Identification and Experimental Validation of Parkinson's Disease with Major Depressive Disorder Common Genes

Huiqing Wang
Southeastern University
Shanshan Dou
Jining Medical University
Wenming Gao
Jining Medical University
Baohua Cheng
Jining Medical University
Fuling Yan (jyfykong@163.com)
Southeast University Medical College https://orcid.org/0000-0001-5632-8679

Research Article

Keywords: Parkinson's disease, major depressive disorder, neutrophil, immune process, bioinformatics

Posted Date: December 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2377915/v1

Version of Record: A version of this preprint was published at Molecular Neurobiology on July 7th, 2023.
See the published version at https://doi.org/10.1007/s12035-023-03451-3.

Abstract

Parkinson's disease (PD) is the second most common neurodegenerative disease that affects about 10 million people worldwide. Non-motor and motor symptoms usually accompany PD. Major depressive disorder (MDD) is one of the non-motor manifestations of PD it remains unrecognized and undertreated effectively. MDD in PD has complicated pathophysiologies and remains unclear. The study aimed to explore the candidate genes and molecular mechanisms of PD with MDD. PD (GSE6613) and MDD (GSE98793) gene expression profiles were downloaded from Gene Expression Omnibus (GEO). Above all, the data of the two datasets were standardized separately, and differentially expressed genes (DEGs) were obtained by using the Limma package of R. Take the intersection of the two differential genes and remove the genes with inconsistent expression trends. Subsequently, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were investigated to explore the function of the common DEGs. Additionally, the construction of the protein-protein interaction (PPI) network was to search the hub genes, and then the least absolute shrinkage and selection operator (LASSO) regression was used to further identify the key genes. GSE99039 for PD and GSE201332 for MDD were performed to validate the hub genes by the violin plot and receiver operating characteristic (ROC) curve. Last but not least, immune cell dysregulation in PD was investigated by immune cell infiltration. As a result, a total of 45 common genes with the same trend. Functional analysis revealed that they were enriched in neutrophil degranulation, secretory granule membrane, and leukocyte activation. LASSO was performed on 8 candidate hub genes after CytoHubba filtered 14 node genes. Finally, AQP9, SPI1, and RPH3A were validated by GSE99039 and GSE201332. Additionally, the three genes were also detected by the qPCR *in vivo* model with the same phenomenon. The co-occurrence of PD and MDD can be attributed to AQP9, SPI1, and RPH3A genes. Neutrophils and monocytes infiltration play important roles in the development of PD and MDD. Novel insights may be gained from the findings for the study of mechanisms.

Introduction

Parkinson’s disease is the second most common neurodegenerative disorder with an increasing incidence rate worldwide(1). PD is characterized by age-related loss of dopamine-containing neurons in the substantia nigra pars compacta (SNpc) region and the formation of Lewy bodies which consisted of α-synuclein(2). These pathological characteristics manifested with motor symptoms of bradykinesia, rigidity, postural instability, and static tremor, as well as non-motor symptoms, such as sleep disturbances, depression, hyposmia, and constipation(3). Studies have reported that non-motor symptoms are earlier and more serious than motor symptoms(4; 5). Therefore, it is important to pay early attention to the non-motor symptoms.

Depression(Major depressive disorder, MDD) is one of the non-motor symptoms and is considered the most common neuropsychiatric disorder with PD. An increase in death risk is associated with MDD due to poor cognitive performance, poor quality of life, and a worse functional status (6). Studies have demonstrated that patients with PD are more susceptible to MDD compared with healthy controls(7). Cross-sectional studies have found that 2.7–90% of PD patients have MDD(8). It is crucial to diagnose

PD patients early since MDD significantly impacts their quality of life. However, it is challengable to diagnose timely and accurate MDD in patients with PD, because many features of depression, such as weight loss, insomnia, psychomotor retardation, and fatigue, overlap with the primary symptoms of PD or with the side effects of medication(8).

There are several theories about how MDD occurs in PD, although the exact mechanism remains unclear(9). It is widely accepted that MDD is caused by monoamine oxidase. PD patients with MDD postmortem had substantial destruction in the noradrenaline-producing nucleus coeruleus of the brain. Additionally, morphological changes in the nucleus raphe were found in PD patients with MDD(10). It appears that MDD in PD is caused by the degeneration of monoaminergic neurotransmitter systems and dysfunction of the fronto-cortical system. Besides, the dysfunction of the orbitofrontal cortex is induced by the degeneration of dopaminergic mesocortical and mesolimbic neurons, which affects serotonergic cell bodies in the dorsal raphe nuclei secondarily(11). Neuroinflammation is another aspect that affects the production of serotonin which is linked to MDD and neurodegeneration(12). Above all, the distinctive mechanisms that explain the co-occurrence of PD and MDD remain unclear.

The present study analyzed whole blood transcriptome expression data (GSE98793 and GSE6613) from MDD and PD patients in the GEO database to explore their common molecular mechanisms. Furthermore, gene modules were analyzed by the construction of protein-protein interaction (PPI) nodes and using the search tool for the retrieval of the interacting genes/proteins (STRING) database and Cytoscape software to identify hub genes. LASSO regression was used to further identify the key genes. Finally, hub genes were validated by GSE99039(PD) and GSE201332(MDD). Based on an analysis of independent datasets (GSE99039), we verified the expression level of hub genes and looked at the correlation of PD clinical characteristics. The results of this study will provide new insights into PD and MDD pathogenesis mechanisms.

Materials And Methods

Data collection and preprocessing

GSE6613(13; 14)and GSE99039(15)) of PD datasets as well as (GSE98793(16)) and (GSE201332(17)) of MDD datasets were downloaded from the GEO (https://www.ncbi.nlm.nih.gov/geo/) database(18). Detailed information on the 4 datasets is shown in Table 1, and a flow chart of the study design is shown in Fig. 1. The R software 4.0.5 was used to normalize, log2 transform and convert gene names for each dataset.

Identification Of Differentially Expressed Genes

Microarray data were downloaded from GSE6613 and GSE98793, and box plots and PCA analysis were used to represent the normalized expression matrix. Multiple probes identifying the same gene were averaged to determine its expression level. |log2 Fold change (FC)| > 0 (PD filtration) or |log2 Fold change

(FC)|> 0.585 (MDD filtration) and *p*-value < 0.05 were set as the criteria for identifying DEGs using the “Limma” package of R software(19). A variety of heat maps, volcano maps, and box line plots were created using the 'heatmap' and 'ggplot2' packages of R software (version 3.6.3)(20).

Functional Enrichment Analysis Of Common Degs

Gene Ontology (GO) is a structured, computerized approach to describing the functions of genes and genes products, including cellular components (CC), biological processes (BP), and molecular functions (MF)(21). Additionally, biochemical mechanisms and functions are discovered by using KEGG pathway enrichment analysis (22). An analysis of functional enrichment based on the Sangerbox platform(http://vip.sangerbox.com/) was conducted.

Protein-protein Interaction(Ppi) Network Construction

Based on the identified common DEGs with the same expression trend, STRING(https://string-db.org/) is a database, and used to construct the PPI network with combined scores greater than 0.4(23). The PPI network was then constructed and presented on the Cytoscape platform(24). The significant modules and core genes were identified using CytoHubba, a plugin in Cytoscape. Four different algorithms (Maximum Neighborhood Component (MNC), Maximal Clique Centrality (MCC), DEGREE, and Density of Maximum Neighborhood Component (DNMC)) were used to identify hub genes. Lastly, a Venn diagram was used to demonstrate that the common genes obtained by the four algorithms were reliable hub genes.

Identification Of Common Hub Genes Using Lasso Logistic Regression

LASSO analysis is a regression method that improves prediction accuracy by selecting a variable from high dimensional data that has a strong predictive value and low correlation(25). The hub genes were further identified by LASSO analysis. LASSO logistic regression was built using hub gene expression levels and clinical traits. To differentiate patients from healthy controls, ROC curves were used using the "pROC" R package(26). Moreover, the hub genes were also represented by box line plots.

Validation Of Hub Gene Expression

GSE99039 for PD(15) and GSE201332 for MDD(17) were used to validate the expression levels of the identified hub genes. GSE99039 contains 205 PD and 233 control samples. GSE201332 contains 20 MDD and 20 control samples. The expression difference of hub genes was also represented by a box line plot and ROC curve. Wilcoxon test was used to compare the two datasets. P value < 0.05 was considered significant.

Animals

Male C57BL/6 mice weighing 23-27g at the age of 9–11 weeks were purchased from Pengyue experimental animal Ltd. (Jinan, Shandong, PR China) and acclimated for at least 1 week. All mice were housed under controlled lighting conditions in a 12/12 hours' light/dark cycle and a comfortable temperature of 22–26°C with water available ad libitum but the food required our experiments. Additionally, all animal care and experimental procedures were approved by the Ethics Committee. These animals' conditions are consistent with our previous studies(27).

Mptp Treatment

The *in vivo* model of PD was administrated with 1-methyl-4-phenyl-1,2,3,6- tetrahydropyran (MPTP). MPTP-treated groups received four intraperitoneal (i.p.) injections of MPTP, HCl (20 mg/kg free base) in saline at 2 h intervals in a single day(28). The sham groups were treated with sterile saline in the same dose.

Chronic Unpredictable Mild Stress (Cums) Model

The model of depression was processed with CUMS, the protocol was described as the previous study(29). The animals in various groups were unpredictably subjected to mild stress for 4 weeks except for those animals in the vehicle + no stress group. Individual stressors and the length of time they were applied each day were as follows: (i) food deprivation; (ii) water deprivation; (iii) restraint; (iv) restraint at 4◦C; (v) flashing light for 120–210 min; and (vi) isolation. Stressor stimuli were applied at different times every day, to minimize their predictability.

Immunohistochemistry

Frozen sections of the SN*pc* were conducted as the protocol of the immunohistochemistry kit as our previous study described(27). The number of TH-positive neurons represented dopaminergic degeneration.

Forced Swimming Test (Fst)

FST was conducted followed by the previous study(30). In this study, the FST was carried out to evaluate depressive-like behavior associated with stress or MPTP effects. The FST apparatus consisted of a glass cylinder having a 25 cm diameter and a 35 cm height. Water was maintained at a temperature between 22 and 26°C. The various groups were subject to FST for 6 min to evaluate the immobility, climbing, and swimming times.

Rna Isolation And Quantitative Polymerase Chain Reaction (Qpcr)

Total RNA was extracted from the peripheral blood of multiple groups according to the manufacturer's protocol [Trizol method and miRNeasy Mini kit (Qiagen)]. To reverse transcribe total RNA, the HiScript Q RT SuperMix for qPCR Kit (Vazyme, R123-01) was used. In this experiment, QuantStudio 6 (Applied Biosystems) was used to perform quantitative PCR to measure SYBR Green levels in real-time PCR using the manufacturer's recommended cycling conditions. In the analysis, duplicates of samples with Cq values over 35 were removed. The primers were provided in Table 1. The $2-\Delta\Delta Ct$ method was applied to calculate the relative expression level of mRNA.

Gene Set Enrichment Analysis

Based on gene expression levels for population phenotypes, GSEA (http://software.broadinstitute.org/gsea/index.jsp) was used to assess pathway and molecular mechanisms relationships between the two groups(31). Enriched gene sets with nominal P values of < 0.05, |normalized enrichment scores (NES) | > 1, and false positive rate (FDR) q values of < 0.25 were considered significant.

Immune Infiltration Analysis

CIBERSORT, a computational method for estimating the proportion of diverse immune cells based on gene expression profiles, was used to assess immune cell proportion in patients and healthy control(32). Infiltration analysis using the "Cibersort" R package was performed. Bar plots were used to visualize the proportion of different types of immune cells. Box plots were used to compare the proportions of different types of immune cells in two datasets. The “corrplot” R package was used to create a heatmap illustrating the correlation between 22 types of infiltrating immune cells(33).

Statistical Analysis

All statistical analyses were performed using R software. *P*-value < 0.05 was considered statistically significant.

Results

Identification of Common DEGs between PD and MDD

The expression matrices of the two datasets GSE6613 and GSE98793 were normalized, and box plots showed straight lines as distribution trends (Fig. 2A-D). Data repeatability between the two datasets was assessed using PCA in this study, and the results showed good repeatability between the two datasets(Fig. 2E-2F).

