Identification of a diagnostic model and molecular subtypes of major depressive disorder based on endoplasmic reticulum stress-related genes

Shuwen Huang[1,2], Yong Li[1,2], Jianying Shen[1,2], Wenna Liang[1,2] and Candong Li[1,2,3]*

[1]Research Base of Chinese Medicine Syndrome, Fujian University of Traditional Chinese Medicine, Fuzhou, Fujian, China, [2]FuJian Key Laboratory of TCM Health State, Fuzhou, Fujian, China, [3]LI Candong Qihuang Scholar Studio, Fujian University of Traditional Chinese Medicine, Fuzhou, Fujian, China

Subject: Major depressive disorder (MDD) negatively affects patients' behaviours and daily lives. Due to the high heterogeneity and complex pathological features of MDD, its diagnosis remains challenging. Evidence suggests that endoplasmic reticulum stress (ERS) is involved in the pathogenesis of MDD; however, relevant diagnostic markers have not been well studied. This study aimed to screen for ERS genes with potential diagnostic value in MDD.

Methods: Gene expression data on MDD samples were downloaded from the GEO database, and ERS-related genes were obtained from the GeneCards and MSigDB databases. Differentially expressed genes (DEGs) in MDD patients and healthy subjects were identified and then integrated with ERS genes. ERS diagnostic model and nomogram were developed based on biomarkers screened using the LASSO method. The diagnostic performance of this model was evaluated. ERS-associated subtypes were identified. CIBERSORT and GSEA were used to explore the differences between the different subtypes. Finally, WGCNA was performed to identify hub genes related to the subtypes.

Results: A diagnostic model was developed based on seven ERS genes: KCNE1, PDIA4, STAU1, TMED4, MGST1, RCN1, and SHC1. The validation analysis showed that this model had a good diagnostic performance. KCNE1 expression was positively correlated with M0 macrophages and negatively correlated with resting CD4+ memory T cells. Two subtypes (SubA and SubB) were identified, and these two subtypes showed different ER score. The SubB group showed higher immune infiltration than the SubA group. Finally, NCF4, NCF2, CSF3R, and FPR2 were identified as hub genes associated with ERS molecular subtypes.

Conclusion: Our current study provides novel diagnostic biomarkers for MDD from an ERS perspective, and these findings further facilitate the use of precision medicine in MDD.

KEYWORDS

major depressive disorder, endoplasmic reticulum stress, diagnostic model, molecular subtype, bioinformatics

Highlights

- A diagnostic model based on ERS-related genes was developed.
- This model had good diagnostic performance for MDD.
- KCNE1 correlated with M0 macrophages and resting CD4 memory T cells.
- Two molecular subtypes with different ER scores and immune characteristics were identified.
- NCF4, NCF2, CSF3R, and FPR2 were identified as the hub genes associated with subtypes.

Introduction

Major depressive disorder (MDD) is a common mental disorder with an estimated annual prevalence of 4.4% worldwide, affecting more than 300 million people [1]. It is ranked as the leading cause of disability globally and the third leading cause of the global burden of disease [2]. MDD is chronic or recurrent in nature, usually associated with prolonged periods of depressed mood and anhedonia, and is accompanied by considerable morbidity, suicide risk, and mortality [3]. Despite receiving evidence-based treatment, approximately 30–50% of patients remain unresponsive to therapy, imposing a huge economic burden on society [4]. Depression is associated with many mental and physical disorders and is influenced by an interplay of genetic and environmental factors, suggesting that its underlying mechanisms are complex [5]. Previous studies have suggested that the ineffectiveness of antidepressant drugs may be partly attributed to their failure to address important biological processes involved in the pathogenesis of depression [6]. Therefore, it is essential to explore the molecular mechanisms underlying MDD and identify diagnostic markers and therapeutic targets for this disease.

The endoplasmic reticulum (ER) is the largest organelle in eukaryotic cells and is involved in the regulation of protein synthesis, folding, and transport [7]. ER dysregulation can lead to the accumulation of unfolded and misfolded proteins in the lumen, stimulating the unfolded protein response, a process known as ER stress (ERS) [8]. Gold et al. in 1988 proposed that MDD represents dysregulation of the stress system in a readily inducible stressful environment [9]. Moreover, multiple studies have shown that ERS is involved in the pathophysiology of central nervous system disorders, such as schizophrenia and MDD [10, 11]. Elevated ERS responses in the brain have been observed in many human and animal models of depression. For example, ERS-related proteins (such as GRP78, CHOP, and XBP1) were found to be associated with hippocampal damage and cognitive impairment in a rat model of stress, and the expression levels of ERS-related proteins (such as GRP78 and CHOP) in the hippocampus of patients with MDD were upregulated compared to those in control subjects [12, 13]. Meanwhile, ERS has been proven to be associated with cardiovascular diseases as well as several chronic diseases, such as diabetes and inflammatory bowel disease [14]. Notably, depression is connected with reduced heart rate variability, increased sympathetic nervous system, and platelet aggregation, all of which are risk factors for cardiovascular diseases [15]. Hence, ERS is closely correlated with common comorbidities of depression. Besides, ERS plays a key role in mediating immune and metabolic responses, and these mechanisms also contribute to the development of psychiatric disorders such as depression [16]. Taken together, ERS may be directly or indirectly involved in several key biological processes that alter the course of MDD. Currently, there is evidence that targeting ERS may be a new strategy for the treatment of this disorder [17]. However, the specific biomarkers of the diagnosis of MDD have not been fully explored, especially from the perspective of ERS-related genes.

In this study, we collected the transcriptome data of patients with MDD and ERS-related genes from public databases and identified ERS-related differentially expressed genes (DEGs) between MDD and control samples. Next, genes with a diagnostic value for MDD were screened using LASSO analysis to establish a diagnostic model. The diagnostic performance of this model was assessed, and the correlation between diagnostic genes and immune infiltrates was analysed. Molecular subtypes were screened based on ERS-related diagnostic genes. Our findings may help further explore potential biomarkers of MDD and provide reference targets for drug development.

Materials and methods

Dataset acquisition and pre-processing

The design of the analysis is illustrated in Figure 1. Raw gene expression profiles for four datasets (GSE98793 [18], GSE19738 [19], GSE32280 [20], and GSE38206 [21]) downloaded from GEO were included in this analysis. The sample information and detection platform for each dataset are presented in Table 1. A total of 114 patients with MDD and 115 healthy controls were analysed. GSE98793, GSE19738, and GSE32280 were selected as training cohorts. After merging the data from the three datasets, we used the sva package in R3.6.1[1] to remove batch effects, and ultimately obtained the expression profile data of 211 samples, including 105 MDD and 106 control samples. The GSE38206 dataset was selected as the validation cohort. ERS-related genes were extracted from Genecards and the Molecular Signature Database v7.4 (MSigDB). The downloaded genes were duplicated and merged to obtain 904 genes.

Screening of ERS-related (DEGs in MDD patients vs. controls)

The Limma package (v 3.34.7) [22] was employed to perform differential expression analysis, and genes with value of $p < 0.01$ were considered as DEGs between MDD and control samples. These DEGs were intersected with ERS-related genes to identify differentially expressed ERS-related genes. Correlations between ERS-related DEGs were calculated using the cor function[2] in R.
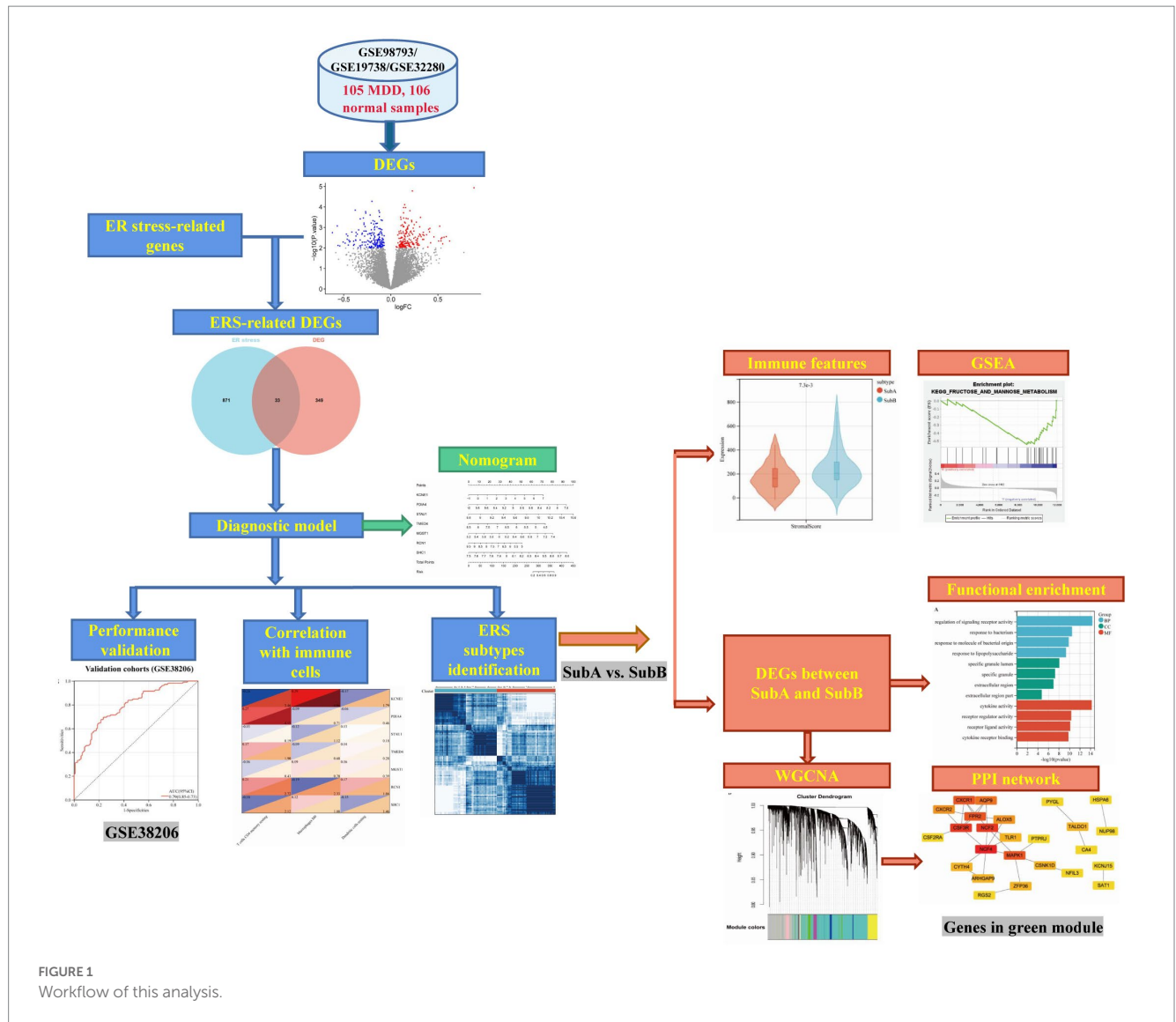
---

1 http://www.bioconductor.org/packages/release/bioc/html/sva.html
2 http://77.66.12.57/R-help/cor.test.html

FIGURE 1
Workflow of this analysis.

