ORIGINAL ARTICLE
Iran J Allergy Asthma Immunol
In press.

Exploring Inflammatory-Related Hub Genes as Therapeutic Targets in Major Depressive Disorder: Implications for Immunological Pathways

Ruiqi Wang[1,2], Hongyuan Fan[1,3], Yu Feng[1,4], Haibo Li[1,3], and Jin Cui[1,2]

[1] Guizhou University of Traditional Chinese Medicine, Guiyang, China
[2] Department of Acupuncture and Moxibustion, The First Affiliated Hospital of Guizhou University of Traditional Chinese Medicine, Guiyang, China
[3] Department of Tuina, The First Affiliated Hospital of Guizhou University of Traditional Chinese Medicine, Guiyang, China
[4] Department of Tuina, The Second Affiliated Hospital of Guizhou University of Traditional Chinese Medicine, Guiyang, China

Received: 9 November 2024; Received in revised form: 14 December 2024; Accepted: 30 December 2024

ABSTRACT

This study explored the mechanisms of action of inflammation related central genes in severe depression (MDD) and analyzes their potential as therapeutic targets. By identifying key genes and establishing the link between immune regulatory mechanisms and depression, we provide a theoretical basis for developing more accurate diagnostic and treatment methods.

Gene expression datasets related to MDD were obtained from the Gene Expression Omnibus (GEO). Differentially expressed genes (DEGs) associated with inflammatory processes were identified and analyzed through Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses. Protein-protein interaction (PPI) networks were constructed to identify hub genes. Additionally, we explored regulatory networks of miRNAs, transcription factors, and potential drug interactions were explored. Immune infiltration analysis was performed to examine immune cell profiles.

Seven key genes—HMGB1, HSP90AB1, MAPK1, MMP9, MYD88, S100A12, and TLR2—were identified as central players in the inflammatory pathways underlying MDD. These genes demonstrated moderate diagnostic accuracy with AUC values ranging from 0.5 to 0.7. Enrichment analyses revealed significant associations with immune signaling pathways, including IL-17 and Toll-like receptor signaling. Immune infiltration analysis highlighted altered abundances of regulatory T cells, neutrophils, and dendritic cells in MDD samples.

Inflammatory-related hub genes play crucial roles in linking immune dysregulation to the pathophysiology MDD pathophysiology. These findings offer insights into the immunological underpinnings of MDD and present potential therapeutic targets for intervention through immune-modulatory approaches.

Keywords: Diagnostic biomarkers; Immune system; Inflammation; Major depressive disorder; Terapeutic targets

Corresponding Author: Jin Cui, MD;
Guizhou University of Traditional Chinese Medicine, Guiyang, China.
Tel: (+98 86) 13985 026278, Fax: (+98 86) 13985 026278 Email: cuijin6311@163.com

Published by Tehran University of Medical Sciences (http://ijaai.tums.ac.ir)

INTRODUCTION

Major depression is a serious mental health disorder. Its high incidence rate and recurrence rate make it a major challenge in the field of global public health. Severe depression not only significantly affects the psychological state of patients but is also closely related to various physical diseases, increasing the overall mortality rate of patients. Due to the long-term effects of depression of patients, which is often accompanied by social dysfunction, increased risk of suicide and significant burden on families and society, it is particularly important to initiate effective clinical treatment for severe depression as soon as possible. The prevalence of depression worldwide is as high as 5–10%. The physical well-being of the patient is closely linked to depression, as it often results in a reduction in immune system performance. The patient's body is progressively going into a state of hyperactivity due to ongoing mental stress and sadness, which adversely affects the immune system and raises the risk of infection and illness. Research indicates that those experiencing significant depression have an increased risk of developing long-term conditions including diabetes mellitus and cardiovascular disease, underscoring the serious danger that depression poses to one's physical well-being. Furthermore, there exists a tight connection between major depressive disorder (MDD) and the physiological processes of patients, whereby the nervous system and hormone levels are adversely affected. Patients may have irregular sleep patterns, changes in appetite, and low energy, all of which may have an impact on energy balance and weight control. These physiological anomalies not only contribute to the patient's total physical load but also complicate the course of therapy. The treatment of MDD has a number of shortcomings and difficulties. First, the primary medication treatment may not always be effective and may have major adverse effects. For instance, long-term usage of antidepressants may result in issues like sleeplessness and sexual dysfunction, which restricts the medication's broad use. Second, although psychotherapy has been somewhat beneficial for certain patients, not all individuals can benefit from it, and it does require a significant time and energy commitment. Furthermore, patients often need long-term or even lifetime therapy, which causes significant difficulties and inconveniences in their personal and professional lives. This is because the present strategies of treatment have not been able to adequately address the issue of the disease's recurrence. Furthermore, the absence of effective treatment options is reflected in the unclearknowledge of the root cause. The precise etiology of MDD is unknown due to the great complexity of its pathophysiology, which includes genetic, physiological, psychological, and environmental factors. Because we do not fully understand the origin of the disorder and it is challenging to identify a threatment that can completely fix the issue, leading to the limits of the therapies that are now available. The connection between inflammatory factors and serious depression has garnered a lot of attention recently as medical research has deepened, and the findings indicate a tight interaction between the two. There is evidence from both scientific studies and clinical experience that the onset and progression of MDD are strongly correlated with aberrant levels of inflammatory markers. By influencing neurotransmitters, neurocircuits, neuroplasticity, and neuroendocrine systems, inflammatory reactions are linked to structural and functional alterations in the brain. These alterations subsequently impact behavioral performance, emotional control, and cognitive function. To be more precise, prior research has shown links between serious depression and many important inflammatory variables. Significant correlations have beenobserved between the frequency and severity of MDD and aberrant levels of inflammatory markers, including interleukin-6 (IL-6), tumor necrosis factor-α (TNF-α), and C-reactive protein (CRP). Research has shown that the levels of inflammatory markers such as IL-6, CRP, and TNF - α are significantly elevated in MDD patients. These factors exacerbate depressive symptoms by affecting neurotransmitter metabolism, neural circuit function, and neuroplasticity.[1] In addition, chronic inflammation may lead to changes in brain structure and function, especially in areas such as the amygdala, prefrontal cortex, and hippocampus, which are closely related to emotional regulation.[2] Research indicates that anti-inflammatory treatment approaches might be useful in reducing MDD symptoms. For instance, it has been shown that some medications that target particular inflammatory variables or therapies that target inflammatory pathways can help patients with their depressive symptoms. As a result, a thorough investigation of the connection between inflammatory variables and severe depression will contribute to our growing understanding of the pathophysiology of this

illness and offer fresh concepts and methods for potential future interventions. By building protein interaction networks and using differential expression analysis to identify important genes, this study examined the connection between MDD and inflammatory variables. It also looked at the significance of these genes within regulatory networks. In order to identify a possible connection between depression and these genes, we conducted a thorough analysis of their interactions with miRNA, transcription factors (TFs), and medications. Finally, we used the ssGSEA algorithm to uncover a possible connection between depression and the immune system, which provided us with more insight into the roles of the chosen Hub genes in immune regulation. Through this study, we unveil the possible roles of these genes in immune regulation. It is anticipated that these discoveries will provide fresh viewpoints on the pathological mechanisms of depression and serve as a helpful guide for the developments in of relevant medications and therapies.

MATERIALS AND METHODS

Data Download

The datasets od GSE 98793 (Major Depression Disorder, MDD)4 and GSE 197384 were supplied by the GEOquery R package[3] and the GEO database.[4] As shown in Table 1, the samples for the datasets GSE98793 and GSE19738 are derived from humans, and the tissue sources used were whole blood. To ensure comparability between different datasets, batch effect removal was performed on the data matrices of GSE98793 and GSE19738 using the SVA package. After batch effect removal, the data standardization effect was verified through distribution box plots and principal component analysis (PCA) plots. We standardized and integrated the two datasets to generate a comprehensive dataset containing 97 MDD samples and 98 healthy control samples for subsequent analysis. The chip platforms for GSE98793 and GSE19738 are GPL570 and GPL6848, respectively. Of these, 64 MDD samples and 64 controls (Control) samples are found in dataset GSE 98793; 33 MDD samples and 34 controls (Control) samples are found in dataset GSE 19738. This research includes samples of MDD and control (Control). Ethical approval was waived because this article does not contain any studies with human or animal subjects performed by any of the authors.

Inflammatory factor-related genes, also known as IFRGs, were gathered from the GeneCards database.[5] Comprehensive data on human genes may be found in this database. A total of 350 inflammatory factor-related genes (IFRGs) were found after restricting the search to protein coding inflammatory factor-related genes (IFRGs) and using the term "Inflammatory Factor" as the keyword. The sva package in R was applied.[6] The integrated GEO dataset (Combined Datasets) was produced from the datasets GSE 98793 and GSE19738. Ultimately, using the R program limma, 97 MDD samples and 98 control (Control) samples are included in the combined GEO data set (Combined Datasets).[7] Both the annotated probes and the combined GEO dataset (Combined Datasets) underwent standardization and normalization.

Differentially Expressed Genes Associated to Major Factors Associated with Major Disorder

The MDD group and the control (Control) groups were created using the samples in the combined GEO data set (Combined Datasets). Using the R program limma, gene expression in the MDD and control (Control) groups was differentially analyzed to identify differences between the two sample groups. Specifically, the R package limma is used to perform differential expression analysis on the comprehensive dataset. The screening criteria for differentially expressed genes were set as $|\log FC|>0$ and p-value<0.05, and Benjamin-Hochberg correction was used to control for false positive rates. By cross comparing the differentially expressed genes (DEGs) selected from the screening with the inflammation related genes (IFRGs) included in the GeneCards database, the inflammation related differentially expressed genes (IFRDEGs) were finally screened, and heatmaps and volcano maps were generated.

Table 1. GEO Microarray Chip Information

	GSE98793	GSE 19738
Platform	GPL570	GPL 6848
Species	Homo sapiens	Homo sapiens
Tissue	Whole Blood	Whole Blood
Samples in MDD group	64	33
Samples in Control group	64	34
Reference	PMID: 28688579	PMID: 20471630

GEO: Gene Expression Omnibus; MDD: Major Depressive Disorder

Gene Ontology (GO) and Pathways (KEGG) Enrichment Analysis

The Gene Ontology (also known as GO)[8] and large-scale functional enrichment studies on biological processes (Biological Process, BP), cellular components (Cell Component, CC), and molecular functions (Molecular Function, MF) are often carried out using this technique. The Kyoto Gene and Genome Encyclopedia (Kyoto Gene and Genome Encyclopedia, KEGG)[9] is a popular database that has data on medications, illnesses, biological processes, and genomes. The clusterProfiler, a R package was utilized.[10] The differentially expressed genes IFRDEGs underwent enrichment analysis using GO and pathways (KEGG), with p value serving as the entrance screening criterion. GO analysis covers three types of functions: biological processes (BP), cellular components (CC), and molecular functions (MF), while KEGG analysis is used to explore gene related biological pathways. The screening criteria are a p-value (adj. p<0.05) and an FDR value<0.25 after correction. The analysis results showed that IFRDEGs were significantly enriched in the IL-17 signaling pathway, NOD like receptor signaling pathway, and other inflammation related pathways.