For the GSE6613 dataset, according to the Limma R method, 797 DEGs were identified in the combined PD dataset, of which 420 were upregulated and 377 were downregulated, and a total of 2145 DEGs were identified as DEGs in MDD including 1384 were upregulated and 761 were downregulated. As shown in Fig. 3A, The volcano plot of PD DEGs was shown. Furthermore, the top 50 upregulated concurrently with the top 50 downregulated DEGs were shown using heatmaps(Fig. 3B-C). Regarding the MDD dataset, the volcano plot and heatmaps were shown in Fig. 3D-F. Besides, The intersection of the Venn diagram yielded 94 common DEGs (Fig. 4A). However, only 45 common DEGs with the same trend. The heatmaps of 45 common DEGs in the PD and MDD datasets were shown in Fig. 4B-C.

Enrichment Analysis Of Common Degs

To investigate the biological functions of the common DEGs, we performed KEGG pathways and GO terms enrichment analyses. Based on the KEGG enrichment analysis, common genes were primarily enriched in the “Metabolic signaling pathway,” “Osteoclast differentiation,” and “Steroid biosynthesis”(Fig. 5A-B). Furthermore, GO analysis showed that genes were enriched in “neutrophil degranulation”, “neutrophil activation involved in immune response” and “B cell activation involved in immune response” (BP); “cytoplasmic vesicle lumen”, “secretory granule membrane” and “ficolin-1 rich granule” (CC); “immune receptor activity”, guanyl-nucleotide exchange factor activity” and “protein phosphatase binding” (MF) (Fig.C-F). As a result of these findings, the progression of both diseases appears to be mediated by inflammatory pathways as well as metabolism pathways.

Figure 5 Enrichment analysis of the intersection of genes. (A) Analysis of the intersection of genes using KEGG pathways. Different colors represent different significant pathways and elated enriched genes. (B) Genes associated with KEGG pathways in a bubble plot. (C) Gene intersection analysis using GO terms. Colors represent significant terms and related enriched genes. (D-F) GO-BP, GO-CC, and GO-MF analysis for the common DEGs showing the specific genes associated with these terms through cent plots.

Ppi Network Analysis And Identification Of Hub Gene

Identifying hub genes that may be involved in the cooccurrence of PD and MDD. STRING was used to analyze PPI networks of communal DEGs to clarify interactions between them (Fig. 6A). As shown in Fig. 6B, modular gene enrichment analysis and significant gene modules (Fig. 6B). MCC, MNC, Degree, and DNMC were used to predict and explore the top 15 hub genes in the PPI network using CytoHubba. 14 candidate hub genes were identified from the intersection of the 15 genes from the four algorithms: S100A11, COASY, SNCA, FECH, UGP2, CXCR1, AQP9, CSF3R, FDFT1, CFP, LILRB2, RPH3A, CYTH4, SPI1(Fig. 6C). Moreover, GO and KEGG enrichment analysis was performed on the 14 hub genes (Fig. 6D-E).

Identification Of Candidate Hub Genes Via Lasso Logistic Regression

Nomogram construction and diagnostic value evaluation were performed using LASSO logistic regression. As shown in Fig. 7A-B, eight potential biomarkers were identified using the LASSO regression algorithm including CYTH4, AQP9, SPI1, CFP, RPH3A, FECH, UGP2, and FDFT1. Multivariate Cox regression analysis revealed that this prognostic model was an independent prognostic parameter of PD (Fig. 7C). Additionally, the heatmap showed the expression of the candidate genes according to the risk score (Fig. 7D). Through the LASSO analysis, the candidate genes were further validated.

Validation Of Hub Genes For Pd With Mdd

The expression of the screened eight genes was tested with GSE99039 of PD to improve the dependability of hub genes. As shown in Fig. 8A, AQP9, CFP, RPH3A, CYTH4, and SPI1 were consistent with the above study. In these three datasets, ROC curves were drawn using the R package with the expression of the four hub genes to assess the diagnostic accuracy. Five candidate genes possess a high diagnostic value (Fig. 8B-D). Additionally, we further applied the external validation of the MDD dataset with GSE201332. The violin plot showed that the expression of AQP9, RPH3A, and SPI1 were differentially expressed in the MDD dataset (Fig. 8E). Meanwhile, the ROC curve of three genes was established to have a significant value for diagnosing MDD. Above all, these results suggested that AQP9, RPH3A, and SPI1 could be promising markers for diagnosing PD and MDD.

Mptp And Cums Model

Similar to some previous studies(27; 34), MPTP-induced neurotoxicity of dopaminergic neurons exhibited as TH-positive neurons. As shown in Fig. 9A, the MPTP group significantly showed the loss of TH-positive neuron populations in the *SNpc* compared to the sham group. Additionally, the MPTP associated with the stress group also showed the same phenomenon. In other words, it also indicated the success of the PD model *in vivo*. As shown Fig. 9B-D is representing data on immobility, climbing, and swimming times in the forced swimming test. In the stress groups, the time of climbing and swimming was significantly decreased while the immobility time was enhanced as well as interaction for stress and MPTP.

Validation Of Hub Genes Using Qpcr

The transcriptional changes of overlapped hub genes AQP9, RPH3A, and SPI1 were detected in the peripheral blood from the MPTP associated with stress groups and control groups by quantitative qPCR. The results indicated that the expression levels of AQP9, RPH3A, and SPI were both increased in the MPTP associated with stress groups in comparison with those in controls (Fig. 10A, B), which was in line with the bioinformatics analysis.

Correlation Analysis With Clinical Features

For PD patients, the Unified Parkinson's Disease Rating Scale (UPDRS) (Fig. 11A–L) provides key indicators of the severity of the disease(35). AQP9 expression levels were positively correlated with UPDRS I (r = 0.11, p < 0.02); UPDRS II (r = 0.21, p < 0.01); UPDRSIII (r = 0.14, $p < 0.01$); UPDRS IV (r = 0.20, $p < 0.01$). RPH3A expression levels were positively correlated with UPDRS I (r = 0.16, $p < 0.01$); UPDRS II (r = 0.14, $p < 0.01$; UPDRSIII (r = 0.04, $p = 0.41$); UPDRS IV (r = 0.25, $p < 0.01$). SPI1 expression levels were positively correlated with UPDRS I (r = 0.26, $p < 0.01$); UPDRS II (r = 0.21, $p < 0.01$; UPDRSIII (r = 0.25, $p < 0.01$); UPDRS IV (r = 0.29, $p < 0.01$). There is a direct link between the severity of PD and AQP9, RPH3A, and SPI1.

Gsea Results Of Hub Genes

Based on GSEA analysis of GSE6613 and GSE98793, differentially regulated pathways between the high and low expression groups were identified to determine the potential functions of hub genes in PD and MDD. GSEA analysis was conducted for the gene sets ‘PANTOTHENATE_AND_COA_BIOSYNTHESIS“GRAFT_VERSUS_HOST_DISEASE’ ‘PROTEASOME’ ‘CHRONIC_MYELOID_LEUKEMIA’ ‘NEUROTROPHIN_SIGNALING_PATHWAY’.

Construction Of Putative Hub Genes Protein-protein Interaction Network And Tf-gene Interaction Network

By employing the tool ‘GeneMANIA’, we constructed a putative PPI network of 20 genes involving hub genes APQ9, RPH3A, and SPI1 (Fig. 13A). Inner circle contains hub genes, while the outer circle contains predicted genes. We visualized the interaction network between candidate genes and TF genes using NetworkAnalyst (http://www.networkAnalyx.ca/). Human-mouse TF target interaction data are currently managed manually in the TRRUST database (https://www.grnpedia.org/trrust/). Submitting target genes can be used to search for key TF(Fig. 13B).

Immune Cell Infiltration Analysis

Microenvironmental factors include immune cells, extracellular matrix, inflammatory factors, and growth factors that influence clinical therapeutic sensitivity and disease diagnosis. The CIBERSORT algorithm was used to estimate the proportion of 22 immune cells in GSE6613 and GSE98793, as shown in the graphs in Figs. 14A and 14B. The correlation of 22 types of immune cells revealed that CD8 T cells were negatively associated with Neutrophils (r = − 0.69) and that naive B cells were negatively related to Plasma cells (r = − 0.61), whereas CD8 T cells were negatively related to naive CD4 T cells (r = − 0.57) in GSE6613 dataset. In addition, CD8 T cells were negatively associated with Neutrophils (r = − 0.67) and that naive B cells were negatively related to memory B cells (r = − 0.69), whereas CD8 T cells were negatively related to naive CD4 T cells (r = − 0.61) in GSE98793 dataset (Fig. 12C-D). The immune cell infiltration of two datasets was compared in boxplots (Fig. 12E-F). The results showed that in the PD

groups only follicular helper T cells were lower than in the control group in the PD dataset. However, naïve B cells, plasma cells, CD8 T cells, and the other five cells were significantly different between the MDD groups and the control groups.

Besides, it was found that all hub genes were significantly correlated with neutrophil cells with $p < 0.05$ in GSE6613 (AQP9, r = 0.66, p = 2.2e-10; RPH3A, r = 0.27, p = 0.02; SPI1, r = 0.40, p = 4.9e-4, Fig. 13A-C), while only AQP9 showed a remarkable correlation with neutrophil cells in GSE98793(AQP9, r = 0.22, p = 0.01; RPH3A, r = 0.12, p = 0.19; SPI1, r = 0.13, p = 0.13, Fig. 13D-F). Interestingly, all hub genes exhibited a significant correlation with monocytes with $p < 0.05$ in GSE6613 (AQP9, r = 0.30, p = 5.0e-4; RPH3A, r = 0.17, p = 0.06; SPI1, r = 0.36, p = 3.0e-5, Fig. 13G-I). These results suggested that the neutrophil cells and monocytes may play important roles in PD with MDD.

Discussion

As a progressive neurodegenerative disease, Parkinson's disease(PD) is characterized by bradykinesia, tremors at rest, rigidity, and postural instability. Depression, insomnia, constipation, and hyposmia, however, are not considered motor symptoms and often precede them(36). The one of most common psychiatric symptoms of PD is depression (major depressive disorder, MDD), which has not been paid high attention to(37). A study of meta-analyses found that MDD occurs in around 23% of patients with PD, which is higher than that seen in other chronic and disabling conditions(8; 38). The incidence of MDD will increase exponentially as PD progress (39). MDD is rarely recognized as a manifestation, which delays diagnosis and treatment, and ultimately results in poor living conditions for patients. However, the associations and biomarkers of PD with MDD have not yet been clarified.

In the present study, we used bioinformatics analysis to construct a nomogram to assess MDD in PD patients and evaluated its diagnostic value. Public datasets of peripheral whole blood from PD and MDD patients are an efficient and practical method for clinical use. In addition to identifying three pivotal candidate genes (AQP9, RPH3A, and SPI1), we developed a nomogram for diagnosing MDD in PD patients.

Aquaporins 9(AQP9), a novel biomarker found in our study, was used in the diagnosis of MDD in patients with PD. It belongs to the family of membrane proteins that mediate the transport of water(40). AQP9 is the aquaporin most closely related to GlpF, a bacterial aquaglyceroporin that fluxes water as well as glycerol(41). AQP9 is the first mitochondrial protein with two novel isoforms in brain tissue that is selectively expressed in dopaminergic neurons(42). Studies in vitro showed that AQP4 and AQP9 transcription and expression changed dynamically after differentiation toward DAergic neurons, resulting in the vulnerability of the human SH-SY5Y cell line(43). Additionally, in vivo study demonstrated that the deletion of the aquaglyceroporin AQP9 is protective in a mouse model of Parkinson’s disease(44). Besides, AQP9 is also involved in the pathogen of epilepsy(45). Above all, AQP9 plays an important role in the central nervous system including PD.

Rabphilin-3A (RPH3A) is a synaptic protein initially known as a synaptic vesicle-associated protein involved in the regulation of exo- and endocytosis processes at presynaptic sites(46). RPH3A retains NMDA receptors at synaptic sites through interaction with GluN2A/PSD-95 complex(47). The previous study has suggested that interfering with the formation of the RPH3A/GluN2A/PSD-95 complex could lead to a more specific and direct approach to tackling the aberrant NMDAR localization and function in levodopa (L-DOPA)-induced dyskinesias(48). Interestingly, a recent study has demonstrated that RPH3A overexpression improves motor behavior in the PD mice model, and is a novel molecular target to counteract α-syn-induced synaptic failure (49). Besides, a study has suggested that the increase in Rph3A in the brain penumbra may be an endogenous protective mechanism against ischemia-reperfusion injury, which is mainly dominated by astrocytes(50). RPH3A loss correlated with dementia severity, cholinergic deafferentation, and increased β-amyloid (Aβ) concentrations(51). Additionally, RPH3A has also been reported that it associated with cognitive resilience via a proteome-wide association study of the human dorsolateral prefrontal cortex(52). Briefly, RPH3A was involved in multiple neuropsychiatric diseases, which may as a target for PD with MDD.

