fraction was relatively lower. All the results are shown in Figure 5.

Hub genes of LOD and functional enrichment analysis

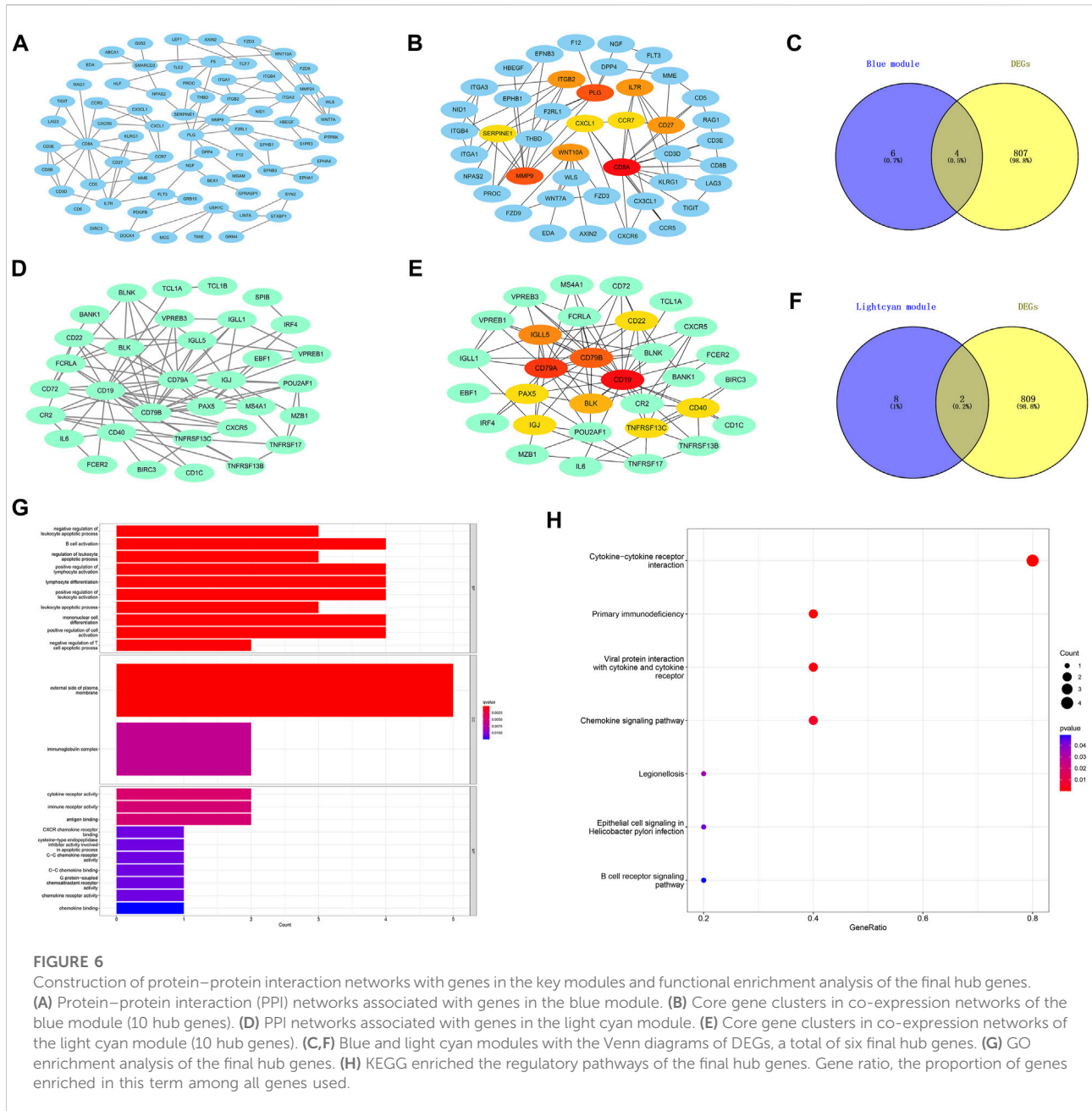
The PPI networks of the first two most clinically significant modules (blue module and light cyan module) are shown in Figure 6. We used the cytoHubba tool to calculate the top 10 hub genes in the blue module and light cyan module, according to the standard of the node degree. As shown in Figure 6, in the blue module, the top 10 hub genes were *CD8A*, *MMP9*, *PLG*, *IL7R*, *CD27*, *ITGB2*, *WNT10A*, *SERPINE1*, *CXCL1*, and *CCR7*. Meanwhile, in the light cyan module, the top 10 hub genes were *CD19*, *CD79A*, *CD79B*, *IGLL5*, *BLK*, *CD22*, *TNFRSF13C*,



