Article

Yu Wu[1,†]
Zhichao Wang[2,†]
Houjia Hu[3]
Tong Wu[4]
Alabed Ali A. Alabed[5]
Zhenghai Sun[4]
Yuchen Wang[4]
Guangcheng Cui[4]
Weiliang Cong[6]
Chengchong Li[4,*]
Ping Li[4,*]

Identification of Immune-Related Gene Signature in Schizophrenia

[1]School of Nursing, Qiqihar Medical University, 161000 Qiqihar, Heilongjiang, China
[2]Department of Academic Research, Qiqihar Medical University, 161000 Qiqihar, Heilongjiang, China
[3]School of Basic Medical Sciences, Nanchang University, 330006 Nanchang, Jiangxi, China
[4]Department of Psychology, Qiqihar Medical University, 161000 Qiqihar, Heilongjiang, China
[5]Community Medicine Department, Faculty of Medicine, Lincoln University College, 47301 Petaling Jaya, Selangor, Malaysia
[6]Department of Anaesthesiology, The Third Affiliated Hospital of Qiqihar Medical University, 161000 Qiqihar, Heilongjiang, China

Abstract

Background: Schizophrenia (SCZ) is a type of psychiatric disorder characterized by multiple symptoms. Our aim is to decipher the relevant mechanisms of immune-related gene signatures in SCZ.

Methods: The SCZ dataset and its associated immunoregulatory genes were retrieved using Gene Expression Omnibus (GEO) and single-sample gene set enrichment analysis (ssGSEA). Co-expressed gene modules were determined through weighted gene correlation network analysis (WGCNA). To elucidate the functional characteristics of these clusters, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were used. Additionally, gene set enrichment analysis (GSEA) and Gene Set Variation Analysis (GSVA) were conducted to identify enriched pathways for the immune subgroups. A protein-protein interaction (PPI) network analysis was performed to identify core genes relevant to SCZ.

Results: A significantly higher immune score was observed in SCZ compared to control samples. Seven distinct gene modules were identified, with genes highlighted in green selected for further analysis. Using the Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT) method, degrees of immune cell adhesion and accumulation related to 22 different immune cell types were calculated. Significantly enriched bioprocesses concerning the immunoregulatory genes with differential expressions included interferon-beta, IgG binding, and response to interferon-gamma, according to GO and KEGG analyses. Eleven hub genes related to immune infiltration emerged as key players among the three top-ranked GO terms.

Conclusions: This study underscores the involvement of immunoregulatory reactions in SCZ development. Eleven immune-related genes (*IFITM1* (interferon induced transmembrane protein 1), *GBP1* (guanylate binding protein 1), *BST2* (bone marrow stromal cell antigen 2), *IFITM3* (interferon induced transmembrane protein 3), *GBP2* (guanylate binding protein 2), *CD44* (CD44 molecule), *FCER1G* (Fc epsilon receptor Ig), *HLA-DRA* (major histocompatibility complex, class II, DR alpha), *FCGR2A* (Fc gamma receptor IIa), *IFI16* (interferon gamma inducible protein 16), and *FCGR3B* (Fc gamma receptor IIIb)) were identified as hub genes, representing potential biomarkers and therapeutic targets associated with the immune response in SCZ patients.

Keywords

schizophrenia; gene expression profiling; immune-related gene signature; KEGG; WGCNA

Introduction

Schizophrenia (SCZ) is a psychiatric disorder characterized by diverse manifestations, including delusions, dis-

*Corresponding author details: Chengchong Li, Department of Psychology, Qiqihar Medical University, 161000 Qiqihar, Heilongjiang, China. Email: lcchong@qmu.edu.cn; Ping Li, Department of Psychology, Qiqihar Medical University, 161000 Qiqihar, Heilongjiang, China. Email: qyliping@qmu.edu.cn
†These authors contributed equally.

organized speech and behavior, and hallucinations, among others. Additionally, certain SCZ patients may exhibit impairments in executive functioning and attention [1]. The lifetime risk of SCZ is approximately 1%, accompanied by significant morbidity and mortality. Additionally, this condition imposes substantial economic burdens on both families and society [2–4]. Despite the accessibility of Western medicine for schizophrenia, a significant proportion of patients exhibit a poor response to these treatments [5]. Currently, all available drugs for SCZ are recognized to take effect via blocking the type 2 dopaminergic receptor. However, over 60 years after discovering this theory, innovative effective target drugs have not been produced [6,7]. Previous studies have indicated that inherited genetic variants play an influential role in SCZ pathology, as evidenced by its high heritability [8,9]. Therefore, it is essential to employ a bioinformatics approach to identify a novel gene signature associated with the development of SCZ.

The field of immunotherapy offers a novel and promising approach for the treatment of certain malignancies, such as breast cancer and hepatocellular carcinoma [10,11]. Immune responses, including the infiltration of immune cells, actively contribute to the pathological mechanisms underlying a wide range of diseases [12]. SCZ is a complex disorder resulting from the interplay of genetic, immune, and other factors [13]. Previous research has demonstrated that various SCZ phenotypes are associated with the stimulation of the immune-inflammatory response system (IRS) [14,15]. The relationship between the immune system and the incidence and progression of mental illnesses has become an increasingly important topic in psychiatry. Relevant investigations have revealed a strong role of the immune response in psychiatric disorders [16]. Likewise, this study found that immunomodulatory genes are closely linked to the onset and progression of SCZ. However, the use of an immune gene expression-based signature in SCZ has not been fully explored.

The objective of this study was to utilize a bioinformatics approach to identify a gene signature correlated with immune infiltration and its role in SCZ development. The Gene Expression Omnibus (GEO) and single-sample gene set enrichment analysis (ssGSEA) were employed to obtain the SCZ dataset, calculate immune-related gene modules, and grade the samples accordingly. Based on the proportions of immune infiltrations, the samples were classified into high immunity and low immunity groups. The researchers then identified genes with differential expressions related to these subgroups and intersected them with genes that had strong immunological associations from the weighted gene correlation network analysis (WGCNA). This allowed for the classification of individuals with SCZ into two clusters of immune-related differentially expressed genes (IDEGs), labeled A and B. The differentially expressed genes (DEGs) between these two clusters were further analyzed using gene set enrichment analysis (GSEA) for function annotations. Additionally, protein-protein interaction (PPI) networks of IDEGs were constructed to identify core genes associated with SCZ. Finally, these hub genes were utilized to predict regulatory networks involved in the etiology of SCZ. These findings pave the way for a better understanding of the immunoregulatory role in SCZ development.

Materials and Methods

GEO Data Acquisition

The original data for SCZ was obtained from three datasets: GSE17612 [17], GSE21138 [18], and GSE21935 [19], all available on the GEO (https://www.ncbi.nlm.nih.gov/geo/) [20], and processed using the GEOquery R package (version 3.1.3, https://bioconductor.org/packages/release/bioc/html/GEOquery.html) [21]. The GSE17612 dataset consists of a total of 51 samples, including 23 control and 28 SCZ samples, utilizing the GPL570 data platform. The GSE21138 dataset comprises 59 samples, with 29 control and 30 SCZ samples, and uses the GPL1570 data platform. Similarly, the GSE21935 dataset includes 42 samples, with 19 control and 23 SCZ samples, utilizing the GPL1570 data platform. All data from these samples was inputted into the limma R package (version 4.0.2, The R Foundation for Statistical Computing, Vienna, Austria) [22] for normalization and standardization of the expression profiles. To mitigate batch effects, the ComBat function from the sva package [23] was employed to integrate the three datasets.

Immune Score Analysis of Samples

The ssGSEA technique [24] and the ImmPort platform (https://immport.niaid.nih.gov) [25] were utilized to assess the contents of immunomodulatory genes from samples with SCZ. The ssGSEA approach enabled the calculation of immune gene expression levels, grading of the samples, and quantitative analysis of immune gene expression levels across the three datasets. By estimating the immune infiltration enrichment scores of disease samples, two subgroups with high and low immune levels were distinguished. For the ssGSEA analysis, the Gene Set Variation Analysis (GSVA) R package was employed. The ImmPort database served as the source of gene sets related to the immune system to be analyzed.

