

# Exploration Of Signature Genes and Their Correlation with Immune Cell Infiltration in Cirrhosis

Yuxin Li<sup>1</sup>, Wenjie Zhang<sup>1</sup>, Yuanyuan Zhao<sup>1</sup>, \*

## Abstract

Early diagnosis of cirrhosis is crucial for improving patient prognosis. This study aims to investigate signature genes and their correlation with immune cell infiltration in cirrhosis. We utilized a liver cirrhosis patient dataset obtained from the GEO database to identify differentially expressed genes (DEGs). Weighted gene co-expression network analysis (WGCNA), least absolute shrinkage and selection operator (LASSO), and random forest analysis were employed to identify signature genes, including RGS1, DEFB1, and ANOS2P. Subsequently, gene set enrichment analysis (GSEA) revealed that these signature genes are associated with positive correlations with allograft rejection and the focal adhesion pathway. Moreover, CIBERSORT analysis suggested potential involvement of these signature genes in immune cell infiltration within cirrhotic conditions. This study enhances our understanding of cirrhosis pathogenesis and may contribute to the development of early diagnostic tools and therapeutic strategies.

## BACKGROUND

Liver cirrhosis represents a pervasive health challenge globally, affecting both low-income, middle-income, and high-income countries alike, and is associated with substantial morbidity and mortality rates. Presently ranked as the 11th leading cause of death worldwide, cirrhosis is responsible for approximately 2 million fatalities annually.<sup>1</sup> Liver cirrhosis is a consequence of chronic inflammation causing progressive hepatic fibrosis.<sup>2</sup> Cirrhosis evolves through progressive hepatic fibrosis and ultimately advances to a decompensated stage marked by life-threatening complications, such as gastrointestinal bleeding, ascites, bacterial infections, hepatic encephalopathy (HE), and hepatorenal syndrome (HRS).<sup>3</sup> These complications often lead to elevated morbidity and mortality rates.<sup>4</sup> Cirrhosis stands out as a serious manifestation of chronic liver disease, imposing a significant global health burden due to its

increasing prevalence and profound impact on public health.<sup>5</sup> Characterized by extensive fibrosis, architectural distortion, and the transformation of normal liver structure into regenerative nodules, cirrhosis signifies the advanced scarring process in response to persistent liver injury.<sup>6</sup> The primary causes of cirrhosis encompass a range of factors, including chronic viral hepatitis (hepatitis B virus and hepatitis C virus), alcoholic liver disease, non-alcoholic fatty liver disease (NAFLD), autoimmune hepatitis, and various metabolic disorders.<sup>7</sup> Chronic HCV infection emerges as one of the most prevalent causes of chronic hepatitis, leading to severe liver conditions such as steatosis, cirrhosis, and hepatocellular carcinoma (HCC).<sup>8,9</sup> In many cases, hepatitis C presents with no overt symptoms or signs, yet the risk of chronicity post-infection remains relatively high.<sup>10</sup> Without standardized antiviral treatment, 15% - 30% of chronic hepatitis C patients may progress to cirrhosis within 20 years

<sup>1</sup> Institute of Biomedical Engineering, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin, 300192, China.

**Address for correspondence to:** Yuanyuan Zhao, Ph.D, Institute of Biomedical Engineering, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin, 300192, China.

(zhaoyy95@foxmail.com). Mobile no: +86 188 1156 3919

**Financial Disclosures:** This work was funded by China Postdoctoral Science Foundation (2022M720503) and Fundamental Research Funds for the Central Universities (3332023071).

0024-7758 © Journal of Reproductive Medicine®, Inc.

*The Journal of Reproductive Medicine*®

after infection.<sup>11</sup> Given the asymptomatic nature of hepatitis C in many cases, patients often seek medical attention only when cirrhosis symptoms are apparent. Some may even reach the decompensated stage of cirrhosis, missing the optimal treatment window and severely impacting their quality of life. Consequently, the identification of new therapeutic targets for liver cirrhosis resulting from hepatitis C assumes a crucial role in enhancing the quality of life for patients grappling with hepatitis C-induced cirrhosis. The intricate interplay between the immune system and HCV in the development of cirrhosis or liver cancer constitutes a complex and continually evolving field of study.<sup>12</sup>

The initial interaction between HCV and the immune system elicits both innate and adaptive immune responses.<sup>13</sup> However, HCV has developed sophisticated strategies to elude immune detection, resulting in chronic infection. Viral persistence emerges as a crucial factor in the transition from chronic hepatitis C to cirrhosis. The ongoing presence of HCV in the liver establishes a pro-inflammatory microenvironment, creating a conducive environment for the emergence of liver cancer.

Persistent HCV infection induces chronic inflammation, characterized by the infiltration of immune cells into the liver parenchyma.<sup>14</sup> This prolonged inflammatory response contributes to the activation of hepatic stellate cells and the deposition of extracellular matrix, ultimately leading to fibrosis—a precursor to cirrhosis and HCC.<sup>15</sup> Inflammatory mediators, including cytokines and chemokines, play a pivotal role in orchestrating the immune response within the hepatic microenvironment.<sup>16, 17</sup>

In our study, we applied multiple bioinformatic methodologies to identify signature genes implicated in liver cirrhosis induced by hepatitis C. These genes exhibited notable diagnostic efficacy. Moreover, we conducted a comprehensive assessment of the enrichment signaling pathways associated with these genes and their roles in immune cell infiltration. Our findings aim to offer fresh insights to clinicians for improved diagnosis and treatment strategies in the context of liver cirrhosis caused by hepatitis C.