CD40, *IGJ*, and *PAX5*. The hub genes were screened by intersecting DEGs and the 10 hub genes of the blue module and light cyan module. Finally, the six hub genes were *CD27*, *IL7R*, *CXCL1*, *CCR7*, *IGLL5*, and *CD79A*. In addition, we performed functional enrichment analysis on the six hub genes. As shown in the GO results of the hub genes (Figure 6), the BP was mainly enriched in B-cell activation and positive regulation of lymphocyte activation. The CC was mainly enriched in the external side of the plasma membrane and immunoglobulin complex. The MF was mainly enriched in the cytokine receptor activity, and immune receptor activity. Meanwhile, the KEGG pathway was mainly enriched in the cytokine–cytokine receptor interaction and primary immunodeficiency.

ROC verification

We evaluated the diagnostic accuracy of the six hub genes (*CD27*, *IL7R*, *CXCL1*, *CCR7*, *IGLL5*, and *CD79A*) through ROC verification. As shown in Figure 7, the respective AUC values of the six hub genes were 0.875, 0.800, 0.908, 0.752, 0.833, and 0.758 (Figure 7). All AUC values were greater than 0.700. Meanwhile, ROC results for cross-validation show that the respective AUC values of the six hub genes (*CD27*, *IL7R*, *CXCL1*, *CCR7*, *IGLL5*, and *CD79A*) were 0.809, 0.622, 0.644, 0.417, 0.801, and 0.737 (Supplementary Figure S2). The AUCs of the ROC of *CD27*, *IGLL5*, and *CD79A* were still higher than 0.700. Taken together, our results suggest *CD27*, *IGLL5*, and *CD79A* can be used as the accurate diagnostic indicators of LOD.



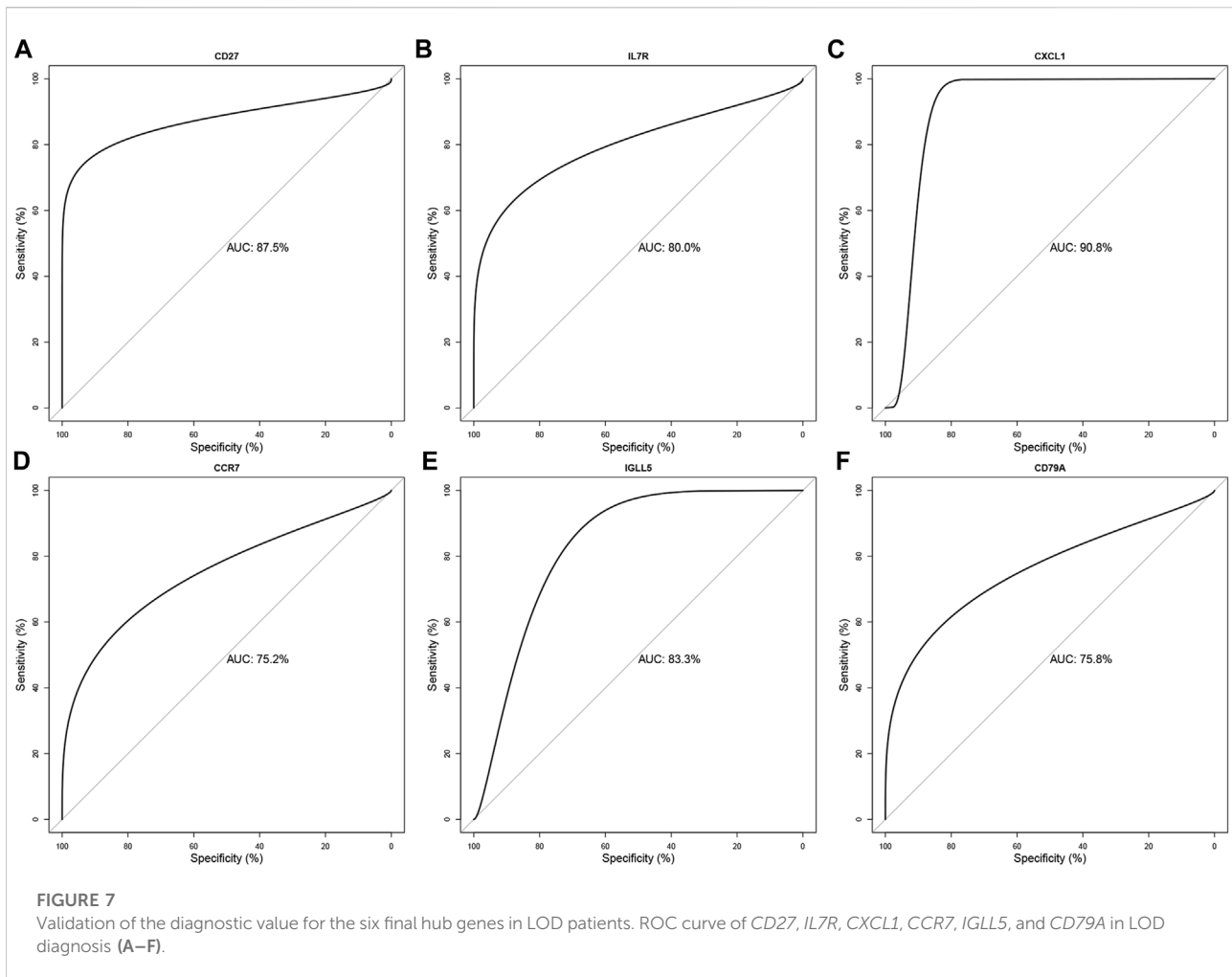
LASSO regression model and correlation between genes and immune cells