Gene-set Enrichment Analysis (GSEA)

The GSEA was employed in this study to examine the variations in biological processes across groups within the GEO data collection. The GSEA is particularly appropriate for this study because it allows for the evaluation of gene expression changes at the pathway level, which provides a broader understanding of the biological processes underlying MDD. This approach is crucial in understanding the broader biological processes and pathways that may be involved in MDD, as it identifies whether entire sets of genes associated with specific pathways are significantly different between MDD and control groups. Given the complex and multifactorial nature of MDD, analyzing individual genes may not capture the full scope of molecular interactions involved in the disorder. GSEA helps identify whether entire sets of genes, associated with specific pathways, are significantly different between MDD and control groups. By focusing on gene sets rather than individual genes, GSEA offers a more comprehensive view of the molecular mechanisms that could be driving MDD, thus enhancing the study's ability to uncover potential therapeutic targets. The GEO data collection (Combined Datasets) is based on integrated gene enrichment analysis (Gene collection Enrichment Analysis, GSEA) to examine the variations in biological processes across groups.[11]

Using the R program clusterProfiler, gene set enrichment analysis (GSEA) was first carried out on each gene in the data set. The GSEA uses the following parameters: 2020 is the seed, each gene set has a minimum of 10 genes and a maximum of 500 genes. Furthermore, via the MSigDB (Molecular Signatures Database) database.[12] Get the c2.cp.all.v2022.1.Hs.symbols gene set here. For the GMT, gene set enrichment analysis (GSEA) was carried out, and p was the screening criterion for substantial enrichment in GSEA. Benjamini-Hochberg (BH) was used as the p value correction procedure when the adj<0.05 and the FDR value (q value)<0.25.

Protein-protein Interaction (PPI) Network and Hub Gene Screening

Proteins interacting with one another make up the Protein-Protein Interaction Network (PPI). Understanding the functional relationships, underlying ideas, and reaction processes between proteins is possible thanks to the protein-protein interaction network (PPI). The STRING database is a resource[13] that looks for protein interactions between predicted and known proteins. Using the STRING database for differentially expressed genes (IFRDEGs) with a minimum needed interaction score of 0.4 (medium confidence),, a protein-protein interaction network (PPI Network) was built. Cytoscape was used after the protein-protein interaction network (PPI Network) was constructed.[14] The program utilized the cytoHubba plugin, which showed the protein-protein interaction network (PPI Network).[15] The plug-in contains five algorithms: Edge Percolated Component (EPC), Degree, Maximum Neighborhood Component (MNC), Maximal Clique Centrality (MCC), and Closeness.[16] The differentially expressed genes (IFRDEGs) were scored one at a time. The top 10 IFRDEGs were then determined based on their scores and displayed using the protein-protein interaction network (PPI Network), respectively. The last step involves creating a Wayn diagram by comparing the genes produced by the five distinct algorithms. The hub genes associated with inflammatory factors are the intersection genes.

Construction of Control Network

miRNA is a key regulator of mRNA, controlling processes including translation inhibition and mRNA degradation. Different target genes may be controlled by miRNA, and different miRNAs can also regulate the same target gene. The mRNA-miRNA regulatory network (mRNA-miRNA Regulatory Network) is visualized using Cytoscape software. The miRNAs related to inflammatory factor-related hub genes (hub genes) were obtained through the TarBase database.[17] With "H. sapiens", "Official GeneSymbol", and "Gene-miRNA Interactions," the TarBase v8.0 data was searchedThrough interactions with target genes (mRNA) at the post-transcriptional stage, transcription factors (TFs) regulate the expression of genes using hTFtarget[18] and the ChIPBase database.[19] The union of all transcription factors associated with hub genes related to inflammatory factors was identified by searching the database. The mRNA-TF regulatory network (mRNA-TF Regulatory Network) was then visualized using Cytoscape software.

In order to complete the network construction, the inflammatory factor-related hub genes (hub direct and indirect drug targets of genes) will be examined, along with drugs.[20] Chemical drugs that have a Reference Count of>5 will be screened, and the mRNA-Drug Regulatory Network will be visualized using Cytoscape software.

Expression differences and ROC Analysis of Hub Genes

In the combined GEO data set (Combined Datasets), the inflammatory factor-related hub genes (hub genes) were identified based on the gene expression level group comparison chart in order to determine the expression difference of the hub genes in the MDD group and the control group (Control). Furthermore, the R package pROC was utilized to create the diagnostic ROC curve of inflammatory factor-related hub genes (hub genes) and to determine the AUC in order to assess the diagnostic accuracy of the expression of hub genes related to inflammatory factors for MDD. The ROC curve's AUC typically ranges from 0.5 to 1. AUC>0.5 indicates that the molecule's expression tends to encourage the occurrence of events. The higher the diagnostic impact, the closer the AUC is to 1. AUC values in the range of 0.5 to 0.7 indicate lesser accuracy, 0.7 to 0.9 indicate definite accuracy, and 0.9 or more indicates better accuracy.

Immune Infiltration Analysis

One technique for measuring the relative abundance of each immune cell infiltration is single-sample gene set enrichment analysis, or ssGSEA (Single-Sample Gene-Set Enrichment Analysis).[21] First, label all immune cell types that are penetrating the body, including gamma delta T cells, natural killer cells, regulatory T cells, activated CD8 T cells, activated dendritic cells, and other human immune cell subtypes. The relative abundance of each immune cell infiltration in each sample was then represented by the enrichment scores in the MDD and control (Control) groups using ssGSEA. The immune cell infiltration matrix was then derived by filtering the to include samples where the infiltration abundance was greater than 0 and p-value<0.05. A correlation heatmap is created using the R package pheatmap to show immune cells and visualize the difference in immune cell infiltration abundance

Published by Tehran University of Medical Sciences (http://ijaai.tums.ac.ir)

between the MDD group and the control (Control) group of the combined GEO data set (Combined Datasets) using violin plots. The findings of the correlation study will explore the relationships between immune cells and inflammatory factors, as well as between inflammatory factor-related hub genes (hub genes) and inflammatory factors. If the correlation coefficient (r value) is less than 0.3, it is weak or unimportant; if it is between 0.3 and 0.5, it is moderate; if it is between 0.5 and 0.8, it is strong; and if it is over 0.8, it is considered very strong.

Statistical Analysis

This article's entire data processing and analysis were built using R program (Version 4.3.0). For group comparisons, the Mann-Whitney U test, also known as the Wilcoxon Rank Sum Test, was used since the data did not meet normality requirements and is a non-parametric test. If not otherwise noted, the correlation coefficient between several molecules is determined using Pearson correlation analysis, with a p value of less than 0.05 serving as the threshold for findings that vary substantially. Pearson correlation was employed for its ability to measure the strength and direction of the linear relationship between two continuous variables, assuming a normal distribution of the data. This method is particularly effective when the relationship between the variables is expected to be linear, as it provides a straightforward and widely recognized metric for such associations. While other correlation methods, such as Spearman's rank correlation, could be used when data do not meet normality assumptions, the selection of Pearson correlation in this study is justified by the underlying linearity and distribution characteristics of the data being analyzed.

RESULTS

Merging of Data Sets

Technology Roadmap was shown in Figure 1. First, use the R package SVA to remove batch effects from the MDD data sets GSE98793 and GSE19738 to obtain integrated GEO data sets (Combined Datasets), and use distribution box plots and PCA (Principal Component Analysis) plots compare the data sets before and after batch effect removal (Figure 2A-D). The results of distribution boxplots and PCA plots show that the batch effect of samples in the MDD data set is substantially eliminated after batch effect removal.

Differentially Expressed Genes Related to Inflammatory Factors related to Major Depression

In the integrated GEO data set (Combined Datasets), the samples were divided into the MDD group and the control (Control) group. Differential analysis was performed using the R package limma to obtain the differentially expressed genes of the two sets of data. The analysis results showed that there was a total of 880 differentially expressed genes (DEGs) that met the threshold of logFC>0.00 and p value<0.05, and a volcano plot was drawn based on the results (Figure 3A). In order to obtain differentially expressed genes related to MDD and IFRGs, all DEGs obtained by differential analysis were intersected with IFRGs to obtain 30 inflammatory factor-related genes. Differentially expressed genes (IFRDEGs) were identified and a Venn diagram was generatd (Figure 3B). Finally, a heat map was created to show the gene expression of 30 IFRDEGs in the MDD group and the control (Control) group.

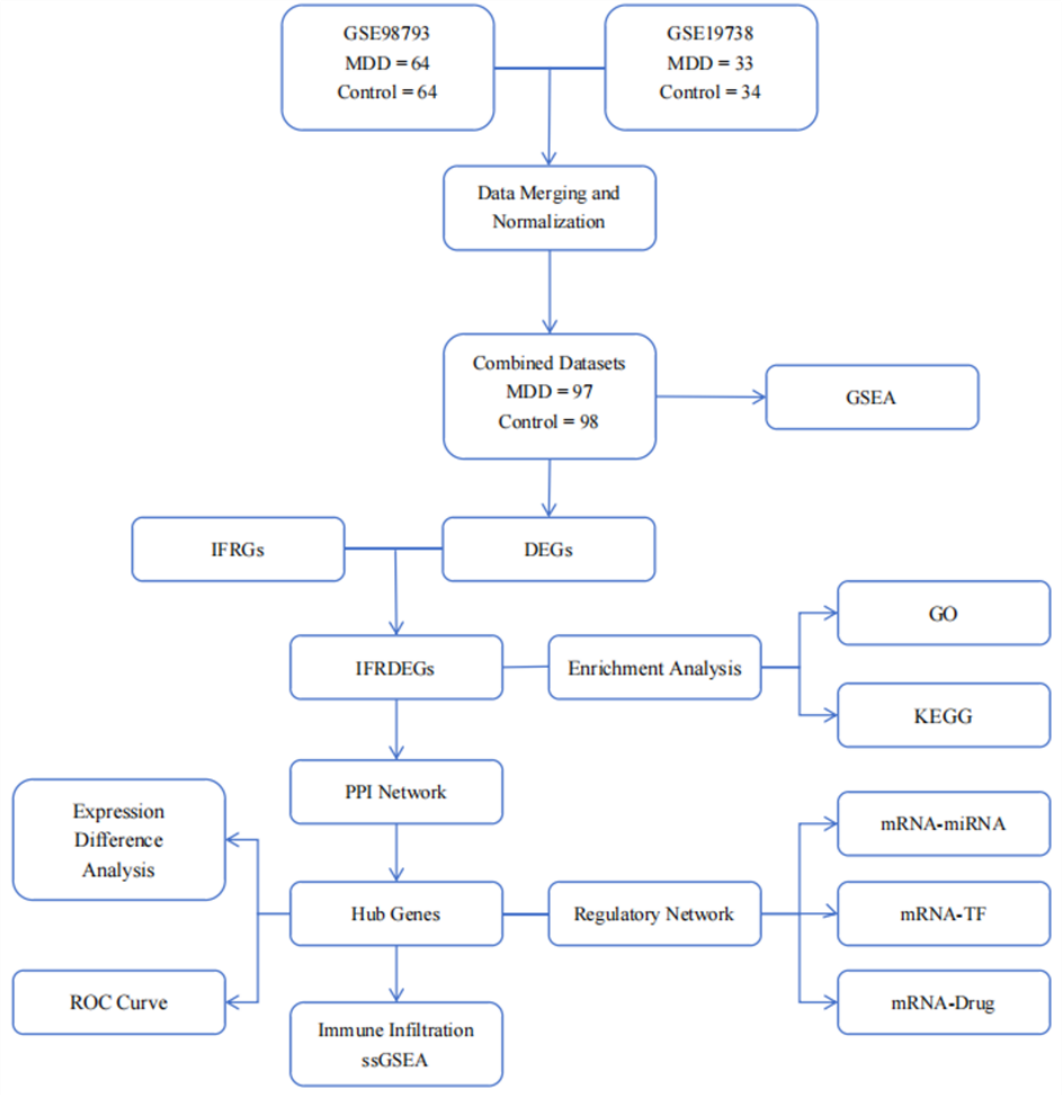


Figure 1. Flow Chart for the Comprehensive Analysis of IFRDEGs.
MDD: Major Depression Disorder; DEGs: Differentially Expressed Genes; IFRGs: Inflammatory Factor-Related Genes; IFRDEGs: Inflammatory Factor-Related Differentially Expressed Genes; GSEA: Gene Set Enrichment Analysis; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI: Protein-protein Interaction; ssGSEA: single-sample Gene-Set Enrichment Analysis; ROC Curve: Receiver Operating Characteristic Curve; TF: Transcription Factor.