## METHOD AND MATERIAL

### Data sources

For the current study, a dataset, namely GSE14323 (Platforms: GPL96 [HG- U133A] Affymetrix Human Genome U133A Array, GPL571 [[HG- U133A\_2] Affymetrix Human Genome U133A 2.0 Array), have been downloaded from Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>). GSE14323 contained 60 non-tumor patients, including 41 patients with cirrhosis and 19 normal controls.

### Identification of DEGs

Using R software's limma package,<sup>18</sup> differentially expressed genes (DEGs) between septic shock cohort and control cohort were analyzed, with the following criterion: adjust p value <0.05 and  $|\log \text{fold change (FC)}| > 1$ . The volcano plot was generated to show these DEGs, while the top100 DEGs were displayed by the heatmap.

### Functional and pathway enrichment analyses

Functional enrichment analyses of differentially expressed genes (DEGs) were performed utilizing the clusterProfiler package in R, focusing on Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) annotations.<sup>19</sup> In the context of the GO analysis, three distinct categories, namely biological process (BP), cellular component (CC), and molecular function (MF), were discerned to comprehensively investigate the biological relevance of these DEGs. Additionally, the exploration of potential signaling pathways was conducted through KEGG analysis.

### Weighted gene co-expression network analysis

Utilizing the scale-free topology criterion, we constructed a co-expression network in the GSE14323 cohort through weighted gene co-expression network analysis (WGCNA).<sup>20</sup> The determination of the soft threshold power and adjacencies was carried out using the pick Soft Threshold function within the WGCNA package.

Subsequently, the adjacency matrix underwent conversion into a topological overlap matrix (TOM), with the corresponding dissimilarity computed for hierarchical clustering analysis. The dynamic tree cutting method, employing a minimum module size of 50, was applied to discern co-expressed gene modules. To establish a link between the gene modules and patients with cirrhosis, we measured the connection using gene significance (GS) values and module membership (MM) values, thereby identifying the key modules that exhibit substantial relevance in the context of cirrhosis.

### **Signature gene identification**

Candidate hub genes were discerned through the intersection of Differentially Expressed Genes (DEGs) and key module genes. Subsequently, two machine learning algorithms, namely the Least Absolute Shrinkage and Selection Operator (LASSO) and Random Forest, were employed for the screening of hub genes. LASSO analysis was conducted using the *glmnet* package, incorporating penalty parameters for 10-fold cross-validation—a method recognized for its superiority in evaluating high-dimensional data.<sup>21</sup> Additionally, we utilized the R package "randomforest" to classify the DEGs and identify hub genes. The Random Forest model determined the optimal number of variables by calculating the average error rate of candidate hub genes.<sup>22</sup> Subsequently, we computed the error rate for each range from one to 500 trees, determining the optimal number of trees based on the lowest error rate. With the identified parameters, a Random Forest tree model was constructed. Finally, the feature importance scores of each candidate hub gene were determined, and the top 50 genes with the highest importance values were selected. The intersection of genes obtained from these two machine learning algorithms constituted the signature genes associated with cirrhosis patients. The diagnostic efficiency of these signature genes was assessed using the area under the curve (AUC) of the Receiver Operating Characteristic curve (ROCs). An AUC greater than 0.7 indicated a favorable diagnostic performance.

### **Gene set enrichment analysis**

To elucidate the association between the signature genes and signaling pathways, we stratified the cirrhosis cohort based on the median expression values of hub genes. Subsequently, gene set enrichment analysis (GSEA) was conducted on distinct subgroups, with a significance threshold set at adjusted  $p < 0.05$ .<sup>23</sup>

### **Immune cell infiltration**

The CIBERSORT method, employing the principles of linear support vector regression to deconvolute the expression matrix of 22 human immune cell subtypes, was applied to investigate variations in immune cell composition between children with cirrhosis and their healthy counterparts.<sup>24</sup> Subsequently, we identified immune cells exhibiting significant differences in infiltration between cirrhotic and normal individuals. We then conducted an analysis of the correlation between these immune cells and the signature genes utilizing the Spearman method.

### **Statistical analysis**

All statistical analyses in the present study were implemented using R software (version 4.1.3). Unless otherwise stated,  $p < 0.05$  was deemed as statistically significant, and all  $p$  values were two-tailed. The flow chart of this research was shown in Figure 1.

## **RESULTS**

### **Identification of DEGs between cirrhosis and control**

Differentially expressed genes (DEGs) between individuals diagnosed with cirrhosis and their normal counterparts underwent analysis using the "limma" package. A meticulous examination identified a total of 880 DEGs, with 637 genes demonstrating up-regulation and 159 genes exhibiting down-regulation in the cirrhosis group compared to the normal control cohort (Figure 2A). The visual representation of the top 100 DEGs, as depicted in the heatmap, delineates the

distinctive gene expression patterns between patients with cirrhosis and the healthy control group (Figure 2B).