Identifying DEGs and Analyzing Enrichment

To identify DEGs related to the aforementioned two subgroups, this study employed the limma R package and established the criteria of log2FC > 0.5 and *p* adj < 0.05 for significance. The DEGs were visualized using the ggplot2 and pheatmap R packages. Additionally, Gene Ontology (GO) [25] and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment [26], as well as gene set enrichment analysis (GSEA), were performed using the ClusterProfiler package [27] in R software (version 3.6.3, University of Auckland, Auckland, New Zealand). The MSigDB database (http://www.gsea-msigdb.org/gsea/msigdb) [28] was utilized to obtain reference data including "Hallmark Gene sets", "KEGG Gene sets", and "GO Gene sets" for GSEA. A False Discovery Rate (FDR) value below 0.25 and a *p*-value smaller than 0.05 indicated statistical significance. The study also conducted Gene Set Variation Analysis (GSVA) [29] using the Tidyverse package within R software, selecting "C2.cp.all.v7.5.1.symbols" as the reference.

The WGCNA R package [30] was employed for weighted gene co-expression network analysis (WGCNA). To determine the soft threshold, the pickSoftThreshold function was utilized, revealing that a value of 3 was the best fit. With the soft threshold established, a scale-free network was constructed, a topology matrix was created, and hierarchical clustering was conducted. This exploration identified gene modules by dynamically cutting the hierarchical tree using a minimum module size of 30, and then calculated the Eigengenes. Next, a module correlation network on the basis of the Eigengenes was constructed, and hierarchical clustering was performed to generate 7 distinct modules. Finally, the Pearson statistical method was selected to determine the correlation between the modules and clinical features.

Distinguishing Pattern Based on Immune Related Genes

Consensus clustering serves as an approach to identify the characteristics of potential clusters from datasets, such as a dataset on microarray. To more accurately distinguish different immune subtypes of SCZ, ConsensusClusterPlus was utilized to conduct such an analysis of the target dataset. Specifically, we used the overlapping genes having the strongest immune association with WGCNA and the differential expressions between subgroups of high and low immunity. The figure of clusters was set to 9, and the process was repeated 1000 times while extracting 80% of the total samples. The k-means algorithm was performed for clustering, and the Spearman distance metric was used to calculate the similarity between samples.

PPI Network Construction

The STRING platform (https://cn.string-db.org/) [31] was selected to acquire the PPI network concerning hub genes. Subsequently, Cytoscape software (Version 3, Institute for Systems Biology, San Diego, CA, USA) was utilized for visualization. The importance of related genes was determined using CytoHubba [32], with those ranking at the top and having scores above 10 considered as core genes.

Functional Similarities Analysis

The GOSemSim package [33] in R software was utilized for performing functional similarity analysis of the core genes.

Immune Infiltration Analysis

The Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT) method [34] was employed to investigate the immune infiltration of different tissues from the three datasets. This facilitated the identification of any abundance dissimilarity of specific immune cell categories. CIBERSORT has been demonstrated as a reliable method for quantifying the presence of 22 different immune cell categories, including B cells, T cells, natural killer cells, mast cells, plasma cells, dendritic cells, neutrophils, eosinophils, and macrophages.

Statistics

R programming (https://www.r-project.org/, version 4.1.2) [35] was utilized for statistical analyses. Co-gene identification was performed using the Ggvenn R package. To compare two groups of continuous variables, if normally distributed, the group *t*-test was conducted, otherwise, the Mann-Whitney U test (i.e., Wilcoxon rank-sum test) was employed. Two-tailed *p* values below 0.05 indicated statistical significance.

Results

Identifying 326 Immune Infiltration-Related Genes in SCZ

The GEO database served as a source of an RNA expression matrix from three datasets, namely GSE17612, GSE21138, and GSE21935. The data was normalized using the limma package, and the resulting table is presented in Table 1 (Ref. [19,36,37]). To demonstrate the standardiza-

Table 1. The materials of SCZ GEO data sets.

GEO datasets	Platforms	Sample (SCZ/CTL)	Rows per platform
GSE17612 (Maycox *et al.* [36], 2009)	GPL570 [HG-U133_Plus_2]	28/23	22,647
GSE21138 (Narayan *et al.* [37], 2008)	GPL570 [HG-U133_Plus_2]	30/29	22,647
GSE21935 (Barnes *et al.* [19], 2011)	GPL570 [HG-U133_Plus_2]	23/19	22,647

SCZ, schizophrenia; GEO, Gene Expression Omnibus; CTL, control.

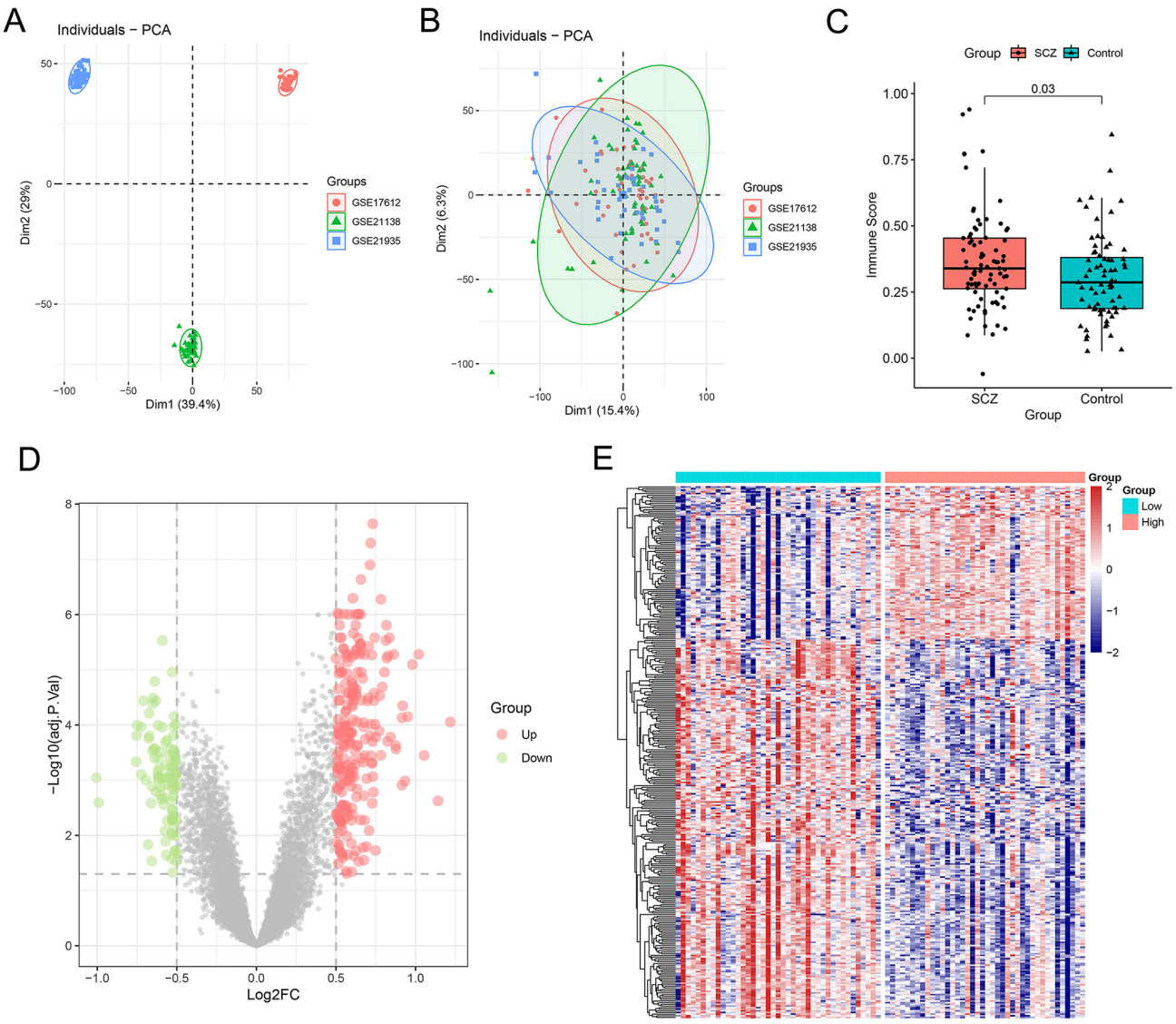

Fig. 1. Principal component analysis (PCA) analysis and Immune infiltration-related differentially expressed genes (DEGs). (A) PCA was conducted for the GSE17612, GSE21138, and GSE21935 gene datasets prior to normalization and batch effect adjustment. (B) PCA was performed on the three gene expression datasets post normalization and batch effect adjustment. (C) The difference in immune scores between SCZ and control groups was analyzed. (D) The volcano plot displays the DEGs in the high immunity and low immunity subgroups. Red dots symbolize the up-regulated genes, green dots for down-regulated genes, and gray dots represent the genes with no statistically differential expressions. The DEGs were identified based on a statistical threshold of adjusted *p*-value below 0.05 and a fold change above 0.5. (E) A heatmap shows the DEGs within the high immunity and low immunity subgroups. The subject IDs are shown on the x-axis, while genes with different expressions are shown on the y-axis. The genes with up-regulated expressions are reflected by red color, while the genes with down-regulated expressions are shown by blue color.