As shown in Figure 8, the error of the LASSO regression model is the smallest when $\log(\lambda)$ is between 2 and 4. Through the LASSO regression model, the obtained hub genes were *IGLL5*, *CD27*, *ITGB2*, and *CXCL1*. Next, we intersected the genes obtained from LASSO regression and the genes obtained from the ROC to obtain *IGLL5* and *CD27*, which had good diagnostic values for LOD. In addition, we found that the expressions of *CD27* and *IGLL5* were significantly

decreased in the blood of LOD patients. Furthermore, we found that *CD27* had a significant correlation with CD4 memory resting T cells, neutrophils, and resting mast cells.

Discussion

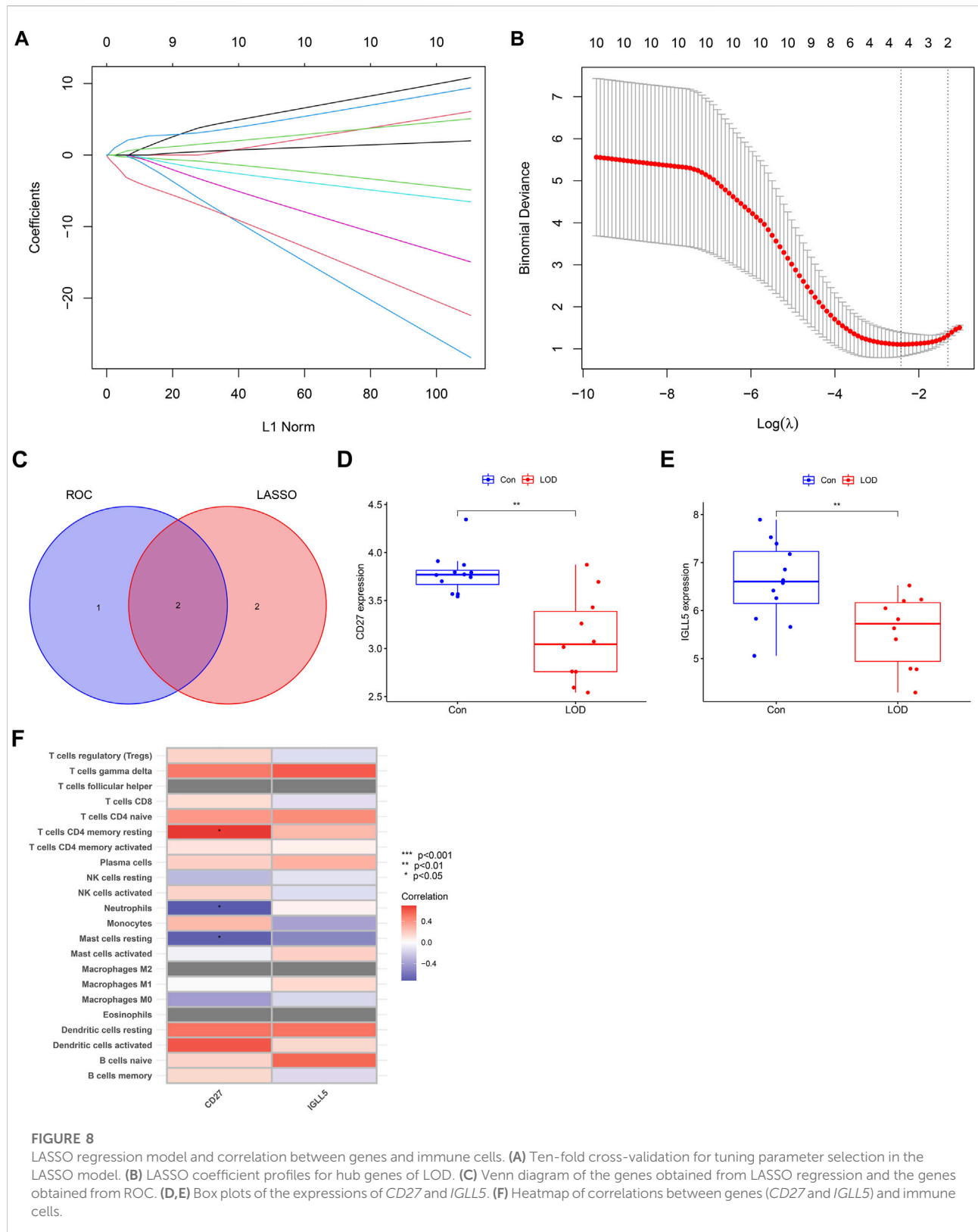
MDD is a highly prevalent psychiatric disorder (Sato and Yeh, 2013). MDD is the primary cause of disability, even suicide, which brings heavy pressure and burden to the family and society (Salomon et al., 2012). Growing evidence suggests the patients