Subsequent to the identification of DEGs, functional enrichment analyses were conducted to gain comprehensive insights into the underlying biological processes (BP), cellular components (CC), and pathways associated with these genes. The BP analysis uncovered significant involvement in positive regulation of lymphocyte activation, positive regulation of leukocyte activation,

and lymphocyte-mediated immunity.

In the CC analysis, predominant terms included MHC protein complex, MHC class II protein complex, and chemokine activity (Figure 2C). Furthermore, the KEGG analysis revealed the top three enriched pathways as Rheumatoid arthritis, Type I diabetes mellitus, and Antigen processing and presentation, shedding light on the potential molecular mechanisms underlying cirrhosis (Figure 2D).

Figure 1. The flow chart of this research.

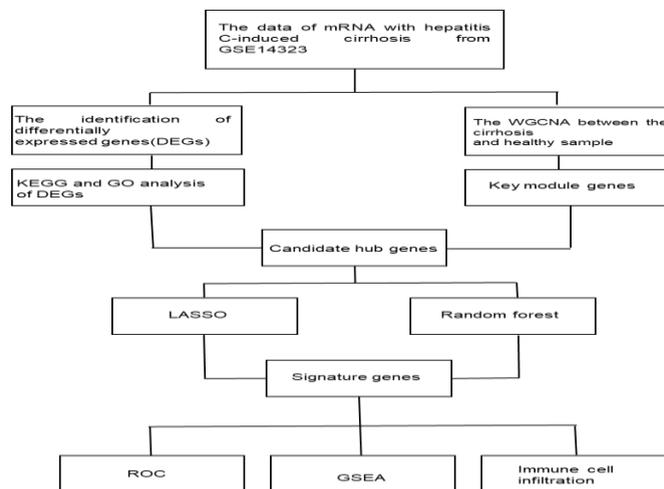
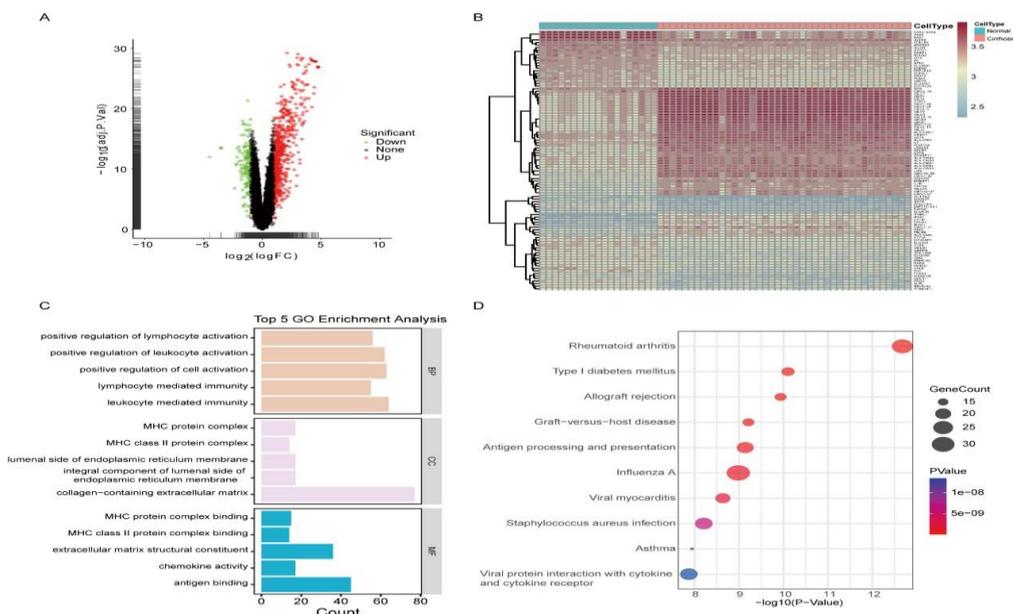


Figure 2. Identification of the DEGs. (A) Volcano showed expression of DEGs between the patients with cirrhosis and healthy cohort. (B) The heatmap showed the top 100 DEGs. (A) The top 5 functional enrichment in BP, CC, and MF analysis, respectively. (B) The KEGG analysis of DEGs.