TABLE 1 Specific information for four major depressive disorder-related datasets.

GO accession	Platform	Sample number		Sample source	Purpose
		Control	MDD		
GSE98793	GPL570	64	64	Whole blood	Training cohort
GSE19738	GPL6848	34	33	Whole blood	Training cohort
GSE32280	GPL570	8	8	Peripheral blood	Training cohort
GSE38206	GPL13607	9	9	Peripheral blood	Validation cohort

Establishment of protein−protein interaction (PPI) network

These ERS-related DEGs were imported into the STRING database (v 11.0) (23) to retrieve the interaction relationship between gene-encoded proteins, and PPI pairs with interaction scores ≥0.4 were retained to construct the PPI network.

Univariate logistic regression analysis of ERS-related DEGs

Rms (v 6.3−0, https://cran.r-project.org/web/packages/rms/index.html) (24) in R was used to conduct univariate logistic regression analysis of ERS-related DEGs, and genes with value of $p < 0.05$ were retained for further analyses.

Development of a diagnostic model based on ERS-related genes

The LASSO algorithm in the R lars package (v 1.2) (25) was used to further screen the selected ERS-related DEGs. Next, RMS in R was used to perform multivariate logistic regression analysis, and the optimal ERS-related gene signature was selected as the diagnostic gene. The risk score (RS) was calculated based on the expression and coefficient of each diagnostic gene, followed by construction of a diagnostic model.

Performance evaluation of diagnostic model

The receiver operating characteristic (ROC) curve method in the pROC package (v. 1.12.1) (26) was used to evaluate the efficacy of the diagnostic model constructed in the training and validation cohorts.

Correlation analysis of diagnostic genes and immune states

In the training cohort, CIBERSORT was used to assess the proportion of immune cells in the samples, and the Kruskal–Wallis test in R was applied to compare the differences in the distribution of immune cells between the MDD and control groups. Meanwhile, the correlations between ERS-related diagnostic genes and immune cells with significant differences in MDD vs. control samples were calculated using the cor function in the R software.

Construction and evaluation of the nomogram model

To predict the incidence of MDD, a diagnostic nomogram model was established using the rms package. Meanwhile, calibration curve and decision curve analysis (DCA) were employed to evaluate the prediction ability and practical utility of the model, respectively.

Prediction of molecular subtype based on ERS-related genes

After extracting the expression levels of ERS-related diagnostic genes in the training cohort, ConsensusClusterPlus (v 1.54.0) (27) was used to determine the molecular subtype of the MDD samples. Furthermore, the ER score of each case was calculated using the GSVA (v 1.36.3) package, and the Kruskal–Wallis test was used to compare the differences in ER scores of different subtype groups.

Comprehensive analysis of immune cells and molecules in different subgroups

CIBERSORT (28) was used to evaluate the distribution of immune cells in MDD samples in the training set, and the estimate package was then employed to calculate the ESTIMATE, immune, and stromal scores of patients with MDD. In addition, differences in immune cells and scores between different subgroups were analysed using the Kruskal–Wallis test.

We also compared the expression levels of immune checkpoint genes (CD27, CD274, and CD40) in different molecular subtypes using the Kruskal–Wallis test.

Gene set enrichment analysis of different ERS subtypes

Based on the genome-wide expression levels in the training cohort, the GSEA database (29) was used to identify the KEGG signalling pathways that were significantly associated with the subtypes. We selected value of $p < 0.05$ as the threshold for significant enrichment of the KEGG pathway.

Screening of DEGs related to ERS-related molecular subtypes

Differential gene analysis was conducted on samples between subtypes, and genes with value of $p < 0.05$ and $|\log_2$ Fold change (FC)$| > 0.263$ (also means FC > 1.2) were regarded as DEGs between different subtypes. To understand the biological functions of these genes, clusterProfiler was used to perform Gene Ontology (GO) (30) and KEGG pathway (31) enrichment analyses. A value of $p < 0.05$ was defined as a significant enrichment result.

Weighted gene co-expression network analysis

The MDD-related expression matrix was analysed using the R package WGCNA (v 1.6.1) (32) to identify highly covariant gene set modules. The specific analysis methods were as follows: first, the optimal power and connectivity k were selected to convert the gene expression matrix into a topological overlap matrix (TOM); second, the highly corrected genes were clustered into modules using clustering and dynamic pruning with these parameters (minModuleSize = 30 and mergeCutHeight = 0.25), and finally, the correlations between modules and molecular subtypes were calculated. The module with value of $p < 0.05$ and with the highest connection with the subtype was selected for subsequent analysis. The hub genes in this module were screened using the gene significance (GS) and module membership (MM) indices; MM > 0.8 and GS > 0.2 were regarded as screening thresholds.

Construction of PPI network based on hub genes

The STRING database was used to search for interactions between hub genes to build networks, and this network was visualised using Cytoscape (v 3.9.0) (33).
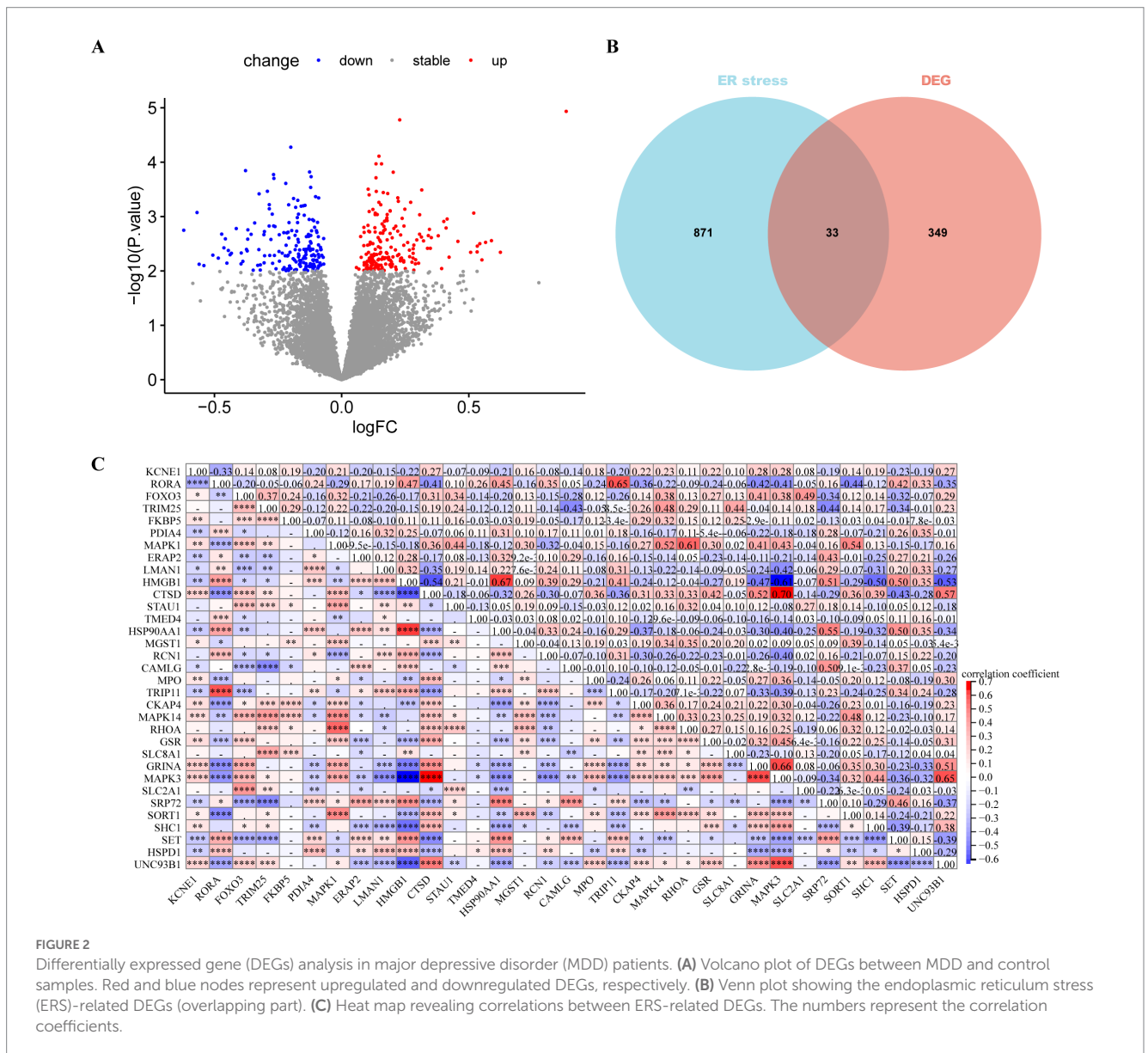
Results

Screening of 33 ERS-related DEGs between MDD and control samples

Differential expression analysis revealed 382 DEGs (200 upregulated and 182 downregulated genes) between the MDD and control samples. The DEGs were visualised using a volcano plot (Figure 2A). After integration with the ERS-related genes, 33 ERS-related DEGs were identified (Figure 2B). We also observed

a significant correlation between gene expression (Figure 2C). For example, MAPK3 was positively correlated with CTSD and negatively correlated with HMGB1. STRING was used to search for interactions between gene-encoded proteins and a PPI containing 23 ERS-related DEGs was established (Figure 3). We found that MAPK3 and MAPK1 were linked to a large number of genes.