who had the first onset of MDD at different ages show different depressive symptoms (Naismith et al., 2012; Park et al., 2014). This may explain the different responses to anti-depressants in MDD patients with different ages. In addition, the healthy risk of LOD patients increases greatly with age because of the limitations to diagnosis and the treatment level. Therefore, improving the early clinical diagnosis accuracy of LOD patients is particularly critical for the treatment and protection of patients. We used the WGCNA and CIBERSORT algorithm to identify hub genes and immune infiltrating cells in the plasma of LOD patients. Genes in the first two most clinically significant modules (blue and light cyan modules) to LOD were analyzed for feature and functional enrichment, respectively. Connectedness of PPI network neighborhoods identifies regulatory top 10 hub genes of the blue and light cyan modules. The six hub genes (*CD27*, *IL7R*, *CXCL1*, *CCR7*, *IGLL5*, and *CD79A*) were obtained by overlapping the top 10 hub genes in the two modules with DEGs. The functional enrichment results of six hub genes of LOD were mainly related to the immune function, indicating that the six hub genes may be crucial to the immune process of LOD.

Meanwhile, we confirmed the immune infiltrating cells to LOD through the gene expression profiles of the GSE98793 dataset. Finally, we obtained the final genes (*CD27* and *IGLL5*) through cross-validation of ROC and LASSO regression and found that *CD27* was significantly related to the immune correlation in LOD.

It has long been confirmed that inflammation plays an important role in the process of MDD (Berk et al., 2013). The decrease in the body immune function is a kind of physiological change that gradually appears with increasing age. Under the condition of decreased immune function, the inflammatory response in the body is likely to be one of the important promoters of the development of LOD, which may provide an insight into the underlying pathological mechanism of LOD. The immune response process in the body involves many processes, including the infiltration of immune cells and the release of various cytokines and chemokines. PPI networks can provide information on the strength of the relationship between different genes, being used to identify hub genes among multiple genes (Nangraj et al., 2020). In the present study, the results of the



functional enrichment analysis with the genes in the first two most clinically significant modules suggest the pathological mechanism of LOD involves the immune cell response and regulation of cytokines and chemokines. The CIBERSORT algorithm confirmed that immune cell infiltration occurred in the blood of LOD patients. In addition, the functional enrichment results of six hub genes (*CD27*, *IL7R*, *CXCL1*, *CCR7*, *IGLL5*, and *CD79A*) of LOD, which were obtained by intersecting the results of DEGs and the PPI network, were mainly related to the immune function. Taken together, these results indicate that the immune response may be crucial to the development of LOD, and six hub genes may be the key to these immune response mechanisms.

CD27 is a member of the TNF receptor superfamily, which is involved in regulating T cells, B cells, and immunoglobulins to mediate the inflammatory response. It has been reported that *CD27* may lead to anxiety or depression in patients with systemic lupus erythematosus by regulating T cells (Gu et al., 2021). As a receptor of interleukin-7 (IL-7), *IL7R* plays a key role in immune cell development and immune system homeostasis. Previous studies found the notable relevance of *IL7R* allele polymorphism with the onset age of multiple sclerosis (Salomon et al., 2012; Sato and Yeh, 2013). In addition, recent evidence suggests *IL7R*, which is associated with lymphoid, regulates the development of tissue-resident macrophages (Park et al., 2014). Recent studies have revealed that the *IL7R* content in the blood of the elderly is at a low level, which is related to the aging speed and health status of the elderly (Passtoors et al., 2015). In our study, we found the decrease of the *IL7R* content in LOD, which was in accordance with a previous study (Passtoors et al., 2015).