### Construction of the weighted gene co-expression network

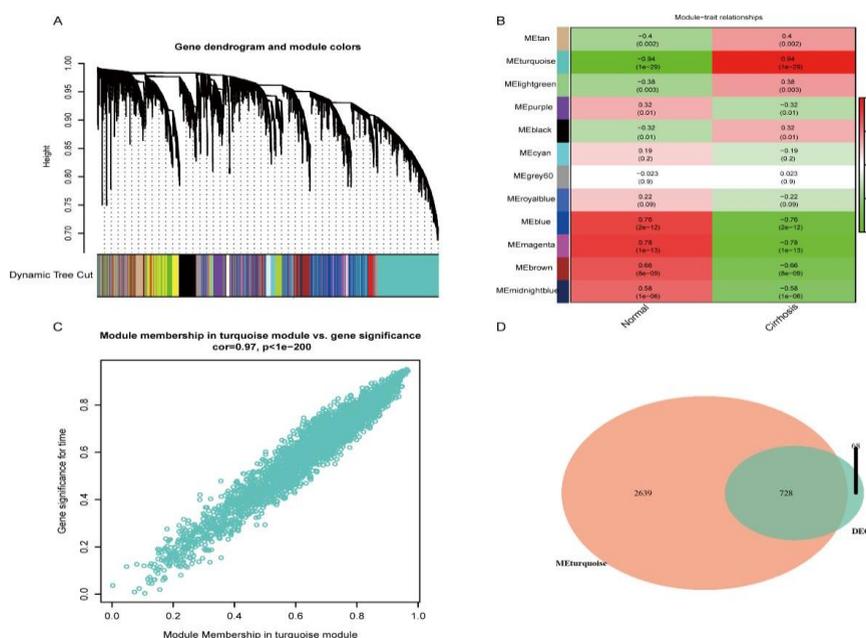
The establishment of a weighted gene co-expression network involved analyzing patients with cirrhosis and healthy subjects using the WGCNA package in R software, resulting in the creation of a scale-free co-expression network. The cluster dendrogram illustrating this network is presented in Figure 3A. Subsequently, the data were systematically clustered into 12 distinct modules (Figure 3B).

To assess the correlation with patients with cirrhosis, the correlation between each module and the clinical condition was calculated. Notably, the

MEturquoise module emerged as significantly associated with patients with cirrhosis ( $cor=0.94$ ,  $p<0.0001$ ). A robust positive correlation between Module Membership (MM) and Gene Signature (GS) is depicted in Figure 3C, further highlighting the module's relevance. The MEturquoise module, encompassing 3367 genes, was identified as a pivotal module closely linked to patients with cirrhosis.

The intersection between the differentially expressed genes (DEGs) and the genes within the MEturquoise module is visually represented in Figure 3D, providing insight into the shared genetic components between the two datasets.

**Figure 3.** The WGCNA analysis of GSE14323 and identification of candidate hub genes. (A) The cluster dendrogram of WGCNA. (B) The clustered modules of WGCNA. (C). A scatterplot of gene significance (GS) for weight vs module membership in MEturquoise module. (D) The veen plot showed the interaction between DEGs and genes in MEturquoise module.



### Selection of signature genes via LASSO and random forest algorithms

Two machine algorithms were applied to screen out signature genes from candidate key genes in patients with cirrhosis. For the LASSO analysis selected 17 signature genes (Figures 4A, B), while in the random forest analysis, 17 signature genes were determined with relative importance more than 0.18 (Figures 4C, D).

These screened out signature genes were

displayed in Table S1. Three signature genes were finally determined via the interaction of these two algorithms, containing RGS1, DEFB1, ANOS2P (Figure 4E).

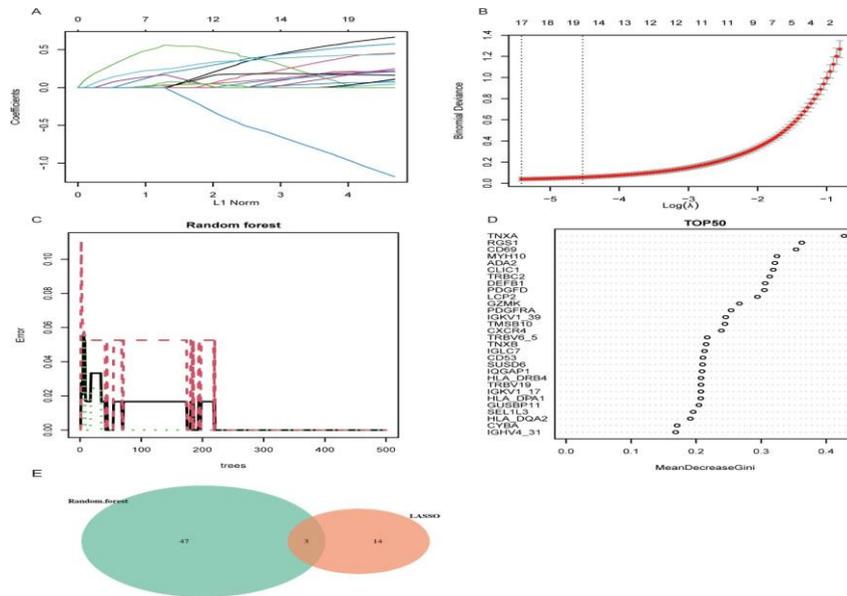
### Diagnostic efficacy of signature genes in predicting cirrhosis and control

The screened signature genes were highly expressed in patients with cirrhosis than those in healthy samples, suggesting that these genes may