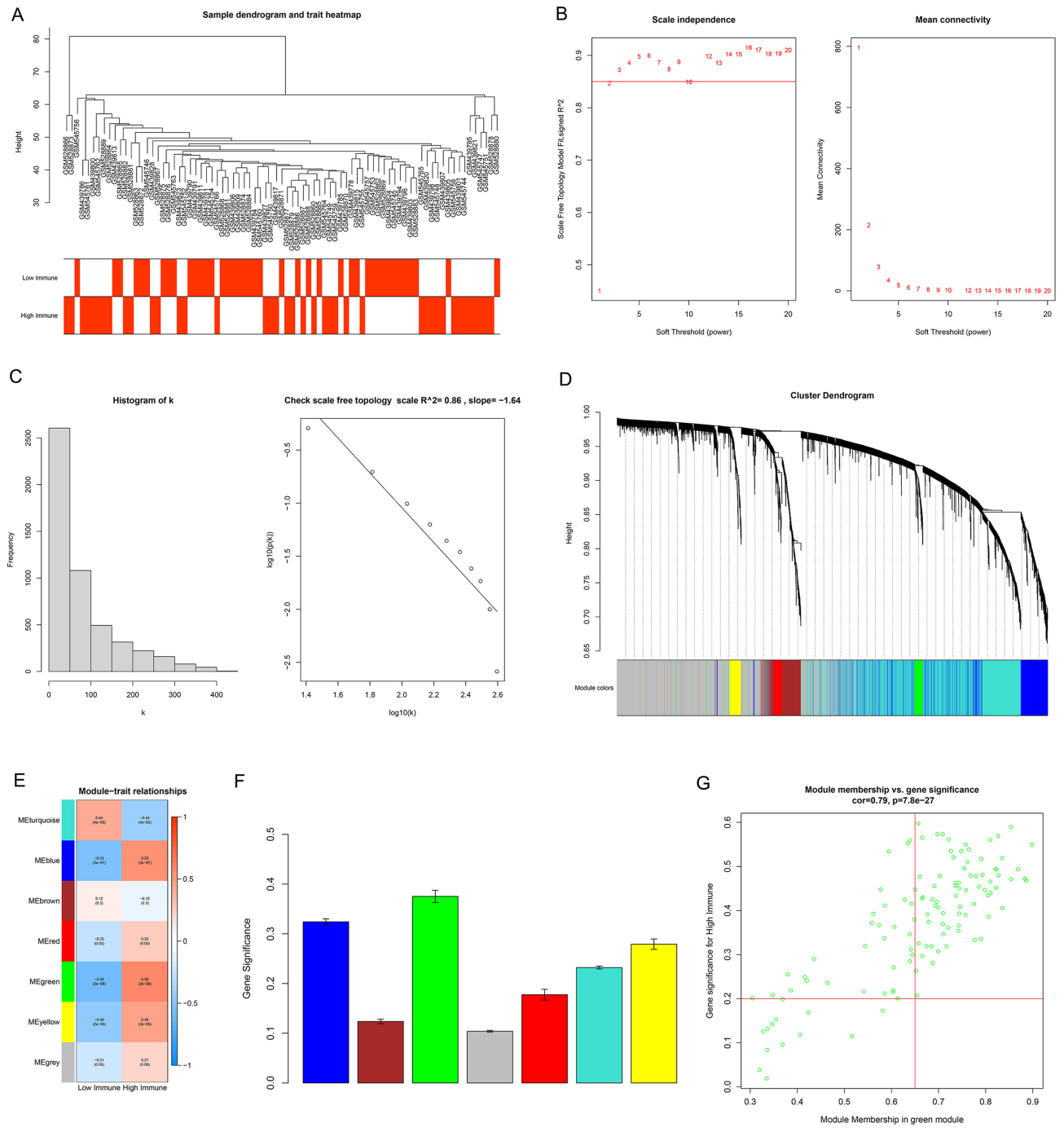

Fig. 2. Weighted correlation network analysis. (A) The grouping of SCZ samples based on their traits is presented in a dendrogram and heatmap. (B) The selection of the soft-thresholding power value was visualized by a plot. (C) The number of modules was determined and the scale-free topology was checked using a histogram when β was set to 3. (D) A dendrogram of genes and their respective module colors is displayed. Each color represents a different module. (E) The correlation between different modules and the immunophenotype is displayed in the heatmap. (F) The profile of average gene significance and errors in the modules related to the immune system of SCZ samples is presented. (G) The correlation between high immune phenotype and genes in the green module is displayed in a scatter plot.

tion and correction of batch effects across the three datasets, a principal component analysis (PCA) was conducted accordingly (Fig. 1A,B).

ssGSEA was conducted to explore the role of immunoregulatory genes in SCZ. The results revealed a significantly elevated immune score in SCZ compared to controls