Construction of a diagnostic model based on seven ERS-related DEGs

Using univariate logistic regression analysis and LASSO regression algorithm, candidate ERS-related genes were screened from 33 ERS-related genes to predict the occurrence of MDD disease. These results indicate that the 18 genes could be used as potential diagnostic markers (Figures 4A–C). Multivariate regression analysis identified seven diagnostic biomarkers to

FIGURE 2

Differentially expressed gene (DEGs) analysis in major depressive disorder (MDD) patients. **(A)** Volcano plot of DEGs between MDD and control samples. Red and blue nodes represent upregulated and downregulated DEGs, respectively. **(B)** Venn plot showing the endoplasmic reticulum stress (ERS)-related DEGs (overlapping part). **(C)** Heat map revealing correlations between ERS-related DEGs. The numbers represent the correlation coefficients.
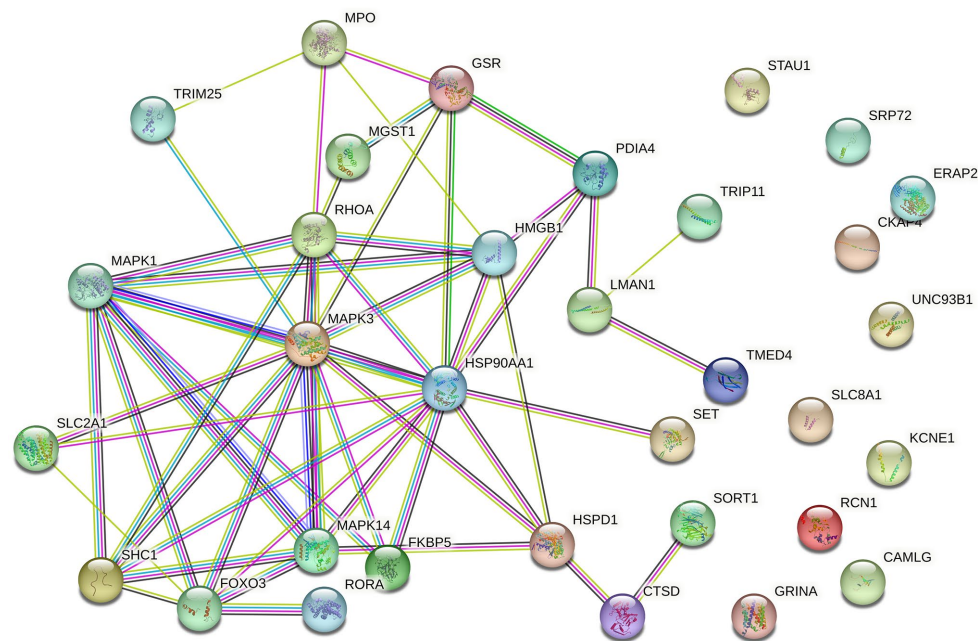
FIGURE 3
Protein−protein interaction (PPI) network of ERS-related DEGs.

construct the model (Figure 4D): KCNE1, PDIA4, STAU1, TMED4, MGST1, RCN1, and SHC1. RS values were calculated to establish diagnostic models based on the LASSO regression coefficients and the expression level of each gene.

Diagnostic model had reliable predictive power for MDD in the training and validation cohorts

Furthermore, we used training and validation cohorts to assess the predictive ability of the established model. The AUC values of ROC for models in the training and validation (GSE38206) cohorts were 0.79 (95% CI:0.73–0.85) and 0.94 (95% CI,0.83–1.00), indicating that this model had reliable predictive performance for MDD diagnosis (Figures 5A,B). The RS of MDD samples was significantly higher than that of control samples in both cohorts (Figures 5C,D). We also observed differences in the expression of these genes between the MDD and control groups in the training cohort (Figures 5E,F).

Diagnostic nomogram model construction

A nomogram model based on seven biomarkers was generated to predict the risk of MDD. As shown in Figure 6A, each predictive marker was projected upward to the "point" of the value at the top of the nomogram to obtain a score of 0 to 100 points, and the total score of seven points was calculated to predict the probability of MDD risk. The calibration cure displayed that the predicted risk of MDD was in

good agreement with the actual risk (Figure 6B). Moreover, the DCA revealed that the hub gene curves were above the grey line, indicating that the use of a nomogram to predict MDD risk had a significant net benefit (Figure 6C).

Selected ERS-related DEGs were significantly associated with immune cell infiltration

To observe the immune statuses of the control and MDD groups, we compared the infiltration levels of immune cells between the two groups using CIBERSORT. The results indicated that the levels of infiltration of the three types of immune cells were significantly different between the two groups (Figure 7A). In brief, compared with the control samples, MDD samples had lower levels of resting CD4 memory *T* cells and resting dendritic cells, but had higher level of M0 macrophages. Next, the correlations between the seven diagnostic genes and the three immune cell types were analysed. KCNE1 expression showed the highest positive correlation with M0 macrophages ($r = 0.29$) and the highest negative correlation with resting CD4 memory T cells ($r = -0.24$) (Figure 7B).

Two molecular subtypes of MDD patients identified using seven diagnostic markers

Based on the expression levels of seven selected diagnostic markers, the molecular subtypes of MDD patients were screened.
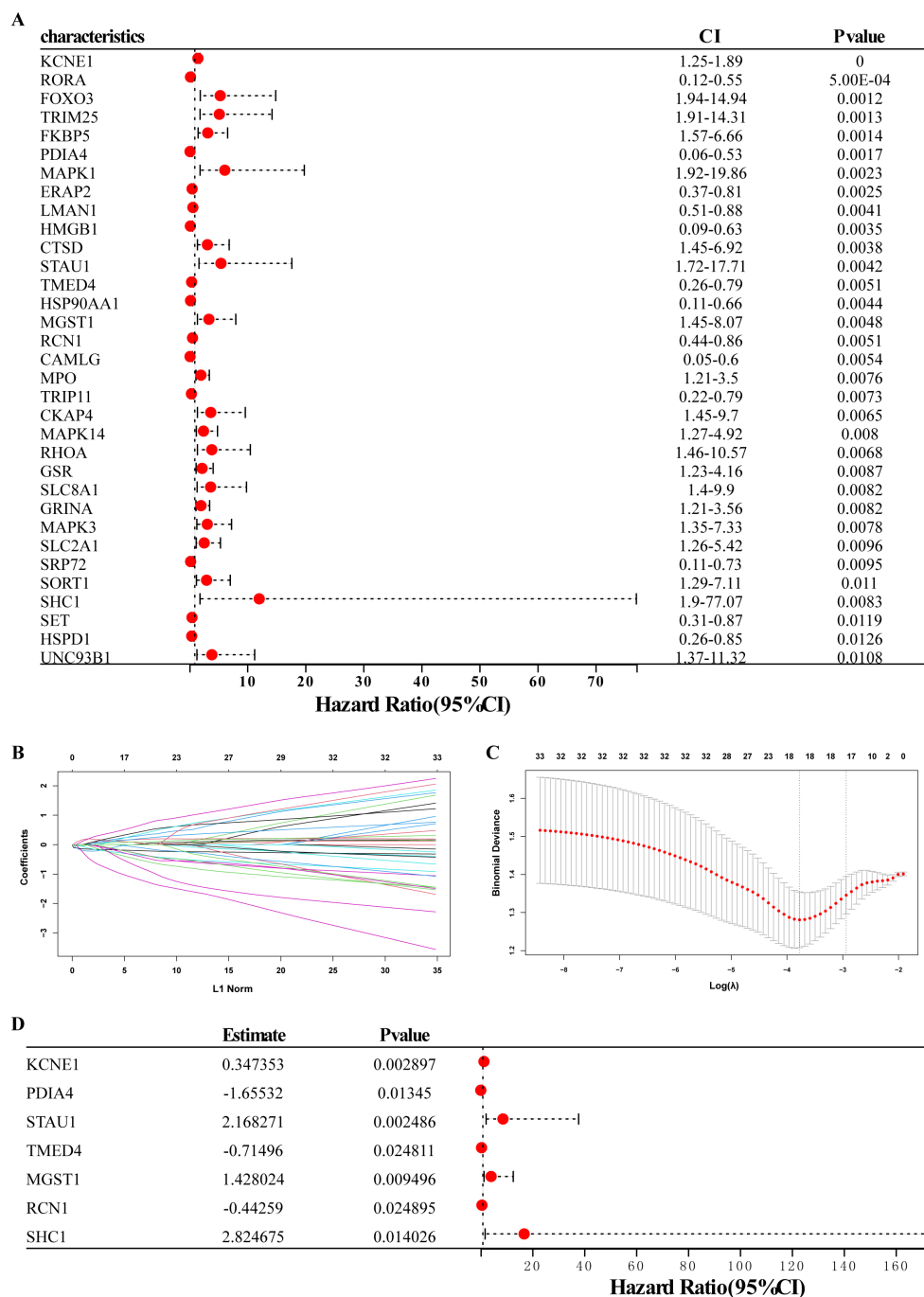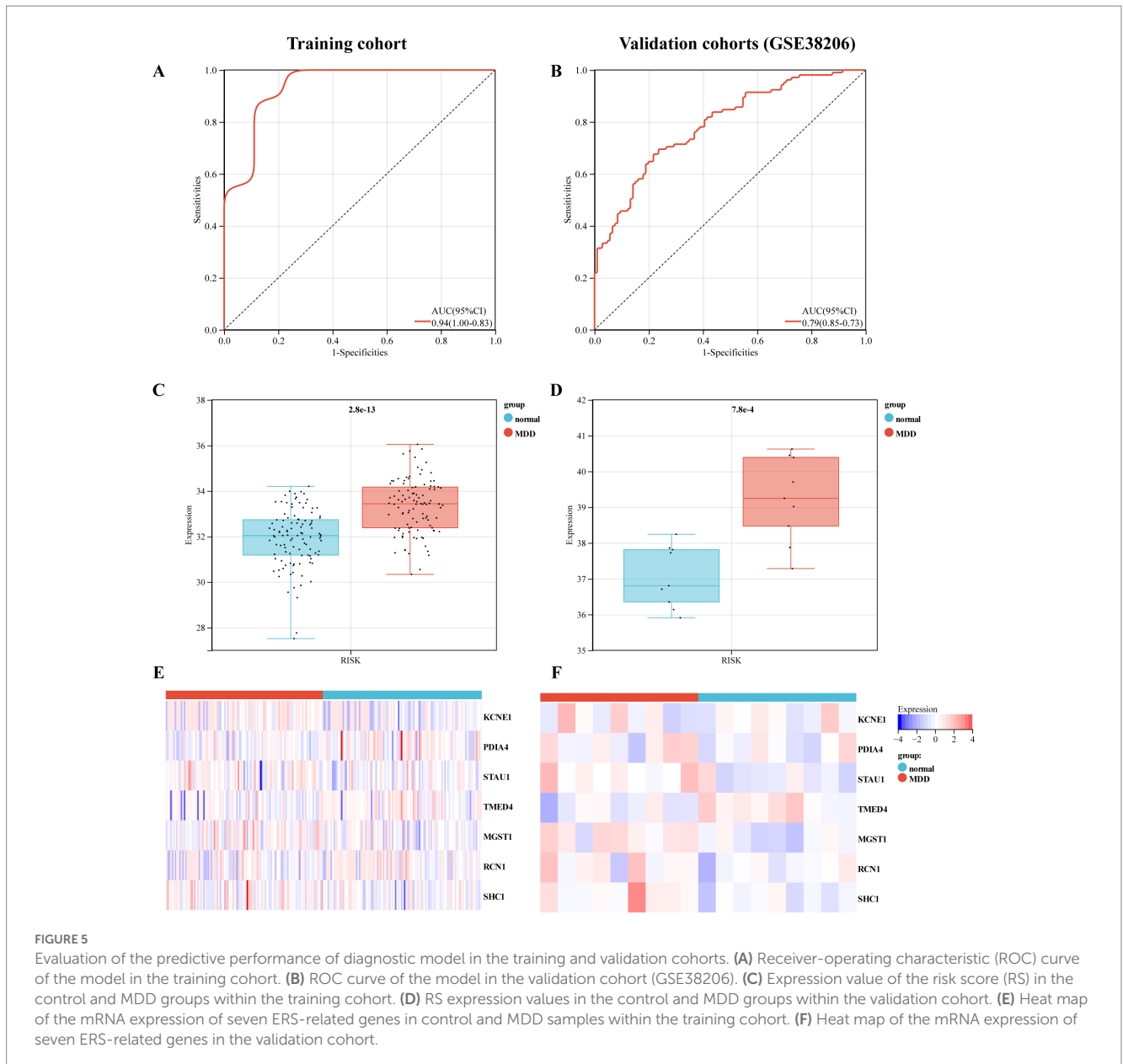
FIGURE 4