Chemokines and chemokine receptors play an important role in neurobiological changes in major depressive disorder (Miller and Raison, 2016). It has been demonstrated that the upregulation of the chemokine *CXCL1* in the hippocampus participated in depression development (Song et al., 2020). *CXCL1* overexpression activates the GSK3 β /BDNF/CREB signaling pathway to induce depression-like behavior (Chai et al., 2019). Interestingly, a decline in chemokine gene expression is associated with suicidal behavior among depressive patients. Compared with healthy subjects, patients with suicidal depression had significantly decreased mRNA expressions of chemokines in the prefrontal cortex, including *CXCL1* (Pandey et al., 2021). In addition, changed expression of the chemokine *CXCL1* was found easily in the plasma of MDD, especially in the elderly (Fanelli et al., 2019). As a homeostatic chemokine receptor, *CCR7* was induced by the inflammatory response, playing an important role in neuropsychiatric disorders (Wilson et al., 2009). The central expression of *CCR7* proved to help lymph nodes clear memory T cells in the brain to the peripheral nervous system (Charo and Ransohoff, 2006). It has been demonstrated that the neuroprotective effects of *CCR7* expression in astrocytes in

inflammation during infection (Gomez-Nicola et al., 2010) increased susceptibility to anxiety and depression in *CCR7*-knockout mice (Harrison et al., 2014). It seems that the decreased expressions of *CXCL1* and *CCR7* will increase the risk of depression. In this study, low levels of *CXCL1* and *CCR7* expression were observed in LOD patients, which are in accordance with the results of a previous study (Fanelli et al., 2019).

The *IGLL5* gene encodes one of the immunoglobulin lambda-like polypeptides that is involved in tumor progression. *IGLL5* was mainly found to frequently occur in the mutation in chronic lymphoblastic leukemia, lymphoma, and multiple myeloma (Kasar et al., 2015; Pérez-Carretero et al., 2020; Fan et al., 2020; Li et al., 2020; Panea et al., 2019; D'Agostino et al., 2020; White et al., 2018). In the cortex of patients with chronic traumatic encephalopathy, *IGLL5* could be a significant diagnosis factor. The *CD79A* gene encodes the Ig-alpha protein of the B-cell antigen component, which is typically associated with the B-cell receptor signaling complex. *CD79A* was often used as the marker of the B-cell-relevant tumor, such as Hodgkin's lymphoma (Sakatani et al., 2020). Frequent downregulation or deletion of *CD79A* could be potential clues for the diagnosis of plasma cell myeloma in the elderly (Tanaka et al., 2009). Previous studies have not reported the association of MDD with *IGLL5* and *CD79A*. In the present study, we confirmed *IGLL5* and *CD79A* in LOD with a low expression level for the first time, which may provide some clues for the future diagnosis and treatment of LOD.

The ROC is widely used to evaluate the diagnostic accuracy of hub genes for diseases based on the characteristics of comprehensive sensitivity and specificity (Zhao et al., 2021). LASSO regression is characterized by variable screening and complexity adjustment by constructing a penalty function while fitting the generalized linear model (Wang et al., 2019). The joint analysis of the ROC and LASSO can further accurately filter hub genes. In our study, we confirmed *CD27* and *IGLL5* as the accurate predictors of LOD through cross-validation of ROC and LASSO regression models. Meanwhile, the correlation analysis with immune cells showed that *CD27* and not *IGLL5* had significant correlation with immune infiltrating cells in LOD, indicating that *CD27* may be the target of immune function changes in LOD.

Conclusion

In summary, we explore the difference between healthy subjects and LOD patients by using the WGCNA and immune infiltration. We confirmed the six hub genes and the immune infiltrating cells in LOD. Among the six hub genes, the role of three genes (*CD27*, *CXCL1*, and *CCR7*) in depression has been reported. In addition, there are few

studies on IL7R, IGLL5, and CD79A in depression. ROC and LASSO helped us screen CD27 and IGLL5 as the accurate predictors of LOD. In addition, we found that CD27 may be the target of immune function changes in LOD. Our research shows that the changes in immune function may be an important mechanism in the development of LOD. However, there are several limitations to the present study. The detailed mechanism of the final hub genes found in LOD needs to be further studied. In addition, the dominant immune infiltrating cells in LOD need to be confirmed, which may interfere with the development of LOD.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Author contributions

P-AY: participated in the research conception and design, collected and analyzed the experimental data, and wrote and revised the manuscript. H-JS: collected and analyzed the experimental data and wrote and revised the manuscript. X-YL: participated in the research conception and design, collected and analyzed the experimental data, and wrote and revised the manuscript. The manuscript was reviewed and approved by the all authors.

