Uncorrected Proof

Hepat Mon. 2021 October; 21(10):e118535. doi: 10.5812/hepatmon.118535.

Published online 2022 January 19. **Research Article**

Identification of Molecular Subgroups in Liver Cirrhosis by Gene Expression Profiles

Ying-Xue Zhang [1], Feng-Xia Sun [1, *], Xiao-Ling Li [1], Qing-Hua Liu [2], Zi-Meng Chen [1] and Yu-Fei Guo [1]

[1] Department of Infections, Beijing Hospital of Traditional Chinese Medicine, Affiliated to Capital Medical University, Beijing, China
[2] Beijing University of Chinese Medicine, Beijing, China

[*] *Corresponding author*: Department of Infections, Beijing Hospital of Traditional Chinese Medicine, Affiliated to Capital Medical University, 100010, Beijing, China. Email: sunfengxia1969@126.com

Received 2021 August 14; **Revised** 2021 December 19; **Accepted** 2021 December 26.

Abstract

Background: Liver cirrhosis is characterized by high mortality, bringing a serious health and economic burden to the world. The clinical manifestations of liver cirrhosis are complex and heterogeneous. According to subgroup characteristics, identifying cirrhosis has become a challenge.
Objectives: The purpose of this study was to evaluate the difference between different subgroups of cirrhosis. The ultimate goal of research on these different phenotypes was to discover groups of patients with unique treatment characteristics, and formulate targeted treatment plans that improve the prognosis of the disease and improve the patients' quality of life.
Methods: We obtained the relevant gene chip by searching the gene expression omnibus (GEO) database. According to the gene expression profile, 79 patients with liver cirrhosis were divided into four subgroups, which showed different expression patterns. Therefore, we used weighted gene coexpression network analysis (WGCNA) to find differences between subgroups.
Results: The characteristics of the WGCNA module indicated that subjects in subgroup I might exhibit inflammatory characteristics; subjects in subgroup II might exhibit metabolically active characteristics; arrhythmogenic right ventricular cardiomyopathy and neuroactive ligand-receptive somatic interaction pathways were significantly enriched in subgroup IV. We did not find a significantly upregulated pathway in the third subgroup.
Conclusions: In this study, a new type of clinical phenotype classification of liver cirrhosis was derived by consensus clustering. This study found that patients in different subgroups may have unique gene expression patterns. This new classification method helps researchers explore new treatment strategies for cirrhosis based on clinical phenotypic characteristics.

Keywords: Liver Cirrhosis, Gene Expression Profile, Classification of Subgroups, Weighted Gene Coexpression Network Analysis Module

1. Background

Liver cirrhosis is a pathological stage characterized by diffuse liver fibrosis, pseudolobular formation, and blood vessel proliferation in and outside the liver as the characteristics of various chronic liver diseases. About two million people die of liver disease, and one million people die of complications of liver cirrhosis in the world every year. Liver cirrhosis is currently the 11th most common cause of death in the world, bringing a serious health and economic burden to the world (1, 2). Liver cirrhosis can lead to gastrointestinal bleeding, ascites, hepatic encephalopathy, and even life-threatening. Thus far, there is still no clear consensus on the treatment of liver cirrhosis. At present, it is mainly etiological treatment, anti-inflammatory, anti-hepatic fibrosis, and active prevention and treatment of complications. We may need a liver transplant if the liver is severely damaged (3). Patients with early liver disease usually have no symptoms and may have reached the advanced stage of the disease when they are found. Cirrhosis can lead to irreversible damage, but the early stage of chronic liver disease can be reversed to a healthy state. Therefore, it is essential to find biomarkers to identify the early stage of chronic liver disease to prevent severe liver damage. The discovery of biomarkers for liver cancer and liver cirrhosis has promoted sequencing technology development, and high-throughput sequencing is one of the representative technological innovations in the biological field in recent decades (4). Recently, many studies have been conducted to clarify the pathogenesis of liver cirrhosis based on the experimental data of microarray and high-throughput sequencing (5-8).

The cost reduction of high-throughput technology and

Copyright © 2022, Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/) which permits copy and redistribute the material just in noncommercial usages, provided the original work is properly cited.

Uncorrected Proof

the development of bioinformatics have brought new opportunities to reveal the pathogenesis of liver cirrhosis. Although most studies only pay attention to the differences between cirrhotic cases and normal controls, they rarely pay attention to the differences between cirrhotic cases. To reveal the heterogeneity between tumors, cancer cases are usually divided into multiple subgroups based on gene expression patterns (9). Inspired by cancer research, we used the same method to divide cirrhosis cases into four subgroups, the differences between subgroups were found by annotating the corresponding coexpression functional modules with the path of Kyoto Encyclopedia of Genes and Genomes (KEGG), thereby deepening understanding of the differences between cases of cirrhosis.

2. Methods

2.1. Data Collection

2.1.1. Data Download and Annotation

We used R/Bioconductor package GEOquery (10) to extract "gene expression omnibus" (GEO) objects. Searched the GEO database (https://www.ncbi.nlm.nih.gov/geo/), the search term was "cirrhosis", the gene chip that meets the requirements was incorporated into our research. According to the annotation information of the platform file, converted the probe matrix into a gene matrix.

2.1.2. Inclusion and Exclusion Criteria

Inclusion criteria: (1) populations diagnosed with cirrhosis, regardless of the severity of the disease; (2) study type is expression profiling by array; (3) entry type is series; (4) research contains samples from the cirrhosis disease group and the normal group.

Exclusion criteria: (1) there is no comparison between the disease group and the normal group; (2) the sample size of each group is less than six.

2.2. Removal of Batch Effect

Firstly, the installation package in R/Bioconductor was used to merge the obtained gene matrix, and batch correction was carried out (11). When merging data, log2 was taken for the data with large value for conversion. When there is a difference between batch effects in the data, ComBat-seq can obtain better statistical ability and control the false positive rate compared with other available methods (12). Therefore, the combat method was used to eliminate the batch effect. Finally, principal component analysis (PCA) was performed to evaluate the results.

2.3. Consensus Clustering

First, we used the "consensus cluster plus" package in R/Bioconductor for consensus clustering (13). The included cases of liver cirrhosis were clustered using the K-means algorithm with Spearman distance, the clustering results were determined by referring to the consistency score (> 0.8).

2.4. Extraction of Specific Upregulated Genes and Hub Genes in Subtypes

The subtypes were compared with other subtypes to determine the subgroup-specific upregulated genes. Wilcoxon rank-sum test was used to test the differential expression. The corrected P-value was < 0.05, and the absolute difference of means > 0.2. The difference of means is the mean of the subgroup minus the mean of the other subgroups. The specific upregulated genes in the subgroup were imported into the STRING database, and the species condition was set to human, the confidence level was greater than 0.9, and the protein-protein interaction (PPI) network was constructed. The PPI was imported into Cytoscape 3.8.0, and the Degree algorithm in the cytoHubba plug-in was used to obtain the most significant top 10 hub genes.