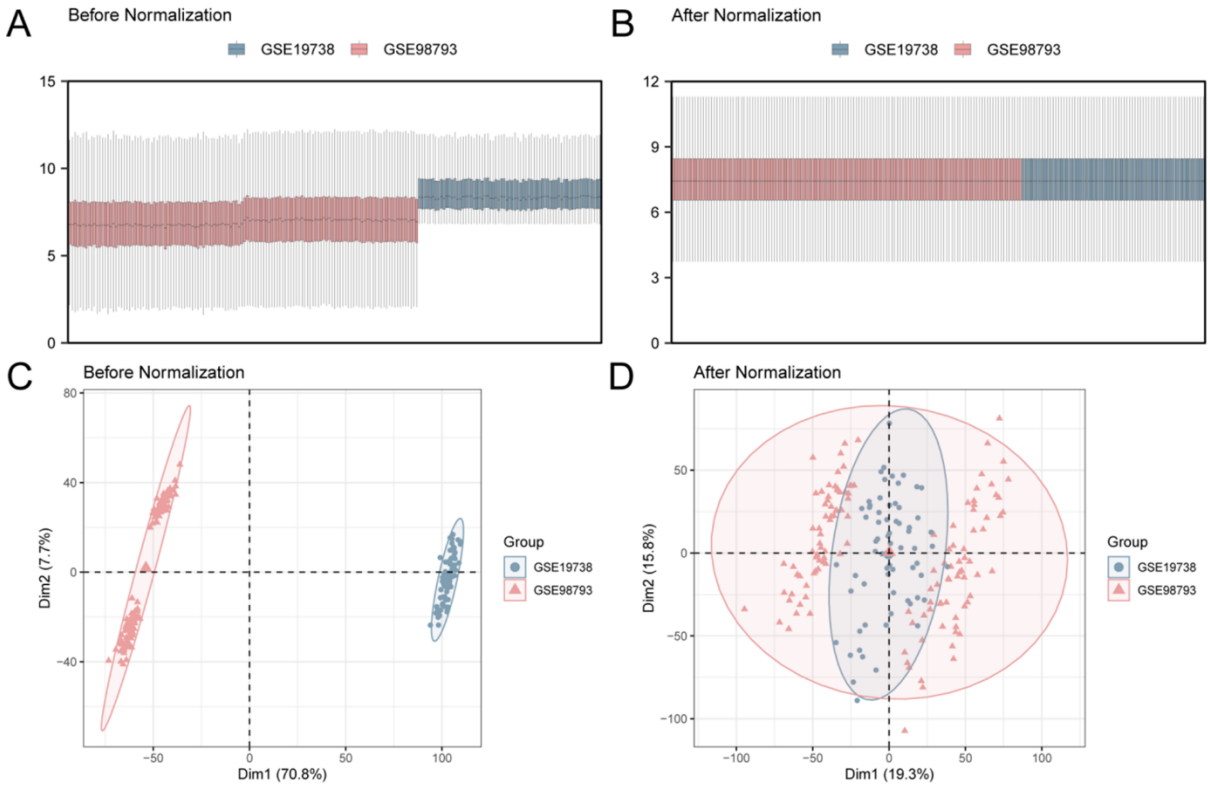
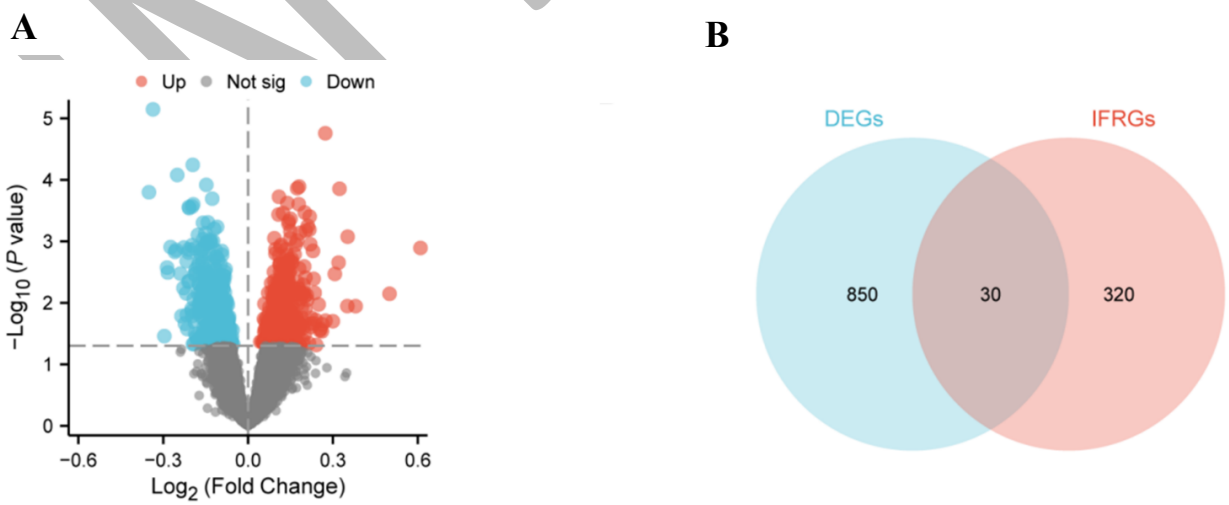


Figure 2. Batch Effects Removal of GSE 98793 and GSE 19738; (A) Box plot of the distribution of the data set before de-batch processing. (B) Distribution box plot of the integrated GEO data set (Combined Datasets) after de-batch processing. (C) PCA plot of the two GEO data sets before batch processing. (D) PCA plot of the integrated GEO data set (Combined Datasets) after de-batch processing. PCA, Principal Component Analysis. The red color is the major depressive disorder (MDD) data set GSE 98793, and the blue color is the MDD data set GSE 19738.



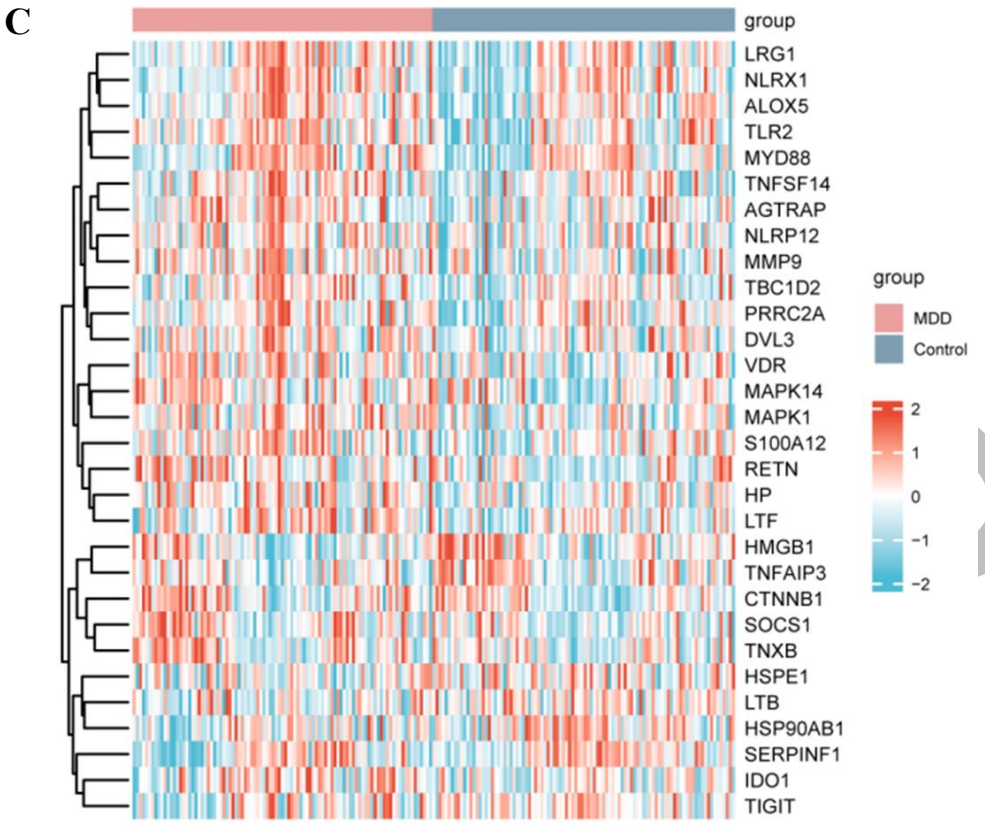


Figure 3. Differential Gene Expression Analysis; (A) Volcano plot of differentially expressed genes analysis between the major depressive disorder (MDD) group and the control (Control) group in the combined GEO data sets (Combined Datasets). Red represents up-regulated differentially expressed genes, blue represents down-regulated differentially expressed genes, and gray represents insignificantly expressed genes. (B) Venn diagram of differentially expressed genes (DEGs) and IFRGs in combined GEO data sets (Combined Datasets). (C) Heatmap of expression values of IFRDEGs in the Combined GEO datasets (Combined Datasets). Purple is the control group, and pink is the MDD group. MDD, Major Depression Disorder; DEGs, Differentially Expressed Genes; IFRGs, Inflammatory Factor -Related Genes; IFRDEGs, Inflammatory Factor -Related Differentially Expressed Genes.

Gene Ontology (GO) and Pathway (KEGG) Enrichment Analyses

We further investigated the biological processes (BP), cellular components (CC), molecular functions (MF), and pathways of the 30 IFRDEGs using GO and KEGG pathway enrichment analyses. The enriched pathways and processes identified in the GO and KEGG analyses provide insights into the underlying molecular mechanisms that may contribute to the pathophysiology of MDD. Several enriched biological processes, such as the regulation of inflammatory responses and interleukin-12 production, align with the growing evidence that chronic inflammation plays a significant role in the development and progression of MDD. The involvement of these pathways suggests that dysregulated immune responses might be a contributing factor to this disorder. The enrichment of pathways, such as the IL-17 signaling pathway and the NOD-like receptor signaling pathway, further underscores the connection between immune system dysregulation and MDD. These pathways are known to influence neuroinflammation, which has been implicated in the alteration of brain function associated with depressive symptoms. Moreover, enriched cellular components such as the secretory granule lumen and molecular functions, such as Toll-like receptor binding, are also noteworthy. These components and functions are involved in the activation of immune responses, which could exacerbate or perpetuate depressive states by promoting neuroinflammatory processes. This study

Published by Tehran University of Medical Sciences (http://ijaai.tums.ac.ir)

screened a total of 312 DEGs through differential expression analysis, including 175 upregulated genes and 137 down regulated genes. By cross referencing with inflammation related genes (IFRG) included in the GeneCards database, 47 inflammatory cytokine genes (IFRDEG) related to MDD were ultimately screened. The volcano plot shows a significant distribution of differentially expressed genes, with upregulated and downregulated genes exhibiting distinct differential expression patterns between patients and healthy controls. Statistical analysis further revealed that these genes were significantly enriched in biological processes such as "immune response" (aj. $p<0.01$), "inflammatory response regulation" (aj. $p<0.01$), as well as important pathways such as IL-17 signaling pathway (aj. $p=0.003$) and NF - κ B signaling pathway (aj. $p=0.021$).