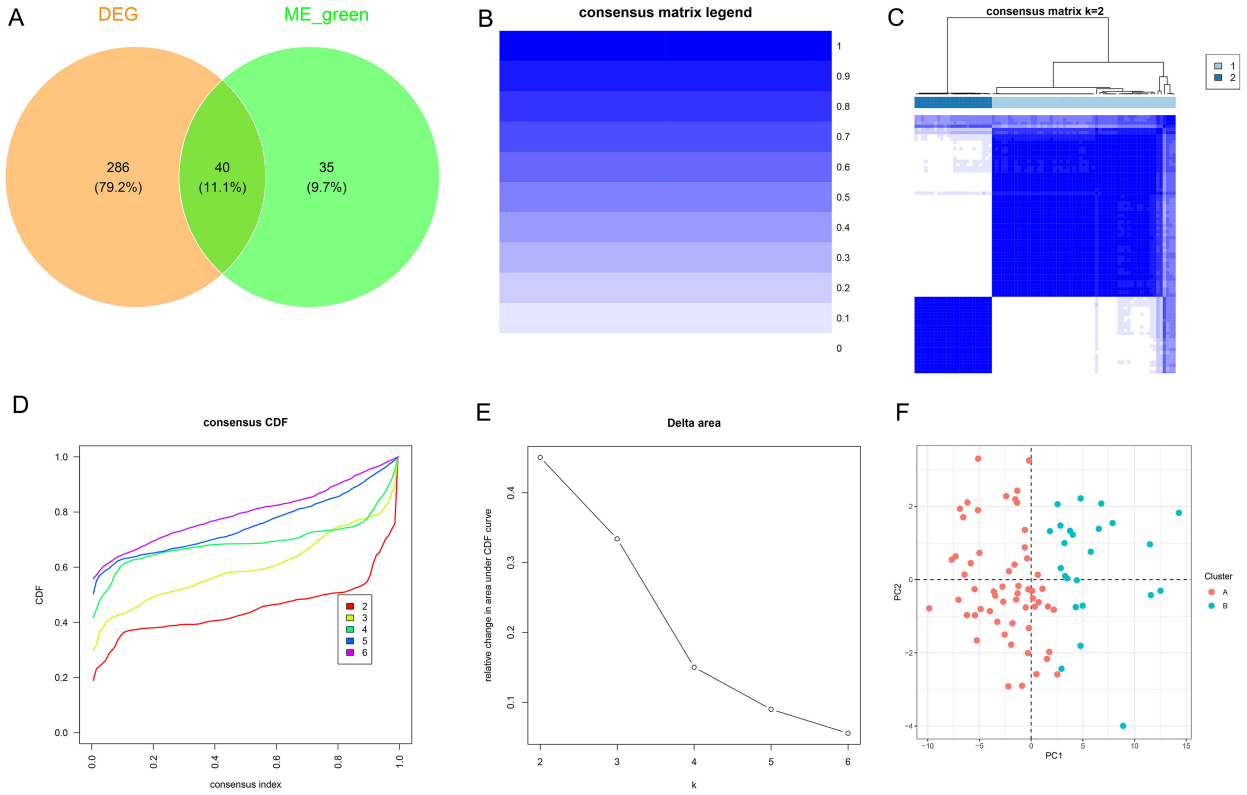

Fig. 3. Constructing immune-related gene clusters. (A) A Venn diagram was used to show the overlapping genes between the green module genes and DEGs regarding the above subgroups. (B) The color legend for the consistency unsupervised clustering. (C) The consistency matrix of all datasets with k = 2. (D) A consistent cumulative profile was used to show the cumulative profile function with various k values to obtain the best k. (E) A delta area map was utilized to show the change of delta area with different values of k. (F) PCA was leveraged to determine the differences between the two subtypes identified by the consistent unsupervised cluster analysis. The 81 SCZ samples were categorized into cluster A and B (immune-related differentially expressed genes (IDEGs)-related clusters) on the basis of the results of the clustering analysis.

($p = 0.03$), as depicted in Fig. 1C. Subsequently, the 81 SCZ samples were divided into high immunity and low immunity subgroups based on the median of the ssGSEA score. The corresponding DEGs were identified using adjusted p value below 0.01 and |log2FC| above 0.5 cut-offs, resulting in 326 DEGs as illustrated in the volcano plot (Fig. 1D). Among these, 232 genes were upregulated and 94 genes were downregulated in the high versus low group, as shown in Fig. 1D,E.

Identifying 75 Genes in the Green Modules Relevant to Immune Scores

WGCNA was selected to identify co-expressed gene modules in the three datasets and investigate the role of the immune response in the gene networks. The dendrogram and traits of all samples were illustrated (Fig. 2A). We selected the top 5000 variance genes and constructed a network using a soft-thresholding power of $\beta = 3$, resulting in a scale-free network with high connectivity (scale-free R2 = 0.86, slope = –1.64) (Fig. 2B,C). A total of 7 modules were identified after merging similar ones, and clustering dendrograms were presented (Fig. 2D,E). The green module emerged with the highest Module Significance (MS) (Fig. 2F). Additionally, the relationship between Module Membership (MM) and Gene Significance (GS) in this module was highly significant (cor = 0.79, $p = 7.8 \times 10^{-27}$), as shown in Fig. 2G. Therefore, the 75 genes in the green module were selected for further analysis.

Identifying 2 Immune-Related Differentially Expressed Genes (IDEGs)-Related Clusters

To explore the dissimilarities in immune infiltration patterns, we identified the co-expressed genes (co-genes) within the green module and DEGs of the aforementioned subgroups (Fig. 3A). A total of 40 co-genes were identified, and an unsupervised cluster analysis was performed

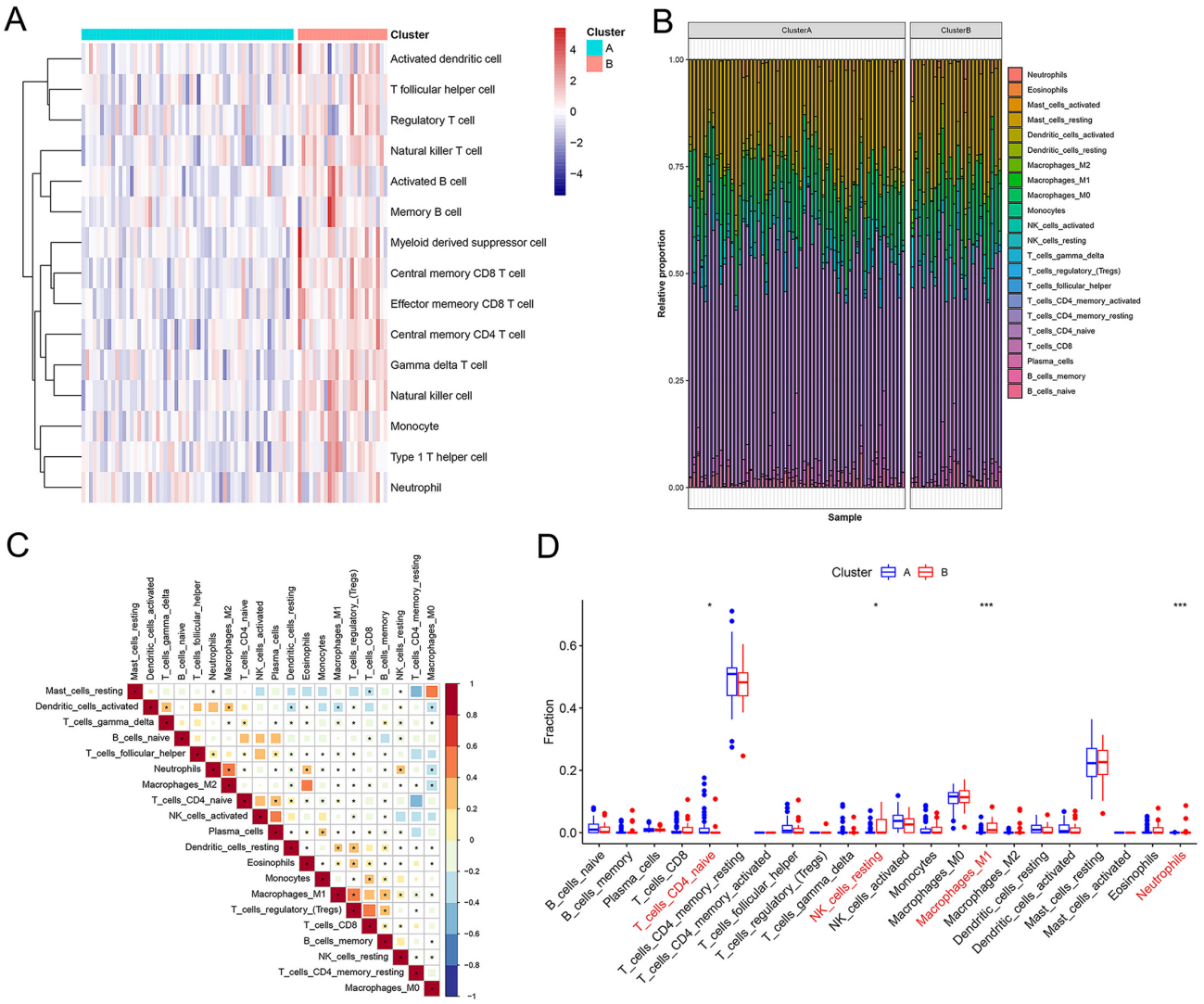

Fig. 4. Immune infiltration analysis of IDEG subtypes. (A) The single-sample gene set enrichment analysis (ssGSEA) was leveraged for comparing the immune infiltration status regarding the two immune-related clusters, and the heatmap indicated that immune infiltration appeared drastically severer in cluster B compared to in cluster A. Red indicates stimulated immune cells, with blue representing inhibited immune cells. (B) The CIBERSORT method was used for calculating the immune infiltration levels about the 22 immune cells from SCZ patients, and the landscape of immune cell infiltration proportions was shown. It showed mast cells, T cells, and macrophages were the most prevalent types of 22 forms of cells. (C) The 22 forms of immune cells' interaction was illustrated. (D) The Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT) algorithm was used to compare the infiltration proportion of the 22 immune cells from the two clusters related to immune system in SCZ patients. The percents of M1 macrophages, natural killer (NK) cells, and neutrophils turned out to be drastically increased in cluster B, while the T cells CD4 naive's proportions showed a remarkable increase in cluster A (*$p \leq 0.05$; ***$p \leq 0.001$).

using these genes. The findings show that the delta area of subtype aggregation reduced drastically and reached a plateau when k = 2 (Fig. 3B–E). PCA indicated that the two subtypes were distinct (Fig. 3F). Therefore, the 81 patient samples were divided into two clusters, cluster A and B, using consistent unsupervised cluster analyses, which were termed as immune-related differentially expressed genes (IDEGs)-related clusters.