2.5. Gene Set Enrichment Analysis

Gene set enrichment analysis (GSEA) was implemented in GseaPrerank mode in GSEA 4.1.0. It was used to judge whether the specific differential genes in each subtype are also different from normal samples. This operation requires two files, one is the gene set file (type-specific gene file) composed of subgroup-specific upregulated genes, and the other is the gene list file, which is obtained by comparing the specific type with the normal sample. We compared these two files to see how similar they are.

2.6. Weighted Gene Coexpression Network Analysis

Weighted gene coexpression network analysis (WGCNA) is often used to find modules of highly relevant genes (14, 15). We used WGCNA to type the gene and got gene's module. Firstly, a distance matrix can be obtained by finding the optimal power value and calculating the distance between genes. It was convenient for gene clustering and dynamic identification module. After merging similar modules, we finally determined six functional modules. At the same time, we used the labeled heat map function option in the WGCNA package to draw the heat map, visualized the data, set the low expression of the gene as blue, the middle expression as white, and the high expression as red.

Uncorrected Proof

2.7. GO Enrichment Analysis and KEGG Enrichment Analysis

GO enrichment analysis, and KEGG enrichment analysis were performed for each functional module of WGCNA, in which GO was divided into three modules: (1) biological process (BP); (2) cellular component (CC); and (3) molecular function (MF) (16). KEGG is a database that integrates genomic, chemical, and system function information (17). We compared the pathways between modules through KEGG enrichment analysis. During KEGG enrichment analysis, set p-value < 0.05 for filtration. The enrichment results were visualized, and the bubble diagram was drawn.

3. Results

3.1. Microarray Data Characteristics

This study included four independent microarray information, involving four independent clinical trials, all from the GEO database, GSE14323, GSE77627, GSE123932, GSE128726, a total of 127 samples (including 79 patients with liver cirrhosis and 48 healthy subjects). GSE123932 provided clinical information of gender and age.

3.2. Removal of Batch Effect

Four gene chips detected a total of 11,152 genes. In order to eliminate the batch effect, we used the combat method. The PCA chart before correction showed that there were large sample differences between experiments before the batch effect was eliminated (Figure 1A). In contrast, the samples in the corrected PCA chart were randomly distributed, and the results showed that the cross-platform normalization successfully eliminated the batch effect (Figure 1B).

3.3. Consensus Clustering of Cirrhosis Cases

After eliminating the batch effect, 79 patients with liver cirrhosis were divided into subgroups. According to the consistency scoring situation in the data statistics, the consistency cluster analysis of gene expression profiles was divided into four subgroups, among which the number of cases in subgroups I, II, III and IV were 21, 30, 18, and 10, respectively. The internal relationship of subgroups is close, while the correlation between subgroups is weak (Figure 2A). Generally speaking, the stability of the type is positively correlated with the consistency score. According to the results of this study, when divided into 2, 3, or 4 groups, the cluster agreement score of each subgroup was higher than 0.8 (Figure 2B), which indicated that these classifications were more robust than other clusters. When considering more groups better, 79 patients with liver cirrhosis were divided into four subgroups.

3.4. GSEA Enrichment Analysis and WGCNA Analysis Results

The 2033, 2134, 804, and 1892 genes that were specifically upregulated in subgroups I, II, III, and IV were determined by differential analysis. The subgroup differential genes were imported into the STRING database, and a PPI network was constructed. Imported it into Cytoscape 3.8.0 software and arranged the nodes in an orderly manner according to the degree value. The top 10 hub genes of each subgroup are shown (Figure 3A - D). The results of GSEA enrichment analysis showed that the specific upregulated genes in the subtypes were also significantly different when compared with normal samples (Figure 4A - D, FDR < 0.05). It still needs to be emphasized that subgroup III had the least subgroup-specific upregulated genes, and compared with the normal group, there were less specific genes (Table 1).

We typed 6,863 specific genes upregulated in four subgroups by WGCNA, and identified six WGCNA modules. The subgroups corresponding to WGCNA modules are summarized in Table 1. The genes of six modules were analyzed by KEGG. The results showed that TGF-β signaling pathway, viral protein interaction with cytokines, and cytokine receptors, including a variety of chemokines and inflammatory factors, were only enriched in the black module significantly. At the same time, the PI3K/Akt signaling pathway was also significantly enriched in the black module, and there were many upregulated genes. The genes in the black module were upregulated in subgroup I, indicating that subjects in subgroup I may exhibit inflammatory characteristics. Fatty acid metabolism and biosynthesis of cofactors were only significantly enriched in the brown module, and carbon metabolism and protein processing pathways in the endoplasmic reticulum were only significantly enriched in the brown and purple module. The genes in the brown and purple module were in subgroup II. They were upregulated, indicating that subjects in subgroup II may show metabolic activity; arrhythmogenic right ventricular cardiomyopathy and neuroactive ligand-receptor interaction pathways were only significantly enriched in the turquoise module, and the genes in the turquoise module were upregulated in subgroup IV. We did not find a significant pathway of upregulation in the third subgroup. Based on the above results, different subgroups regulate the occurrence or progression of liver cirrhosis through specific pathways (Figure 5A - B).

After GO enrichment analysis (Figure 5C - E), the biological process (BP): the black module mainly included the regulation of cell-cell adhesion, positive regulation of cell adhesion, positive regulation of cell-cell adhesion, blood coagulation, and hemostasis; the brown module mainly included small molecule catabolic process, cellular amino acid metabolic process, cellular ketone metabolic

Uncorrected Proof

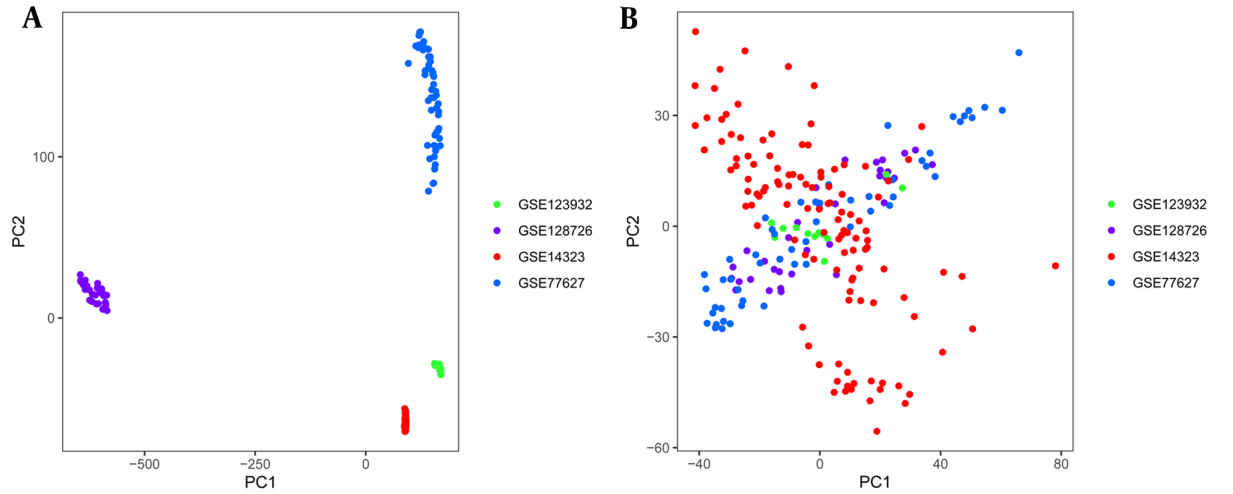


Figure 1. PCA of gene expression data set. scatters of different colors represent samples from four different data sets. A, PCA diagram before batch correction; B, PCA diagram after batch correction.