Table 2 displays the particular outcomes of GO employing IFRDEGs, and Table 3 displays the unique outcomes of pathway (KEGG) enrichment analysis. The findings indicate that the genes associated with IFRDEGs in MDD are primarily enriched in inflammatory response regulation, lipopolysaccharide-mediated signaling pathway, interleukin-12 production, regulation of interleukin-12 production, and positive regulation of biological processes, including interleukin-12 production (BP); cellular components, including the secretory granule lumen, cytoplasmic vesicle lumen, vesicle lumen, ficolin-1-rich granule (CC);

Table 2. Result of GO Enrichment Analysis for IFRDEGs

ONTOLOGY	ID	GeneRatio	BgRatio	*p*	p.adjust	q value
BP	GO:0050727	11/30	394/18800	1.14E-11	2.02E-08	1.07E-08
BP	GO:0031663	6/30	62/18800	5.60E-10	2.74E-07	1.45E-07
BP	GO:0032615	6/30	63/18800	6.19E-10	2.74E-07	1.45E-07
BP	GO:0032655	6/30	63/18800	6.19E-10	2.74E-07	1.45E-07
BP	GO:0032735	5/30	42/18800	5.95E-09	2.11E-06	1.12E-06
CC	GO:0034774	10/30	322/19594	2.81E-11	1.40E-09	9.17E-10
CC	GO:0060205	10/30	325/19594	3.07E-11	1.40E-09	9.17E-10
CC	GO:0031983	10/30	327/19594	3.27E-11	1.40E-09	9.17E-10
CC	GO:1904813	7/30	124/19594	6.18E-10	1.99E-08	1.30E-08
CC	GO:0101002	7/30	185/19594	1.01E-08	2.61E-07	1.70E-07
MF	GO:0001530	3/30	34/18410	2.26E-05	3.59E-03	1.97E-03
MF	GO:0032813	3/30	49/18410	6.84E-05	5.33E-03	2.93E-03
MF	GO:0050786	2/30	10/18410	1.15E-04	5.33E-03	2.93E-03
MF	GO:0035325	2/30	12/18410	1.68E-04	5.33E-03	2.93E-03
MF	GO:1990226	2/30	12/18410	1.68E-04	5.33E-03	2.93E-03

GO: Gene Ontology; BP: Biological Process; CC: Cell Component, MF: Molecular Function; IFRDEGs: Inflammatory Factor -Related Differentially Expressed Genes

Table 3. Result of KEGG Enrichment Analysis for IFRDEGs

ONTOLOGY	ID	GeneRatio	BgRatio	*p*	p.adjust	q value
KEGG	hsa04621	7/21	184/8164	2.35E-07	2.15E-05	1.11E-05
KEGG	hsa05145	6/21	112/8164	2.67E-07	2.15E-05	1.11E-05
KEGG	hsa04657	5/21	94/8164	3.20E-06	1.72E-04	8.86E-05
KEGG	hsa05417	6/21	215/8164	1.21E-05	4.88E-04	2.52E-04
KEGG	hsa05132	6/21	249/8164	2.80E-05	9.01E-04	4.65E-04

KEGG: Kyoto Encyclopedia of Genes and Genomes; IFRDEG s: Inflammatory Factor -Related Differentially Expressed Genes

lipopolysaccharide binding, and tumor Necrosis factor receptor superfamily binding, RAGE receptor binding, Toll-like receptor binding, and histone methyltransferase binding are a few examples of molecular functions (MF). In addition, biological pathways including the toxoplasmosis, IL-17 signaling route, NOD-like receptor signaling pathway, lipid and atherosclerosis, and salmonella infection are also concentrated in it. The findings of the enrichment analysis for GO and pathway (KEGG) were shown using bubble charts and histograms (Figure 4A, B). Furthermore, the findings of the combined logFC enrichment analysis for GO and pathway (KEGG) are shown in the chord diagram (Figure 4C) and the circle diagram (Figure 4D). According to the findings of the circle diagram (Figure 4D), the biological process (BP) that is most significantly negatively regulated is the synthesis of interleukin-12, while the most significantly positively regulated cellular component (CC) is this one.

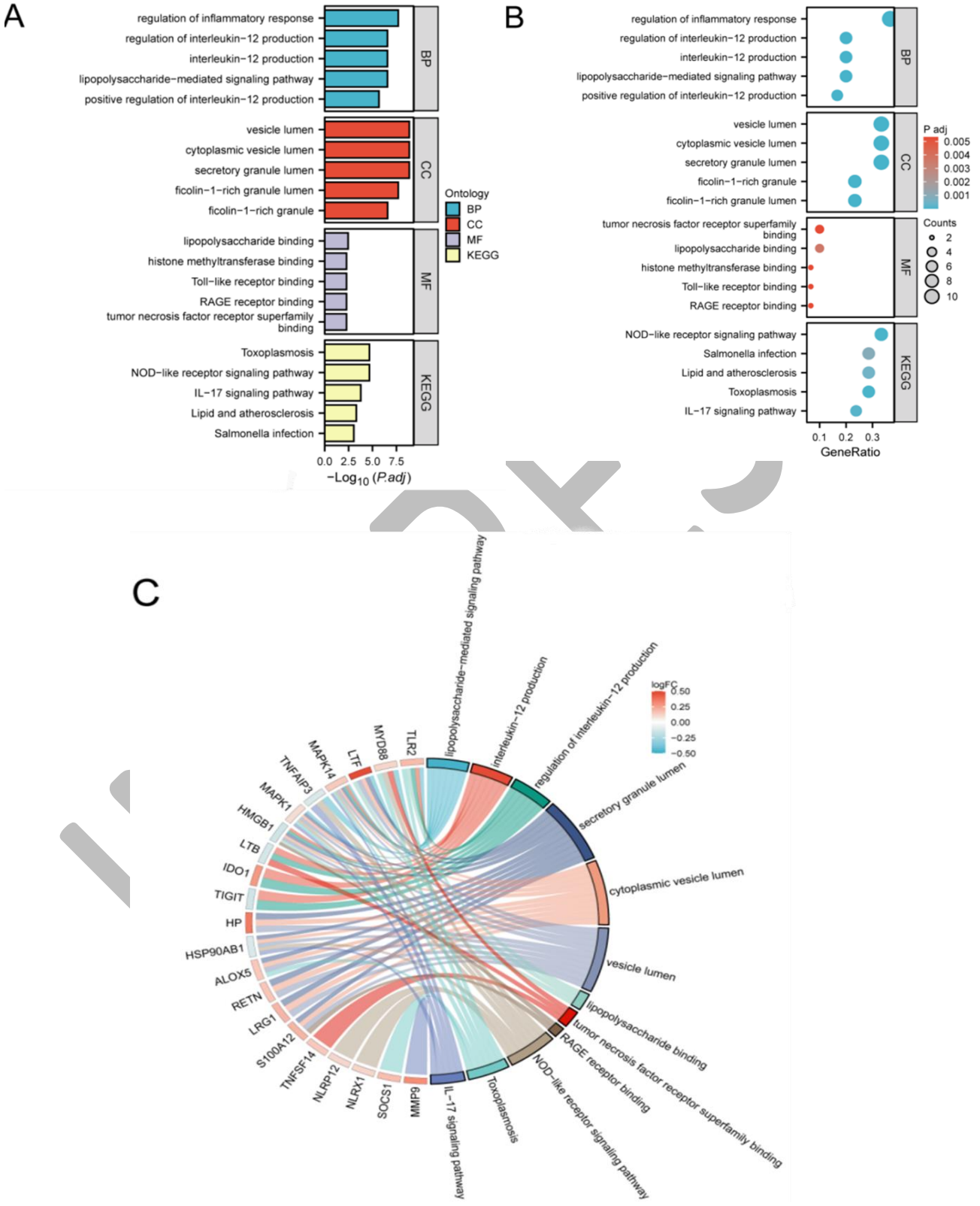
Gene Set Enrichment Analysis (GSEA)

In order to determine the impact of the expression levels of all genes in the Combined GEO Datasets (Combined Datasets) on MDD, the expression of all genes in the Combined GEO Datasets (Combined Datasets) was studied by gene set enrichment analysis (GSEA). The relationship between it and the biological processes it participates in, the cellular components it affects, and the molecular functions it exerts (Figure 5A), the specific results are shown in Table 4. The results showed that all genes in the combined GEO data sets (Combined Datasets) were significantly enriched in MAPK Signaling Pathway (Figure 5B), Notch Signaling Pathway (Figure 5C), IL4 Signaling Pathway (Figure 5D), IL6 Signaling Pathway (Figure 5E), Interleukin1 IL1 Structural Pathway (Figure 5F) and other biologically relevant functions and signaling pathways.

Construction of Protein-protein Interaction Network and Screening of Hub Genes

First, protein-protein interaction analysis was performed, and a protein-protein interaction network (PPI Network) of 30 IFRDEGs using the STRING database (Figure 6A). Protein-protein interaction network (PPI Network) results show that 26 IFRDEGs are related, namely: ALOX5, TLR2, MMP9, CTNNB1, MAPK1, MAPK14, DVL3, HMGB1, HSP90AB1, IDO1, TNFAIP3, SOCS1, MYD88, CYP27B1, HP, S100A12, RETN, LTF, LRG1, HSPE1, NLRP12, TIGIT, TNFSF14, LTB, SERPINF1, NLRX1. Subsequently, the 5 algorithms of the cytoHubba plug-in of Cytoscape software were used to calculate the scores of 26 IFRDEGs, and the top 10 IFRDEGs were arranged in order according to the scores. The five algorithms are: Maximal Clique Centrality (MCC), Degree, Maximum Neighborhood Component (MNC), Edge Percolated Component (EPC), and Closeness. Next, the IFRDEGs of Top10 in 5 algorithms were used to draw the protein-protein interaction network, which includes: MCC (Figure 6B), MNC (Figure 6C), Degree (Figure 6D), EPC (Figure 6E) and Closeness (Figure 6F). The color of the circle from red to yellow represents the score from high to low. Finally, the genes of the five algorithms were intersected and a Venn diagram (Figure 6G) was drawn for analysis. The intersection genes of the algorithms are the hub genes of MDD. The 9 hub genes are respectively: MMP9, MYD88, CTNNB1, TLR2, HSP90AB1, HMGB1, MAPK1, S100A12, IDO1.