The High Immune Cluster (Cluster B) Having a Higher Degree of Immune Infiltration

To investigate the immune infiltration patterns regarding the two immunoregulatory subtypes, ssGSEA was utilized. The resulting heatmap showed a significant increase in immune cell infiltration in cluster B compared to cluster A, indicating that cluster B was associated with

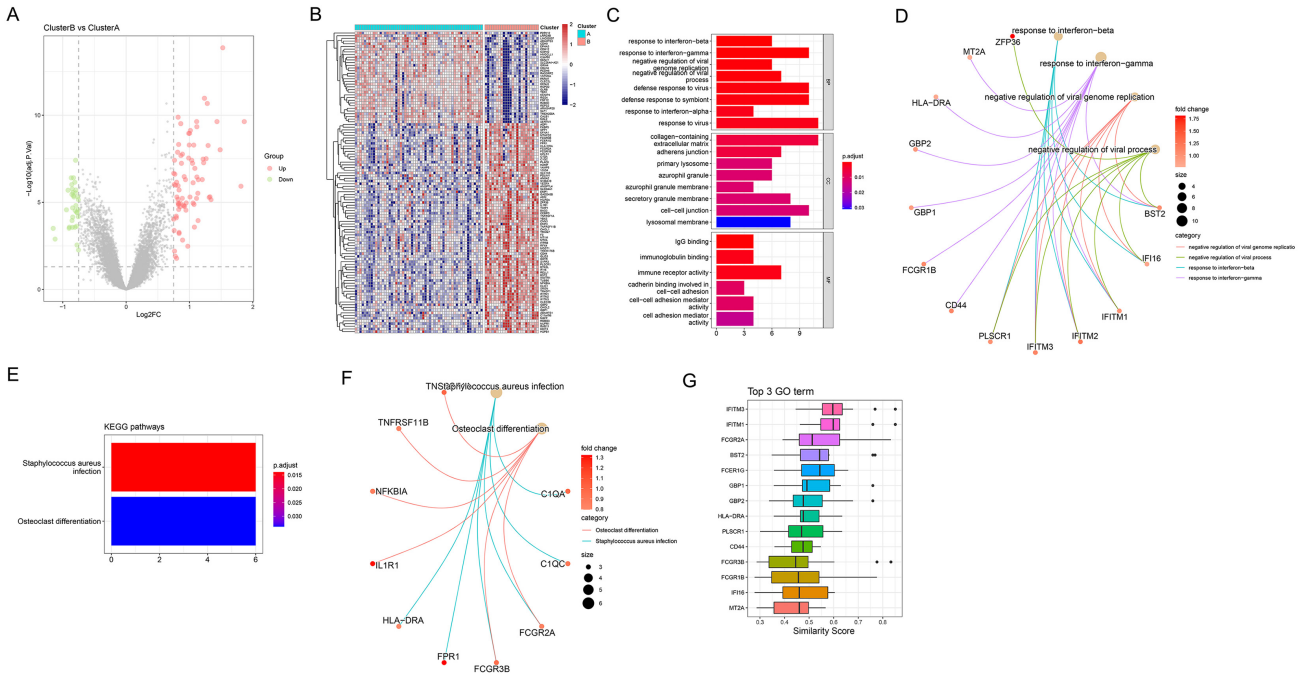

Fig. 5. The immune-related differentially expressed genes (IRDEGs) enrichment analysis. (A) The volcano plot displays the genes that are differentially expressed in cluster B relative to cluster A. Red color indicates the genes with upregulation, green for genes with downregulation, and gray for genes with no significant difference. Differentially expressed genes (DEGs) were defined as ones with an adjusted *p*-value below 0.05 and fold change less than 0.75. (B) The heatmap shows the DEGs in cluster B versus cluster A. (C) The bubble plot illustrates that the top 10 Gene Ontology (GO) terms are enriched in the IRDEGs. (D) The network profile displays the association between the genes and the top four bioprocesses. The nodes' colors represent the gene expressional contents, among which red nodes reflect genes with upregulation and blue nodes for genes with downregulation. The circles' sizes symbolize the association. (E) The bubble plot illustrates that the top 10 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways are enriched within the IRDEGs. (F) The network profile shows the association of the genes with the KEGG pathways. The nodes' colors represent the gene expressional contents, where red nodes signify genes with upregulation and blue nodes for genes with downregulation. The circles' sizes symbolize the association. (G) The functional similarities analysis identifies the core genes participating in the top three GO terms. The black dots in the image represent outliers.

a high immune signal (Fig. 4A). Additionally, the study further examined the immune infiltration regarding SCZ using the CIBERSORT method, which revealed that mast cells, T cells, and macrophages were the most prevalent immune cell types among the 22 immune cell forms (Fig. 4B). The relationship coefficients within SCZ-related immune cells were also calculated and found to be positively correlated (Fig. 4C). Furthermore, the proportions of M1 macrophages, resting natural killer (NK) cells, and neutrophils were drastically increased in SCZ patients in cluster B relative to cluster A, while the proportions of naive CD4+ T cells were remarkably elevated in SCZ patients in cluster A relative to cluster B (Fig. 4D). These findings suggest that clusters A and B were associated with distinct immune cell infiltration patterns, with the high immune cluster (cluster B) having a higher degree of immune infiltration.

Interferon-Beta, IgG Binding, and Response to Interferon-Gamma are Key Biological Processes Involved in the Different Immune Gene Subgroups

To better understand the biological characteristics of the two immune gene subgroups, 112 immune-related differentially expressed genes (IRDEGs) between the two clusters were identified (Fig. 5A,B). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed to determine the enriched bioprocesses, molecular functions, cellular components, and pathways related to IRDEGs (Fig. 5C–F, **Supplementary Table 1**). It was observed that IRDEGs had significant enrichments in bioprocesses related to response to interferon-beta and interferon-gamma, and molecular functions associated with IgG binding. Enriched GO terms were linked to inflammatory response, virus receptor effect, interferon response, and exogenous protein integration. Addition-