Table 1. The Number of Differentially Expressed Genes in Each Subgroup Was Obtained By Comparing the Control Group and the Case Group and the Weighted Gene Coexpression Analysis Module

Subtype	Specific Genes were Compared with The Normal Group	Specific Genes Compared with Other Subgroups	Specific Upregulated Genes in Subgroup	Modular
I	6,021	4,848	2,033	Black
II	3,387	3,169	2,134	Brown, purple
III	3,929	2,274	804	
IV	6,404	6,407	1,892	Turquoise

process, and response to endoplasmic reticulum stress; the purple module mainly included nucleotide-excision repair, respiratory electron transport chain, ATP synthesis coupled electron transport, mitochondrial ATP synthesis coupled electron transport, cellular respiration, etc.; and the turquoise module mainly included axonogenesis, gland development, regulation of trans-synaptic signaling, modulation of chemical synaptic transmission, embryonic organ development, sensory organ morphogenesis, embryonic organ morphogenesis, etc. Cell components (CC): the black module concentrated on focal adhesion, cell-substrate junctions, membrane rafts, collagen-containing extracellular matrix, membrane microdomain, and endocytosis vesicles; the brown module concentrated on mitochondrial matrix, mitochondrial protein complex, mitochondrial inner membrane, focal adhesion, cell-substrate junction, collagen-containing extracellular matrix; the purple module concentrated on mitochondrial matrix, mitochondrial protein complex, mitochondrial inner membrane, organellar ribosome, respirasome, and oxidoreductase complex. The turquoise module mainly concentrated on transmembrane transporter complex, transporter complex, synaptic membrane, neuron to neuron synapse, postsynaptic membrane, and ion channel complex. Molecular function (MF): the black module mainly included receptor ligand activity, signaling, receptor activator activity, DNA-binding, transcription factor binding, etc.; the brown module mainly included DNA-binding, transcription factor binding, oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor, lyase activity, carbon-oxygen lyase activity, electron transfer activity; the purple module mainly included NADH dehydrogenase activity, NADH dehydrogenase (quinone) activity, NADH dehydrogenase (ubiquinone) activity; the turquoise module mainly included signaling receptor activator activity, receptor ligand activity, metal ion transmembrane transporter activity, channel activity, neurotransmitter receptor activity, etc.

4. Discussion

Cancer molecular subtypes were inspired this research. The current molecular subtypes guide the diagno-

Uncorrected Proof

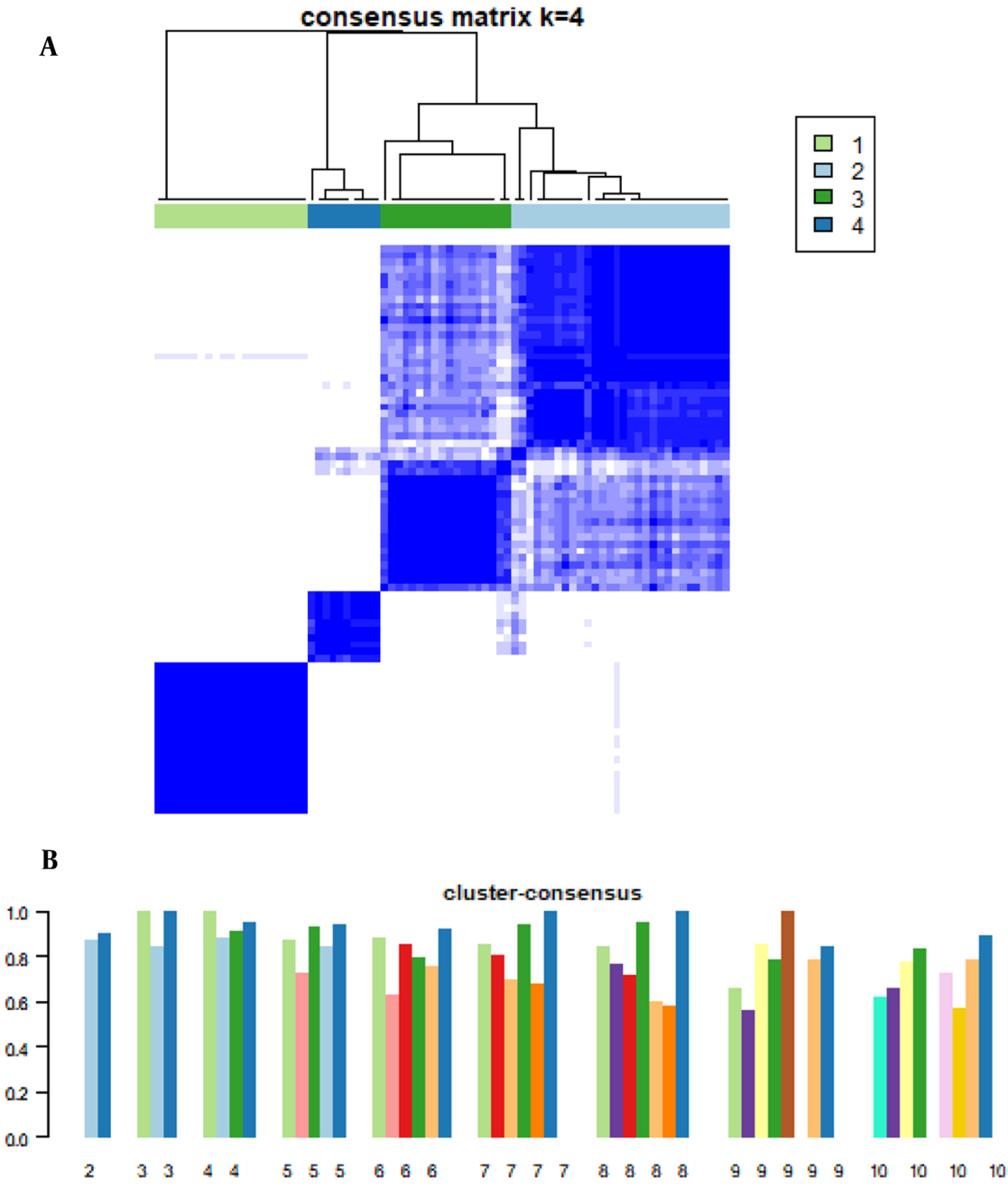


Figure 2. Consensus clustering of gene expression profiles in patients with liver cirrhosis. A, the consistency matrix, when the number of clusters, is 4 that is determined by the minimum consistency score (> 0.8) of the subgroup; B, consistency scores for subgroups with cluster numbers between 2 and 10.

sis and clinical treatment of many cancer types (18-20). In addition to cancer research, non-cancerous diseases can also identify new and clinically relevant molecular subtypes of diseases based on gene expression profiles (21, 22). Although these molecular subtypes are not perfect, they help for the diagnosis and treatment of diseases to a certain extent.