SPI-1 proto-oncogene (SPI1, also known as PU.1) was found by Moreau-Gcherin et al, known as one of the erythrocyte transformation-specific (Ets) family of transcription factors (TF). A meta-analysis of the PD public dataset showed that SPI1 is the TF of the hub gene(53). Meanwhile, SPI1 is also an important TF of the hub gene in PD with bipolar disorder(54). Besides, a study has demonstrated that SPI1 is a key gene associated with postpartum depression (PPD)(55). In addition, abnormalities in SPI1 in the brain-spleen axis might, in part, play a role in the pathophysiology of MDD(56). Therefore, SPI may as a key gene in PD with MDD.

One limitation of the present study was the lack of the common hub genes of PD and MDD that have been validated in patients with either disease, but not both. As PD and MDD patients are not included in the current datasets, a more comprehensive validation can be attempted in the future.

Conclusion

AQP9, RPH3A, and SPI1 have been identified as common hub genes in PD and MDD in this study. Neutrophil and monocyte infiltration plays a central role in the development of PD and MDD, suggesting they may be potential targets for diagnosis and treatment.

Declarations

CRediT authorship contribution statement

HQW analyzed the data and drafted manuscripts. SSD and WMG participated in the revision of the manuscript and figures. BHCand FLYdesigned a significant research topic and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding:

The present work was supported by the National Nature Science Foundation of China(No. 82071325).

Declaration of Competing Interest

The authors declare no conflict of interest.

Data availability:

The data that support the findings of this study are available from public databases.

Consent to Participate

All participants involved in this study provided written informed consent.

Consent for Publication

All co-authors approved the final version of the manuscript and agreed to submit it to Molecular Neurobiology.

Acknowledgments

Not applicable

References

1. Chagas MH, Linares IM, Garcia GJ, Hallak JE, Tumas V, Crippa JA (2013) Neuroimaging of depression in Parkinson's disease: a review. Int Psychogeriatr 25:1953–1961
2. Goedert M, Compston A (2018) Parkinson's disease - the story of an eponym. Nat reviews Neurol 14:57–62
3. Schapira AHV, Chaudhuri KR, Jenner P (2017) Non-motor features of Parkinson disease. Nat Rev Neurosci 18:435–450
4. Berg D, Postuma RB, Adler CH, Bloem BR, Chan P, Dubois B et al (2015) MDS research criteria for prodromal Parkinson's disease. Mov disorders: official J Mov Disorder Soc 30:1600–1611
5. van Uem JM, Marinus J, Canning C, van Lummel R, Dodel R, Liepelt-Scarfone I et al (2016) Health-Related Quality of Life in patients with Parkinson's disease–A systematic review based on the ICF model. Neurosci Biobehav Rev 61:26–34
6. Schrag A, Taddei RN (2017) Depression and Anxiety in Parkinson's Disease. Int Rev Neurobiol 133:623–655
7. Becker C, Brobert GP, Johansson S, Jick SS, Meier CR (2011) Risk of incident depression in patients with Parkinson disease in the UK. Eur J Neurol 18:448–453
8. Goodarzi Z, Mrklas KJ, Roberts DJ, Jette N, Pringsheim T, Holroyd-Leduc J (2016) Detecting depression in Parkinson disease: A systematic review and meta-analysis. Neurology 87:426–437

9. Hussain M, Kumar P, Khan S, Gordon DK, Khan S (2020) Similarities Between Depression and Neurodegenerative Diseases: Pathophysiology, Challenges in Diagnosis and Treatment Options. Cureus 12:e11613
10. Lemke MR, Fuchs G, Gemende I, Herting B, Oehlwein C, Reichmann H et al (2004) Depression and Parkinson's disease.Journal of neurology251 Suppl 6:Vi/24 – 7
11. Mayberg HS, Solomon DH (1995) Depression in Parkinson's disease: a biochemical and organic viewpoint. Adv Neurol 65:49–60
12. Wirleitner B, Neurauter G, Schröcksnadel K, Frick B, Fuchs D (2003) Interferon-gamma-induced conversion of tryptophan: immunologic and neuropsychiatric aspects. Curr Med Chem 10:1581–1591
13. Scherzer CR, Eklund AC, Morse LJ, Liao Z, Locascio JJ, Fefer D et al (2007) Molecular markers of early Parkinson's disease based on gene expression in blood. Proc Natl Acad Sci USA 104:955–960
14. Scherzer CR, Grass JA, Liao Z, Pepivani I, Zheng B, Eklund AC et al (2008) GATA transcription factors directly regulate the Parkinson's disease-linked gene alpha-synuclein. Proc Natl Acad Sci USA 105:10907–10912
15. Shamir R, Klein C, Amar D, Vollstedt EJ, Bonin M, Usenovic M et al (2017) Analysis of blood-based gene expression in idiopathic Parkinson disease. Neurology 89:1676–1683
16. Leday GGR, Vértes PE, Richardson S, Greene JR, Regan T, Khan S et al (2018) Replicable and Coupled Changes in Innate and Adaptive Immune Gene Expression in Two Case-Control Studies of Blood Microarrays in Major Depressive Disorder. Biol Psychiatry 83:70–80
17. Xiu J, Li J, Liu Z, Wei H, Zhu C, Han R et al (2022) Elevated BICD2 DNA methylation in blood of major depressive disorder patients and reduction of depressive-like behaviors in hippocampal Bicd2-knockdown mice. Proc Natl Acad Sci USA 119:e2201967119
18. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M et al (2013) NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res 41:D991–D995
19. Davis S, Meltzer PS (2007) GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinf (Oxford England) 23:1846–1847
20. Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinf (Oxford England) 32:2847–2849
21. The Gene Ontology Resource: 20 years and still GOing strong.Nucleic acids research47:D330-d8
22. Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M (2019) New approach for understanding genome variations in KEGG. Nucleic Acids Res 47:D590–d5
23. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J et al (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43:D447–D452
24. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ (2019) Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. J Proteome Res 18:623–632

25. Yang C, Delcher C, Shenkman E, Ranka S (2018) Machine learning approaches for predicting high cost high need patient expenditures in health care. Biomed Eng Online 17:131
26. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC et al (2011) pROC: an open-source package for R and S + to analyze and compare ROC curves. BMC Bioinformatics 12:77
27. Wang H, Dou S, Zhu J, Shao Z, Wang C, Cheng B (2020) Ghrelin protects dopaminergic neurons against MPTP neurotoxicity through promoting autophagy and inhibiting endoplasmic reticulum mediated apoptosis. Brain Res 1746:147023
28. Karuppagounder SS, Brahmachari S, Lee Y, Dawson VL, Dawson TM, Ko HS (2014) The c-Abl inhibitor, nilotinib, protects dopaminergic neurons in a preclinical animal model of Parkinson's disease. Sci Rep 4:4874
29. Willner P, Muscat R, Papp M (1992) Chronic mild stress-induced anhedonia: a realistic animal model of depression. Neurosci Biobehav Rev 16:525–534
30. Porsolt RD, Anton G, Blavet N, Jalfre M (1978) Behavioural despair in rats: a new model sensitive to antidepressant treatments. Eur J Pharmacol 47:379–391
31. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 102:15545–15550
32. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y et al (2015) Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 12:453–457
33. Hu K (2020) Become Competent within One Day in Generating Boxplots and Violin Plots for a Novice without Prior R Experience. Methods and protocols 3
34. Zhu J, Gao W, Shan X, Wang C, Wang H, Shao Z et al (2020) Apelin-36 mediates neuroprotective effects by regulating oxidative stress, autophagy and apoptosis in MPTP-induced Parkinson's disease model mice. Brain Res 1726:146493
35. Hoehn MM, Yahr MD (2001) Parkinsonism: onset, progression, and mortality. 1967. Neurology 57:S11-26
36. Ravina B, Camicioli R, Como PG, Marsh L, Jankovic J, Weintraub D et al (2007) The impact of depressive symptoms in early Parkinson disease. Neurology 69:342–347
37. Grover S, Somaiya M, Kumar S, Avasthi A (2015) Psychiatric aspects of Parkinson's disease. J neurosciences rural Pract 6:65–76
38. Chwastiak L, Ehde DM, Gibbons LE, Sullivan M, Bowen JD, Kraft GH (2002) Depressive symptoms and severity of illness in multiple sclerosis: epidemiologic study of a large community sample. Am J Psychiatry 159:1862–1868
39. Cooney JW, Stacy M (2016) Neuropsychiatric Issues in Parkinson's Disease. Curr Neurol Neurosci Rep 16:49
40. Agre P, King LS, Yasui M, Guggino WB, Ottersen OP, Fujiyoshi Y et al (2002) Aquaporin water channels–from atomic structure to clinical medicine. J Physiol 542:3–16

41. Maurel C, Reizer J, Schroeder JI, Chrispeels MJ, Saier MH Jr (1994) Functional characterization of the Escherichia coli glycerol facilitator, GlpF, in Xenopus oocytes. J Biol Chem 269:11869–11872
42. Amiry-Moghaddam M, Lindland H, Zelenin S, Roberg BA, Gundersen BB, Petersen P et al (2005) Brain mitochondria contain aquaporin water channels: evidence for the expression of a short AQP9 isoform in the inner mitochondrial membrane. FASEB journal: official publication of the Federation of American Societies for Experimental Biology 19:1459–1467
43. Avola R, Graziano ACE, Pannuzzo G, Albouchi F, Cardile V (2018) New insights on Parkinson's disease from differentiation of SH-SY5Y into dopaminergic neurons: An involvement of aquaporin4 and 9. Molecular and cellular neurosciences 88:212 – 21
44. Stahl K, Rahmani S, Prydz A, Skauli N, MacAulay N, Mylonakou MN et al (2018) Targeted deletion of the aquaglyceroporin AQP9 is protective in a mouse model of Parkinson's disease. PLoS ONE 13:e0194896
45. Salman MM, Sheilabi MA, Bhattacharyya D, Kitchen P, Conner AC, Bill RM et al (2017) Transcriptome analysis suggests a role for the differential expression of cerebral aquaporins and the MAPK signalling pathway in human temporal lobe epilepsy. Eur J Neurosci 46:2121–2132
46. Bourgeois-Jaarsma Q, Miaja Hernandez P, Groffen AJ (2021) Ca(2+) sensor proteins in spontaneous release and synaptic plasticity: Limited contribution of Doc2c, rabphilin-3a and synaptotagmin 7 in hippocampal glutamatergic neurons. Mol Cell Neurosci 112:103613
47. Stanic J, Carta M, Eberini I, Pelucchi S, Marcello E, Genazzani AA et al (2015) Rabphilin 3A retains NMDA receptors at synaptic sites through interaction with GluN2A/PSD-95 complex. Nat Commun 6:10181
48. Stanic J, Mellone M, Napolitano F, Racca C, Zianni E, Minocci D et al (2017) Rabphilin 3A: A novel target for the treatment of levodopa-induced dyskinesias. Neurobiol Dis 108:54–64
49. Ferrari E, Scheggia D, Zianni E, Italia M, Brumana M, Palazzolo L et al (2022) Rabphilin-3A as a novel target to reverse α-synuclein-induced synaptic loss in Parkinson's disease. Pharmacol Res 183:106375
50. Zhu X, Li H, You W, Yu Z, Wang Z, Shen H et al (2022) Role of Rph3A in brain injury induced by experimental cerebral ischemia-reperfusion model in rats. CNS Neurosci Ther 28:1124–1138
51. Tan MG, Lee C, Lee JH, Francis PT, Williams RJ, Ramírez MJ et al (2014) Decreased rabphilin 3A immunoreactivity in Alzheimer's disease is associated with Aβ burden. Neurochem Int 64:29–36
52. Yu L, Tasaki S, Schneider JA, Arfanakis K, Duong DM, Wingo AP et al (2020) Cortical Proteins Associated With Cognitive Resilience in Community-Dwelling Older Persons. JAMA psychiatry 77:1172–1180
53. Su L, Wang C, Zheng C, Wei H, Song X (2018) A meta-analysis of public microarray data identifies biological regulatory networks in Parkinson's disease. BMC Med Genom 11:40
54. Hossain MB, Islam MK, Adhikary A, Rahaman A, Islam MZ (2022) Bioinformatics Approach to Identify Significant Biomarkers, Drug Targets Shared Between Parkinson's Disease and Bipolar Disorder: A Pilot Study. Bioinform Biol insights 16:11779322221079232