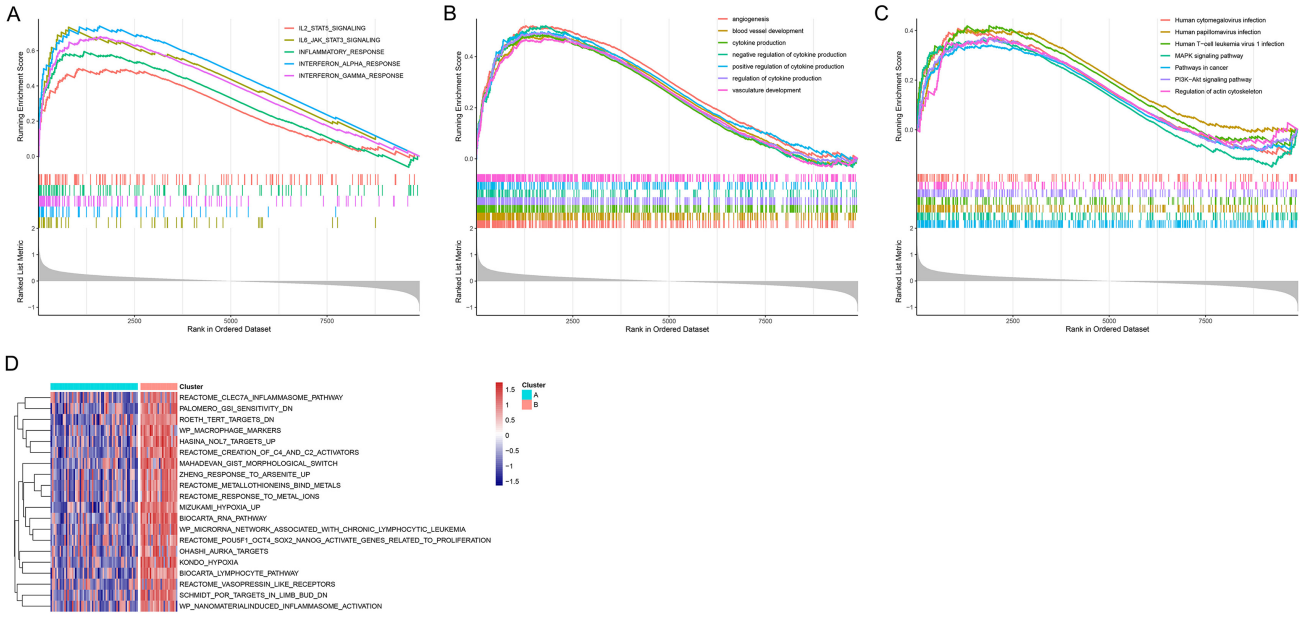

Fig. 6. Gene set enrichment analysis (GSEA) and Gene Set Variation Analysis (GSVA) enrichment analysis of the two immune-related clusters. (A) GSEA showing the enriched pathways in HLLMARK set of cluster B. (B) GSEA showing the enriched pathways in GO set of cluster B. (C) GSEA showing the top 7 enriched pathways in KEGG set of cluster B. Each line represents a particular term in a unique color. The up-regulated genes located near the left, while the down-regulated genes located near the right. The threshold was False Discovery Rate (FDR) $q < 0.05$. (D) GSVA showing the top 20 pathways enriched in the cluster B.

ally, IRDEGs showed enrichments within pathways such as Staphylococcus aureus infection and osteoclast differentiation. The latter was enriched by nine IRDEGs and was the most remarkable pathway. In order to identify core genes related to these processes, a functional analogy analysis was conducted using GO, revealing 11 hub-genes closely associated with response to interferon-beta, IgG binding, and response to interferon-gamma. These hub genes included *PLSCR1* (phospholipid scramblase 1), *FCGR1B* (Putative high affinity immunoglobulin gamma Fc receptor IB), *MT2A* (metallothionein 2A), *IFITM1* (interferon induced transmembrane protein 1), *GBP1* (guanylate binding protein 1), *BST2* (bone marrow stromal cell antigen 2), *IFITM3* (interferon induced transmembrane protein 3), *GBP2* (guanylate binding protein 2), *CD44* (CD44 molecule), *FCER1G* (Fc epsilon receptor Ig), *HLA-DRA* (major histocompatibility complex, class II, DR alpha), *FCGR2A* (Fc gamma receptor IIa), *IFI16* (interferon gamma inducible protein 16), and *FCGR3B* (Fc gamma receptor IIIb) (Fig. 5G, **Supplementary Table 2**). These results suggest that interferon-beta, IgG binding, and response to interferon-gamma are key biological processes involved in the different immune gene subgroups.

Analyzing GSEA and GSVA Enrichment of the Two Clusters Relevant to Immune System

GSEA and GSVA were conducted to analyze the enriched biological pathways for the two immune subtypes. The results of GSEA using the Hallmark gene sets showed enrichment of IL6-JAK-STAT3 signaling, interferon alpha response, interferon gamma response, inflammatory response, and IL2-STAT5 signaling in cluster B (Fig. 6A; **Supplementary Table 3**). Furthermore, GSEA using the GO gene sets revealed enrichment of angiogenesis, blood vessel development, cytokine production, positive regulation, and negative regulation of cytokine production, as well as vasculature growth in cluster B (Fig. 6B; **Supplementary Table 4**). Additionally, GSEA using the KEGG gene sets demonstrated enrichment of human papillomavirus infection, human T-cell leukemia virus 1 infection, MAPK signaling pathway, PI3K-AKT signaling pathway, regulation of actin cytoskeleton, and human cytomegalovirus infection in cluster B (Fig. 6C; **Supplementary Table 5**). These results indicate that cluster B is associated with immune-related pathways.

The GSVA outcomes revealed that immunoregulatory genes largely impact the following pathways: “blanco melo respiratory syncytial virus infection a594 cells up”, “reactome interleukin 10 signaling”, “reactome interleukin 4

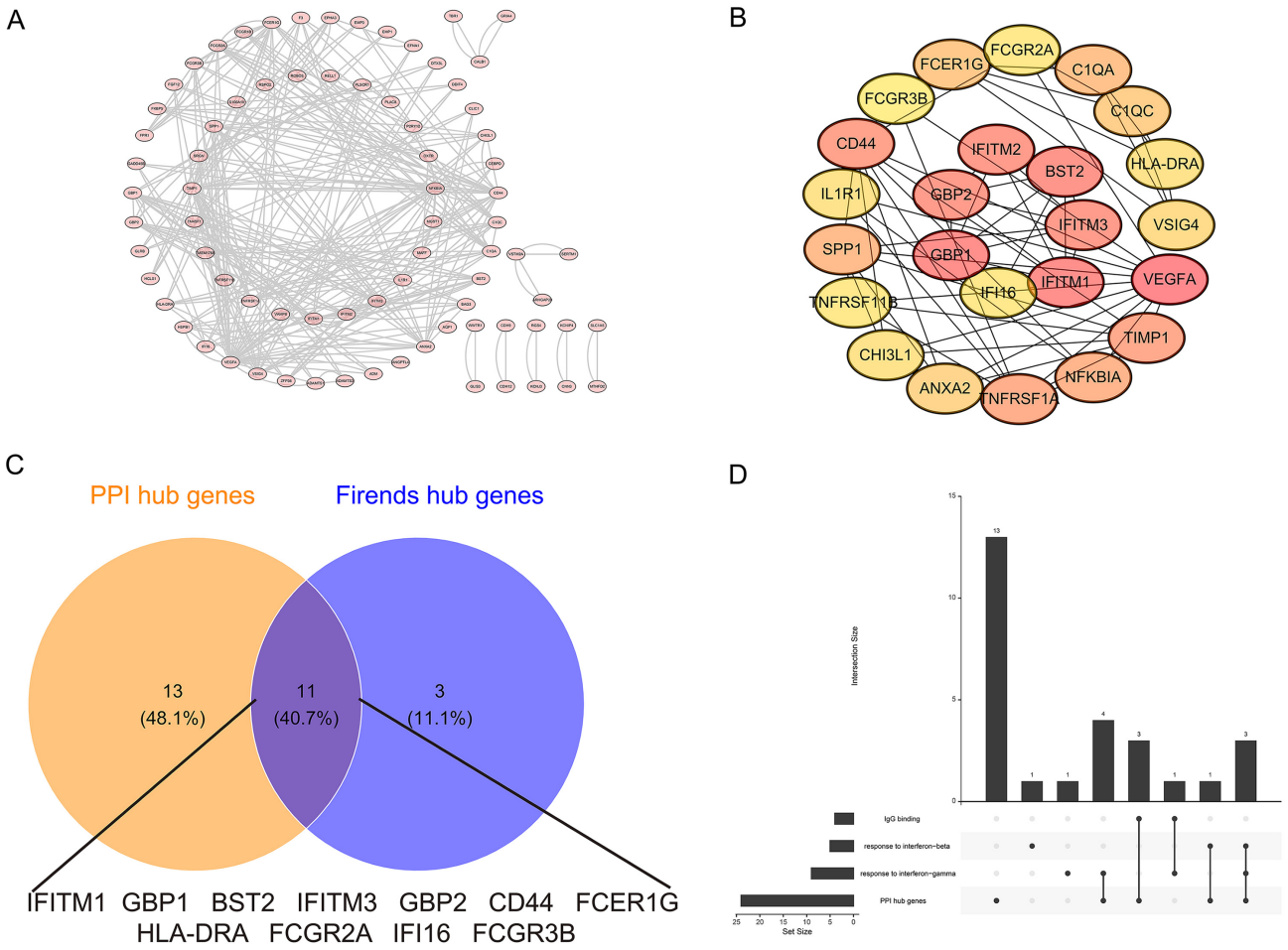

Fig. 7. Protein-protein interaction (PPI) network analyses of the IRDEGs. (A) IRDEG-encoded protein-protein interactions are depicted in a PPI network. (B) Highly-ranked PPI hub genes. (C) A Venn diagram depicting the overlap between PPI core genes and hub genes with analogic functions, including 11 genes. (D) Genes that are both PPI hubs and among the top 3 in GO keywords.

and interleukin 13 signaling", "blanco melo human parainfluenza virus 3 infection a594 cells up", "salvador martin pediatric tbd anti-TNF therapy non-responder post-treatment up", and "reactome creation of c4 and c2 activators" (Fig. 6D). These pathways showed activation within cluster B.