Liver cirrhosis is relatively complex and shows clinical heterogeneity. Previous studies mostly compared the gene expression profiles of liver cirrhosis cases with normal con-

Uncorrected Proof

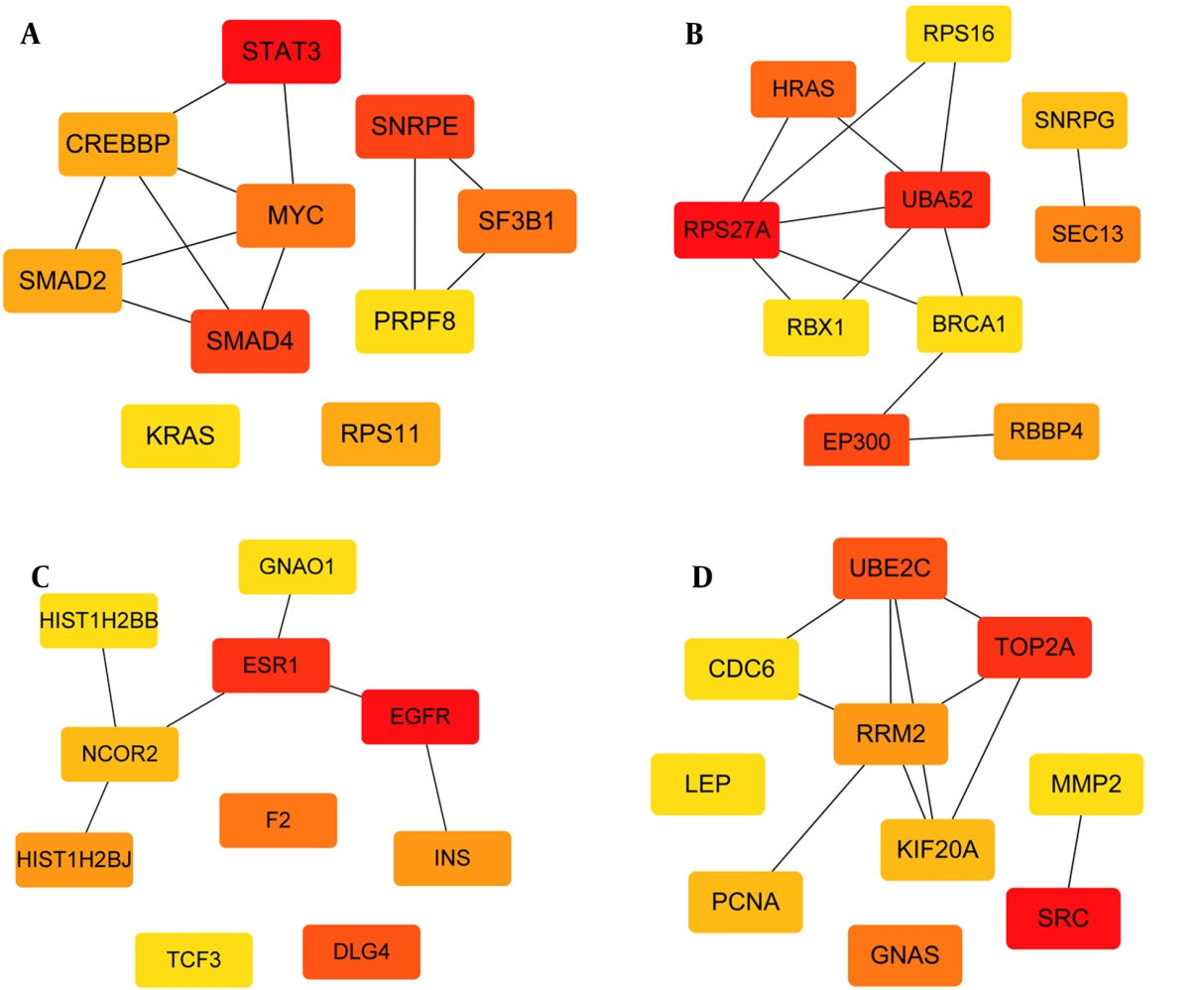


Figure 3. Top 10 hub genes in the four subgroups.

trols. In this study, we further divided liver cirrhosis cases into subgroups. Unlike previous studies, we linked specific pathways with specific subgroups of liver cirrhosis.

The hub genes in the first subgroup are mainly STAT3, SMAD2, SMAD4, MYC. STAT3 regulates basic cell processes, including inflammation, cell growth, proliferation, differentiation, etc. Studies have shown that STAT3 integrates a variety of profibrotic signals and may be the core mediator of fibrosis (23). Moreover, TGF-β is one of the key drivers of fibrosis. The production of TGF-β is related to the progression of liver fibrosis. In many experimental models, inhibiting the TGF-β1 signaling pathway has been shown to reduce the development of fibrosis (24). In TGF-β1 signaling, the Smads protein is the main mediator of receptor-induced phosphorylation and nuclear translocation (25, 26). In chronic liver disease, inflammatory mediators interact with c-Myc, leading to fibrosis, liver cirrhosis, and liver tumors (27). In our study, the TGF-β signaling pathway was most significantly enriched in the first subcomponent type. TGF-β signaling is involved in the entire process from liver injury to inflammation and fibrosis, cirrhosis, and cancer, TGF-β plays a key role in hepatocyte cell growth and apoptosis and promotes liver differentiation during embryogenesis and physiological liver regeneration. However, due to chronic liver injury, high levels of TGF-β cause stellate cells to activate myofibroblasts and a large number of hepatocytes to die, promoting liver fibrosis and cirrhosis (28).Viral protein interactions with cytokine and cytokine receptors were also most significantly enriched in the first subcomponent type, in-