55. Deng Z, Cai W, Liu J, Deng A, Yang Y, Tu J et al (2021) Co-expression modules construction by WGCNA and identify potential hub genes and regulation pathways of postpartum depression. Front bioscience (Landmark edition) 26:1019–1030
56. Zhang J, Chang L, Pu Y, Hashimoto K (2020) Abnormal expression of colony stimulating factor 1 receptor (CSF1R) and transcription factor PU.1 (SPI1) in the spleen from patients with major psychiatric disorders: A role of brain-spleen axis. J Affect Disord 272:110–115

Table

Table 1 is available in the Supplementary Files section

Figures

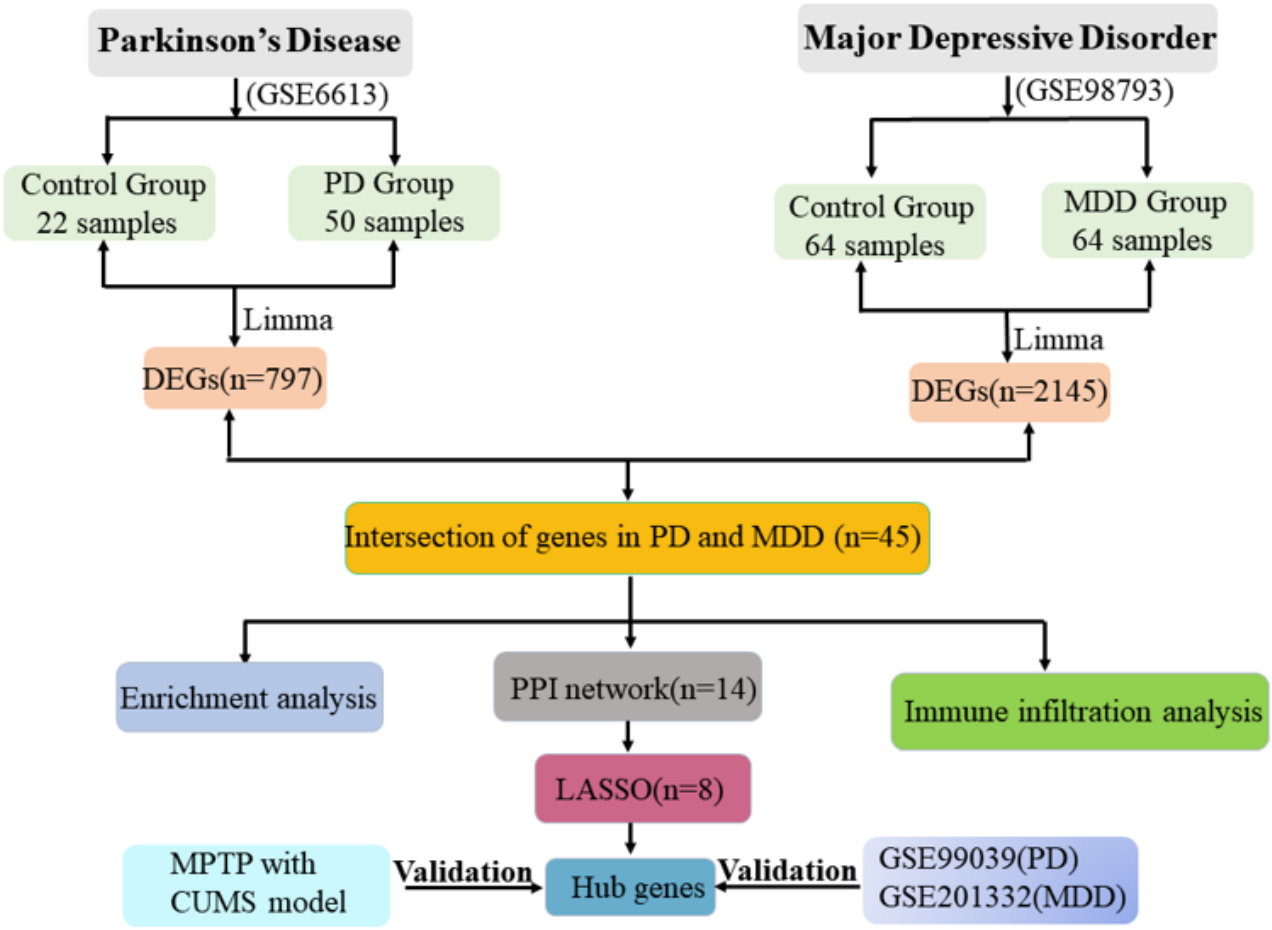

Figure 1

Flowchart of the study design. DEGs, differentially expressed genes; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein-protein interaction.

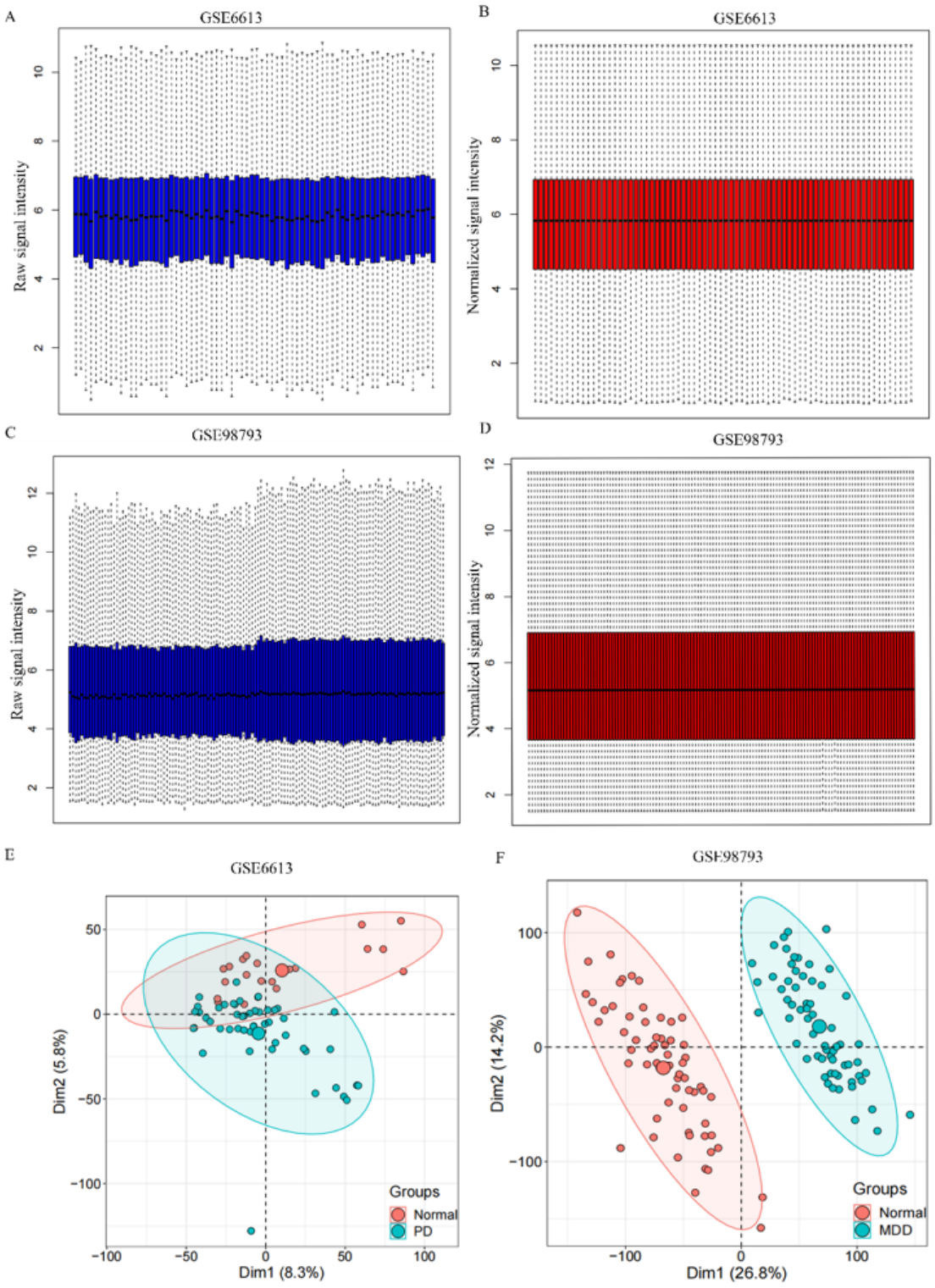

Figure 2

Raw and normalized expression matrices. (A-D) box plots and PCA diagrams (E-F) of the GSE6613 and GSE98793 datasets. PCA, principal component analysis.

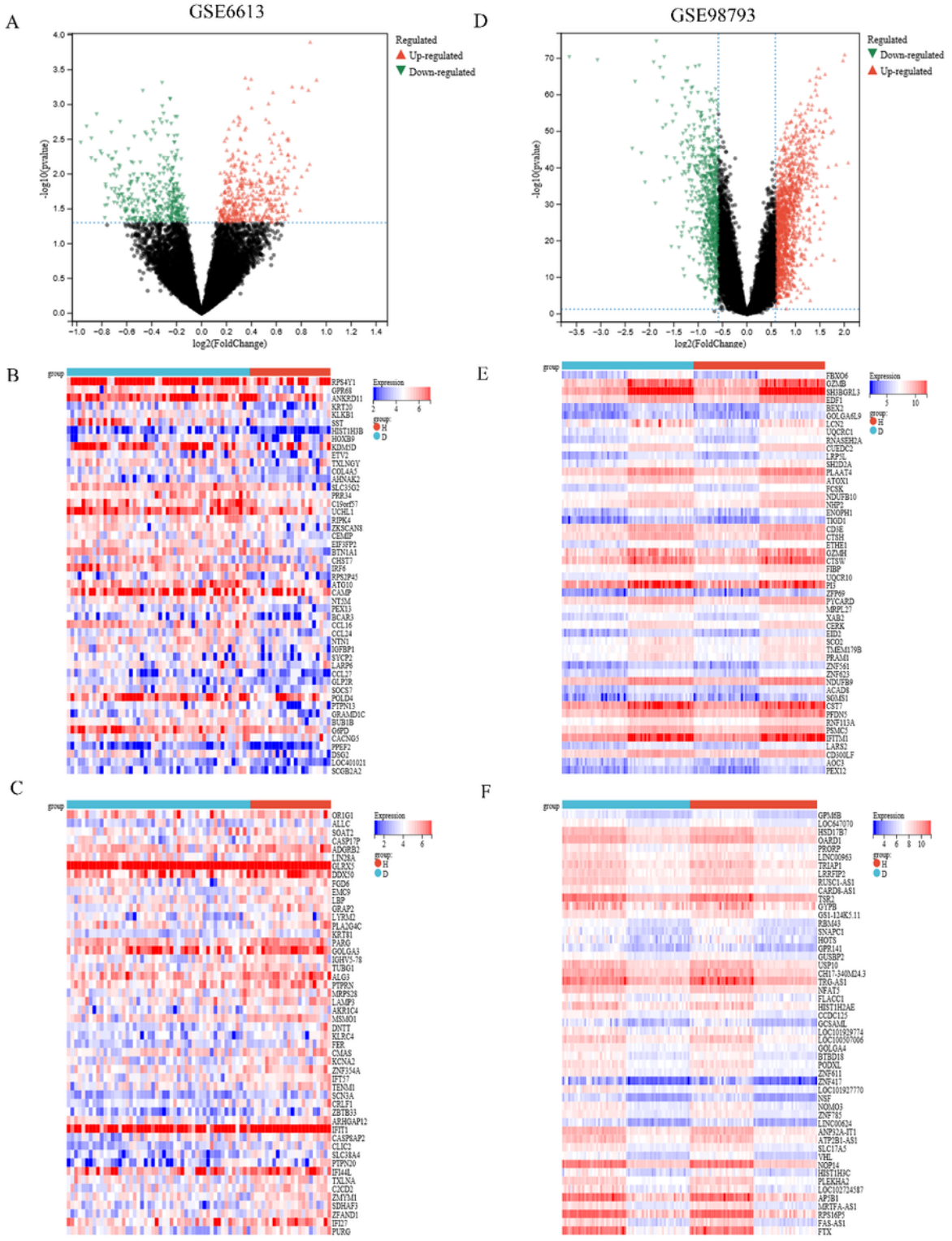

Figure 3

The expression and identification of differentially expressed genes in the two datasets. (A) Volcano map of differentially DEGs between in GSE6613. The threshold was set to |log2FC (fold change)| > 0, and *p*-value < 0.05. (B-C) The top 50 upregulated concurrently with the top 50 downregulated DEGs of the PD dataset(D) Volcano map of differential DEGs in GSE98793. The threshold was set to |log2FC (fold

change)| > 0.585, and *p*-value < 0.05. (E-F) The top 50 upregulated concurrently with the top 50 downregulated DEGs of the MDD dataset.