Published by Tehran University of Medical Sciences (http://ijaai.tums.ac.ir)



Published by Tehran University of Medical Sciences (http://ijaai.tums.ac.ir)

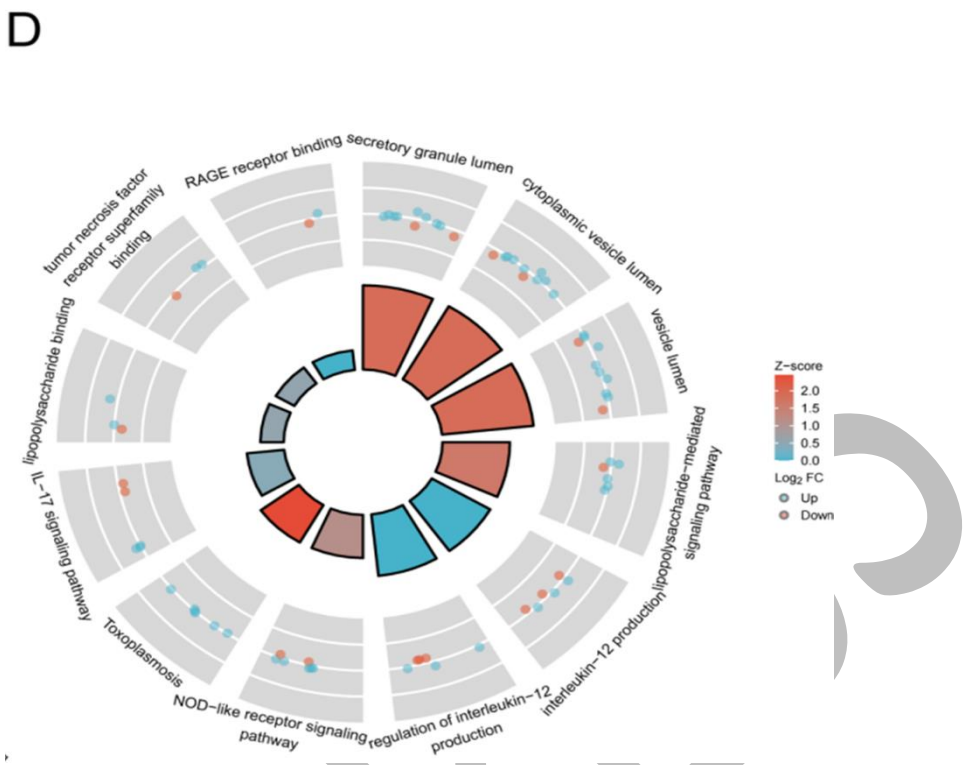


Figure 4. Gene ontology (GO) and KEGG Enrichment Analysis for IF RDEGs; (A) of GO and pathway (KEGG) enrichment analysis results of inflammatory factor-related differentially expressed genes (IFRDEGs). (B) Bubble chart of GO and pathway (KEGG) enrichment analysis results of IFRDEGs. (C) Chord plot of GO and pathway (KEGG) enrichment analysis results of IFRDEGs. (D) Circle diagram of GO and pathway (KEGG) enrichment analysis results of IFRDEGs. In the circle diagram, the outer circle contains molecules and logFC values, with red being up-regulated genes and blue being down-regulated genes; each column in the inner circle corresponds to an entry, the height is the relative size of p.adj, and the color is the Z-score size. If the Z- score is positive, it may be a positive adjustment, if it is negative, it may be a negative adjustment. The larger the absolute value, the higher the degree of adjustment. MDD, Major Depression Disorder; IFRDEGs: Inflammatory Factor - Related Differentially Expressed Genes; GO: Gene Ontology; BP: Biological Process; MF, Molecular Function; KEGG, Kyoto Encyclopedia of Genes and Genomes. The screening criteria for GO and pathway (KEGG) enrichment analysis entries are p.adj<0.05 and FDR value (q value)<0.25 is considered statistically significant, and the *p* value correction method is Benjamini-Hochberg (BH).

Table 4. Results of GSEA for Combined Datasets

ID	S et Size	E nrichment Score	NES	*p*	p.adjust	qvalue
WP_MAPK_SIGNALING_PATHWAY	1.35E+02	3.50E-01	1.50E+00	4.11E-03	4.09E-02	3.27E-02
KEGG_NOTCH_SIGNALING_PATHWAY	3.20E+01	5.26E-01	1.71E+00	2.83E-03	3.23E-02	2.58E-02
WP_IL4_SIGNALING_PATHWAY	4.10E+01	5.57E-01	1.93E+00	2.05E-04	5.31E-03	4.25E-03
WP_IL6_SIGNALING_PATHWAY	3.30E+01	5.22E-01	1.71E+00	2.92E-03	3.31E-02	2.65E-02
WP_INTERLEUKIN1_IL1_STRUCTURAL_PATHWAY	3.70E+01	5.00E-01	1.69E+00	4.23E-03	4.14E-02	3.31E-02

GSEA: Gene Set Enrichment Analysis

Published by Tehran University of Medical Sciences (http://ijaai.tums.ac.ir)

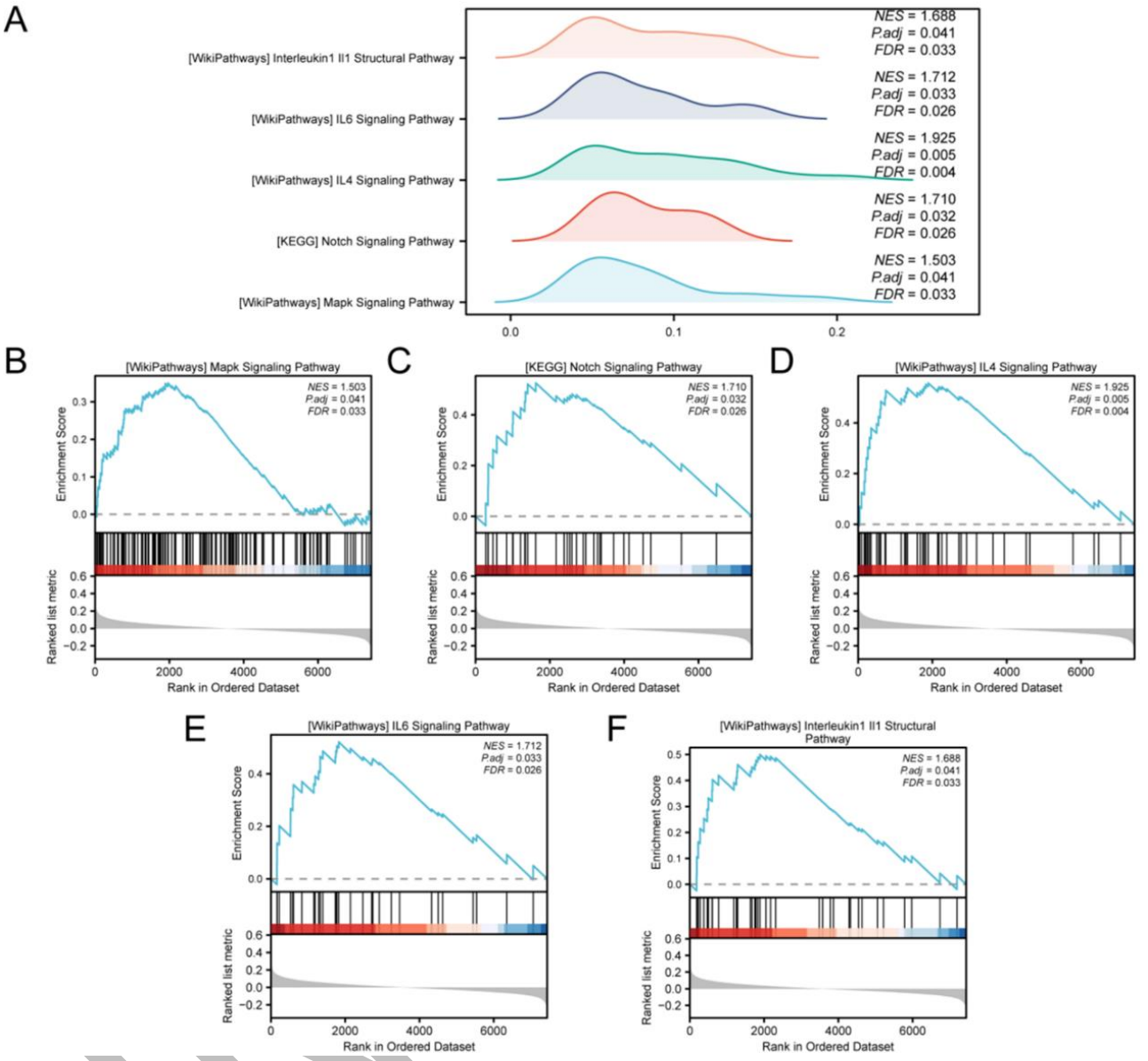


Figure 5. GSEA for Combined Datasets; (A) Gene Set Enrichment Analysis (GSEA) of the Combined GEO Datasets (Combined Datasets). Mountain chart display of 5 biological functions. (B-F) Gene set enrichment analysis (GSEA) shows that major depressive disorder (MDD) is significantly enriched in MAPK Signaling Pathway (B), Notch Signaling Pathway (C), IL4 Signaling Pathway (D), IL6 Signaling Pathway (E), Interleukin1 IL1 Structural Pathway (F). GSEA, Gene Set Enrichment Analyses; MDD, Major Depression Disorder. The screening criteria of gene set enrichment analysis (GSEA) are p.adj<0.05 and FDR value (q value)<0.25, and the p value correction method is Benjamini-Hochberg (BH).

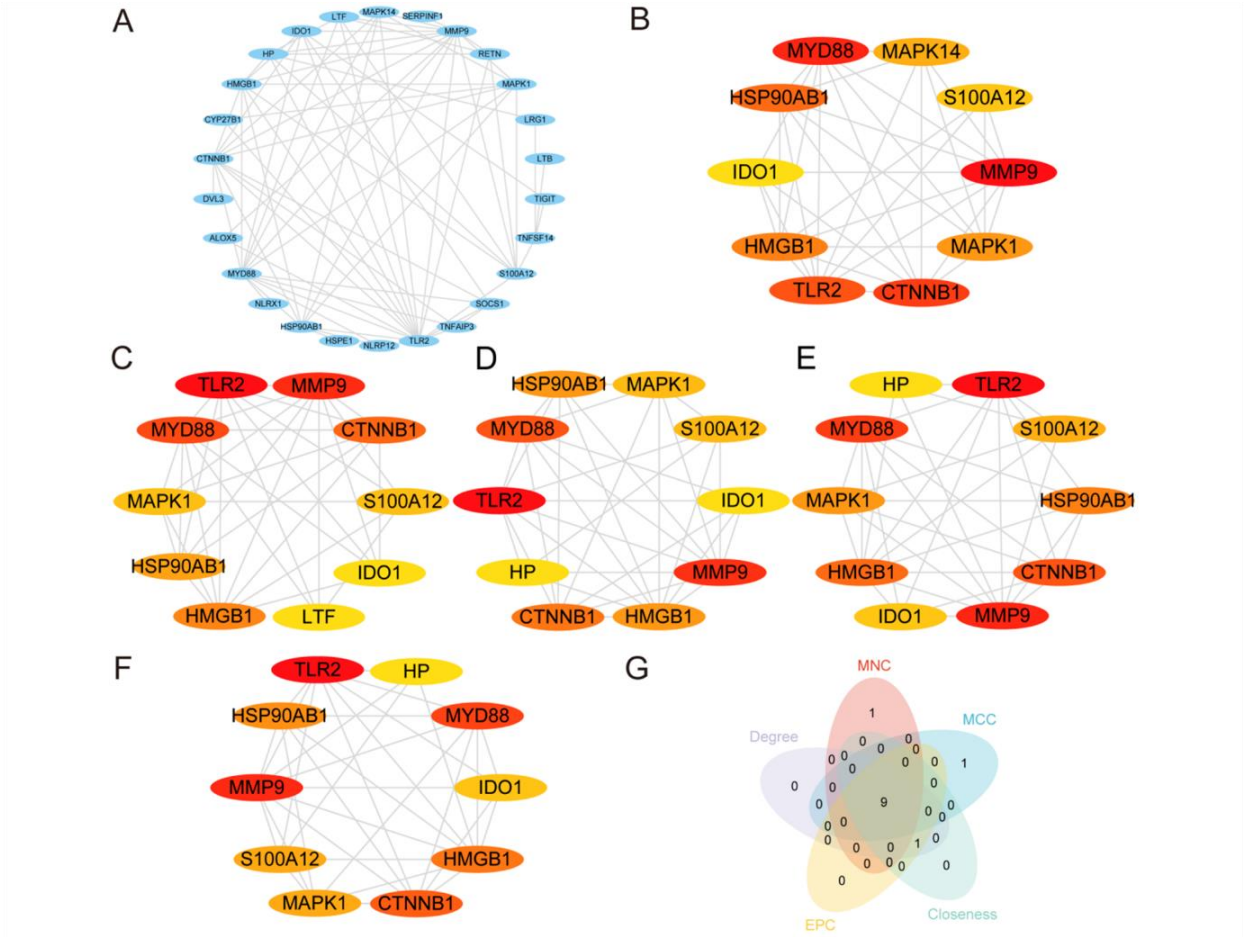


Figure 6. PPI Network and hub genes Analyses; (A) Protein-protein interaction network (PPI Network) of inflammatory factor -related differentially expressed genes (IFRDEGs) calculated from the STRING database. (B–H) Protein -protein interaction network (PPI Network) drawn by the Top 10 IFRDEGs calculated by the 5 algorithms of the cytoHubba plug-in, including MCC (B), MNC (C), Degree (D), EPC (E) and Closeness (F). G. Top 10 IFRDEGs Venn diagram for the 5 algorithms of the cytoHubba plug-in. MDD: Major Depression Disorder; PPI Network: Protein-protein Interaction Network.