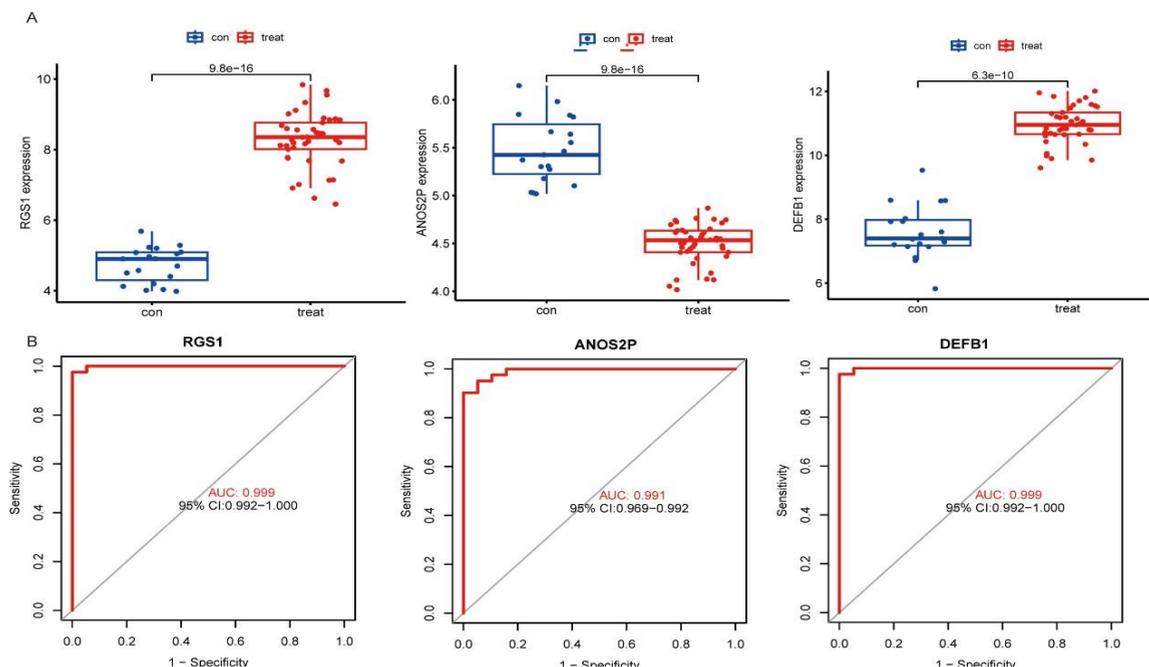
play a potential role in liver cirrhosis (Figures 5A). Furthermore, the area under curve (AUC) of the receiver operating characteristic curve (ROC) of

these signature genes were 0.999 of RGS1, 0.991 of ANOS2P, 0.999 of DEFB1 respectively (Figures 5B).

**Figure 4.** The machine algorithms for signature genes. (A) LASSO plot showed the variations in the size of coefficients for parameters shrank as the value of k penalty increased. (B) Penalty plot of the LASSO model with error bars denoting standard errors. (C) The error rate confidence intervals for random forest model. (D) The relative importance of genes is more than 0.18 in random forest model, (E) The interaction of the LASSO and random forest algorithms.



**Figure 5.** The performance of the signature genes in GSE14323. (A) The expression of signature genes between the patients with cirrhosis and healthy cohort. (B) ROC showed the diagnostic performance of the signature genes.