References

- Alexopoulos, G. S., Young, R. C., Meyers, B. S., Abrams, R. C., and Shamoian, C. A. (1988). Late-onset depression. *Psychiatr. Clin. North Am.* 11, 101–115. doi:10.1016/s0193-953x(18)30519-7
- Berk, M., Williams, L. J., Jacka, F. N., O'Neil, A., Pasco, J. A., Moylan, S., et al. (2013). So depression is an inflammatory disease, but where does the inflammation come from? *BMC Med.* 11, 200. doi:10.1186/1741-7015-11-200
- Chai, H. H., Fu, X. C., Ma, L., Sun, H. T., Chen, G. Z., Song, M. Y., et al. (2019). The chemokine CXCL1 and its receptor CXCR2 contribute to chronic stress-induced depression in mice. *FASEB J.* 33, 8853–8864. doi:10.1096/fj.201802359RR
- Chan, K. L., Cathomas, F., and Russo, S. J. (2019). Central and peripheral inflammation link metabolic syndrome and major depressive disorder. *Physiol. (Bethesda)* 34, 123–133. doi:10.1152/physiol.00047.2018
- Charo, I. F., and Ransohoff, R. M. (2006). The many roles of chemokines and chemokine receptors in inflammation. *N. Engl. J. Med.* 354, 610–621. doi:10.1056/NEJMra052723
- D'Agostino, M., Zaccaria, G. M., Ziccheddu, B., Rustad, E. H., Genuardi, E., Capra, A., et al. (2020). Early relapse risk in patients with newly diagnosed multiple myeloma characterized by next-generation sequencing. *Clin. Cancer Res.* 26, 4832–4841. doi:10.1158/1078-0432.CCR-20-0951
- Das-Munshi, J., Goldberg, D., Bebbington, P. E., Bhugra, D. K., Brugha, T. S., Dewey, M. E., et al. (2008). Public health significance of mixed anxiety and depression: Beyond current classification. *Br. J. Psychiatry* 192, 171–177. doi:10.1192/bjp.bp.107.036707
- Dowlati, Y., Herrmann, N., Swardfager, W., Liu, H., Sham, L., Reim, E. K., et al. (2010). A meta-analysis of cytokines in major depression. *Biol. Psychiatry* 67, 446–457. doi:10.1016/j.biopsych.2009.09.033
- Fan, Z., Pei, R., Sha, K., Chen, L., Wang, T., and Lu, Y. (2020). Comprehensive characterization of driver genes in diffuse large B cell lymphoma. *Oncol. Lett.* 20, 382–390. doi:10.3892/ol.2020.11552
- Fanelli, G., Benedetti, F., Wang, S. M., Lee, S. J., Jun, T. Y., Masand, P. S., et al. (2019). Reduced CXCL1/GRO chemokine plasma levels are a possible biomarker of elderly depression. *J. Affect. Disord.* 249, 410–417. doi:10.1016/j.jad.2019.02.042
- Gomez-Nicola, D., Pallas-Bazarra, N., Valle-Argos, B., and Nieto-Sampedro, M. (2010). CCR7 is expressed in astrocytes and upregulated after an inflammatory injury. *J. Neuroimmunol.* 227, 87–92. doi:10.1016/j.jneuroim.2010.06.018
- Gu, X. X., Jin, Y., Fu, T., Zhang, X. M., Li, T., Yang, Y., et al. (2021). Relevant characteristics analysis using natural language processing and machine learning based on phenotypes and T-cell subsets in systemic lupus erythematosus patients with anxiety. *Front. Psychiatry* 12, 793505. doi:10.3389/fpsy.2021.793505
- Harrison, E. L., Jaehne, E. J., Jawahar, M. C., Corrigan, F., and Baune, B. T. (2014). Maternal separation modifies behavioural and neuroendocrine responses to stress in CCR7 deficient mice. *Behav. Brain Res.* 263, 169–175. doi:10.1016/j.bbr.2014.01.036
- Kasar, S., Kim, J., Improgio, R., Tiao, G., Polak, P., Haradhvala, N., et al. (2015). Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. *Nat. Commun.* 6, 8866. doi:10.1038/ncomms9866

Funding

This work was financially supported by Zhejiang Provincial Medical and Health Science and Technology Plan (No. 2023KY883), Zhejiang TCM Science and Technology Program (No. 2023ZR030) and China Postdoctoral Science Foundation (No. 2022M722853).