Establishment of the seven-gene signature diagnostic model. **(A)** Univariate logistic regression analysis confirmed that ERS-related genes were related to MDD occurrence. **(B,C)** LASSO regression analyses. **(D)** Multivariate regression analysis of seven diagnostic biomarkers for model construction.

ConsensusClusterPlus software was used to calculate the optimal number of clusters. Based on the consistency matrix heatmap and CDF curve, k = 2 was defined as the optimal number (Figures 8A,B). Thus, two subtypes of patients with MDD were identified: SubA (*n* = 51) and SubB (*n* = 54) (Figure 8C). Next, the GSVA algorithm was used to assess the ER scores of each sample. The patients with MDD in the SubA group had significantly higher ER scores than those in the SubB group (Figure 8D).

Differences in immune characteristics and immune checkpoint genes of the two ERS subtypes

Furthermore, the immune characteristics of the two subtypes were estimated using the CIBERSORT software. The infiltration levels of the five immune cell types differed between the two subtypes. Compared to the SubB type, the SubA type had

FIGURE 5
Evaluation of the predictive performance of diagnostic model in the training and validation cohorts. **(A)** Receiver-operating characteristic (ROC) curve of the model in the training cohort. **(B)** ROC curve of the model in the validation cohort (GSE38206). **(C)** Expression value of the risk score (RS) in the control and MDD groups within the training cohort. **(D)** RS expression values in the control and MDD groups within the validation cohort. **(E)** Heat map of the mRNA expression of seven ERS-related genes in control and MDD samples within the training cohort. **(F)** Heat map of the mRNA expression of seven ERS-related genes in the validation cohort.

significantly higher levels of CD8+ T cells, M2 macrophages, and resting dendritic cells, and significantly lower levels of M0 macrophages and neutrophils (Figure 9A). Meanwhile, SubB exhibited significantly higher stromal and ESTIMATE scores (Figure 9B), suggesting a higher degree of infiltration than that of SubA. We also observed that samples in the SubA group had higher expression of CTLA4 and HAVCR2 but lower expression of SIRPA than those in the SubB group (Figure 9C).

GSEA revealing the significantly enriched biological pathways between two subtypes

GSEA was performed based on the two subtypes, and the top five significant biological pathways are shown in Figure 10. The results

indicated that fructose and mannose metabolism, leukocyte transendothelial migration, lysosomes, and *Vibrio cholerae* infection were upregulated in the SubB group, and steroid hormone biosynthesis was upregulated in the SubA group.

Differential expression analysis of two distinct subtypes

A total of 1,054 DEGs, including 468 upregulated and 586 downregulated genes, were identified between SubA and SubB subtypes. To explore the biological function of these DEGs, functional enrichment analysis was conducted. DEGs were mainly enriched in 977 GO terms (Figure 11A), such as regulation of signalling receptor activity (BP), specific granules (CC), and
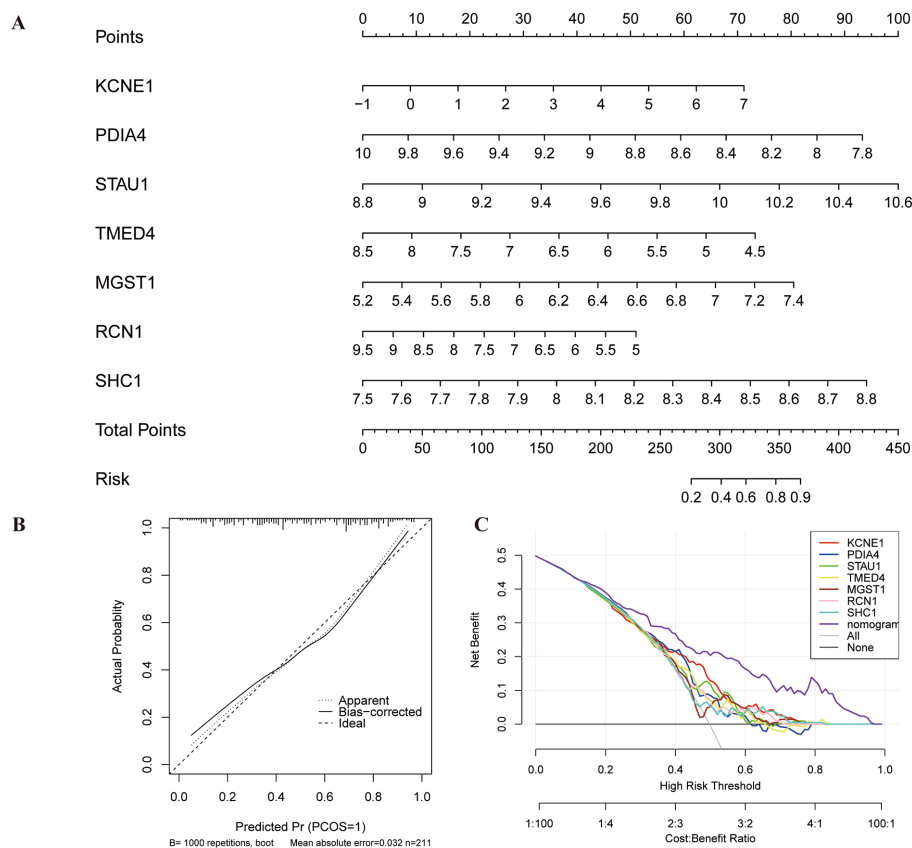
FIGURE 6
Construction of nomogram model for MDD diagnosis. **(A)** Nomogram based on seven genes to predict MDD risk. **(B)** Calibration curve to evaluate the diagnostic potential of the model. **(C)** DCA curve to assess the practical efficacy of the model.

cytokine activity (MF); they were significantly involved in 27 KEGG pathways, such as the TNF signalling pathway and inflammatory bowel disease (Figure 11B).

Identification of hub genes for ERS subtypes using WGCNA

The identified DEGs were selected for WGCNA to screen hub genes. Based on the scale free value and mean connectivity index, the optimal soft threshold was $\beta = 9$ when correlation coefficient > 0.85 (Figure 11C). In total, 85 modules were merged (Figure 11D). Next, the correlation between module genes and ERS subtypes was analysed, and the green module exhibited the most significant positive correlation with the SubB subtype ($r = 0.45$, value of $p < 0.05$) (Figure 11E). A total of 320 genes in the green module served as potential key genes related to these subtypes. Further, genes with MM > 0.8 and GS > 0.2 were considered as hub genes, and 48 genes were obtained (Figure 11F).