Identifying 11 Hub Genes in Schizophrenia

To identify hub genes in schizophrenia, a PPI network was constructed for the 112 differentially expressed genes associated with SCZ (Fig. 7A). CytoHubba, a Cytoscape plugin, was utilized to analyze the network and identify the top 24 genes as core genes (Fig. 7B), referred to as PPI hub genes to distinguish them from the previously identified core genes obtained through functional analogy analysis (hub genes with analogic function) in Fig. 6C. Subsequently, these two types of hub genes were combined using a Venn diagram to identify the 11 overlapping genes related to immune infiltration (Fig. 7C). These genes included *IFITM1*, *GBP1*, *BST2*, *IFITM3*, *GBP2*, *CD44*, *FCER1G*, *HLA-DRA*, *FCGR2A*, *IFI16*, and *FCGR3B*. The UpSet plot displays the co-genes of the PPI network as well as the top 3 GO terms (Fig. 7D).

Discussion

SCZ is a severe and intricate mental illness. Despite the expanded scale of SCZ research, its underlying etiology remains elusive. With the growing proportion of studies focusing on immune function and psychiatric disorders, inflammation and immunology have been implicated in the pathogenesis of SCZ. Research by Cullen AE *et al.* [38] demonstrated a higher prevalence of inflammatory and autoimmune diseases in individuals with schizophrenia, and both SCZ patients and their families are at an increased risk

of developing immune disorders [39]. Previous study has suggested that immunotherapy targeting antibodies may alleviate schizophrenic symptoms in some cases [40]. Additionally, investigations into immune system abnormalities in psychiatric diseases have shown that postpartum SCZ in some women may result from disruptions in immune response during pregnancy [41,42]. These findings support the involvement of immune responses in the development of SCZ, yet the potential mechanism remains unclear. Further studies are needed to expand our understanding of immune system signatures relevant to the pathogenesis of SCZ.

In this study, the impact of immune-related gene expression on SCZ was analyzed using ssGSEA. The median ssGSEA score was utilized to classify the 81 SCZ samples into high immunity and low immunity subgroups. Two hundred and thirty-two genes were increased, and ninety-four genes were decreased in the high immunity subgroup compared to the low immunity subgroup. Using the same unsupervised cluster analyses, the 81 patient samples were then divided into clusters A and B. Differences in the extent of genome sequence across clusters were observed. There were 112 genes whose expression levels varied significantly between the two categories of immune genes. *PLSCR1*, *FCGR1B*, *MT2A*, *IFITM1*, *GBP1*, *BST2*, *IFITM3*, *GBP2*, *CD44*, *FCER1G*, *HLA-DRA*, *FCGR2A*, *IFI16*, and *FCGR3B* are all tightly associated with the response to interferon-beta, IgG binding, and response to interferon-gamma, as indicated by GO and KEGG enrichment studies. Finally, the PPI network was used to focus on eight SCZ hub genes involved in immune infiltration. These genes include *IFITM1*, *GBP1*, *BST2*, *IFITM3*, *GBP2*, *CD44*, *FCER1G*, *HLA-DRA*, *FCGR2A*, *IFI16*, and *FCGR3B*.

Immune response defects have been implicated in the onset of SCZ. In this study, two hub genes, *IFITM1* and *IFITM3*, were discovered. Genes belonging to the interferon-inducible transmembrane (Ifitm/Fragilis) family produce similar small proteins primarily found in the plasma and endolysosomal membranes [43–45]. Previous studies have shown complex, field-specific alterations in the prefrontal cortex regarding the transcripts of the IFITM family and SERPINA3 [46–48]. These findings align closely with the disorders observed in the brains of SCZ patients.

Conclusions

In summary, this study identified 24 potential immune-related genes in SCZ through bioinformatics analysis. Among them, 11 hub genes, including *IFITM1*, *GBP1*, *BST2*, *IFITM3*, *GBP2*, *CD44*, *FCER1G*, *HLA-DRA*, *FCGR2A*, *IFI16*, and *FCGR3B*, were identified by constructing PPI networks. Future research is warranted to comprehensively understand the significance of immunological responses as potential clinical indicators or therapeutic targets, as well as the regulatory functions they play in SCZ.

Availability of Data and Materials

The datasets used or analysed during the current study are available from the corresponding authors on reasonable request.

Author Contributions

Concept—YWu, ZW, TW, WC, Design—YWu, ZW, ZS, YWa, GC, Supervision—WC, Resources—YWu, Materials—PL, Data Collection and/or Processing—YWu, ZW, WC, Analysis and/or Interpretation—YWu, ZW, HH, AA, Literature Search—ZW, CL, PL, Writing—All authors, Critical Review—All authors. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

Not applicable.

Funding

This work is supported by Natural Science Foundation of Heilongjiang Province under grant number (LH2020H130), Project of Education Department of Heilongjiang Province, China (2022-KYYWF-0783).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.62641/aep.v52i3.1648.

References

[1] Winship IR, Dursun SM, Baker GB, Balista PA, Kandratavicius L, Maia-de-Oliveira JP, *et al*. An Overview of Animal Models Related to Schizophrenia. Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie. 2019; 64: 5–17.

[2] Ward J, Le NQ, Suryakant S, Brody JA, Amouyel P, Boland A, *et al*. Polygenic risk of major depressive disorder as a risk factor for venous thromboembolism. Blood Advances. 2023; 7: 5341–5350.

[3] Bauer AE, Liu X, Byrne EM, Sullivan PF, Wray NR, Agerbo E, *et al*. Genetic risk scores for major psychiatric disorders and the risk of postpartum psychiatric disorders. Translational Psychiatry. 2019; 9: 288.

[4] Singh R, Stogios N, Smith E, Lee J, Maksyutynsk K, Au E, *et al*. Gut microbiome in schizophrenia and antipsychotic-induced metabolic alterations: a scoping review. Therapeutic Advances in Psychopharmacology. 2022; 12: 20451253221096525.

[5] Farsi Z, Nicolella A, Simmons SK, Aryal S, Shepard N, Brenner K, *et al*. Brain-region-specific changes in neurons and glia and dysregulation of dopamine signaling in Grin2a mutant mice. Neuron. 2023; 111: 3378–3396.e9.

[6] Guo LK, Su Y, Zhang YYN, Yu H, Lu Z, Li WQ, *et al*. Prediction of treatment response to antipsychotic drugs for precision medicine approach to schizophrenia: randomized trials and multiomics analysis. Military Medical Research. 2023; 10: 24.

[7] Fan B, Zhao JV. Genetic proxies for antihypertensive drugs and mental disorders: Mendelian randomization study in European and East Asian populations. BMC Medicine. 2024; 22: 6.

[8] Zheng P, Zeng B, Liu M, Chen J, Pan J, Han Y, *et al*. The gut microbiome from patients with schizophrenia modulates the glutamate-glutamine-GABA cycle and schizophrenia-relevant behaviors in mice. Science Advances. 2019; 5: eaau8317.

[9] Chen Y, Dai J, Tang L, Mikhailova T, Liang Q, Li M, *et al*. Neuroimmune transcriptome changes in patient brains of psychiatric and neurological disorders. Molecular Psychiatry. 2023; 28: 710–721.

[10] Liu Y, Qiao L, Zhang S, Wan G, Chen B, Zhou P, *et al*. Dual pH-responsive multifunctional nanoparticles for targeted treatment of breast cancer by combining immunotherapy and chemotherapy. Acta Biomaterialia. 2018; 66: 310–324.