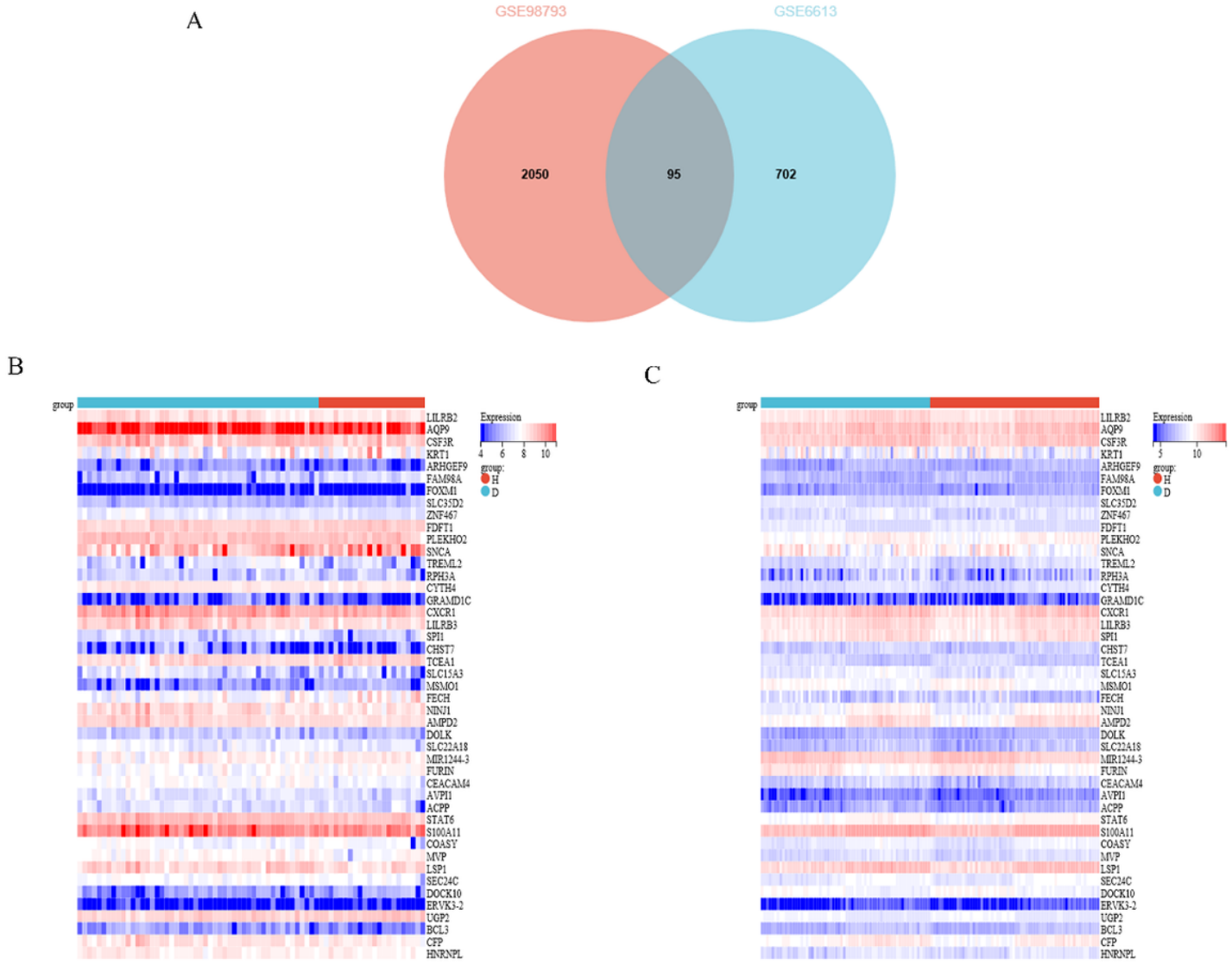

Figure 4

Identification of common DEGs. (A) The intersecting DEGs of the PD dataset and the MDD dataset are represented by a Venn diagram. (B-C) The heatmaps of the common DEGs with the same trend in GSE6613 and GSE98793.

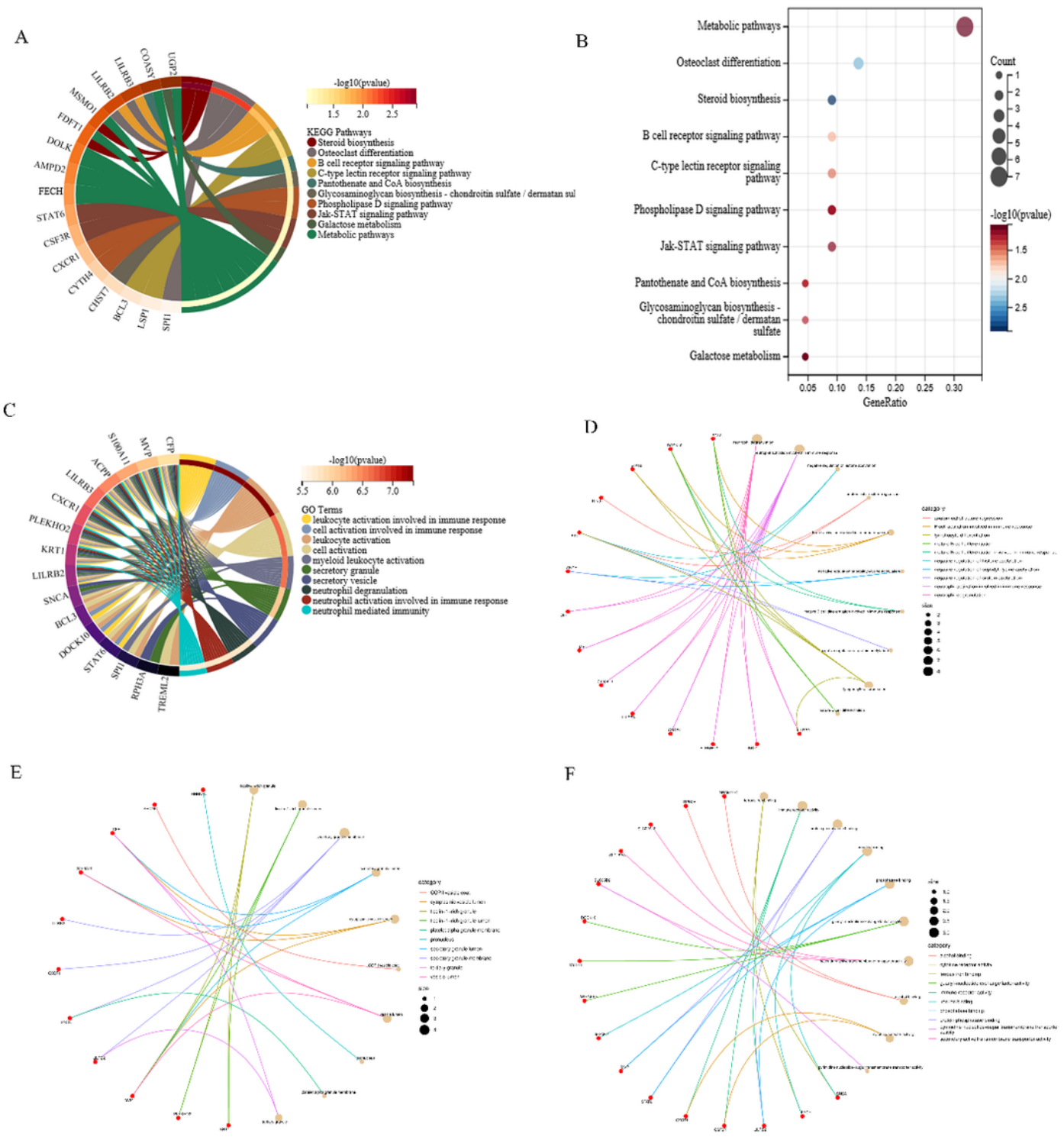

Figure 5

Enrichment analysis of the intersection of genes. (A) Analysis of the intersection of genes using KEGG pathways. Different colors represent different significant pathways and elated enriched genes. (B) Genes associated with KEGG pathways in a bubble plot. (C) Gene intersection analysis using GO terms. Colors represent significant terms and related enriched genes. (D-F) GO-BP, GO-CC, and GO-MF analysis for the common DEGs showing the specific genes associated with these terms through cent plots.

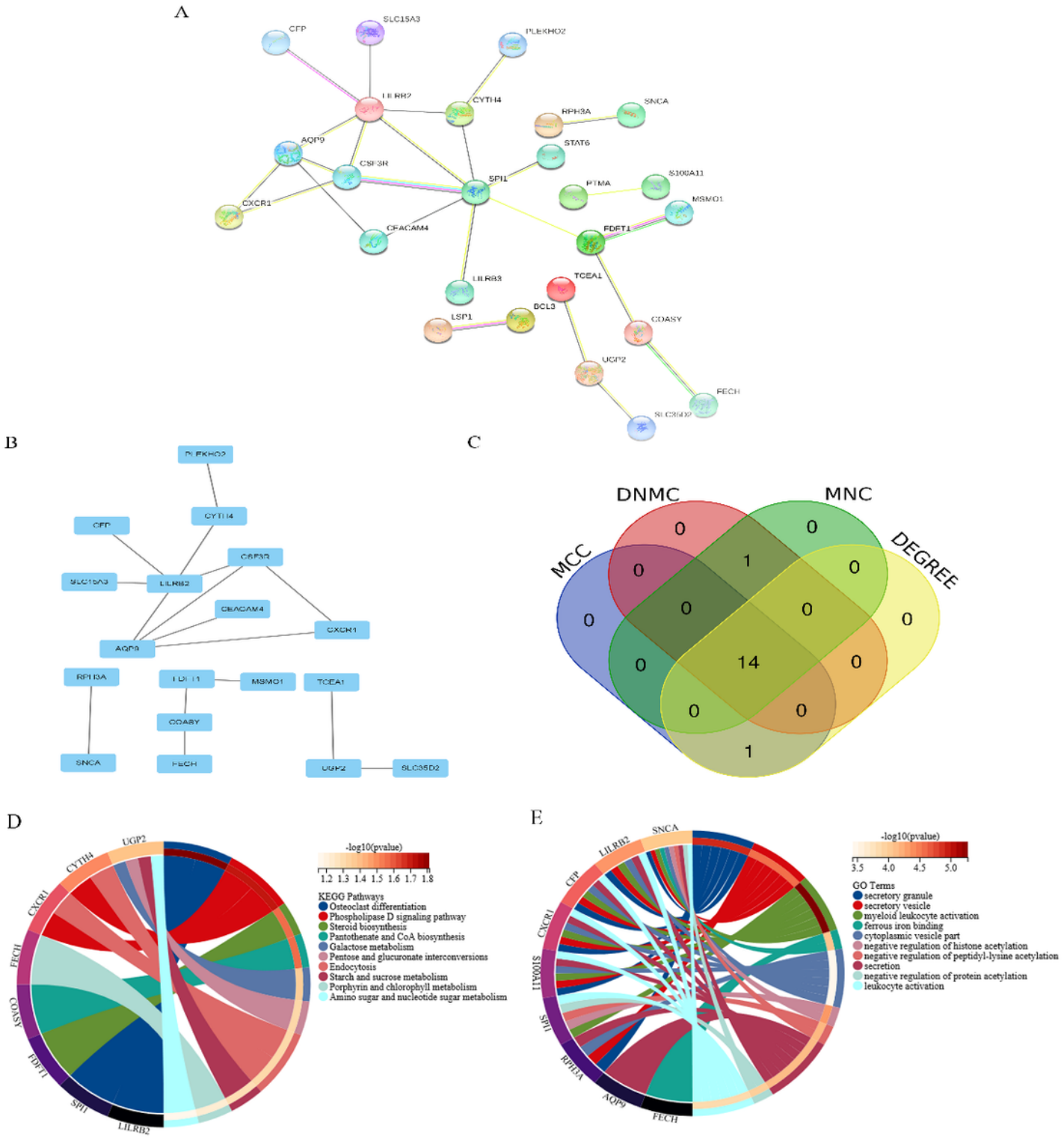

Figure 6

PPI network and significant gene module. (A) The PPI network of the common DEGs has been constructed by STRING. (B) Significant gene modules have been identified and their enrichment analysis has been performed. (C) Four algorithms have been used to identify 14 candidates for hub genes. (D) The KEGG pathway enrichment analysis of the hub genes. (E) The GO term enrichment analysis of the hub genes.