To explore the interactions between these 48 hub genes, a PPI network was constructed. A total of 25 genes were included in this network (Figure 11G). Several nodes with high connectivity were identified, including NFC4, CSF3R, NCF2, and MAPK.
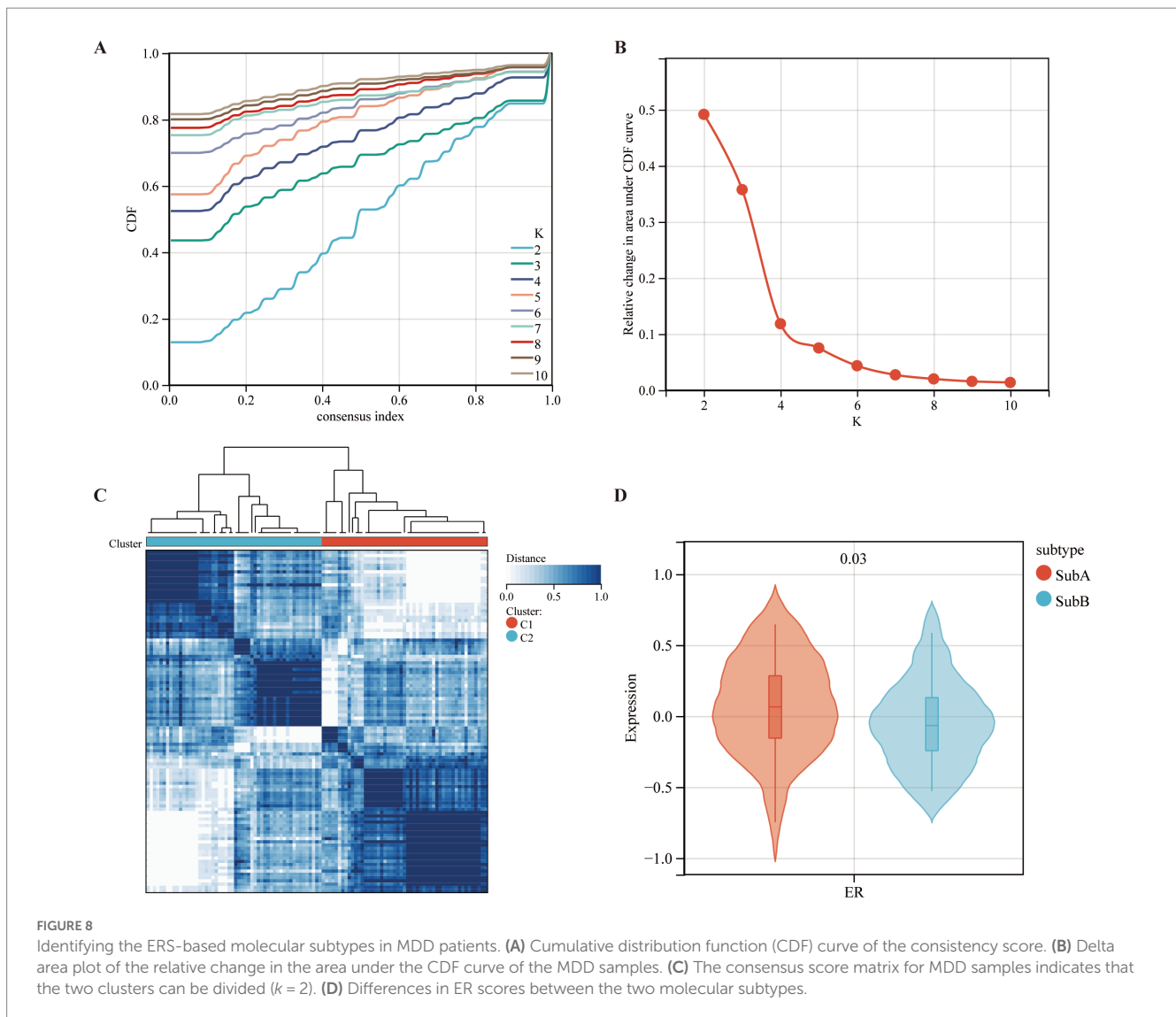
Discussion

MDD is a highly debilitating disorder that affects millions of people worldwide and places a burden on families and communities (34). Given the high heterogeneity and complex pathological features of MDD, its diagnosis remains challenging. Current diagnosis mainly relies on the clinical assessment of patients' self-reported symptoms and lacks objective tests; therefore, MDD still requires the use of specific markers for a definitive diagnosis (35). In addition, evidence suggests that high rates of misdiagnosis may contribute to poor recovery in patients with MDD due to limited knowledge of the diagnostic markers of the disease (36). Therefore, there is an urgent need to develop reliable detection methods for clinical practise. Previous evidence points to the involvement of ERS in the pathogenesis of MDD; however, few relevant diagnostic markers have been identified. This study aimed to screen for ERS genes associated with MDD and explore their potential diagnostic value. In this study, we developed a diagnostic model based on seven ERS-related genes. The validation results indicated that it had good diagnostic performance and closely correlated with the level of immune cell infiltration. Furthermore, two molecular subtypes associated with ERS were identified that differed significantly in terms of ER scores and immune characteristics.
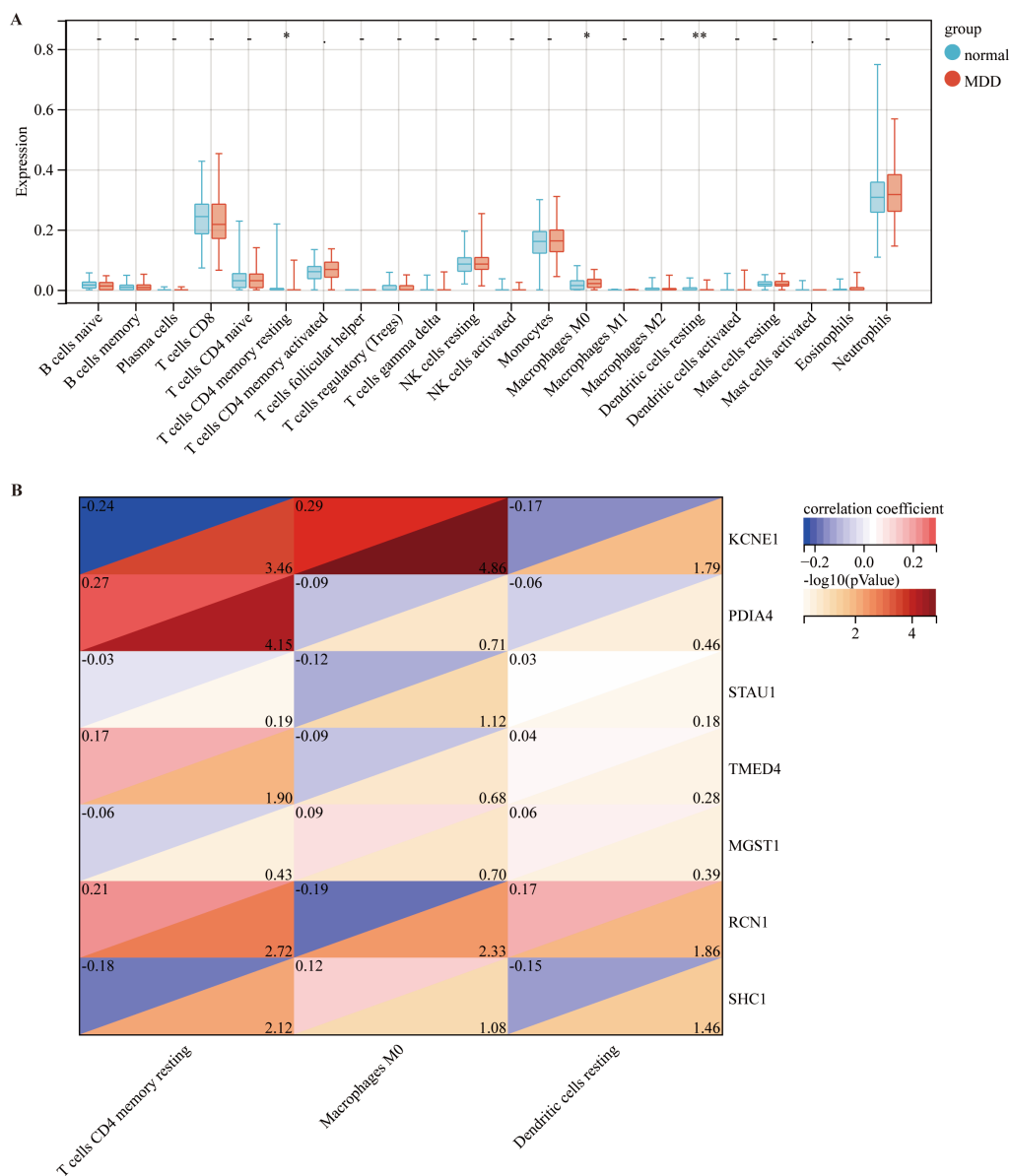
FIGURE 7

Comparison of immune cell infiltration amongst different disease groups. **(A)** Comparison of immune cell infiltration between control (blue box) and MDD (red box) groups. **(B)** Correlation of seven diagnostic markers with three types of immune cell infiltration in MDD.

Seven genes were included in the diagnostic model: KCNE1, PDIA4, STAU1, TMED4, MGST1, RCN1 and SHC1. The product of Potassium Voltage-Gated Channel Subfamily E Regulatory Subunit 1 (KCNE1) belongs to the KCNE family and modulates the function of voltage-gated K(+) channels (37). Previous research has shown that KCNE1 regulates neuronal K(+) channels and resting membrane potential (38). The role of KCNE1 in MDD has not been extensively studied, with only McCaffery et al. proposing that KCNE1 is associated with changes in depressive symptoms over the course of a year (39). Notably, the contribution of K(+) channels is mainly due to the expression of the KCNE, KCNQ, and ERG isoforms (40). Amongst these, the modulators

(retigabine) of KCNQ channels have been shown to improve depressive symptoms and have the potential to treat MDD (41). Based on these findings, we speculate that KCNE1 may also be involved in the neurophysiological mechanisms of stress recovery by modulating neuronal activity, thus presenting a therapeutic effect on depression. However, this hypothesis should be confirmed in future studies. In this study, we observed a close correlation between KCNE1 and resting CD4+ memory T cells and M0 macrophages. However, these results have not yet been investigated and should be explored in future studies. As an ERS gene, protein disulphide isomerase family member 4 (PDIA4) is significantly correlated with the expression levels of inflammatory

FIGURE 8
Identifying the ERS-based molecular subtypes in MDD patients. **(A)** Cumulative distribution function (CDF) curve of the consistency score. **(B)** Delta area plot of the relative change in the area under the CDF curve of the MDD samples. **(C)** The consensus score matrix for MDD samples indicates that the two clusters can be divided (k = 2). **(D)** Differences in ER scores between the two molecular subtypes.

cytokines, a feature of MDD pathology (42, 43). Although the relationship between PDIA4 and MDD has not been reported, we speculate that this gene may be involved in pathological changes in MDD by influencing inflammatory factors. Staufen 1 (STAU1) is generally expressed in mammals, and its downregulation reduces the amplitude and frequency of small postsynaptic excitatory currents, suggesting that STAU1 is important for the processing or transport of dendritic mRNA (44). Interestingly, dendritic mRNA are critical for maintaining changes in functional connectivity, such as hippocampus-dependent learning and memory (45). Microsomal glutathione S-transferase 1 (MGST1) has been proven to be involved in the regulation of oxidative stress (46). Disruption of the insulin pathway in the brain is involved in the pathogenesis of depression (47). It has been reported that Src homology 2 domain containing (Shc1) is activated by the insulin receptor, which in turn activates Grb2 (48). The active Shc1/Grb2 complex stimulates intracellular

signalling pathways, and its disturbance may contribute to spatial memory deficits in rats after water maze training (49). Meanwhile, antidepressant drugs may exert beneficial effects on the insulin receptor phosphorylation pathway through the Shc1/Grb2 complex (50), which further supports the potential use of Sch1 for depression drug development. However, direct evidence of the relationship between these genes and MDD pathogenesis has not been reported, and their functions in MDD require further exploration.