Uncorrected Proof

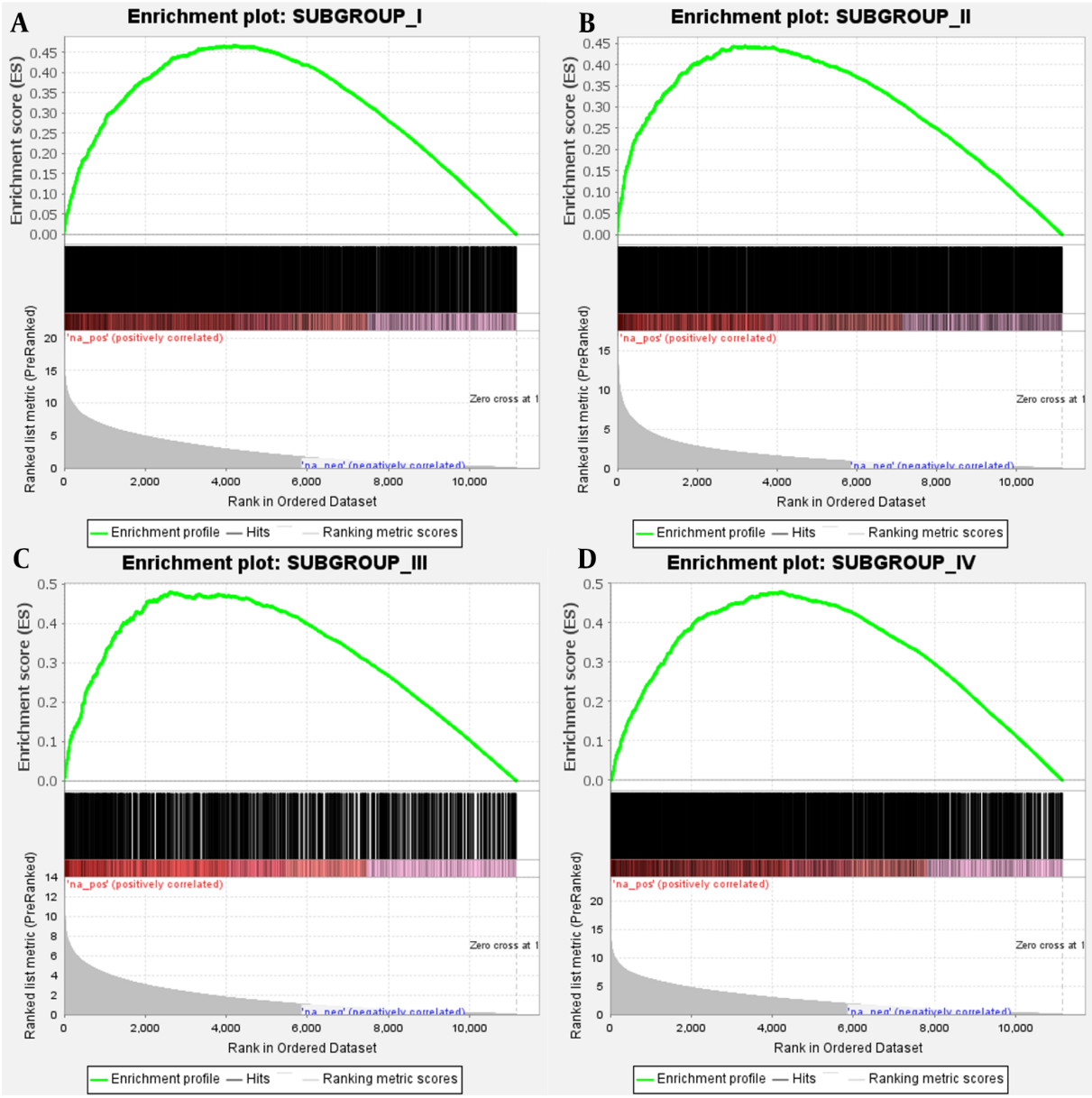


Figure 4. GSEA enrichment analysis. The enrichment plots of A, B, C and D, illustrate the specific upregulated genes in the subgroup that are different from those in the normal samples.

cluding a variety of chemokines and inflammatory factors, such as chemokines CCL2, CX3CL1, etc., biliary epithelial cells (BECs) in damaged small bile ducts in primary biliary cirrhosis (PBC) show senescent features.. Senescent BECs showed upregulation of a variety of chemokines and chemotactic activities. The expression of CCL2 and CX3CL1 increases in senescent BECs in PBC, which may promote the infiltration of corresponding cells expressing CCR2 and CX3CR1, and further aggravate the PBC bile duct inflammation of the lesion (29, 30).

At the same time, PI3K/Akt signaling pathway was also significantly enriched in subgroup I, and there were many upregulated genes. PI3K/Akt signaling pathway regulates the activation of downstream signaling molecules. It plays a key role in regulating the immune response and the release of inflammatory factors in vivo and in vitro (31). Combined with the results of GO functional enrichment analysis, the first subcomponent mainly included the regula-