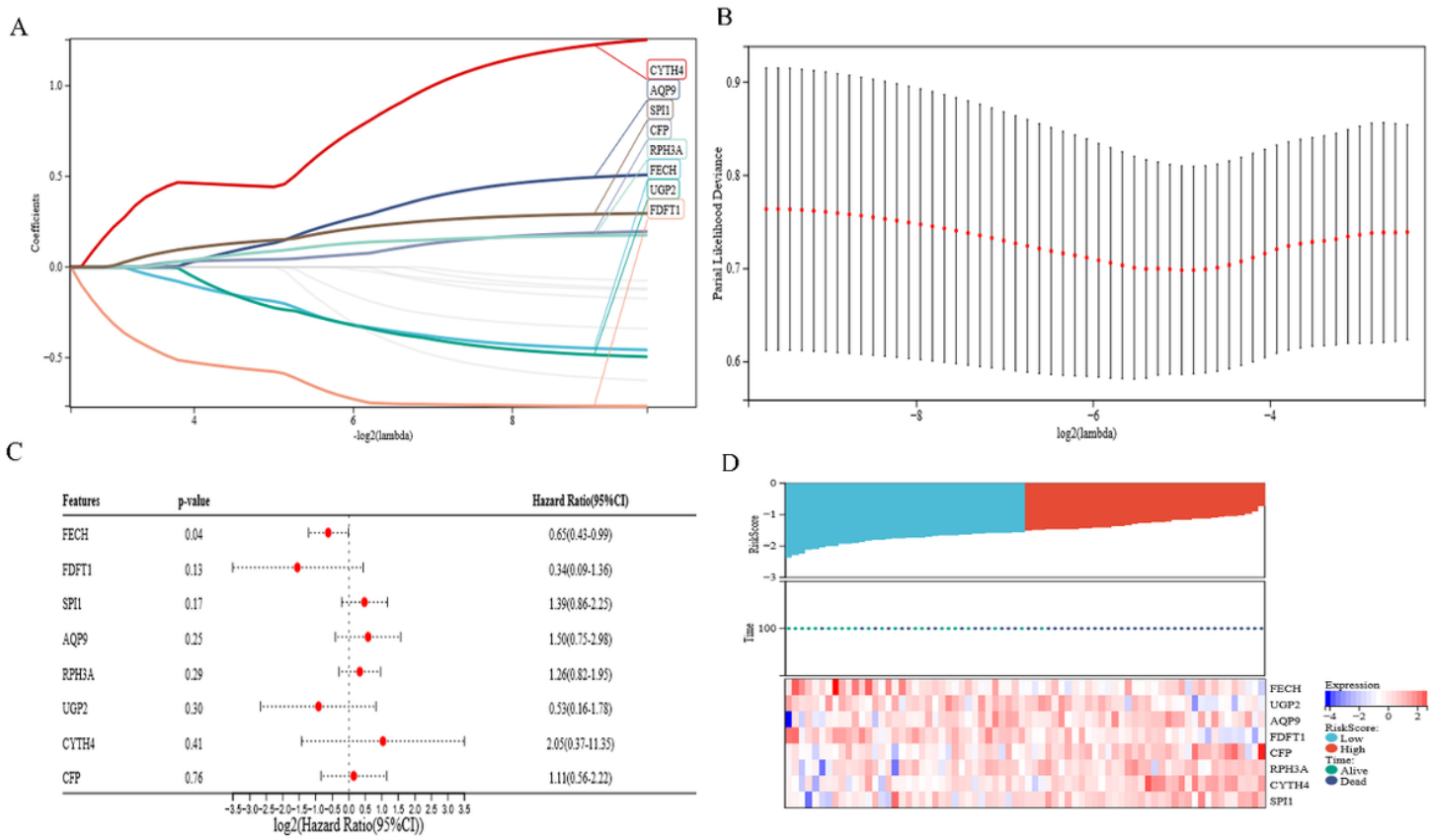

Figure 7

Candidate genes were screened in the Lasso model. (A-B) PD with MDD diagnosis is best represented by the number of genes (n = 8) that correspond to the lowest point of the curve. (C) Multivariate Cox regression analysis revealed that this prognostic model was an independent prognostic parameter of PD. (D) The heat map showed the expression of the eight genes based on the risk score.

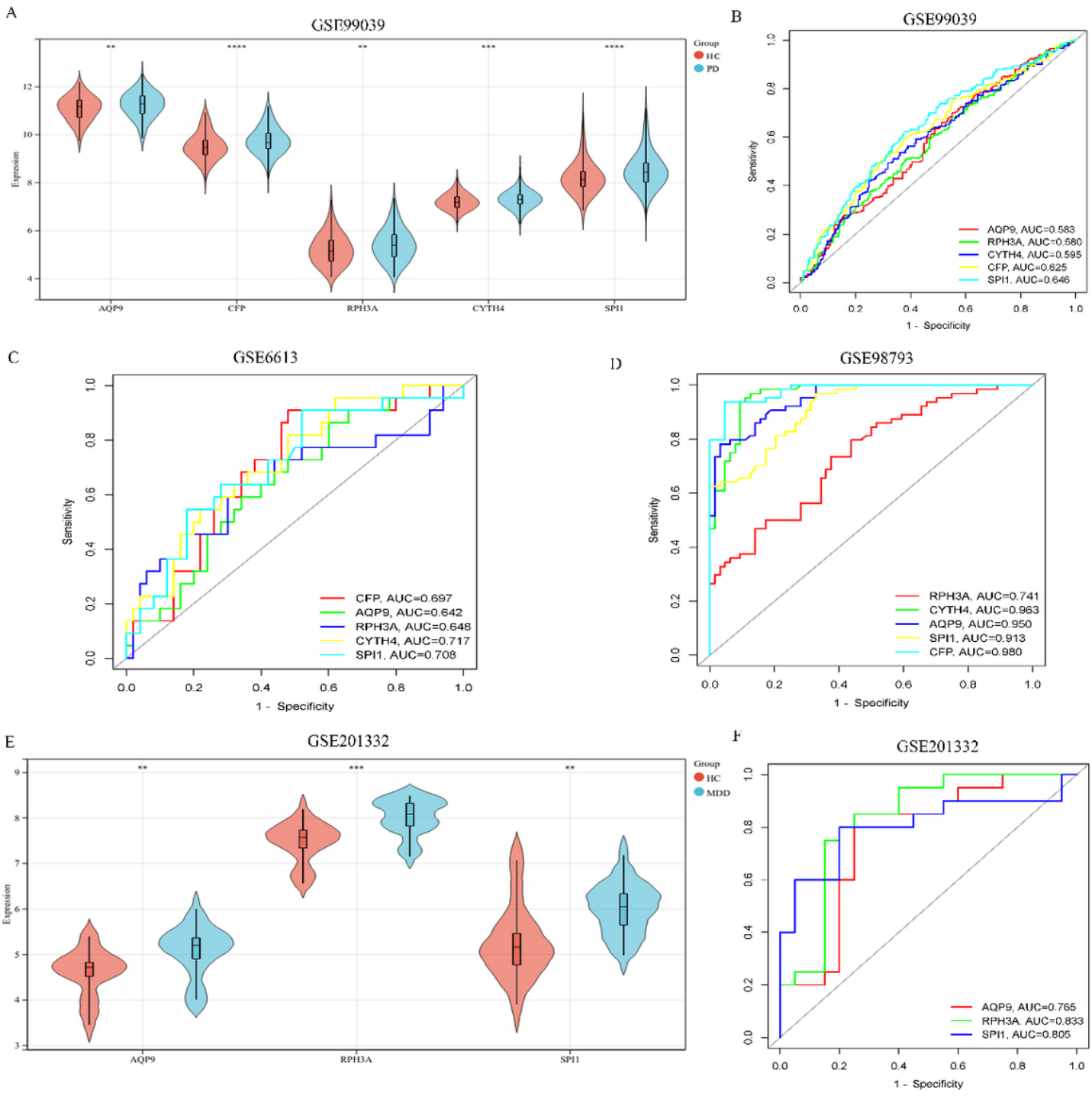

Figure 8

Validation of candidate genes. (A, E) The expression level of significant hub genes in GSE99039 and GSE201332. (B-D) The diagnostic accuracy of hub genes for PD(B-C) or MDD (D, F) was evaluated by ROC curves.

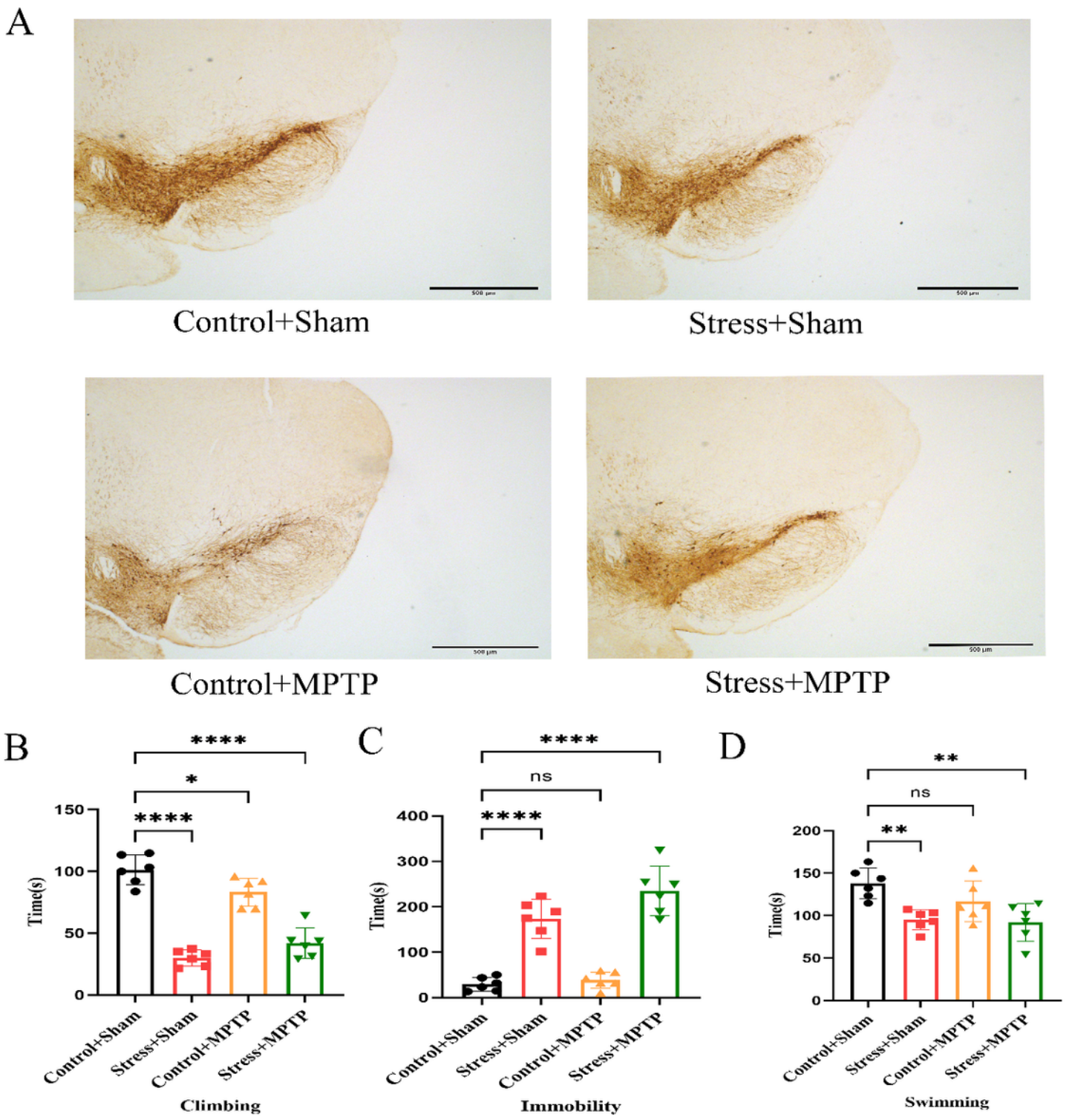

Figure 9

Validation of MPTP and CUMS model. (A) C57Bl/6 mice were treated with intraperitoneal injections of MPTP. The number of SN*pc* TH-positive neurons was determined by immunohistochemistry. Scale bar=500μm. n=6. (B) Fig.9B is representing data on immobility, climbing, and swimming times in the forced swimming test.