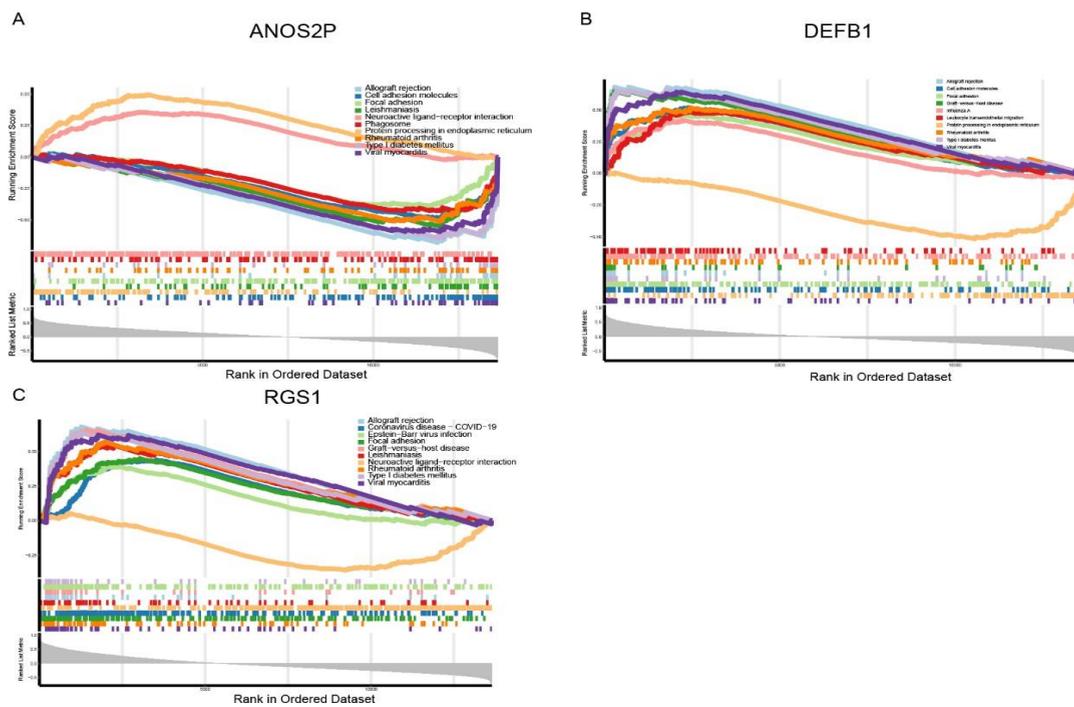


## GSEA analysis

We assessed signaling pathways associated with signature genes via GSEA analysis. The top 10 signaling pathways were displayed in Figure 6. The results showed that ANOS2P was significantly correlated with allograft rejection, cell adhesion molecules, focal adhesion, leishmaniasis, neuroactive ligand-receptor interaction, phagosome, protein processing in endoplasmic reticulum, rheumatoid arthritis, type 1 diabetes mellitus, viral myocarditis (Figure 6A). The expression of DEFB1 significantly correlated with allograft rejection, cell adhesion molecules, focal adhesion, graft-versus-host disease, influenza A,

leukocyte transendothelial migration, protein processing in endoplasmic reticulum, rheumatoid arthritis, type 1 diabetes mellitus, viral myocarditis (Figure 6B). The expression of RGS1 significantly correlated with allograft rejection, coronavirus disease -COVID-19, Epstein-barr virus infection, focal adhesion, graft-versus-host disease, leishmaniasis, neuroactive ligand-receptor interaction, rheumatoid arthritis, type 1 diabetes mellitus, viral myocarditis (Figure 6C). Taken together, these genes all positively correlated allograft rejection and focal adhesion pathway as well as type 1 diabetes mellitus and viral myocarditis pathway.

**Figure 6.** The GSEA of the signature genes in cirrhosis. (A) The GSEA of ANOS2P in cirrhosis. (B) The GSEA of DEFB1 in cirrhosis. (C) The GSEA of RGS1 in cirrhosis.



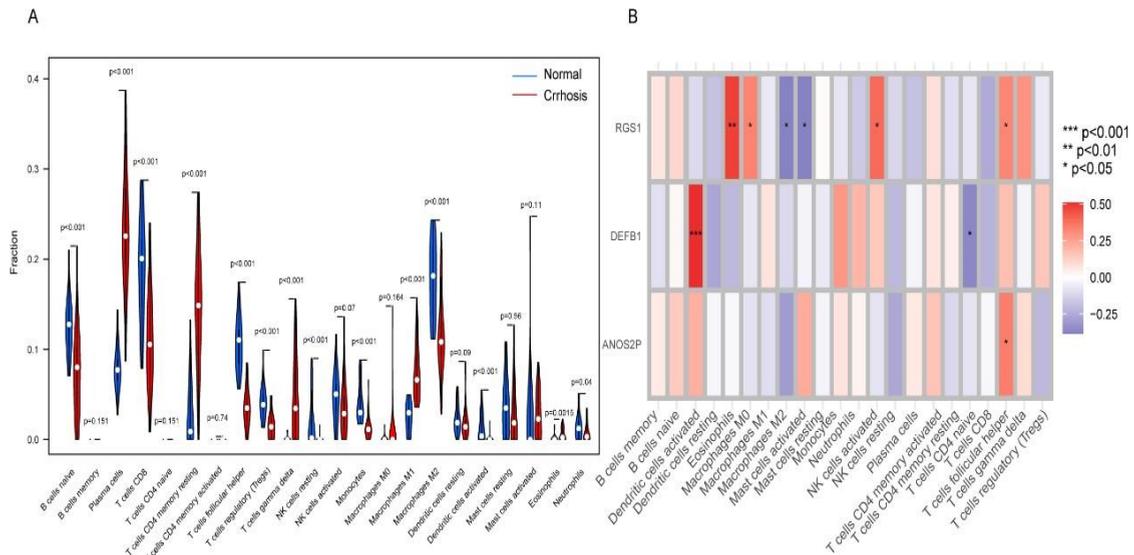
## Immune cell infiltration

Immunological features were evaluated according to immune cell infiltration. Compared with normal samples, patients with cirrhosis have higher plasma cells, CD4 memory resting T cells, gamma delta T cells, macrophages M1, eosinophils infiltration and lower naive B cells, CD8 T cells, follicular helper T cells, gamma delta T cells, Tregs cell infiltration (Figure 7A). The hub gene RGS1 were positively

correlated with the infiltration of Eosinophils and Macrophages M0, activated NK cells, follicular helper T cells, and negatively correlated with the infiltration of Macrophages M2 and activated mast cells. DEFB1 was positively correlated with the infiltration of activated dendritic cells, negatively correlated with the infiltration of naive CD4 T cells. ANOS2P was positively correlated with the infiltration of follicular helper T cells (Figure 7B).

**Figure 7.** The immune cell infiltration association with signature genes. (A) The immune cell infiltration between the cirrhosis and healthy cohort. (B) The association between signature genes and significantly different immune cell infiltration. "ns" means  $P \geq 0.01$ .

\* $P < 0.01$ , \*\* $P < 0.001$ , and \*\*\* $P < 0.0001$ .