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1048761/full#supplementary-material>

- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., et al. (2003). The epidemiology of major depressive disorder: Results from the national comorbidity survey replication (NCS-R). *Jama* 289, 3095–3105. doi:10.1001/jama.289.23.3095
- Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559
- Li, M., Su, X., Wang, Y., Fan, L., Chai, J., Li, P., et al. (2020). Lineage-negative lymphoma with a helper innate lymphoid cell phenotype. *Virchows Arch.* 476, 285–293. doi:10.1007/s00428-019-02658-x
- Miller, A. H., and Raison, C. L. (2016). The role of inflammation in depression: From evolutionary imperative to modern treatment target. *Nat. Rev. Immunol.* 16, 22–34. doi:10.1038/nri.2015.5
- Mishra, D., Sardesai, U., and Razdan, R. (2018). C-Reactive protein level in late-onset depression: A case-control study. *Indian J. Psychiatry* 60, 467–471. doi:10.4103/psychiatry.IndianJPsychiatry_127_17
- Miyata, S., Kurachi, M., Okano, Y., Sakurai, N., Kobayashi, A., Harada, K., et al. (2016). Blood transcriptomic markers in patients with late-onset major depressive disorder. *PLoS One* 11, e0150262. doi:10.1371/journal.pone.0150262
- Naismith, S. L., Norrie, L. M., Mowszowski, L., and Hickie, I. B. (2012). The neurobiology of depression in later-life: Clinical, neuropsychological, neuroimaging and pathophysiological features. *Prog. Neurobiol.* 98, 99–143. doi:10.1016/j.pneurobio.2012.05.009
- Nangraj, A. S., Selvaraj, G., Kalamurthi, S., Kaushik, A. C., Cho, W. C., and Wei, D. Q. (2020). Integrated PPI- and WGCNA-retrieval of hub gene signatures shared between Barrett's esophagus and esophageal adenocarcinoma. *Front. Pharmacol.* 11, 881. doi:10.3389/fphar.2020.00881
- Pandey, G. N., Rizavi, H. S., Bhaumik, R., and Zhang, H. (2021). Chemokines gene expression in the prefrontal cortex of depressed suicide victims and normal control subjects. *Brain Behav. Immun.* 94, 266–273. doi:10.1016/j.bbi.2021.01.033
- Panea, R. I., Love, C. L., Shingleton, J. R., Reddy, A., Bailey, J. A., Moormann, A. M., et al. (2019). The whole-genome landscape of Burkitt lymphoma subtypes. *Blood* 134, 1598–1607. doi:10.1182/blood.2019001880
- Park, S. C., Hahn, S. W., Hwang, T. Y., Kim, J. M., Jun, T. Y., Lee, M. S., et al. (2014). Does age at onset of first major depressive episode indicate the subtype of major depressive disorder?: The clinical research center for depression study. *Yonsei Med. J.* 55, 1712–1720. doi:10.3349/ymj.2014.55.6.1712
- Passtoors, W. M., Van Den Akker, E. B., Deelen, J., Maier, A. B., Van Der Breggen, R., Jansen, R., et al. (2015). IL7R gene expression network associates with human healthy ageing. *Immun. Ageing* 12, 21. doi:10.1186/s12979-015-0048-6
- Pérez-Carretero, C., Hernández-Sánchez, M., González, T., Quijada-Álamo, M., Martín-Izquierdo, M., Hernández-Sánchez, J. M., et al. (2020). Chronic lymphocytic leukemia patients with IGH translocations are characterized by a distinct genetic landscape with prognostic implications. *Int. J. Cancer* 147, 2780–2792. doi:10.1002/ijc.33235
- Ren, C., Li, M., Du, W., Lü, J., Zheng, Y., Xu, H., et al. (2020). Comprehensive bioinformatics analysis reveals hub genes and inflammation state of rheumatoid arthritis. *Biomed. Res. Int.* 2020, 6943103. doi:10.1155/2020/6943103
- Sakatani, A., Igawa, T., Okatani, T., Fujihara, M., Asaoku, H., Sato, Y., et al. (2020). Clinicopathological significance of CD79a expression in classic Hodgkin lymphoma. *J. Clin. Exp. Hematop.* 60, 78–86. doi:10.3960/jslrt.20010
- Salomon, J. A., Vos, T., Hogan, D. R., Gagnon, M., Naghavi, M., Mokdad, A., et al. (2012). Common values in assessing health outcomes from disease and injury: Disability weights measurement study for the global burden of disease study 2010. *Lancet* 380, 2129–2143. doi:10.1016/S0140-6736(12)61680-8