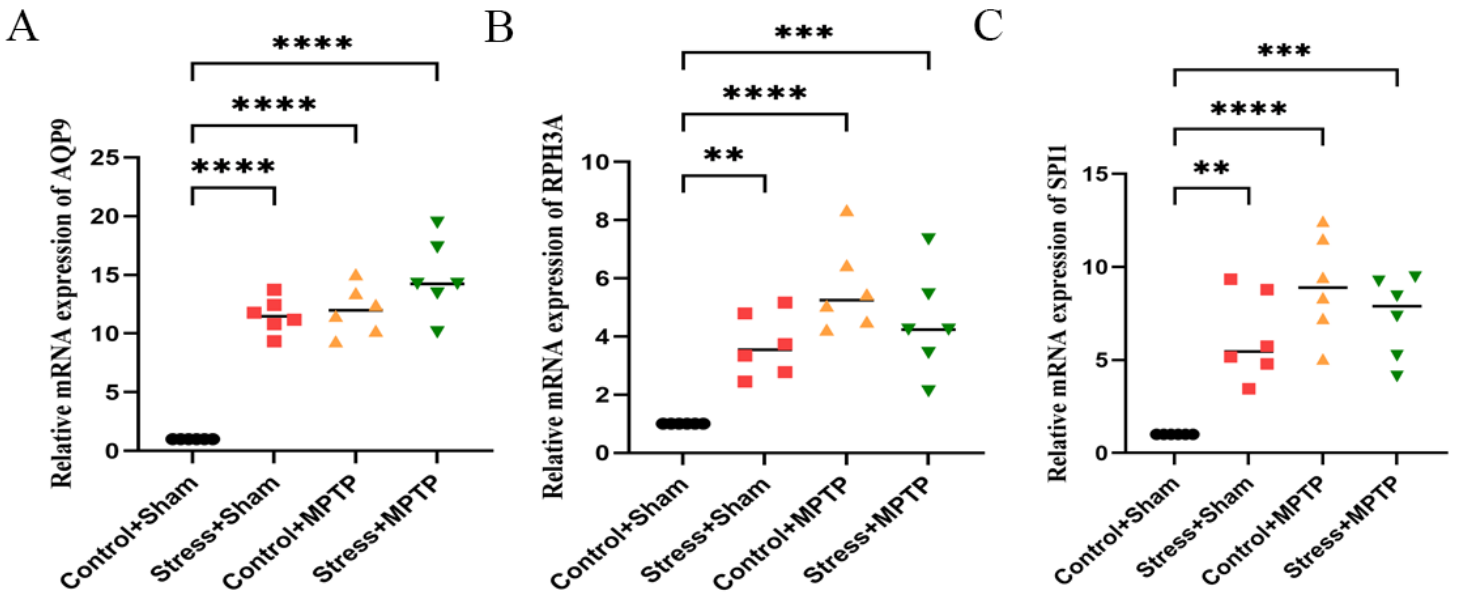

Figure 10

Validation of hub genes. (A) Relative mRNA level of AQP9, RPH3A, and SPI1. (**p< 0.01, *** p < 0.001, **** p < 0.0001).

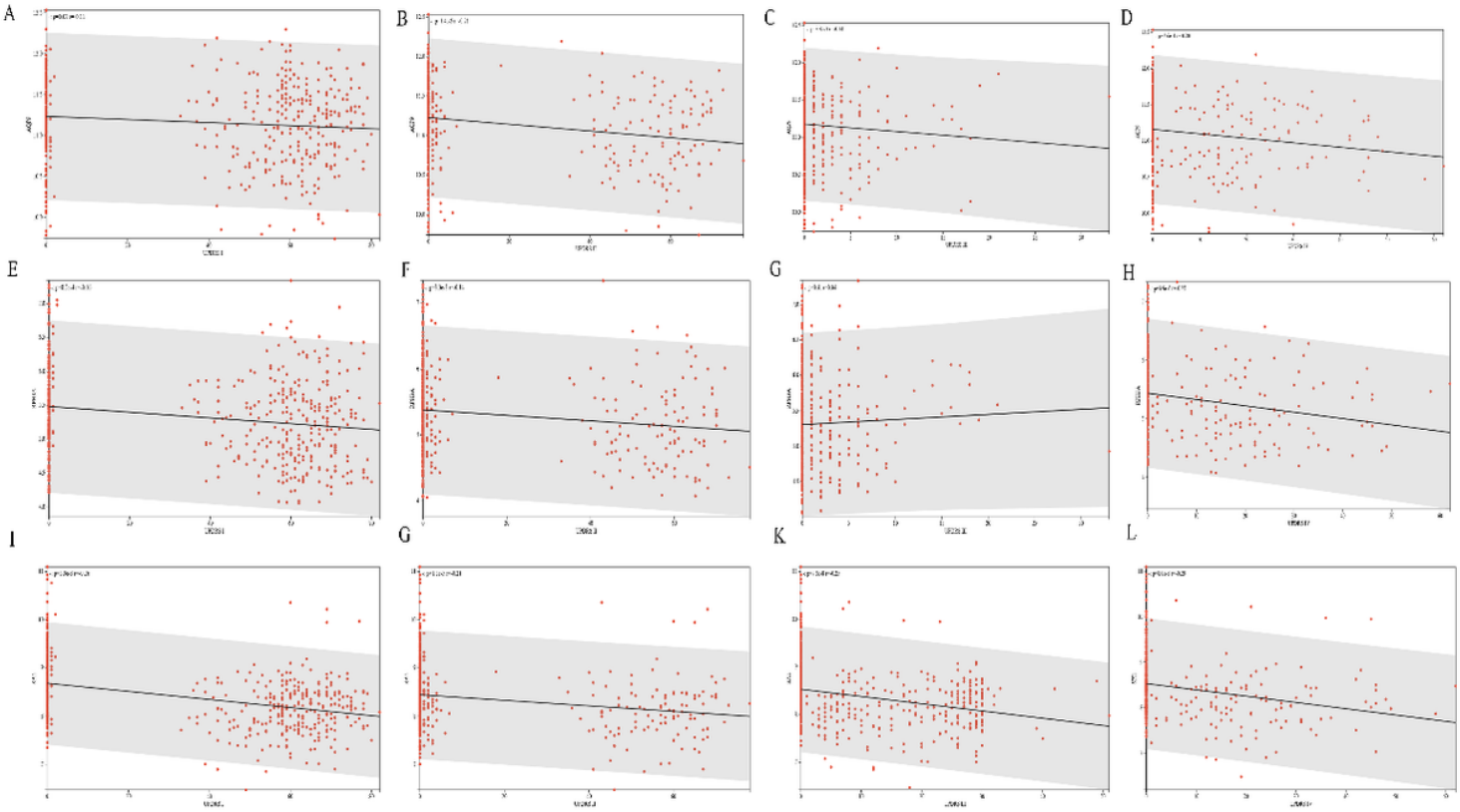

Figure 11

Correlation Analysis with Clinical Features. (A-D) AQP9 expression levels (UPDRS I-IV). (E-H) RPH3A expression levels (UPDRS I-IV). (I-L) SPI1 expression levels (UPDRS I-IV).

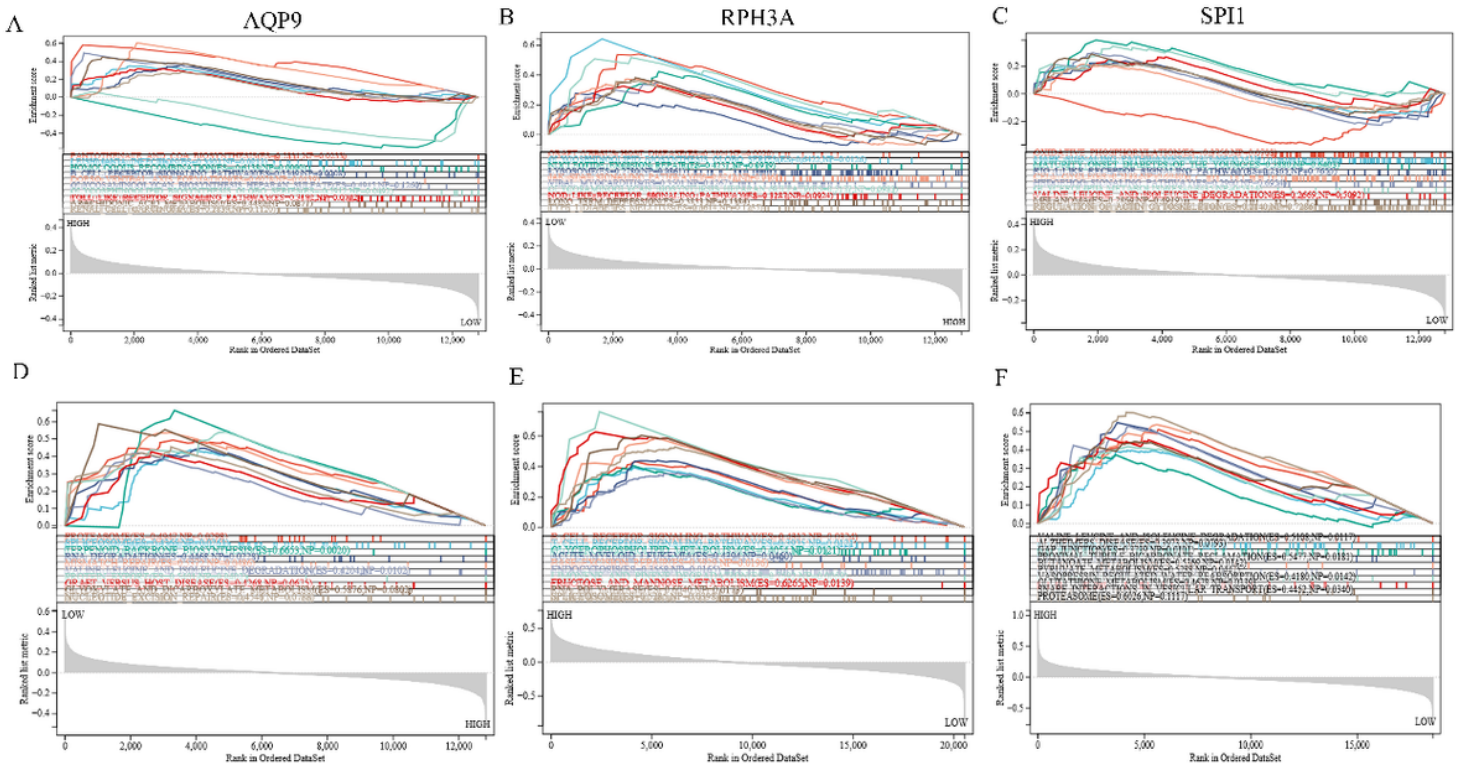

Figure 12

Gene set enrichment analysis. (A-C) These plots showed the APQ9, RPH3A, and SPI1 gene set enrichment scores and gene sets in GSE6613. (D-F) These plots showed the APQ9, RPH3A, and SPI1 gene set enrichment scores and gene sets in GSE98793.