## DISCUSSION

Cirrhosis is recognized as a systemic ailment, impacting various organs and systems, including the immune system.<sup>25</sup> Timely diagnosis and intervention are imperative for enhancing the prognosis of individuals afflicted with liver cirrhosis. Emerging evidence underscores the significant role of immune cell infiltration in cirrhosis.<sup>26</sup> Intriguingly, recent studies propose that impeding extracellular matrix accumulation to inhibit inflammatory cytokines may present a promising therapeutic avenue for cirrhosis.<sup>14, 27</sup> In our investigation, a total of 880 Differentially Expressed Genes (DEGs) were identified between cirrhotic and normal cohorts. The mechanistic insights into these DEGs were elucidated through enrichment function analysis. GO enrichment exploration revealed marked associations of DEGs with positive regulation of lymphocyte activation, positive regulation of leukocyte activation, and lymphocyte-mediated immunity. Furthermore, we employed WGCNA to identify key modules. Signature genes linked to cirrhosis, namely RGS1, ANOS2P, and DEFB1, were identified through LASSO and random forest analyses. Subsequent validation in external datasets confirmed the significance of these signature genes. Gene Set

Enrichment Analysis (GSEA) unraveled signaling pathways correlated with hub genes, and CIBERSORT algorithm quantified immune cell infiltration and its correlation with signature genes in cirrhosis and healthy groups. RGS1, or Regulator of G-Protein Signalling-1, accelerates Gai GTPase activity, downregulating the response to sustained chemokine activation.<sup>28</sup> Genome-wide association studies have implicated RGS1 in polymorphisms associated with the risk of chronic inflammatory diseases such as celiac disease, multiple sclerosis, and type I diabetes.<sup>29, 30</sup> Recent findings suggest RGS1 as a novel marker and promoting factor for CD8 T-cell exhaustion.<sup>31</sup> While RGS1 has been associated with lymphocyte homeostasis control,<sup>32, 33</sup> its role in liver cirrhosis remains unexplored, prompting our hypothesis of its involvement in the immune response associated with cirrhosis. ANOS2P, identified as a pseudogene, warrants further investigation to delineate its specific functions and impacts on HCV infection. Pseudogenes, resembling genes but lacking coding capacity, have been implicated in diverse facets of HCV infection, including immune responses, viral replication, and host cell signaling, as suggested by recent research.<sup>34, 35</sup> However, further in-depth studies are required to

elucidate the specific functions and influences of pseudogenes in the process of HCV infection.

DEFB1,<sup>36</sup> encoding the antimicrobial peptide  $\beta$ -defensin 1, crucially contributes to the immune system's defense against microbial infections. Dysregulation of DEFB1 gene expression has been linked to several cancers.<sup>37</sup> Studies propose that DEFB1 may positively regulate the abundance of tumor-infiltrating CD4 T cells, thereby improving the prognosis of oral squamous cell carcinoma.<sup>34</sup> Consequently, further investigation is warranted to ascertain whether DEFB1 participates in cirrhosis by activating CD4 T cells.

In summary, this study identified three signature genes, RGS1, ANOS2P, and DEFB1, with significant potential for early cirrhosis diagnosis. Additionally, the exploration of immune cell infiltration in cirrhosis and its correlation with signature genes provides a novel perspective on the role of immunity in cirrhosis.

## DECLARATIONS

### Ethical Approval

In this study, no human or animal subjects were involved, and ethical approval was not required as the data were sourced from the GEO database (GSE14323).

### 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.

### Limitation

Several limitations should be highlighted in our study. First, our findings are produced by a microarray and immune-related analysis based on public databases and involved small samples. Second, how these cirrhosis-related signature gene contribute to cirrhosis remains unknown. Therefore, further experiments are needed to verify the biological function of these genes.

## Author contributions

Y.Z. and Y.X. designed the study and drafted the manuscript. Z.W., L.M., Y.M. analyzed data. Q.W. revised the manuscript. All authors contributed to the article and approved the submitted version.

## Acknowledgments

We thank the GEO database for the open access to the data.

## Funding

This work was funded by China Postdoctoral Science Foundation (2022M720503) and Fundamental Research Funds for the Central Universities (3332023071).

## 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.

## REFERENCES

1. Asrani SK, Devarbhavi H, Eaton J, et al. Burden of liver diseases in the world. *J Hepatol* 2019;70:151-171.
2. Ginès P, Krag A, Abraldes JG, et al. Liver cirrhosis. *Lancet* 2021;398:1359-1376.
3. Bernardi M, Moreau R, Angeli P, et al. Mechanisms of decompensation and organ failure in cirrhosis: From peripheral arterial vasodilation to systemic inflammation hypothesis. *J Hepatol* 2015;63:1272-84.
4. The global, regional, and national burden of cirrhosis by cause in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* 2020;5:245-266.
5. Moon AM, Singal AG, and Tapper EB. Contemporary Epidemiology of Chronic Liver Disease and Cirrhosis. *Clin Gastroenterol Hepatol* 2020;18:2650-2666.
6. Nusrat S, Khan MS, Fazili J, et al. Cirrhosis and its complications: evidence-based treatment. *World J Gastroenterol* 2014;20:5442-60.
7. Huang DQ, Terrault NA, Tacke F et al. Global epidemiology of cirrhosis - aetiology, trends and predictions. *Nat Rev Gastroenterol Hepatol* 2023;20:388-398.
8. Chen J, Zhao Y, Zhang C et al. Persistent hepatitis C virus infections and hepatopathological manifestations in immune-competent humanized mice. *Cell Res* 2014;24:1050-66.
9. Chisari FV. Unscrambling hepatitis C virus-host interactions. *Nature* 2005;436:930-2.