Construction of Regulatory Network

First, the miRNAs related to hub genes related to inflammatory factors were obtained from the TarBase database through the Network Analyst tool, and the mRNA-miRNA regulatory network (mRNA-miRNA Regulatory Network) was constructed and visualized using Cytoscape software (Figure 7A). Among them, there are 9 inflammatory factor -related hub genes and 9 miRNAs.

Next, the transcription factors (TF) combined with the hub genes related to inflammatory factors were obtained through the ChIPBase database and the hTFtarget database and the intersection was obtained to construct the mRNA-TF Regulatory Network (mRNA-TF Regulatory Network) and visualize it using Cytoscape software (Figure 7B). Due to the excessive number of related transcription factors (TF), only TF and inflammatory factor-related hub genes (hub genes) that have been combined with inflammatory factor-related hub genes (hub genes) more than 5 times were selected as mRNA- Visualization of TF regulatory networks. Among them, there are 5 inflammatory factor-related hub genes and 17 transcription factors.

Finally, the CTD database was used to identify potential drugs or molecular compounds related to hub genes related to inflammatory factors. An mRNA-Drug Regulatory Network was constructed and visualized using Cytoscape software (Figure 7C). This network contains a total of 4 inflammatory factor-related hub genes and 28 related chemical drugs.

Published by Tehran University of Medical Sciences (http://ijaai.tums.ac.ir)

Expression differences and ROC Analysis of Hub Genes

In order to verify MDD group, the expression difference in the control group is shown in the group comparison chart (Figure 8A). The expression difference of the TLR2 gene is highly statistically significant (p <0.001) in the MDD group and the control group). The expression differences of HSP90AB1, MAPK1, MMP9 and S100A12 genes are highly statistically significant (p<0.01) in the (MDD) group and the control (Control) group (p<0.01). Additionally, the expression differences of HMGB1 and MYD88 genes are statistically significant (p<0.05) in the (MDD) group and the control (Control) group.

We also used ROC curves (Figure 8B-H) to show the diagnostic effect of the expression of 7 hub genes related to inflammatory factors with statistically significant expression differences in the group comparison chart for MDD. The analysis results showed that the AUC values for HMGB1 (AUC=0.605, Figure 8B), HSP90AB1 (AUC=0.619, Figure 8C), MAPK1 (AUC=0.632, Figure 8D), MMP9 (AUC=0.615, Figure 8E), MYD88 (AUC=0.583, Figure 8F), S100A12 (AUC=0.609, Figure 8G) and TLR2 (AUC=0.637, Figure 8H) indicate with lower accuracy, with diagnostic effects ranging between 0.5<AUC<0.7.

Immune Infiltration Analysis

The integrated GEO data set (Combined Datasets) samples were used, and the immune infiltration abundance and correlation of 28 types of immune cells in the integrated GEO data set (Combined Datasets) samples through the ssGSEA algorithm. First, based on the results of immune infiltration analysis, a group comparison chart was drawn to show the expression difference of immune cell infiltration abundance between the MDD group and the control (Control) group in the integrated GEO data set (Combined Datasets). The comparison figure (Figure 9A) shows that 9 types of immune cells are statistically significant (p<0.05). Among them, the infiltration abundance of immune cells CD56 bright natural killer cells and Neutrophils is significantly higher in the MDD group than in the control group with high statistically significant between groups (p<0.01). The expression of activated CD4 T cell, activated CD8 T cell, regulatory T cell, type 1 T helper cell, activated dendritic cell, macrophage, natural killer cell is statistically significant among different groups (p <0.05).

The correlation results of the infiltration abundance of 9 types of immune cells with significant differences in the infiltration abundance in the combined GEO data set samples were displayed through a correlation heatmap (Figure 9B). The results showed that immune cells activated dendritic cells and Neutrophils (r=0.74) show a moderate positive correlation.

Finally, the correlation heatmap shows the correlation between the 9 inflammatory factor-related hub genes (hub genes) and the infiltration abundance of 9 types of immune cells in the combined GEO data set (Figure 9C). The correlation heat map results indicate that the gene MMP9 and immune cells Neutrophil (r value=0.57), gene MMP9 and Macrophage (r value=0.54), gene MMP9 and activated dendritic cells (r=0.52) show moderate positive correlations. Gene MAPK1 and activated dendritic cell (r=0.68), gene MAPK1 and macrophage (r=0.55), gene MAPK1 and immune cell macrophage (r=0.55) show moderate positive correlation. HMGB1 gene and macrophage (r=0.64) shows a moderate negative correlation.

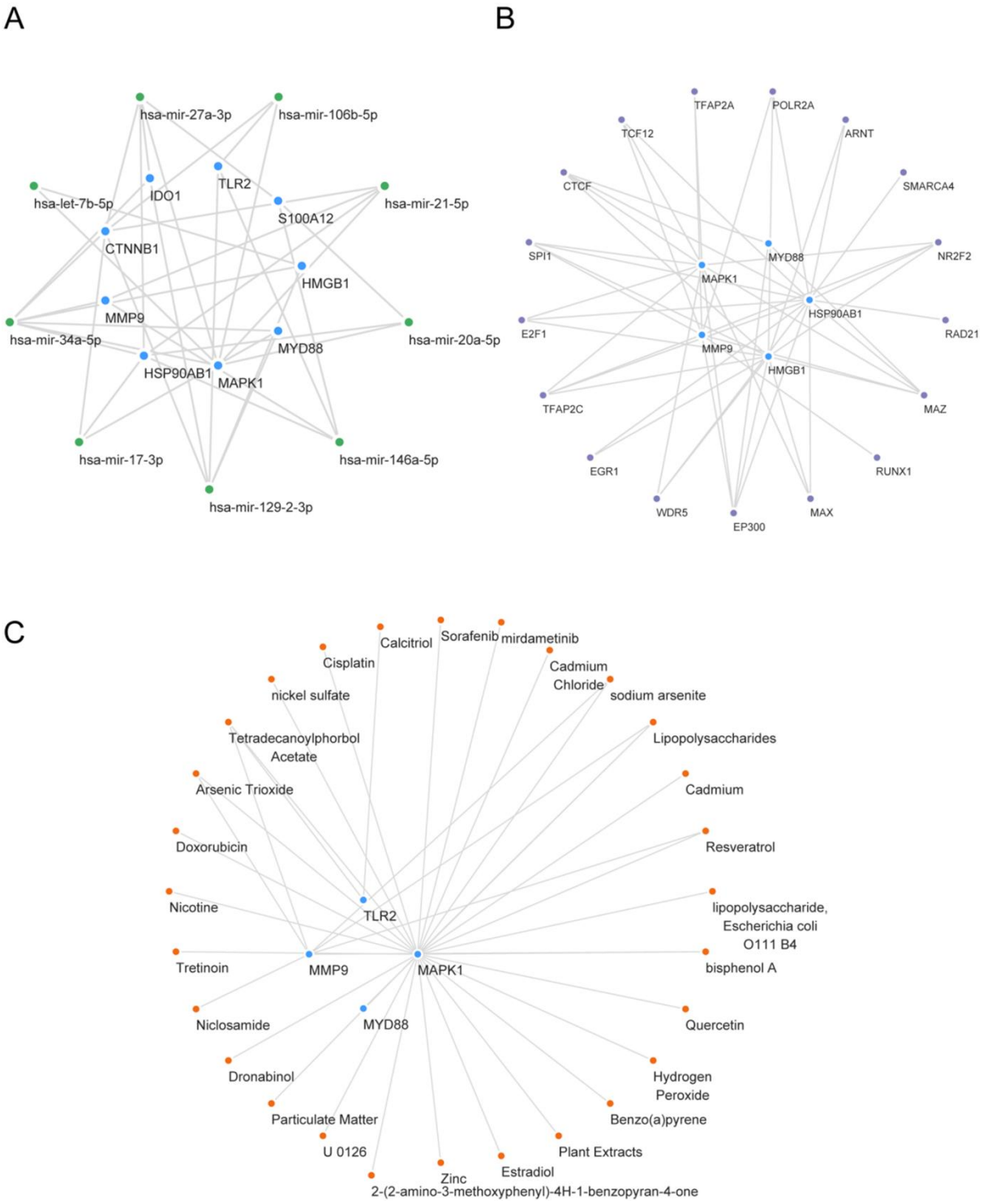


Figure 7. Regulatory Network of hub genes; (A) The mRNA-miRNA Regulatory Network of hub genes related to inflammatory factors. (B) The mRNA-TF Regulatory Network of hub genes related to inflammatory factors. (C) The mRNA-Drug Regulatory Network of hub genes related to inflammatory factors. TF: Transcription Factor. The blue circle is mRNA, the green circle is miRNA, the purple circle is transcription factors, and the orange circle is small drug molecules.