In the validation analysis, we confirmed that the model had good predictive power and that these seven genes could be used as diagnostic markers for MDD. Inflammatory processes are specifically involved in MDD, and the immune profiles of MDD and control groups were analysed. The results revealed that the infiltration level of M0 macrophages was higher in MDD samples than in control samples, which was also observed by Zhang et al. (51). A previous study indicated that a higher production of
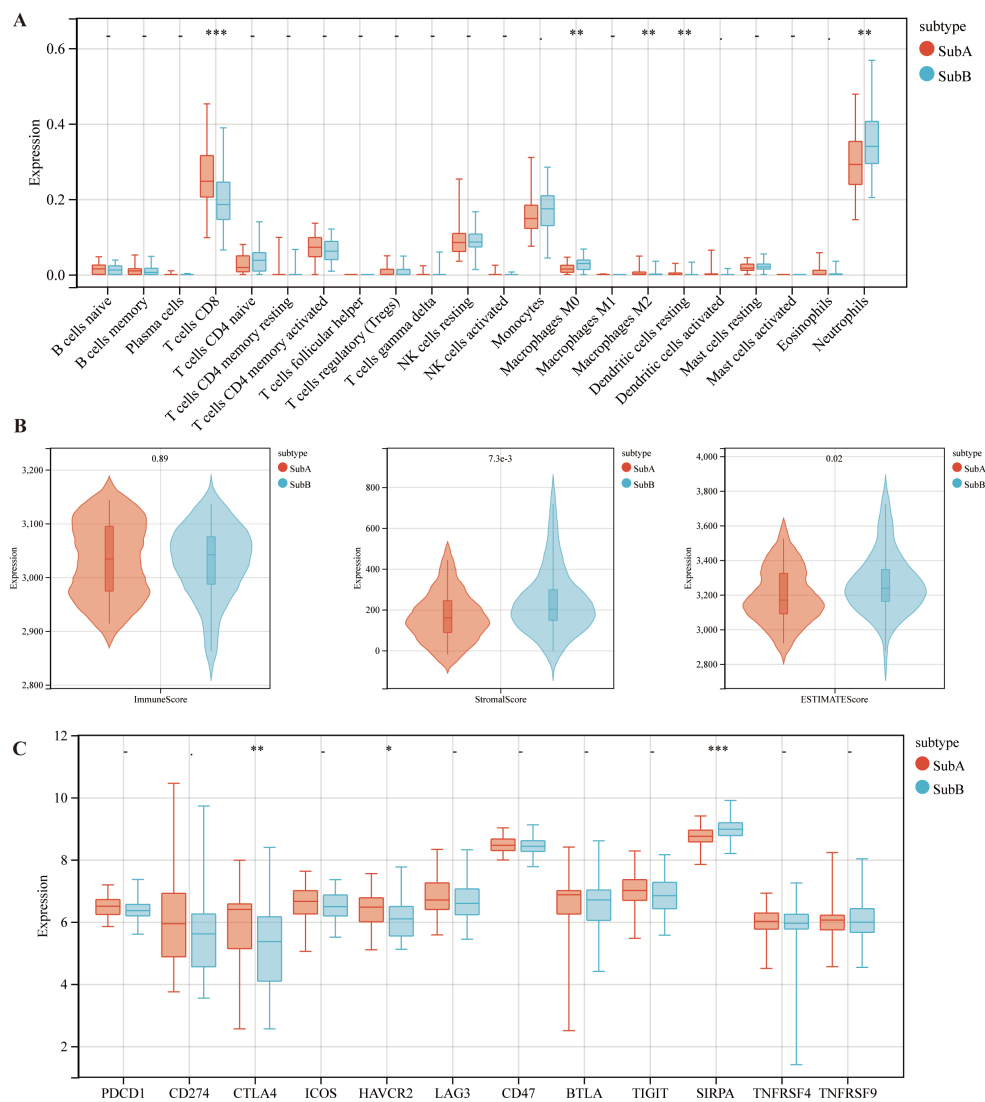
FIGURE 9

The distributions of immune cells and immune checkpoint genes between SubA and SubB subtypes. **(A)** Levels of immune cell infiltration between the two subtypes. **(B)** Differential distribution of immune features between the two subtypes. **(C)** Differential expression patterns of several immune checkpoint genes in the two subtypes.

pro-inflammatory cytokines was detected in M0 macrophages in autologous sera from patients with MDD than in control samples (52). Resting CD4 memory T cells and resting dendritic cells also exhibited different levels of infiltration in the MDD and control groups. However, the mechanisms underlying these complex interactions between the diagnostic genes and immune cells require further investigation.

Furthermore, two ERS-related molecular subtypes for MDD patients were identified. SubA had higher expression value of ER scores than SubB, suggesting that patients in this subtype may be accompanied by a more pronounced activation of ERS. Moreover, the SubB subtype appeared to be immune-infiltrative because of its higher stromal and estimated scores than those of the SubA subtype. Evidence indicates that

depression may damage the immune system and lead to immunosuppression (53). Thus, we speculated that patients in the SubA subtype may have higher severity of MDD. However, due to the lack of clinical information about patients, the correlation between subtypes and clinical characteristics is not evaluated, which is necessary to understand the clinical significance of subtypes. Importantly, we can confirm that ERS can regulate the immune microenvironment of MDD, thus affecting the course of the disease (54). In the future, the role of immune cell dynamics in different subtypes needs to be further explored.

Growing evidence suggests that drugs with major immune targets can ameliorate depressive symptoms (55). In addition, hub genes associated with SubB subtypes were screened by using
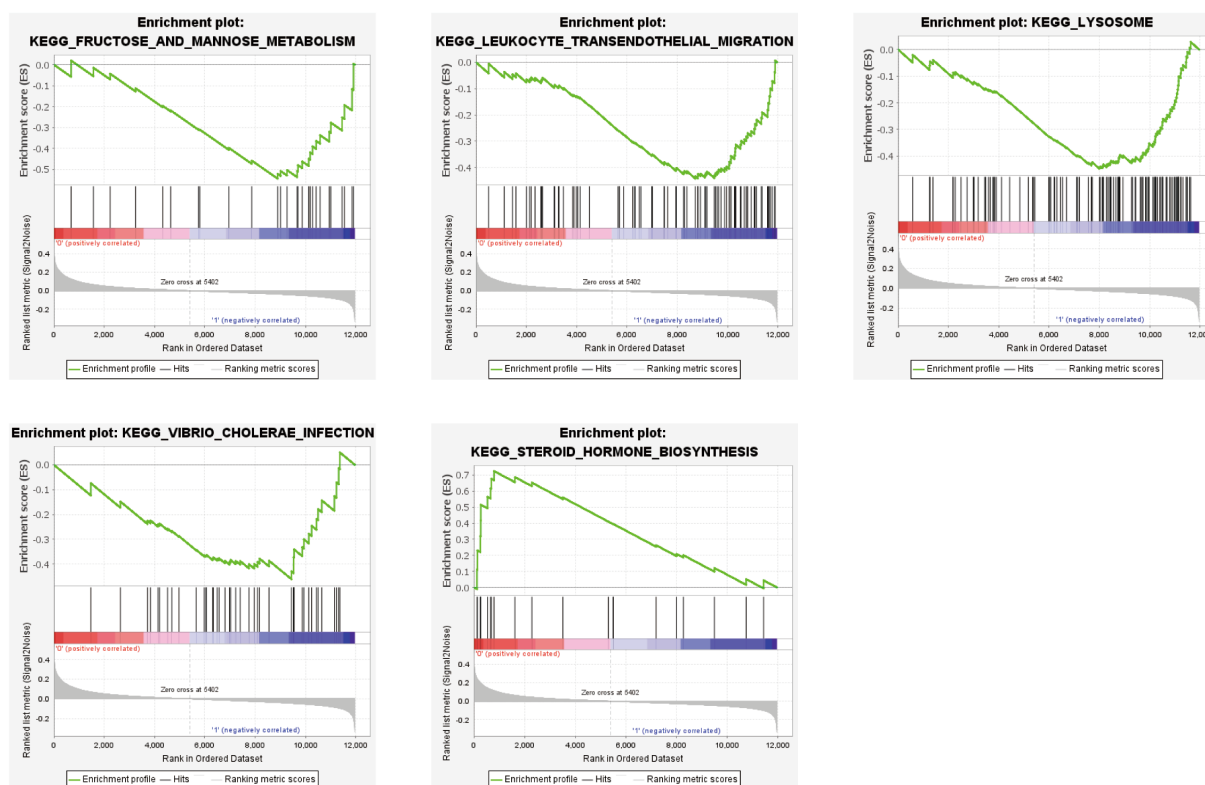
FIGURE 10
GSEA revealing the significantly enriched biological pathways between SubA and SubB subtypes.

WGCNA, such as NCF4, NCF2, CSF3R and FPR2. It has been showed that CSF3R is a cytokine that controls neutrophil expansion and differentiation, and can serve as a biomarker for neuromodulation; FPR2 knockdown may maintain hippocampal homeostasis by preventing depression-related neuronal damage (56, 57). The proteins encoded by NCF4 are cytoplasmic regulatory components of superoxide-producing phagocytic NADPH oxidase, which is a multicomponent enzyme system important for host defence (58). It primarily interacts with NCF2 and binds to NCF1 to form complexes that are transferred to membranes in response to cell stimulation (59). However, the functions of NCF4 and NCF2 have not yet been investigated in patients with MDD.