- Sato, S., and Yeh, T. L. (2013). Challenges in treating patients with major depressive disorder: The impact of biological and social factors. *CNS Drugs* 27, S5–S10. doi:10.1007/s40263-012-0028-8
- Schleifer, S. J., Keller, S. E., Bond, R. N., Cohen, J., and Stein, M. (1989). Major depressive disorder and immunity. Role of age, sex, severity, and hospitalization. *Arch. Gen. Psychiatry* 46, 81–87. doi:10.1001/archpsyc.1989.01810010083011
- Shinko, Y., Otsuka, I., Okazaki, S., Horai, T., Boku, S., Takahashi, M., et al. (2020). Chemokine alterations in the postmortem brains of suicide completers. *J. Psychiatr. Res.* 120, 29–33. doi:10.1016/j.jpsychires.2019.10.008
- Song, A. Q., Gao, B., Fan, J. J., Zhu, Y. J., Zhou, J., Wang, Y. L., et al. (2020). NLRP1 inflammasome contributes to chronic stress-induced depressive-like behaviors in mice. *J. Neuroinflammation* 17, 178. doi:10.1186/s12974-020-01848-8
- Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguéz, P., et al. (2011). The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* 39, D561–D568. doi:10.1093/nar/gkq973
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–d613. doi:10.1093/nar/gky1131
- Tanaka, T., Ichimura, K., Sato, Y., Takata, K., Morito, T., Tamura, M., et al. (2009). Frequent downregulation or loss of CD79a expression in plasma cell myelomas: Potential clue for diagnosis. *Pathol. Int.* 59, 804–808. doi:10.1111/j.1440-1827.2009.02448.x
- Vogelzangs, N., Duivis, H. E., Beekman, A. T., Kluff, C., Neuteboom, J., Hoogendijk, W., et al. (2012). Association of depressive disorders, depression characteristics and antidepressant medication with inflammation. *Transl. Psychiatry* 2, e79. doi:10.1038/tp.2012.8
- Wagner, S., Wollschläger, D., Dreimüller, N., Engelmann, J., Herzog, D. P., Roll, S. C., et al. (2020). Effects of age on depressive symptomatology and response to antidepressant treatment in patients with major depressive disorder aged 18 to 65 years. *Compr. Psychiatry* 99, 152170. doi:10.1016/j.comppsy.2020.152170
- Wang, Z., Zhu, C., Nambi, V., Morrison, A. C., Folsom, A. R., Ballantyne, C. M., et al. (2019). Metabolomic pattern predicts incident coronary heart disease. *Arterioscler. Thromb. Vasc. Biol.* 39, 1475–1482. doi:10.1161/ATVBAHA.118.312236
- White, B. S., Lanc, I., O'Neal, J., Gupta, H., Fulton, R. S., Schmidt, H., et al. (2018). A multiple myeloma-specific capture sequencing platform discovers novel translocations and frequent, risk-associated point mutations in IGLL5. *Blood Cancer J.* 8, 35. doi:10.1038/s41408-018-0062-y
- Wilson, E. H., Harris, T. H., Mrass, P., John, B., Tait, E. D., Wu, G. F., et al. (2009). Behavior of parasite-specific effector CD8+ T cells in the brain and visualization of a kinesis-associated system of reticular fibers. *Immunity* 30, 300–311. doi:10.1016/j.immuni.2008.12.013
- Xu, H., Ge, Y., Liu, Y., Zheng, Y., Hu, R., Ren, C., et al. (2022). Identification of the key genes and immune infiltrating cells determined by sex differences in ischaemic stroke through co-expression network module. *IET Syst. Biol.* 16, 28–41. doi:10.1049/syb2.12039
- Zhao, X., Zhang, L., Wang, J., Zhang, M., Song, Z., Ni, B., et al. (2021). Identification of key biomarkers and immune infiltration in systemic lupus erythematosus by integrated bioinformatics analysis. *J. Transl. Med.* 19, 35. doi:10.1186/s12967-020-02698-x
- Zhou, S., Lu, H., and Xiong, M. (2021). Identifying immune cell infiltration and effective diagnostic biomarkers in rheumatoid arthritis by bioinformatics analysis. *Front. Immunol.* 12, 726747. doi:10.3389/fimmu.2021.726747