[11] Guerra AD, Yeung OWH, Qi X, Kao WJ, Man K. The Anti-Tumor Effects of M1 Macrophage-Loaded Poly (ethylene glycol) and Gelatin-Based Hydrogels on Hepatocellular Carcinoma. Theranostics. 2017; 7: 3732–3744.

[12] Saha S, Chant D, McGrath J. A systematic review of mortality in schizophrenia: is the differential mortality gap worsening over time? Archives of General Psychiatry. 2007; 64: 1123–1131.

[13] Maes M, Plaimas K, Suratanee A, Noto C, Kanchanatawan B. First Episode Psychosis and Schizophrenia Are Systemic Neuro-Immune Disorders Triggered by a Biotic Stimulus in Individuals with Reduced Immune Regulation and Neuroprotection. Cells. 2021; 10: 2929.

[14] Roomruangwong C, Noto C, Kanchanatawan B, Anderson G, Kubera M, Carvalho AF, *et al*. The Role of Aberrations in the Immune-Inflammatory Response System (IRS) and the Compensatory Immune-Regulatory Reflex System (CIRS) in Different Phenotypes of Schizophrenia: the IRS-CIRS Theory of Schizophrenia. Molecular Neurobiology. 2020; 57: 778–797.

[15] Maes M, Vojdani A, Sirivichayakul S, Barbosa DS, Kanchanatawan B. Inflammatory and Oxidative Pathways Are New Drug Targets in Multiple Episode Schizophrenia and Leaky Gut, Klebsiella pneumoniae, and C1q Immune Complexes Are Additional Drug Targets in First Episode Schizophrenia. Molecular Neurobiology. 2021; 58: 3319–3334.

[16] Umeda-Yano S, Hashimoto R, Yamamori H, Weickert CS, Yasuda Y, Ohi K, *et al*. Expression analysis of the genes identified in GWAS of the postmortem brain tissues from patients with schizophrenia. Neuroscience Letters. 2014; 568: 12–16.

[17] Kadakia A, Catillon M, Fan Q, Williams GR, Marden JR, Anderson A, *et al*. The Economic Burden of Schizophrenia in the United States. The Journal of Clinical Psychiatry. 2022; 83: 22m14458.

[18] Na KS, Jung HY, Kim YK. The role of pro-inflammatory cytokines in the neuroinflammation and neurogenesis of schizophrenia. Progress in Neuro-psychopharmacology & Biological Psychiatry. 2014; 48: 277–286.

[19] Barnes MR, Huxley-Jones J, Maycox PR, Lennon M, Thornber A, Kelly F, *et al*. Transcription and pathway analysis of the superior temporal cortex and anterior prefrontal cortex in schizophrenia. Journal of Neuroscience Research. 2011; 89: 1218–1227.

[20] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, *et al*. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics. 2000; 25: 25–29.

[21] Zhang JJ, Shen Y, Chen XY, Jiang ML, Yuan FH, Xie SL, *et al*. Integrative network-based analysis on multiple Gene Expression Omnibus datasets identifies novel immune molecular markers implicated in non-alcoholic steatohepatitis. Frontiers in Endocrinology. 2023; 14: 1115890.

[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 Research. 2015; 43: e47.

[23] Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics (Oxford, England). 2012; 28: 882–883.

[24] Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013; 14: 7.

[25] Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, *et al*. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. Scientific Data. 2018; 5: 180015.

[26] The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Research. 2019; 47: D330–D338.

[27] Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and

genomes. Nucleic Acids Research. 2023; 51: D587–D592.

[28] Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics (Oxford, England). 2010; 26: 1572–1573.

[29] Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Systems. 2015; 1: 417–425.

[30] Ferreira MR, Santos GA, Biagi CA, Silva Junior WA, Zambuzzi WF. GSVA score reveals molecular signatures from transcriptomes for biomaterials comparison. Journal of Biomedical Materials Research. Part a. 2021; 109: 1004–1014.

[31] Tian Z, He W, Tang J, Liao X, Yang Q, Wu Y, *et al*. Identification of Important Modules and Biomarkers in Breast Cancer Based on WGCNA. OncoTargets and Therapy. 2020; 13: 6805–6817.

[32] Ma H, He Z, Chen J, Zhang X, Song P. Identifying of biomarkers associated with gastric cancer based on 11 topological analysis methods of CytoHubba. Scientific Reports. 2021; 11: 1331.

[33] Yu G, Li F, Qin Y, Bo X, Wu Y, Wang S. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. Bioinformatics (Oxford, England). 2010; 26: 976–978.

[34] Craven KE, Gökmen-Polar Y, Badve SS. CIBERSORT analysis of TCGA and METABRIC identifies subgroups with better outcomes in triple negative breast cancer. Scientific Reports. 2021; 11: 4691.

[35] Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, *et al*. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nature Biotechnology. 2019; 37: 773–782.

[36] Maycox PR, Kelly F, Taylor A, Bates S, Reid J, Logendra R, *et al*. Analysis of gene expression in two large schizophrenia cohorts identifies multiple changes associated with nerve terminal function. Molecular Psychiatry. 2009; 14: 1083–1094.

[37] Narayan S, Tang B, Head SR, Gilmartin TJ, Sutcliffe JG, Dean B, *et al*. Molecular profiles of schizophrenia in the CNS at different stages of illness. Brain Research. 2008; 1239: 235–248.

[38] Cullen AE, Holmes S, Pollak TA, Blackman G, Joyce DW, Kempton MJ, *et al*. Associations Between Non-neurological Autoimmune Disorders and Psychosis: A Meta-analysis. Biological Psychiatry. 2019; 85: 35–48.

[39] Cao Y, Ji S, Chen Y, Zhang X, Ding G, Tang F. Association between autoimmune diseases of the nervous system and schizophrenia: A systematic review and meta-analysis of cohort studies. Comprehensive Psychiatry. 2023; 122: 152370.

[40] Siskind D, Orr S, Sinha S, Yu O, Brijball B, Warren N, *et al*. Rates of treatment-resistant schizophrenia from first-episode cohorts: systematic review and meta-analysis. The British Journal of Psychiatry: the Journal of Mental Science. 2022; 220: 115–120.

[41] Pillinger T, McCutcheon RA, Vano L, Mizuno Y, Arumuham A, Hindley G, *et al*. Comparative effects of 18 antipsychotics on metabolic function in patients with schizophrenia, predictors of metabolic dysregulation, and association with psychopathology: a systematic review and network meta-analysis. The Lancet. Psychiatry. 2020; 7: 64–77.

[42] Dutsch-Wicherek MM, Lewandowska A, Zgliczynska M, Szubert S, Lew-Starowicz M. Psychiatric disorders and changes in immune response in labor and postpartum. Frontiers in Bioscience (Landmark Edition). 2020; 25: 1433–1461.

[43] Zhao X, Li J, Winkler CA, An P, Guo JT. IFITM Genes, Variants, and Their Roles in the Control and Pathogenesis of Viral Infections. Frontiers in Microbiology. 2019; 9: 3228.

[44] Liao Y, Goraya MU, Yuan X, Zhang B, Chiu SH, Chen JL. Functional Involvement of Interferon-Inducible Transmembrane Proteins in Antiviral Immunity. Frontiers in Microbiology. 2019; 10: 1097.

[45] Bailey CC, Zhong G, Huang IC, Farzan M. IFITM-Family Proteins: The Cell's First Line of Antiviral Defense. Annual Review of Virology. 2014; 1: 261–283.

[46] Norton ES, Da Mesquita S, Guerrero-Cazares H. SERPINA3 in glioblastoma and Alzheimer's disease. Aging. 2021; 13: 21812–21813.

[47] Klein S, Golani G, Lolicato F, Lahr C, Beyer D, Herrmann A, *et al*. IFITM3 blocks influenza virus entry by sorting lipids and stabilizing hemifusion. Cell Host & Microbe. 2023; 31: 616–633.e20.

[48] Jiménez-Munguía I, Beaven AH, Blank PS, Sodt AJ, Zimmerberg J. Interferon-induced transmembrane protein 3 (IFITM3) and its antiviral activity. Current Opinion in Structural Biology. 2022; 77: 102467.