To the best of our knowledge, this is the first bioinformatic report describing ERS-related diagnostic genes and molecular subtypes of MDD. The present analysis has some unique advantages. Compared with previous diagnostic model construction (51), our analysis employs a larger number of datasets and identifies two molecular subtypes associated with ERS, which may provide a more comprehensive understanding of MDD. Notably, our findings clarify several diagnostic signatures and molecular subtypes from the perspective of ERS, which can provide a deeper understanding of the molecular heterogeneity of depression; meanwhile, it is helpful to assist psychiatrists to formulate accurate and individualised treatment schemes, thus reducing the burden of depression.

Furthermore, previous study has indicated that antidepressant drugs or natural compounds can exert therapeutic effects via reducing ERS (17), and the selected biomarkers in this work could provide information for the development of more effective treatments.

Our study had some limitations. First, MDD is clinically heterogeneous; however, the disease type of the enrolled patients was not recorded in detail in the database. Therefore, this point was not considered in this analysis. In addition, the results were obtained through bioinformatics analysis and are still at the predictive stage, so there is insufficient experimental evidence to confirm our findings. Thus, further experiments are needed to confirm the specific mechanism of action of diagnostic genes in MDD.

Conclusion

Taken together, our results established a diagnostic model based on seven ERS-related genes that exhibited robust and good estimation performance. Simultaneously, two ERS-associated molecular subtypes with different ER scores and immune characteristics were screened. Our findings provide a reliable model for MDD diagnosis and development of individualised treatment plans from an ERS perspective.
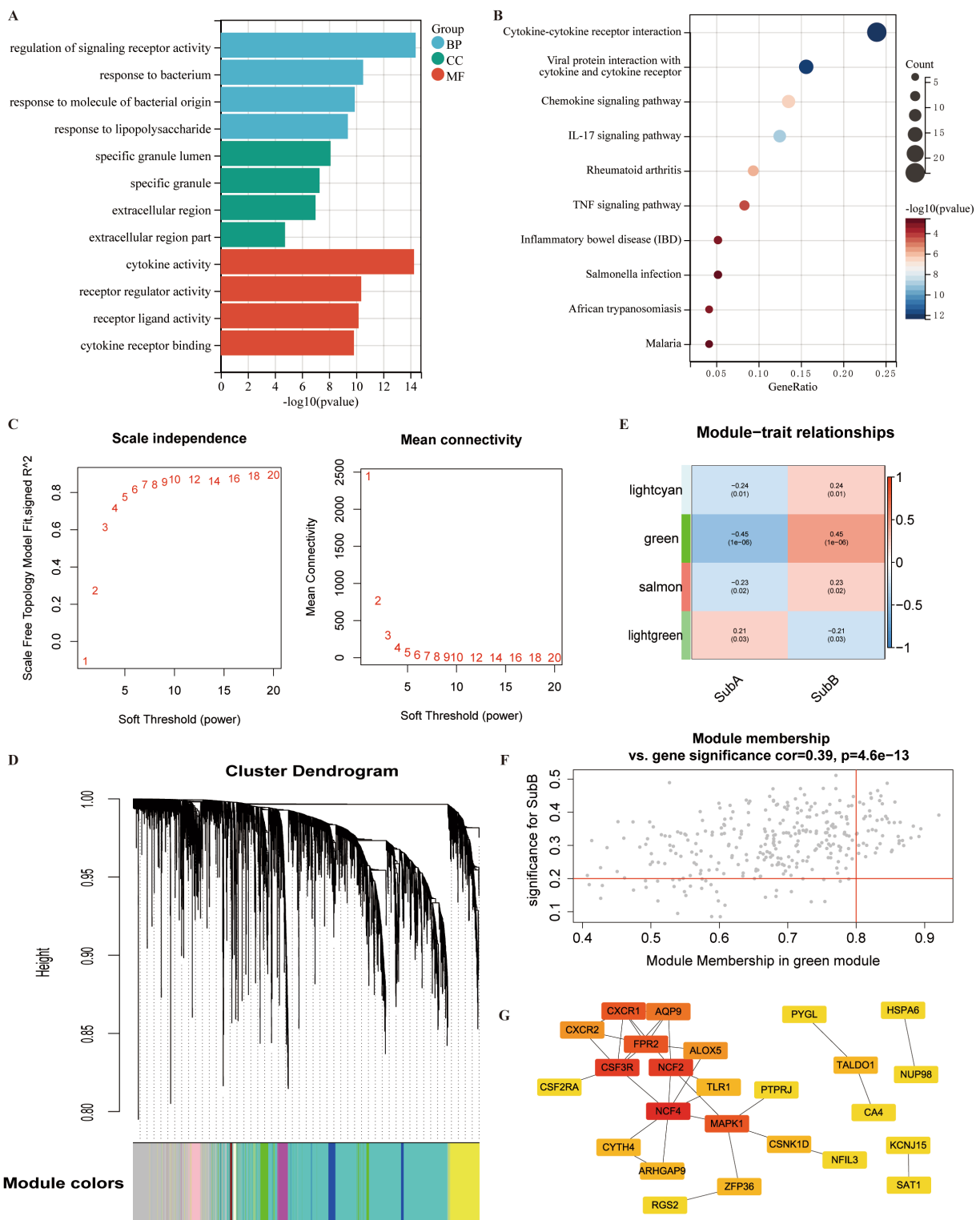
FIGURE 11
Identification of DEGs and hub genes associated with ERS subtypes. **(A)** Gene Ontology (GO) term analysis of DEGs between SubA and SubB subtypes. **(B)** KEGG pathway analysis of DEGs within the SubA vs. SubB groups. **(C)** Analysis of scale-free indices and mean connectivity for various soft-threshold powers. **(D)** Cluster dendrogram developed using weighted correlation coefficients. Each colour represents a module. **(E)** Correlation between modules and ERS subtypes. The upper numbers indicate the correlation coefficients and lower numbers indicate the value of *p*. **(F)** Scatter plot of correlations between gene significance (GS) and module membership (MM). **(G)** PPI network of the hub genes in the green module. Red indicates the nodes with high connectivity.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

SH: conceptualization, data analysis, and draft preparation. YL, JS, and WL: data collection and analysis. CL: supervision, reviewing, and editing. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2023.1168516/full#supplementary-material

References

1. Lim GY, Tam WW, Lu Y, Ho CS, Zhang MW, Ho RC. Prevalence of depression in the community from 30 countries between 1994 and 2014. *Sci Rep*. (2018) 8:2861. doi: 10.1038/s41598-018-21243-x

2. Rehm J, Shield KD. Global burden of disease and the impact of mental and addictive disorders. *Curr Psychiatry Rep*. (2019) 21:10. doi: 10.1007/s11920-019-0997-0

3. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. *Lancet*. (2013) 382:1575–86. doi: 10.1016/S0140-6736(13)61611-6

4. Davis AK, Barrett FS, May DG, Cosimano MP, Sepeda ND, Johnson MW, et al. Effects of psilocybin-assisted therapy on major depressive disorder: a randomized clinical trial. *JAMA Psychiat*. (2021) 78:481–9. doi: 10.1001/jamapsychiatry.2020.3285

5. Gonda X, Petschner P, Eszlari N, Baksa D, Edes A, Antal P, et al. Genetic variants in major depressive disorder: from pathophysiology to therapy. *Pharmacol Ther*. (2019) 194:22–43. doi: 10.1016/j.pharmthera.2018.09.002

6. Blackburn TP. Depressive disorders: treatment failures and poor prognosis over the last 50 years. *Pharmacol Res Perspect*. (2019) 7:e00472. doi: 10.1002/prp2.472

7. Schwarz DS, Blower MD. The endoplasmic reticulum: structure, function and response to cellular signaling. *Cell Mol Life Sci*. (2016) 73:79–94. doi: 10.1007/s00018-015-2052-6

8. Di Conza G, Ho P-C. ER stress responses: an emerging modulator for innate immunity. *Cells*. (2020) 9:695. doi: 10.3390/cells9030695

9. Gold PW, Goodwin FK. Chrousos GP: clinical and biochemical manifestations of depression. Relation to the neurobiology of stress (2). *N Engl J Med*. (1988) 319:413–20. doi: 10.1056/NEJM198808183190706

10. Bown C, Wang JF, MacQueen G, Young LT. Increased temporal cortex ER stress proteins in depressed subjects who died by suicide. *Neuropsychopharmacology*. (2000) 22:327–32. doi: 10.1016/S0893-133X(99)00091-3

11. Patel S, Sharma D, Kalia K, Tiwari V. Crosstalk between endoplasmic reticulum stress and oxidative stress in schizophrenia: the dawn of new therapeutic approaches. *Neurosci Biobehav Rev*. (2017) 83:589–603. doi: 10.1016/j.neubiorev.2017.08.025

12. Liu L, Zhao Z, Lu L, Liu J, Sun J, Wu X, et al. Icariin and icaritin ameliorated hippocampus neuroinflammation via inhibiting HMGB1-related pro-inflammatory signals in lipopolysaccharide-induced inflammation model in C57BL/6 J mice. *Int Immunopharmacol*. (2019) 68:95–105. doi: 10.1016/j.intimp.2018.12.055

13. Nevell L, Zhang K, Aiello AE, Koenen K, Galea S, Soliven R, et al. Elevated systemic expression of ER stress related genes is associated with stress-related mental disorders in the Detroit Neighborhood health study. *Psychoneuroendocrinology*. (2014) 43:62–70. doi: 10.1016/j.psyneuen.2014.01.013

14. Koksal AR, Verne GN, Zhou Q. Endoplasmic reticulum stress in biological processing and disease. *J Investig Med*. (2021) 69:309–15. doi: 10.1136/jim-2020-001570

15. Detweiler-Bedell JB, Friedman MA, Leventhal H, Miller IW, Leventhal EA. Integrating co-morbid depression and chronic physical disease management: identifying and resolving failures in self-regulation. *Clin Psychol Rev*. (2008) 28:1426–46. doi: 10.1016/j.cpr.2008.09.002