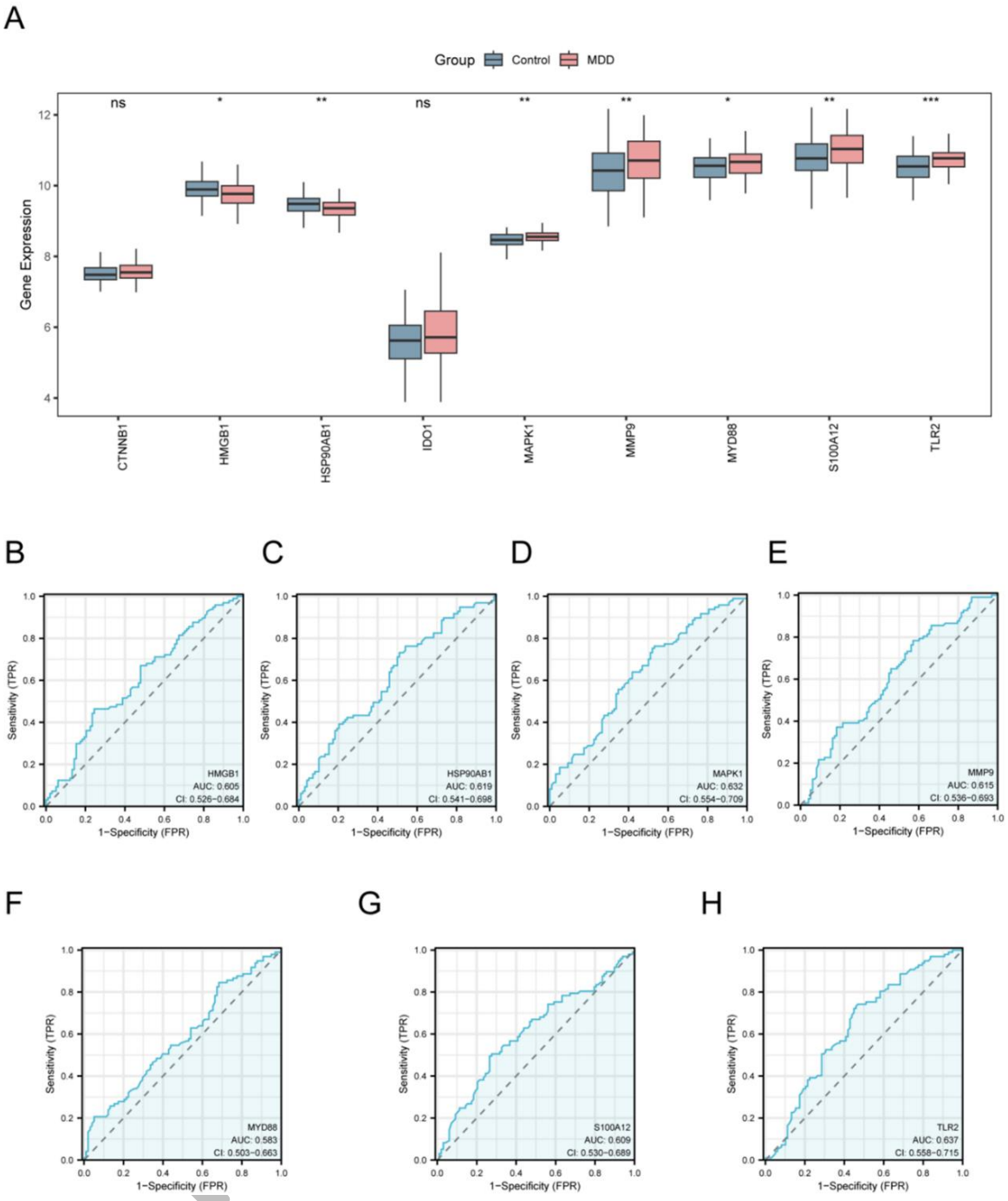


Figure 8. Expression Difference and ROC Curve Analysis; (A) Group comparison diagram of inflammatory factor - related hub genes in integrated GEO data sets (Combined Datasets). The red color in the group comparison chart is the major depressive disorder (MDD) group, and the blue color is the control group. (B-H) Inflammation factor -related hub genes (hub genes) with significantly different expression values in the group comparison chart HMGB1 (B), HSP90AB1 (C), MAPK1 (D), MMP9 (E), MYD88 (F), S100A12 (ROC curves of G) and TLR2 (H). MDD, Major Depression Disorder; ROC, Receiver Operating Character is tic Curve. ns stands for *p* value≥0.05, no statistical significance; * represents *p* value<0.05, statistically significant; ** represents *p* value<0.01, highly statistically significant. * represents *p* value<0.001, which is highly statistically significant. The closer the AUC is to 1, the better the diagnostic effect is. When the AUC is between 0.5 and 0.7, the accuracy is lower.**

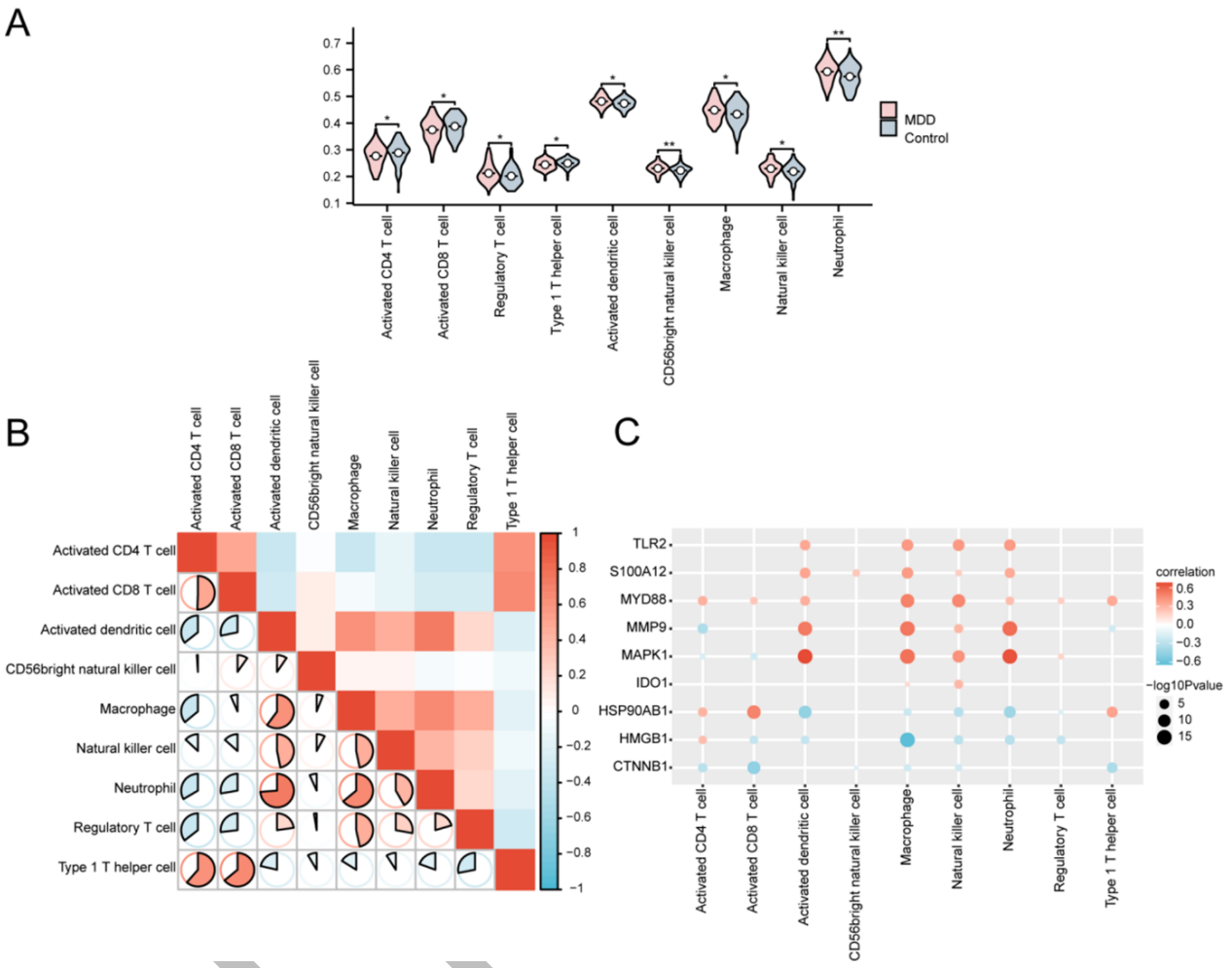


Figure 9. Combined Datasets Immune Infiltration Analysis by ssGSEA Algorithm; (A) immune cells in major depressive disorder (MDD) group and control group. In the group comparison chart, the blue color is the MDD group and the red color is the control group. (B) Correlation heat map of immune cell infiltration abundances with significantly different infiltration abundances in the combined GEO data sets (Combined Datasets) in the group comparison chart. (C) Heat map of the correlation between 9 inflammatory factor - related hub genes and the infiltration abundance of 9 types of immune cells in the combined GEO datasets (Combined Datasets). MDD: Major Depression Disorder; ssGSEA: single-sample Gene-Set Enrichment Analysis. * Represents *p* value<0.05, which is statistically significant; ** represents *p* value<0.01, which is highly statistically significant. The absolute value of the correlation coefficient (r value) is weak or irrelevant when it is below 0.3, it is weak correlation when it is between 0.3 and 0.5, and it is moderate correlation when it is between 0.5 and 0.8.

DISCUSSION

Chronic inflammation, particularly involving elevated levels of interleukin-6 (IL-6) and tumor necrosis factor-α (TNF-α), is believed to disrupt neurotransmitter balance, potentially triggering or worsening of depression symptoms. Hub genes identified in this study play key roles in immune regulation and communication, which may influence the effectiveness of immunotherapy in MDD. Their aberrant expression correlates with MDD outcomes, making them valuable as both prognostic markers and potential therapeutic targets. Despite the established connection between inflammation and MDD, there is a lack of comprehensive studies identifying specific inflammation related genes that could serve as reliable diagnostic markers or therapeutic targets. This study aimed to fill this gap by identifying key genes linked to

Published by Tehran University of Medical Sciences (http://ijaai.tums.ac.ir)

inflammatory processes in MDD and evaluating their potential as diagnostic tools. By concentrating on the intersection of inflammation and MDD, this study identified specific hub genes, providing new insights into the disease's molecular underpinnings. The identification of these genes not only deepens our understanding of MDD's pathophysiology, but also opens up potential avenues for targeted therapies. The expression of these genes related to the onset and progression of MDD may serve as a target for the creation of new MDD therapeutics.

This study identified core hub genes linked to serious depression using a diagnostic model that integrated five algorithms. The mRNA-miRNA regulatory network for these genes was then visualized using Cytoscape. Furthermore, the ChIPBase and hTFtarget databases were used to examine the connections between these hub genes and TF as well as the mRNA-TF regulatory network. The authors investigated the expression variations of the hub genes between MDD and controls by grouping comparisons using the integrated GEO dataset and utilized the ROC curve to assess the diagnostic accuracy of these genes. The effectiveness of the diagnosis was evaluated using the area under the ROC curve (AUC). Ultimately, the CTD database was used to identify possible medications or chemical compounds linked to these Hub genes, and mRNA-Drug regulatory networks were built to provide novel therapeutic targets for future research. GO and KEGG pathway enrichment analyses were employed to explore the association between differentially expressed genes (IFRDEGs) and MDD, providing insights into their roles in biological processes, cellular components, and molecular functions. The seven inflammatory factor-related hub genes—HMGB1, HSP90AB1, MAPK1, MMP9, MYD88, S100A12, and TLR2—received special attention due to their limited diagnostic accuracy in severe depression, as indicated by AUC values ranging from 0.5 to 0.7. While these values suggest limited predictive power individually, they indicated that these genes may still hold significance when considered as part of a broader diagnostic panel. These hub genes provide new directions for targeted therapy and diagnosis of MDD. For example, TLR2 and MAPK1 are both key nodes in the inflammatory signaling pathway, which can alleviate depressive symptoms by reducing inflammation through specific inhibitors such as TLR2 antagonists or MAPK1 inhibitor U0126. In addition, HMGB1, as an inflammatory factor, has shown good anti-inflammatory potential with its inhibitor Glycyrhizin, and its therapeutic effect in MDD can be further explored. In terms of diagnosis, the mRNA expression levels of these genes can serve as molecular markers for MDD, combined with liquid biopsy techniques for early screening and disease monitoring of patients. Future research can further combine gene editing techniques such as CRISPR-Cas9 or RNA interference (siRNA) technology to validate the therapeutic effects of these targets in cell and animal models, and develop personalized and precise treatment plans.