Uncorrected Proof

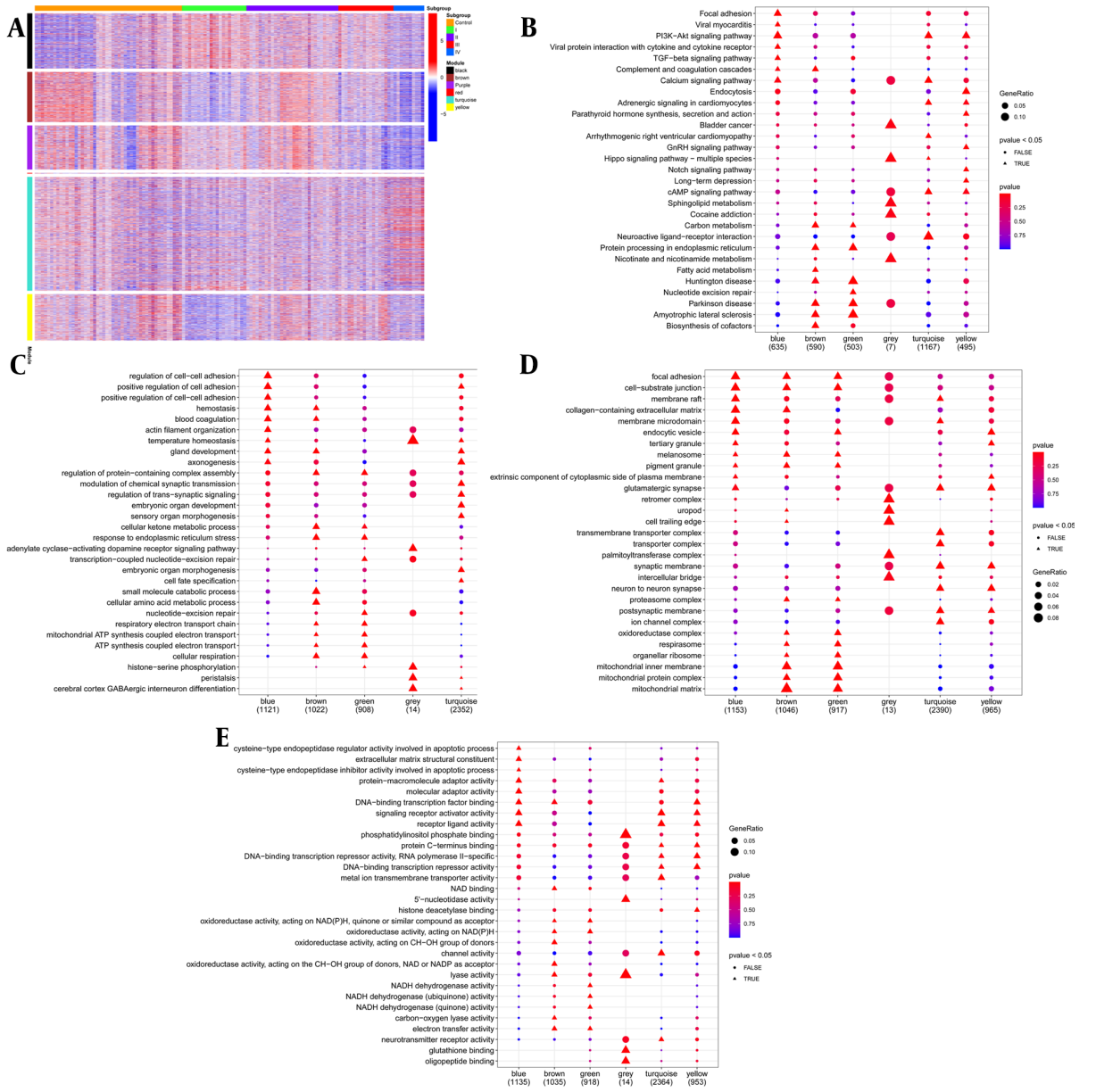


Figure 5. A, the heat map shows the gene expression levels of the six WGCNA modules in each subgroup; B, bResults of KEGG enrichment analysis for genes in each WCGNA module; C - E, the gene enrichment analysis of each WGCNA module by GO pathway.

tion of cell-cell adhesion, positive regulation of cell adhesion, positive regulation of cell-cell adhesion, blood coagulation, and hemostasis, and other biological processes. The combination of cell to extracellular matrix interaction and cell-cell adhesion depends on intercellular adhesion molecules. Cell adhesion molecules participate in various interactions. They can broadly support fibrosis formation by releasing TGF-β, promoting fibrosis, or mediating leukocyte migration. Some candidate cell adhesion molecules have been evaluated as targets for preventing and/or reversing liver inflammation and fibrosis (32). Inflammation triggers clotting, reduces the activity of natural anticoagulant mechanisms, and impairs the fibrinolytic systemz.. Inflammatory mediators can increase platelet count, platelet reactivity, downregulate the natural anticoagulation mechanism, activate the coagulation

Uncorrected Proof

system, promote the spread of coagulation response, and damage fibrinolysis. Similarly, coagulation can increase the inflammatory response by releasing mediators from platelets and activated cells (33). In summary, it was indicated that subjects in subgroup I may exhibit inflammatory features.

In the second subgroup, the hub genes mainly include RPS27A, HRAS, SEC1, UBA52, etc. Among them, RPS27a performs extra-ribosomal functions in addition to playing a role in ribosome biogenesis and post-translational modifications of proteins. RPS27a plays a role in mouse liver cancer and some overexpressed in human tumors. Studies have reported that STAT3 is the key activator of the RPS27a promoter, and upregulation of RPS27a depends on STAT3 phosphorylation (34). Studies have shown that increased HRAS protein level can directly stimulate transduction of the Smad2/3 signal and the accumulation of type I collagen, which is related to the appearance fibrosis (35, 36). Upregulation of metabolism was observed in subgroup II, including fatty acid metabolism, cofactor biosynthesis, carbon metabolism, and protein processing pathways in the endoplasmic reticulum. Combined with the results of the GO functional enrichment analysis, subgroup II mainly included the small molecule catabolic process, the metabolic process of cellular amino acids, the metabolic process of cellular ketones and the response to endoplasmic reticulum stress, repair of nucleotides, respiratory electron transport chain and other biological processes, indicating that subjects in subgroup II may exhibit metabolic activity characteristic. Systems biology studies have shown that carbohydrates, amino acids, and lipid-related pathways have changed in liver fibrosis and cirrhosis, and more studies are needed to verify metabolic characteristics and determine therapeutic targets of liver fibrosis (37).

The hub genes in the fourth subgroup mainly include SRC, TOP2A, UBE2C, and so on. Studies have shown that Src is related to liver fibrosis. The expression of Src mRNA in liver tissues of mice with liver fibrosis is significantly increased, the levels of phosphorylated Src and total Src are significantly increased; Src in liver tissues of patients with liver cirrhosis is significantly higher than in liver tissue from the normal control group. Src inhibition can prevent TAA-induced liver fibrosis, inhibit HSC activation, and is a potential therapeutic target for liver fibrosis (38). Differential gene expression analysis and weighted gene correlation network analysis were performed in tissue samples from normal liver, cirrhosis, and HCC to determine the pathological progress of 20 hub genes from cirrhosis to HCC. TOP2A is one of the pivotal genes in the pathological progression from liver cirrhosis to HCC (39). Research results indicate that UBE2C may be a key gene in the progression of HCC and a promising therapeutic target for the treatment of HCC (40, 41). Arrhythmogenic right ventricular cardiomyopathy and the neuroactive ligand-receptor interaction pathway were only significantly enriched in subgroup IV. Studies have shown that the neuroactive ligand-receptor interaction pathways are closely related to the appearance and development, invasion, and metastasis of liver cancer. Various traditional Chinese medicine and Chinese patent medicine exert their curative effects by influencing this signaling pathway (42-44). However, the relationship between this pathway and its key targets and liver cirrhosis needs further study. In the third subgroup, the hub genes mainly include B4GALT2, FUT9, FUT3, EGFR, and so on. Relevant studies have shown that the expression of EGF in the liver increases during cirrhosis, and inhibition of EGFR can reverse liver fibrosis. The use of FDA-approved inhibitors to inhibit EGFR provides a promising treatment for reducing liver fibrosis and preventing HCC (45). We did not find a significantly upregulated pathway in subgroup III, which may be related to the small sample size.

The above results show that different subgroups of liver cirrhosis have different gene expression patterns and obvious heterogeneity, which may represent different stages of liver cirrhosis, and we seek for large-sample prospective trials to verify it in the future. This study has certain limitations. Because only GSE123932 provided clinical information of gender and age, and the sample size was small, the correlation between clinical characteristics and WGCNA module could not be carried out.

In conclusion, inspired by the study of cancer subgroups, we adopted a similar strategy to reveal the molecular subgroups of liver cirrhosis. The ultimate goal of the study of these different phenotypes is to find patient groups with unique treatment characteristics and formulate targeted treatment plans. Our study suggests that patients in different subgroups may have their unique molecular characteristics. This new classification method is helpful for researchers to explore new treatment strategies for liver cirrhosis according to clinical phenotypic characteristics, to improve the prognosis of the disease and improve the quality of life of patients.

Footnotes

Authors' Contribution: YXZ conceived this research and contributed to data collation, original draft preparation, and draft review. FXS contributed to the writing, review, editing, and Supervision. XLL and QHL performed the image analyses and analyzed the data. ZMC and YFG contributed to the writing-original draft preparation. All authors designed the experiments, wrote the manuscript,

Uncorrected Proof

and revised the paper. All authors read and approved the final version of the manuscript.