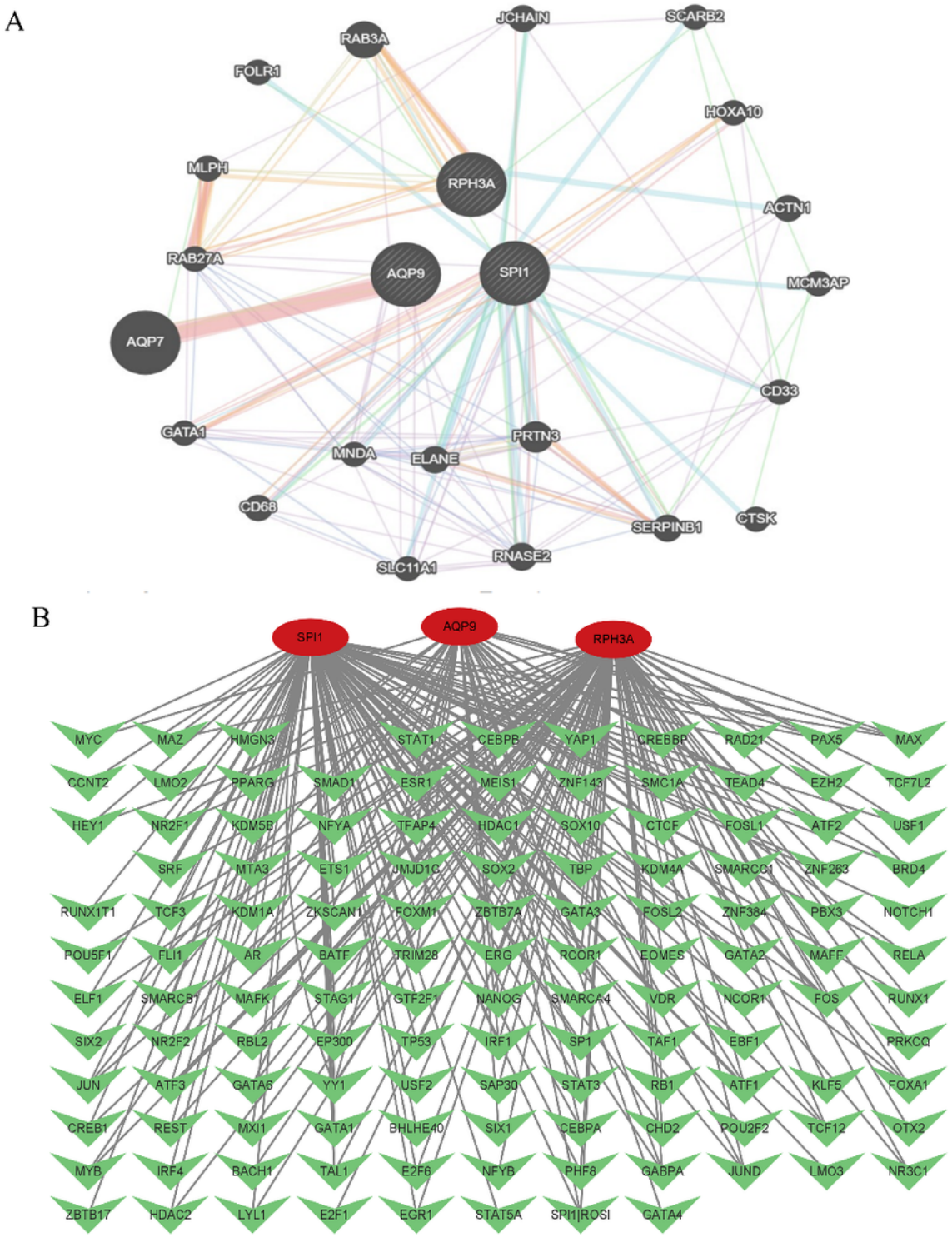

Figure 13

PPI network and TF-gene interaction network. (A) The gene-gene interaction network for hub genes was analyzed using the GeneMANIA database. The 20 most frequently changed neighboring genes are shown. (B) Red nodes represent high-confidence candidate genes, and green nodes represent TF genes in a network for TF-gene interaction.

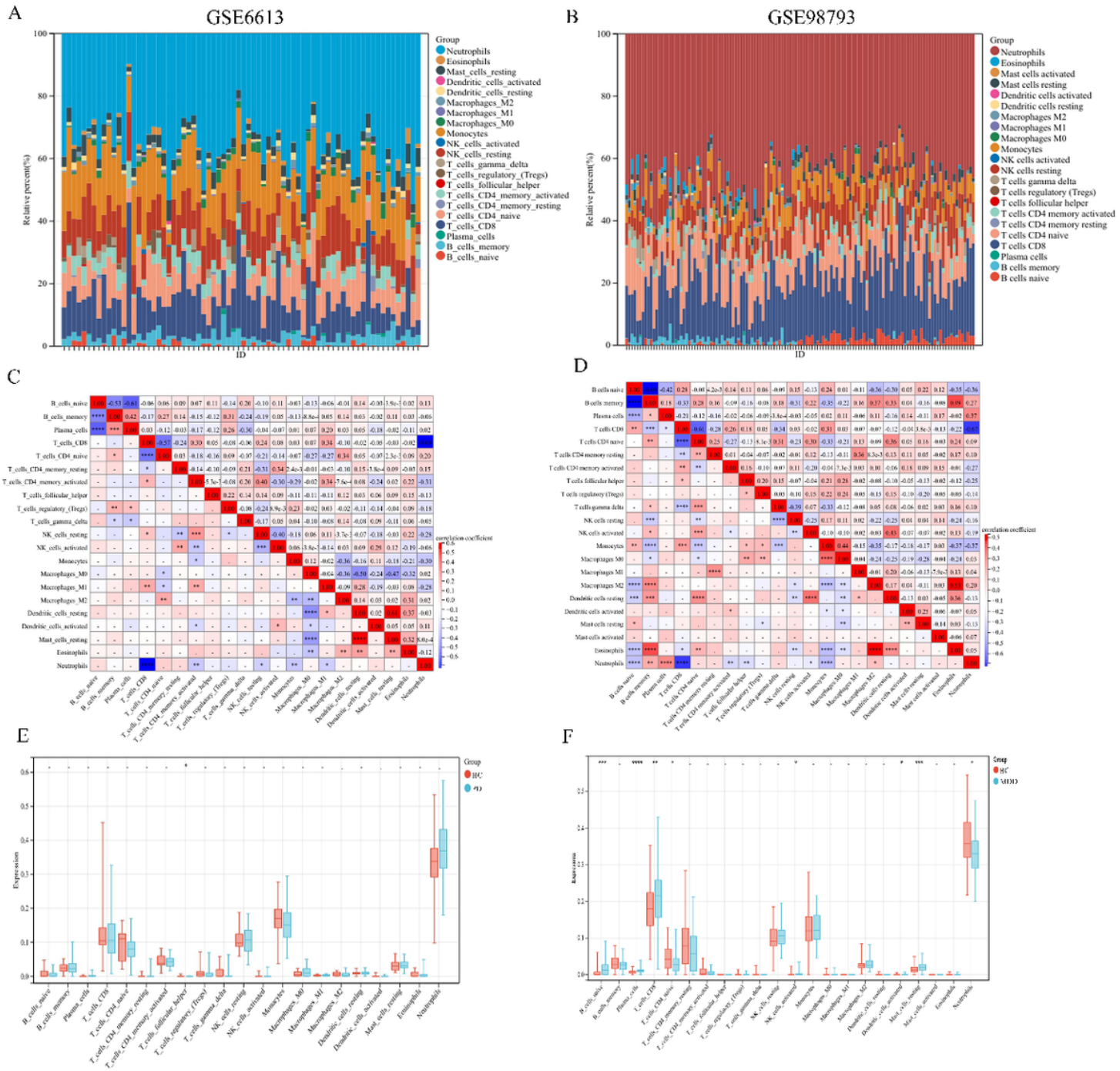

Figure 14

Immune infiltration analysis of GSE6613 and GSE98793. (A-B) Immune cell infiltration map in each sample of GSE6613 and GSE98793. (C-D) Correlation heat map of 22 types of immune cell infiltration of GSE6613 and GSE98793. *, $p < 0.05$, **, $p < 0.01$, ***, $p < 0.001$, ****, $p < 0.0001$. (E-F) Box plots illustrate the comparison of 22 types of immune cells of GSE6613 and GSE98793.

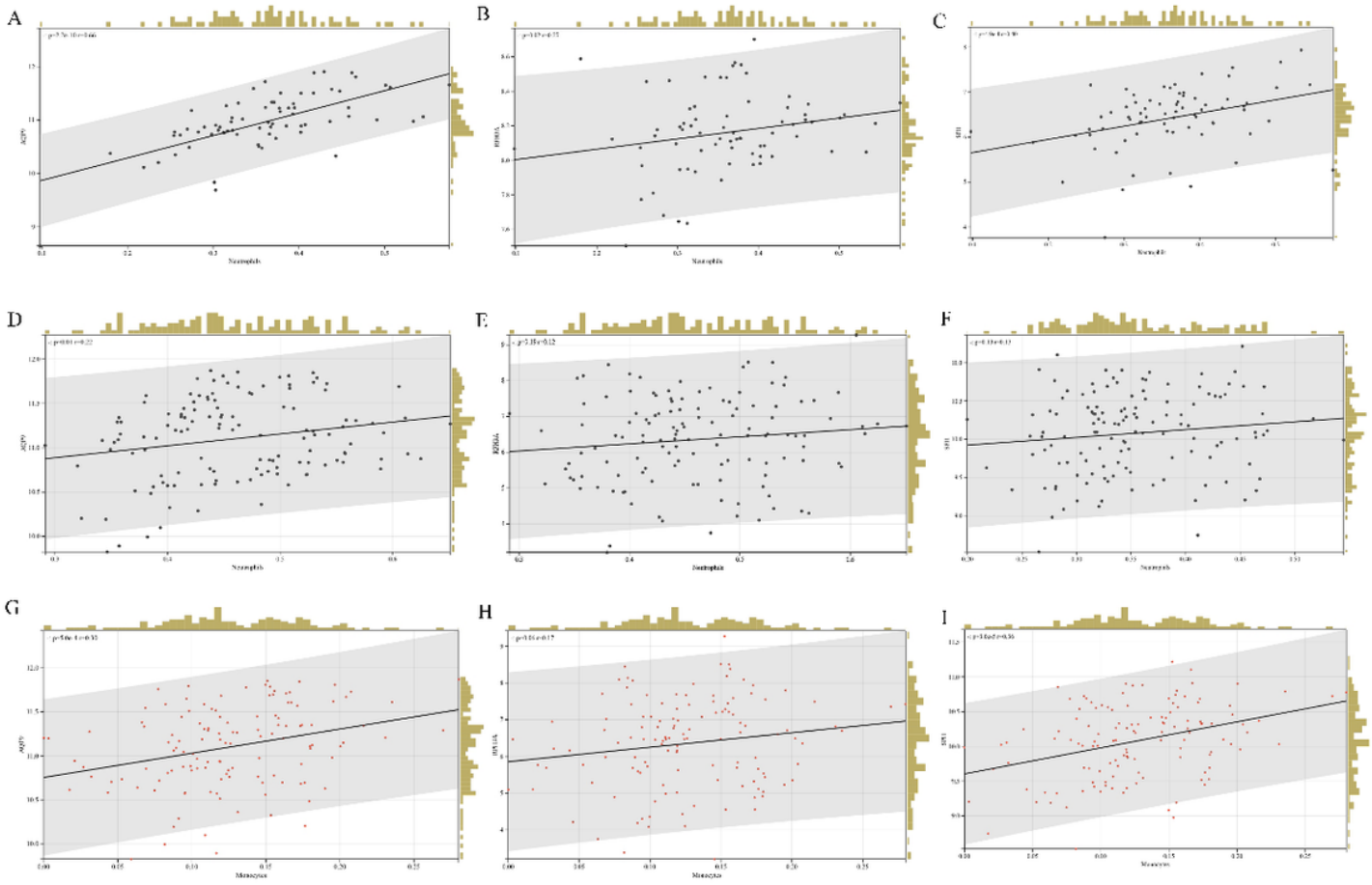

Figure 15

Hub genes correlation with neutrophil cells and monocytes. (A-C) Infiltration of neutrophils in GSE6613 is correlated with hub gene expression. (D-F) Infiltration of neutrophils in GSE98793 is correlated with hub gene expression. (G-I) Infiltration of monocytesin GSE98793 is correlated with hub gene expression.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.png