We have discovered core genes (hub genes) linked to inflammatory factors in this article. These important genes were found by using a Wayn diagram to analyze genes that were acquired using various techniques. The control of miRNA on these Hub genes was then investigated.

Notably, miRNAs can regulate multiple target genes, and conversely, a single gene can be regulated by multiple miRNAs. This study analyzed the differential expression of immune cells, such as dendritic cells and neutrophils, between MDD and control groups. The paper illustrates the relationship between the infiltration abundance of nine immune cells in the integrated GEO dataset and nine hub genes linked to inflammatory factors using a correlation heatmap. For instance, there was a moderate relationship found between the genes HMGB1 and macrophages and MMP9 and neutrophils. Furthermore, the mRNA-miRNA regulatory network was viewed using Cytoscape software to examine the relationship between hub genes and transcription factors. Lastly, the current study created an mRNA-drug regulatory network by using the CTD database to find possible medications or chemical compounds linked to hub genes associated to inflammatory cytokines. Through these studies, the major immune cell types linked to the high-low risk category of Hub genes diagnostic models are identified in this research, offering significant biomarkers and possible therapeutic targets for the identification and management of illnesses connected to inflammatory factors.

The key gene nodes identified in this study, such as MMP9, MYD88, TLR2, may be potential therapeutic targets in the MDD drug interaction network. Researchers have identified drugs associated with these genes using the CTD database and mRNA drug regulatory network analysis, which can be used to treat MDD. By extensively utilizing GO and KEGG, we can further understand the roles of these genes in biological processes, cellular components, molecular functions,

and biological pathways. This understanding provides a crucial biological background and theoretical basis for studying corresponding treatment strategies. This study revealed the core role of inflammation related differentially expressed genes in the pathological mechanism of MDD through GO and KEGG enrichment analysis. In particular, the significant enrichment of IL-17 signaling pathway and NOD like receptor signaling pathway (adj. $p<0.05$) suggests that these genes may be involved in the occurrence and development of MDD by regulating inflammatory response. In addition, key central genes identified by PPI network analysis, such as TLR2 and MAPK1, play an important role in immune cell infiltration dysfunction in MDD patients. The immune analysis results showed that the proportion of macrophages and neutrophils in MDD patients was significantly increased, while Treg cells were significantly reduced ($p<0.05$), further confirming that immune dysfunction is an important characteristic of MDD. These results not only deepen our understanding of the pathogenesis of MDD but also provide potential directions for future treatments targeting inflammation and immune pathways.

Important resources for understanding the mechanisms behind gene regulation include the mRNA-miRNA, mRNA-RBP (RNA binding protein), and mRNA-TF (transcription factor) interaction network. These networks aid in the understanding of gene regulation and the impact of these regulatory mechanisms on pathological states in MDD research.1. The interactions of network between mRNA and miRNA: miRNAs may control the expression of mRNA by, for instance, preventing or degrading translation. The molecular mechanism of MDD may be better understood by identifying the miRNAs that are crucial in the disease and its target mRNA by an analysis of the interaction between mRNA and miRNA.2. The mRNA-RBP interaction network: To control the stability and translation efficiency of mRNA, RNA-binding protein (RBP) interacts to it. Understanding which genes are controlled at the posttranslational stage is crucial for comprehending the regulation of gene expression in MDD, and this may be achieved by examining the interaction between RBP and mRNA.3. The mRNA-TF interaction network: The transcription factor directly influences the transcription of mRNA in the mRNA-TF interaction network. We can uncover patterns of gene regulation pertaining to MDD and discover important transcription factors that regulate the expression of genes linked to the illness by analyzing the connection between TF and mRNA. All things considered, these interaction networks provide us with a means of comprehending the regulatory processes of MDD genes at various levels, assisting in the discovery of the disease's molecular foundation and potentially leading to the development of innovative treatment approaches.

The identified hub genes, including MMP9, MYD88, and CTNNB1, are crucial players in the pathophysiology of MDD, particularly through their roles in inflammation and immune responses. MMP9 is known for its involvement in extracellular matrix remodeling, which can influence the blood-brain barrier's integrity and thereby affect neuroinflammation—a key factor in MDD. Elevated levels of MMP9 have been associated with neuroinflammatory processes that contribute to the onset and progression of depressive symptoms. MYD88 is a central adaptor protein in the TLR signaling pathway, which is crucial for initiating innate immune responses. The activation of TLRs, particularly TLR2 and TLR4, has been linked to increased production of pro-inflammatory cytokines, which are often elevated in MDD patients. MYD88's role in amplifying inflammatory signals may exacerbate neuroinflammation, contributing to the depressive state by altering neurotransmitter systems and neural plasticity. CTNNB1, a key component of the Wnt signaling pathway, is involved in the regulation of gene transcription, cell adhesion, and neural development. Dysregulation of Wnt signaling has been implicated in mood disorders, including MDD, where it can lead to altered neurogenesis and synaptic plasticity—both critical processes in the maintenance of healthy mood regulation. Together, these hub genes highlight the interplay between immune responses and neural function in MDD, suggesting that they contribute to the disorder's pathophysiology by mediating inflammatory processes and affecting brain structure and function. Understanding the roles of these genes not only provides insight into MDD mechanisms but also points to potential therapeutic targets for intervention.

There are some limitations in this study MDD. Firstly, the sample size is quite small and may not be sufficient to accurately capture the complexity and variety of MDD. The data source is restricted to two GEO datasets (GSE98793 and GSE19738). Second, because all of the samples were made from whole blood, there may be more noticeable variations in the depression biomarkers in other tissues or cell types. Moreover, there may still be batch effects and other possible confounders (including

age, sex, and lifestyle). Furthermore, a large number of DEGs may be selected without significant biological relevance and may not take into account multiple test corrections, potentially increasing the risk of false positive findings. This is because the screening criteria for DEGs were set as |logFC|> 0.00 and *p* value<0.05. Lastly, the research omitted to provide follow-up validation studies, which are necessary to determine the biological significance and therapeutic applicability of the study results. Specific experimental approaches should be proposed for future research to validate the identified hub genes and regulatory networks. For instance, *in vitro* studies involving human cell lines could be conducted in this study to explore the functional roles of these hub genes by performing gene knockdown or overexpression experiments. Additionally, animal models, such as transgenic mice, could be employed to investigate the in vivo effects of modulating these genes on depressive behaviors and related molecular pathways. These approaches would not only help validate the findings of this study, but also provide deeper insights into the biological relevance and therapeutic potential of the identified genes and networks. To address these limitations, future research should consider expanding the sample size to enhance the robustness and generalizability of the findings. Additionally, it would be beneficial to incorporate follow-up validation studies, such as in vitro experiments or animal models, to confirm the biological significance and therapeutic potential of the identified differentially expressed genes. Expanding the analysis to include samples from other tissues or cell types could also provide a more comprehensive understanding of MDD biomarkers. These steps will help further validate and extend the findings of this study.

The pathophysiology of MDD was thoroughly investigated in this research, primarily using the following analytic techniques: 1. Data download and integration: Sample data from MDD patients and controls were derived by the MDD-related datasets GSE98793 and GSE19738, which were obtained from the GEO database.2. Differentially expressed gene analysis: The R package limma was used to perform differential analysis of the integrated GEO data set in order to find genes that are expressed differently in the MDD and control groups. The features of IFRDEGs linked to inflammatory factors in biological processes, cellular components, molecular activities, and biological pathways were investigated through GO and Pathway (KEGG) enrichment analysis.3. Gene set enrichment analysis (GSEA): The relationship between the molecular functions, biological processes, and impacted cellular components and expression levels of all genes was investigated.4. ROC curve analysis: This analysis demonstrated how multiple hub genes connected to inflammatory factors affected the diagnosis of severe depression.[5] Analysis of immune infiltration: Based on the findings of this analysis, we plotted differences in immune cell infiltration abundance between the MDD and control groups. This research thoroughly examined the pathophysiology of MDD using these in-depth studies, paying particular attention to the role that inflammatory variables play in the development of the illness. This research advanced the current understanding of MDD by identifying specific inflammatory-related hub genes that could serve as critical biomarkers or therapeutic targets. These findings address a significant gap in the literature and provide a foundation for future studies aimed at developing more effective diagnostic and treatment strategies for MDD.

STATEMENT OF ETHICS

Not applicable.

ACKNOWLEDGMENTS

This work was supported by Special subject of scientific and technological research on traditional Chinese medicine and ethnic medicine in Guizhou (QZYY-20121-121); Guizhou Provincial Health Commission Project (gzwkj2024-132) and The Doctoral Initiation Fund Project of the First Affiliated Hospital of Guizhou University of Traditional Chinese Medicine [GYZYYFY-BS-2023(02)].

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

AI ASSISTANCE DISCLOSURE

Not applicable.

REFERENCES

1. Xue J, Suarez JS, Minaai M, Li S, Gaudino G, Pass HI, et al. HMGB1 as a therapeutic target in disease. J Cell Physiol. 2021 May;236(5):3406-19.
2. Pearson G, Robinson F, Beers Gibson T, Xu BE, Karandikar M, Berman K, et al. Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions. Endocr Rev. 2001;22(2):153-83.
3. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007;23(14):1846-7.
4. Spijker S, Van Zanten JS, De Jong S, Penninx BW, van Dyck R, Zitman FG, et al. Stimulated gene expression profiles as a blood marker of major depressive disorder. Biol Psychiat. 2010;68(2):179-86.
5. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Curr Protoc Bioinformatics. 2016;54:1-30.
6. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28(6):882-3.
7. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
8. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res. 2019;47(D1):D419-26.
9. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27-30.
10. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics. 2012;16(5):284-7.
11. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. P Natl Acad Sci Usa. 2005;102(43):15545-50.
12. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011;27(12):1739-40.
13. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607-13.
14. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-504.
15. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014;8 Suppl 4(Suppl 4):S11.
16. Yang X, Li Y, Lv R, Qian H, Chen X, Yang CF. Study on the Multitarget Mechanism and Key Active Ingredients of Herba Siegesbeckiae and Volatile Oil against Rheumatoid Arthritis Based on Network Pharmacology. Evid-Based Compl Alt. 2019;2019:8957245.
17. Vlachos IS, Paraskevopoulou MD, Karagkouni D, Georgakilas G, Vergoulis T, Kanellos I, et al. DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. Nucleic Acids Res. 2015;43(Database issue):D153-9.
18. Zhang Q, Liu W, Zhang HM, Xie GY, Miao YR, Xia M, et al. hTFtarget: A Comprehensive Database for Regulations of Human Transcription Factors and Their Targets. Genom Proteom Bioinf. 2020;18(2):120-8.
19. Zhou KR, Liu S, Sun WJ, Zheng LL, Zhou H, Yang JH, et al. ChIPBase v2.0: decoding transcriptional regulatory networks of non-coding RNAs and protein-coding genes from ChIP-seq data. Nucleic Acids Res. 2017;45(D1):D43-50.
20. Grondin CJ, Davis AP, Wiegers JA, Wiegers TC, Sciaky D, Johnson RJ, et al. Predicting molecular mechanisms, pathways, and health outcomes induced by Juul e-cigarette aerosol chemicals using the Comparative Toxicogenomics Database. Curr Res Toxicol. 2021;2:272-81.
21. Xiao B, Liu L, Li A, Xiang C, Wang P, Li H, et al. Identification and Verification of Immune-Related Gene Prognostic Signature Based on ssGSEA for Osteosarcoma. Front Oncol. 2020;10:607622.