10. Campollo O, Amaya G, and McCormick PA. Milestones in the discovery of hepatitis C. *World J Gastroenterol* 2022;28:5395-5402.
11. Khullar V and Firpi RJ. Hepatitis C cirrhosis: New perspectives for diagnosis and treatment. *World J Hepatol* 2015;7:1843-55.
12. Gremion C and Cerny A. Hepatitis C virus and the immune system: a concise review. *Rev Med Virol* 2005;15:235-68.
13. Karamichali E, Foka P, Papadopoulou, et al. Hepatitis Viruses Control Host Immune Responses by Modifying the Exosomal Biogenesis Pathway and Cargo. *Int J Mol Sci* 2022;23:10862.
14. Kisseleva T and Brenner D. Molecular and cellular mechanisms of liver fibrosis and its regression. *Nat Rev Gastroenterol Hepatol* 2021;18:151-166.
15. Hammerich L and Tacke F. Hepatic inflammatory responses in liver fibrosis. *Nat Rev Gastroenterol Hepatol* 2023;20:633-646.
16. Abdulkhaleq LA, Assi MA, Abdullah R, et al. The crucial roles of inflammatory mediators in inflammation: A review. *Vet World* 2018;11:627-635.
17. Lim HK, Jeffrey GP, Ramm GA, et al. Pathogenesis of Viral Hepatitis-Induced Chronic Liver Disease: Role of Extracellular Vesicles. *Front Cell Infect Microbiol* 2020;10:587628.
18. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47.
19. Yu G, Wang L-G, Han Y et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic* 2012;16:284-7.
20. Langfelder P and Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008;9:559.
21. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med* 1997;16:385-95.
22. Izmirlan G. Application of the random forest classification algorithm to a SELDI-TOF proteomics study in the setting of a cancer prevention trial. *Ann N Y Acad Sci* 2004;1020:154-74.
23. Subramanian A, Tamayo P, Mootha VK 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.
24. Newman AM, Liu CL, Green MR et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 2015;12:453-7.
25. Albillos A, Martin-Mateos R, Merwe SVd et al. Cirrhosis-associated immune dysfunction. *Nat Rev Gastroenterol Hepatol* 2022;19:112-134.
26. Kisseleva T, Uchinami H, Feirt N et al. Bone marrow-derived fibrocytes participate in pathogenesis of liver fibrosis. *J Hepatol* 2006;45:429-38.
27. Parola M. and Pinzani M. Liver fibrosis: Pathophysiology, pathogenetic targets and clinical issues. *Mol Aspects Med* 2019;65:37-55.
28. Han SB, Moratz C, Huang N-N et al. Rgs1 and Gnai2 regulate the entrance of B lymphocytes into lymph nodes and B cell motility within lymph node follicles. *Immunity* 2005;22:343-54.
29. Smyth DJ, Plagnol V, Walker NM et al. Shared and distinct genetic variants in type 1 diabetes and celiac disease. *N Engl J Med* 2008;359:2767-77.
30. Johnson BA, Wang J, Taylor EM, et al. Multiple sclerosis susceptibility alleles in African Americans. *Genes Immun* 2010;11:343-50.
31. Bai Y, Hu M, Chen Z, et al. Single-Cell Transcriptome Analysis Reveals RGS1 as a New Marker and Promoting Factor for T-Cell Exhaustion in Multiple Cancers. *Front Immunol* 2021;12:767070.
32. Gibbons DL, Abeler-Dörner L, Raine T, et al. Cutting Edge: Regulator of G protein signaling-1 selectively regulates gut T cell trafficking and colitic potential. *J Immunol* 2011;187:2067-71.
33. Hunt KA, Zhernakova A, Turner G. et al. Newly identified genetic risk variants for celiac disease related to the immune response. *Nat Genet* 2008;40:395-402.
34. Wu J, Zhang T, Xiong H, et al. Tumor-Infiltrating CD4(+) Central Memory T Cells Correlated with Favorable Prognosis in Oral Squamous Cell Carcinoma. *J Inflamm Res* 2022;15:141-152.
35. Mizokami M, Imanishi T, Ikeo K, et al. Mutation patterns for two flaviviruses: hepatitis C virus and GB virus C/hepatitis G virus. *FEBS Lett* 1999; 450:294-8.
36. Ling YM, Chen JY, Guo L, et al.  $\beta$ -defensin 1 expression in HCV infected liver/liver cancer: an important role in protecting HCV progression and liver cancer development. *Sci Rep* 2017;7:13404.
37. Donald CD, Sun CQ, Lim SD, et al. Cancer-specific loss of beta-defensin 1 in renal and prostatic carcinomas. *Lab Invest* 2003;83:501-5.