Conflict of Interests: The authors declare that there is no conflict of interest regarding the publication of this paper.

Data Reproducibility: The data presented in this study are openly available in one of the repositories or will be available on request from the corresponding author by this journal representative at any time during submission or after publication. Otherwise, all consequences of possible withdrawal or future retraction will be with the corresponding author.

Funding/Support: The work was supported by Special Research Project of Capital Health Development (NO.2018-2-2173).

Informed Consent: In this study, the relevant data are downloaded from the GEO database for integrated analysis. It does not involve interventions on humans and animals; so there is no informed consent.

References

1. Xiaoyuan X, Huiguo D, Wengang L, Jinghang X, Ying H, Jidong J, et al. Guidelines for diagnosis and treatment of liver cirrhosis. *Journal of Practical Hepatology*. 2019;**22**(6):770–86.
2. Asrani SK, Devarbhavi H, Eaton J, Kamath PS. Burden of liver diseases in the world. *J Hepatol*. 2019;**70**(1):151–71. doi: 10.1016/j.jhep.2018.09.014. [PubMed: 30266282].
3. Lee NY, Suk KT. The Role of the Gut Microbiome in Liver Cirrhosis Treatment. *Int J Mol Sci*. 2020;**22**(1). doi: 10.3390/ijms22010199. [PubMed: 33379148]. [PubMed Central: PMC7796381].
4. Oh S, Jo Y, Jung S, Yoon S, Yoo KH. From genome sequencing to the discovery of potential biomarkers in liver disease. *BMB Rep*. 2020;**53**(6):299–310. [PubMed: 32475383]. [PubMed Central: PMC7330805].
5. Fan Y, Zheng Y, Wang J, Zhao T, Liang T. Differential expression and bioinformatic analysis of circRNA in nonalcoholic steatohepatitis cirrhosis. *Int J Clin Exp Pathol*. 2020;**13**(11):2820–30. [PubMed: 33284878]. [PubMed Central: PMC7716119].
6. Chan KM, Wu TH, Wu TJ, Chou HS, Yu MC, Lee WC. Bioinformatics microarray analysis and identification of gene expression profiles associated with cirrhotic liver. *Kaohsiung J Med Sci*. 2016;**32**(4):165–76. doi: 10.1016/j.kjms.2016.03.008. [PubMed: 27185598].
7. Zhu M, Lu T, Jia Y, Luo X, Gopal P, Li L, et al. Somatic Mutations Increase Hepatic Clonal Fitness and Regeneration in Chronic Liver Disease. *Cell*. 2019;**177**(3):608–621 e12. doi: 10.1016/j.cell.2019.03.026. [PubMed: 30955891]. [PubMed Central: PMC6519461].
8. Gong Z, Tang J, Xiang T, Lin J, Deng C, Peng Y, et al. Genomewide identification of long noncoding RNAs in CCl4induced liver fibrosis via RNA sequencing. *Mol Med Rep*. 2018;**18**(1):299–307. doi: 10.3892/mmr.2018.8986. [PubMed: 29749545]. [PubMed Central: PMC6059682].
9. Giordano TJ. The cancer genome atlas research network: A sight to behold. *Endocr Pathol*. 2014;**25**(4):362–5. doi: 10.1007/s12022-014-9345-4. [PubMed: 25367656].
10. Davis S, Meltzer PS. GEOquery: A bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics*. 2007;**23**(14):1846–7. doi: 10.1093/bioinformatics/btm254. [PubMed: 17496320].
11. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012;**28**(6):882–3. doi: 10.1093/bioinformatics/bts034. [PubMed: 22257669]. [PubMed Central: PMC3307112].
12. Zhang Y, Parmigiani G, Johnson WE. ComBat-seq: Batch effect adjustment for RNA-seq count data. *NAR Genom Bioinform*. 2020;**2**(3):lqaa078. doi: 10.1093/nargab/lqaa078. [PubMed: 33015620]. [PubMed Central: PMC7518324].
13. Wilkerson MD, Hayes DN. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010;**26**(12):1572–3. doi: 10.1093/bioinformatics/btq170. [PubMed: 20427518]. [PubMed Central: PMC2881355].
14. 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. [PubMed: 19114008]. [PubMed Central: PMC2631488].
15. Yin L, Cai Z, Zhu B, Xu C. Identification of Key Pathways and Genes in the Dynamic Progression of HCC Based on WGCNA. *Genes (Basel)*. 2018;**9**(2). doi: 10.3390/genes9020092. [PubMed: 29443924]. [PubMed Central: PMC5852588].
16. Du Z, Zhou X, Ling Y, Zhang Z, Su Z. agriGO: A GO analysis toolkit for the agricultural community. *Nucleic Acids Res*. 2010;**38**(Web Server issue):W64–70. doi: 10.1093/nar/gkq310. [PubMed: 20435677]. [PubMed Central: PMC2896167].
17. Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res*. 2019;**47**(D1):D590–5. doi: 10.1093/nar/gky962. [PubMed: 30321428]. [PubMed Central: PMC6324070].
18. Collisson EA, Bailey P, Chang DK, Biankin AV. Molecular subtypes of pancreatic cancer. *Nat Rev Gastroenterol Hepatol*. 2019;**16**(4):207–20. doi: 10.1038/s41575-019-0109-y. [PubMed: 30718832].
19. Prat A, Pineda E, Adamo B, Galvan P, Fernandez A, Gaba L, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast*. 2015;**24**(Suppl 2):S26–35. doi: 10.1016/j.breast.2015.07.008. [PubMed: 26253814].
20. Choi W, Ochoa A, McConkey DJ, Aine M, Hoglund M, Kim WY, et al. Genetic Alterations in the Molecular Subtypes of Bladder Cancer: Illustration in the Cancer Genome Atlas Dataset. *Eur Urol*. 2017;**72**(3):354–65. doi: 10.1016/j.eururo.2017.03.010. [PubMed: 28365159]. [PubMed Central: PMC5764190].
21. Peng XY, Wang Y, Hu H, Zhang XJ, Li Q. Identification of the molecular subgroups in coronary artery disease by gene expression profiles. *J Cell Physiol*. 2019. doi: 10.1002/jcp.28324. [PubMed: 30805932].
22. Chang Y, Glass K, Liu YY, Silverman EK, Crapo JD, Tal-Singer R, et al. COPD subtypes identified by network-based clustering of blood gene expression. *Genomics*. 2016;**107**(2-3):51–8. doi: 10.1016/j.ygeno.2016.01.004. [PubMed: 26773458]. [PubMed Central: PMC4761317].
23. Chakraborty D, Sumova B, Mallano T, Chen CW, Distler A, Bergmann C, et al. Activation of STAT3 integrates common profibrotic pathways to promote fibroblast activation and tissue fibrosis. *Nat Commun*. 2017;**8**(1):1130. doi: 10.1038/s41467-017-01236-6. [PubMed: 29066712]. [PubMed Central: PMC5654983].
24. Wynn TA, Ramalingam TR. Mechanisms of fibrosis: Therapeutic translation for fibrotic disease. *Nat Med*. 2012;**18**(7):1028–40. doi: 10.1038/nm.2807. [PubMed: 22772564]. [PubMed Central: PMC3405917].
25. Massague J, Seoane J, Wotton D. Smad transcription factors. *Genes Dev*. 2005;**19**(23):2783–810. doi: 10.1101/gad.1350705. [PubMed: 16322555].
26. Heldin CH, Miyazono K, ten Dijke P. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature*. 1997;**390**(6659):465–71. doi: 10.1038/37284. [PubMed: 9393997].
27. Liu T, Zhou Y, Ko KS, Yang H. Interactions between Myc and Mediators of Inflammation in Chronic Liver Diseases. *Mediators Inflamm*. 2015;**2015**:276850. doi: 10.1155/2015/276850. [PubMed: 26508814]. [PubMed Central: PMC4609837].