16. Martinon F. The endoplasmic reticulum: a sensor of cellular stress that modulates immune responses. *Microbes Infect*. (2012) 14:1293–300. doi: 10.1016/j.micinf.2012.07.005

17. Mao J, Hu Y, Ruan L, Ji Y, Lou Z. Role of endoplasmic reticulum stress in depression (review). *Mol Med Rep*. (2019) 20:4774–80. doi: 10.3892/mmr.2019.10789

18. Leday GGR, Vértes PE, Richardson S, Greene JR, Regan T, Khan S, et al. 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*. (2018) 83:70–80. doi: 10.1016/j.biopsych.2017.01.021

19. Spijker S, Van Zanten JS, De Jong S, Penninx BWJH, van Dyck R, Zitman FG, et al. Stimulated gene expression profiles as a blood marker of major depressive disorder. *Biol Psychiatry*. (2010) 68:179–86. doi: 10.1016/j.biopsych.2010.03.017

20. Yi Z, Li Z, Yu S, Yuan C, Hong W, Wang Z, et al. Blood-based gene expression profiles models for classification of subsyndromal symptomatic depression and major depressive disorder. *PLoS One*. (2012) 7:e31283. doi: 10.1371/journal.pone.0031283

21. Belzeaux R, Bergon A, Jeanjean V, Loriod B, Formisano-Tréziny C, Verrier L, et al. Responder and nonresponder patients exhibit different peripheral transcriptional signatures during major depressive episode. *Transl Psychiatry*. (2012) 2:e185. doi: 10.1038/tp.2012.112

22. 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:e47. doi: 10.1093/nar/gkv007

23. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. (2021) 49:D605–12. doi: 10.1093/nar/gkaa1074

24. Pan X, Jin X, Wang J, Hu Q, Dai B. Placenta inflammation is closely associated with gestational diabetes mellitus. *Am J Transl Res*. (2021) 13:4068–79.

25. Goeman JJ. L1 penalized estimation in the cox proportional hazards model. *Biom J*. (2010) 52:70–84. doi: 10.1002/bimj.200900028

26. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. (2011) 12:77. doi: 10.1186/1471-2105-12-77

27. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. (2010) 26:1572–3. doi: 10.1093/bioinformatics/btq170

28. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol.* (2018) 1711:243–59. doi: 10.1007/978-1-4939-7493-1_12

29. 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. *Proc Natl Acad Sci U S A.* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102

30. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *Gene Ontol Consortium Nat Genet.* (2000) 25:25–9. doi: 10.1038/75556

31. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* (2000) 28:27–30. doi: 10.1093/nar/28.1.27

32. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* (2008) 9:559. doi: 10.1186/1471-2105-9-559

33. 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) 13:2498–504. doi: 10.1101/gr.1239303

34. Smith K. Mental health: a world of depression. *Nature.* (2014) 515:181. doi: 10.1038/515180a

35. Alexopoulos GS. Mechanisms and treatment of late-life depression. *Transl Psychiatry.* (2019) 9:188. doi: 10.1038/s41398-019-0514-6

36. Tomasik J, Han SYS, Barton-Owen G, Mirea D-M, Martin-Key NA, Rustogi N, et al. A machine learning algorithm to differentiate bipolar disorder from major depressive disorder using an online mental health questionnaire and blood biomarker data. *Transl Psychiatry.* (2021) 11:41. doi: 10.1038/s41398-020-01181-x

37. Abbott GW. KCNE1 and KCNE3: the yin and yang of voltage-gated K(+) channel regulation. *Gene.* (2016) 576:1–13. doi: 10.1016/j.gene.2015.09.059

38. Melman YF, Um SY, Krumerman A, Kagan A, McDonald TV. KCNE1 binds to the KCNQ1 pore to regulate potassium channel activity. *Neuron.* (2004) 42:927–37. doi: 10.1016/j.neuron.2004.06.001

39. McCaffery JM, Papandonatos GD, Faulconbridge LF, Erar B, Peter I, Wagenknecht LE, et al. Genetic predictors of depressive symptoms in the look AHEAD trial. *Psychosom Med.* (2015) 77:982–92. doi: 10.1097/PSY.0000000000000242

40. Ohya S, Asakura K, Muraki K, Watanabe M, Imaizumi Y. Molecular and functional characterization of ERG, KCNQ, and KCNE subtypes in rat stomach smooth muscle. *Am J Physiol Gastrointest Liver Physiol.* (2002) 282:G277–87. doi: 10.1152/ajpgi.00200.2001

41. Costi S, Han M-H, Murrough JW. The potential of KCNQ Potassium Channel openers as novel antidepressants. *CNS Drugs.* (2022) 36:207–16. doi: 10.1007/s40263-021-00885-y

42. Lee C-H, Chiang C-F, Lin F-H, Kuo F-C, Su S-C, Huang C-L, et al. PDIA4, a new endoplasmic reticulum stress protein, modulates insulin resistance and inflammation in skeletal muscle. *Front Endocrinol (Lausanne).* (2022) 13:1053882. doi: 10.3389/fendo.2022.1053882

43. Shi R, Gwee X, Chua DQ, Tan CT, Yap KB, Larbi A, et al. Inflammatory markers and incident depression: evidence in a population-based prospective study. *Psychoneuroendocrinology.* (2022) 142:105806. doi: 10.1016/j.psyneuen.2022.105806

44. Lebeau G, Maher-Laporte M, Topolnik L, Laurent CE, Sossin W, Desgroseillers L, et al. Staufen1 regulation of protein synthesis-dependent long-term potentiation and synaptic function in hippocampal pyramidal cells. *Mol Cell Biol.* (2008) 28:2896–907. doi: 10.1128/MCB.01844-07

45. Bramham CR, Wells DG. Dendritic mRNA: transport, translation and function. *Nat Rev Neurosci.* (2007) 8:776–89. doi: 10.1038/nrn2150

46. Dai H, Lu X. MGST1 alleviates the oxidative stress of trophoblast cells induced by hypoxia/reoxygenation and promotes cell proliferation, migration, and invasion by activating the PI3K/AKT/mTOR pathway. *Open Med (Wars).* (2022) 17:2062–71. doi: 10.1515/med-2022-0617

47. Moulton CD, Pickup JC, Ismail K. The link between depression and diabetes: the search for shared mechanisms. *Lancet Diabetes Endocrinol.* (2015) 3:461–71. doi: 10.1016/S2213-8587(15)00134-5

48. Zheng Y, Zhang C, Croucher DR, Soliman MA, St-Denis N, Pasculescu A, et al. Temporal regulation of EGF signalling networks by the scaffold protein Shc1. *Nature.* (2013) 499:166–71. doi: 10.1038/nature12308

49. Zhao Q, Williams BL, Abraham RT, Weiss A. Interdomain B in ZAP-70 regulates but is not required for ZAP-70 signaling function in lymphocytes. *Mol Cell Biol.* (1999) 19:948–56. doi: 10.1128/MCB.19.1.948

50. Głombik K, Ślusarczyk J, Trojan E, Chamera K, Budziszewska B, Lasoń W, et al. Regulation of insulin receptor phosphorylation in the brains of prenatally stressed rats: new insight into the benefits of antidepressant drug treatment. *Eur Neuropsychopharmacol.* (2017) 27:120–31. doi: 10.1016/j.euroneuro.2016.12.005

51. Zhang J, Xie S, Chen Y, Zhou X, Zheng Z, Yang L, et al. Comprehensive analysis of endoplasmic reticulum stress and immune infiltration in major depressive disorder. *Front Psych.* (2022) 13:1008124. doi: 10.3389/fpsyt.2022.1008124

52. Cosma NC, Üsekes B, Otto LR, Gerike S, Heuser I, Regen F, et al. M1/M2 polarization in major depressive disorder: disentangling state from trait effects in an individualized cell-culture-based approach. *Brain Behav Immun.* (2021) 94:185–95. doi: 10.1016/j.bbi.2021.02.009

53. Zhou D, Yu H, Yao H, Yuan S, Xia Y, Huang L, et al. A novel joint index based on peripheral blood CD4+/CD8+ T cell ratio, albumin level, and monocyte count to determine the severity of major depressive disorder. *BMC Psychiatry.* (2022) 22:248. doi: 10.1186/s12888-022-03911-5

54. Ciobanu LG, Sachdev PS, Trollor JN, Reppermund S, Thalamuthu A, Mather KA, et al. Co-expression network analysis of peripheral blood transcriptome identifies dysregulated protein processing in endoplasmic reticulum and immune response in recurrent MDD in older adults. *J Psychiatr Res.* (2018) 107:19–27. doi: 10.1016/j.jpsychires.2018.09.017

55. Breit S, Mazza E, Poletti S, Benedetti F. White matter integrity and pro-inflammatory cytokines as predictors of antidepressant response in MDD. *J Psychiatr Res.* (2023) 159:22–32. doi: 10.1016/j.jpsychires.2022.12.009

56. Wang Y, Shao Y, Zhang H, Wang J, Zhang P, Zhang W, et al. Comprehensive analysis of key genes and pathways for biological and clinical implications in thyroid-associated ophthalmopathy. *BMC Genomics.* (2022) 23:630. doi: 10.1186/s12864-022-08854-5

57. Peritore AF, Crupi R, Scuto M, Gugliandolo E, Siracusa R, Impellizzeri D, et al. The role of Annexin A1 and formyl peptide receptor 2/3 Signaling in chronic corticosterone-induced depression-like behaviors and impairment in hippocampal-dependent memory. *CNS Neurol Disord Drug Targets.* (2020) 19:27–43. doi: 10.2174/1871527319666200107094732

58. Nunoi H, Nakamura H, Nishimura T, Matsukura M. Recent topics and advanced therapies in chronic granulomatous disease. *Hum Cell.* (2022) 36:515–27. doi: 10.1007/s13577-022-00846-7

59. Tarazona-Santos E, Machado M, Magalhães WCS, Chen R, Lyon F, Burdett L, et al. Evolutionary dynamics of the human NADPH oxidase genes CYBB, CYBA, NCF2, and NCF4: functional implications. *Mol Biol Evol.* (2013) 30:2157–67. doi: 10.1093/molbev/mst119