Uncorrected Proof

28. Fabregat I, Moreno-Caceres J, Sanchez A, Dooley S, Dewidar B, Giannelli G, et al. TGF-beta signalling and liver disease. *FEBS J*. 2016;**283**(12):2219–32. doi: 10.1111/febs.13665. [PubMed: 26807763].
29. Sasaki M, Miyakoshi M, Sato Y, Nakanuma Y. Modulation of the microenvironment by senescent biliary epithelial cells may be involved in the pathogenesis of primary biliary cirrhosis. *J Hepatol*. 2010;**53**(2):318–25. doi: 10.1016/j.jhep.2010.03.008. [PubMed: 20570384].
30. Sasaki M, Miyakoshi M, Sato Y, Nakanuma Y. Chemokine-chemokine receptor CCL2-CCR2 and CX3CL1-CX3CR1 axis may play a role in the aggravated inflammation in primary biliary cirrhosis. *Dig Dis Sci*. 2014;**59**(2):358–64. doi: 10.1007/s10620-013-2920-6. [PubMed: 24185682].
31. Zhou T, Shi J, Li X. [Role of PI3K/Akt signaling pathway in the innate immune of sepsis]. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue*. 2018;**30**(11):1091–4. Chinese. doi: 10.3760/cma.j.issn.2095-4352.2018.011.016. [PubMed: 30541652].
32. Hintermann E, Christen U. The Many Roles of Cell Adhesion Molecules in Hepatic Fibrosis. *Cells*. 2019;**8**(12). doi: 10.3390/cells8121503. [PubMed: 31771248]. [PubMed Central: PMC6952767].
33. Esmon CT. The interactions between inflammation and coagulation. *Br J Haematol*. 2005;**131**(4):417–30. doi: 10.1111/j.1365-2141.2005.05753.x. [PubMed: 16281932].
34. Wang H, Xie B, Kong Y, Tao Y, Yang G, Gao M, et al. Overexpression of RPS27a contributes to enhanced chemoresistance of CML cells to imatinib by the transactivated STAT3. *Oncotarget*. 2016;**7**(14):18638–50. doi: 10.18632/oncotarget.7888. [PubMed: 26942564]. [PubMed Central: PMC4951316].
35. Marut W, Kavian N, Servettaz A, Hua-Huy T, Nicco C, Chereau C, et al. Amelioration of systemic fibrosis in mice by angiotensin II receptor blockade. *Arthritis Rheum*. 2013;**65**(5):1367–77. doi: 10.1002/art.37873. [PubMed: 23335130].
36. Smaldone S, Olivieri J, Gusella GL, Moroncini G, Gabrielli A, Ramirez F. Ha-Ras stabilization mediates pro-fibrotic signals in dermal fibroblasts. *Fibrogenesis Tissue Repair*. 2011;**4**(1):8. doi: 10.1186/1755-1536-4-8. [PubMed: 21362163]. [PubMed Central: PMC3059295].
37. Seo HY, Lee SH, Lee JH, Kang YN, Hwang JS, Park KG, et al. Src Inhibition Attenuates Liver Fibrosis by Preventing Hepatic Stellate Cell Activation and Decreasing Connetive Tissue Growth Factor. *Cells*. 2020;**9**(3). doi: 10.3390/cells9030558. [PubMed: 32120837]. [PubMed Central: PMC7140470].
38. Lin Y, Liang R, Ye J, Li Q, Liu Z, Gao X, et al. A twenty gene-based gene set variation score reflects the pathological progression from cirrhosis to hepatocellular carcinoma. *Aging (Albany NY)*. 2019;**11**(23):11157–69. doi: 10.18632/aging.102518. [PubMed: 31811111]. [PubMed Central: PMC6932912].
39. Xiong Y, Lu J, Fang Q, Lu Y, Xie C, Wu H, et al. UBE2C functions as a potential oncogene by enhancing cell proliferation, migration, invasion, and drug resistance in hepatocellular carcinoma cells. *Biosci Rep*. 2019;**39**(4). doi: 10.1042/BSR20182384. [PubMed: 30914455]. [PubMed Central: PMC6470407].
40. Wu M, Liu Z, Zhang A, Li N. Identification of key genes and pathways in hepatocellular carcinoma: A preliminary bioinformatics analysis. *Medicine (Baltimore)*. 2019;**98**(5). e14287. doi: 10.1097/MD.0000000000014287. [PubMed: 30702595]. [PubMed Central: PMC6380748].
41. Chang ML, Yang SS. Metabolic Signature of Hepatic Fibrosis: From Individual Pathways to Systems Biology. *Cells*. 2019;**8**(11). doi: 10.3390/cells8111423. [PubMed: 31726658]. [PubMed Central: PMC6912636].
42. Rensong H. *Screening and analysis of lncRNA related differential expression in liver cancer*. Fuzhou, China: Fujian Medical University; 2019.
43. Minggang W, Qi C, Xiang G, Zhihua Y, Mingzhong X, Hanmin L. Bioinformatics analysis of Diwuyanggan Capsules in the prevention and treatment of liver cancer based on expression profile chips. *Journal of Integrated Traditional Chinese and Western Medicine Liver Diseases*. 2019;**29**(3):244–7.
44. Zexin Z, Bingbing L, Xianfeng Q, Jing L, Chong Z, Lizhu L. Based on network pharmacology, research on the mechanism of action of tricornu-zedoary turmeric on the treatment of hepatocellular carcinoma. *Journal of Shandong University of Traditional Chinese Medicine*. 2021;**45**(1):81–9.
45. Fuchs BC, Hoshida Y, Fujii T, Wei L, Yamada S, Lauwers GY, et al. Epidermal growth factor receptor inhibition attenuates liver fibrosis and development of hepatocellular carcinoma. *Hepatology*. 2014;**59**(4):1577–90. doi: 10.1002/hep.26898. [PubMed: 24677197]. [PubMed Central: PMC4086837].