Investigation of the pathogenesis of liver cirrhosis associated with type 2 diabetes mellitus via bioinformatic analysis

zhiyu xiong
Renmin Hospital of Wuhan University

Mengqin Yuan
Renmin Hospital of Wuhan University

Lichao Yao
Renmin Hospital of Wuhan University

Zheng Wang
Renmin Hospital of Wuhan University

Pingji Liu
Renmin Hospital of Wuhan University

Yingan Jiang (jiangya_cn@aliyun.com)
Renmin Hospital of Wuhan University

Kai Dai
Renmin Hospital of Wuhan University

Research Article

Keywords: Liver cirrhosis, T2DM, Bioinformatics, Hub genes, Differentially expressed genes, TFs, Drug

Posted Date: February 8th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2545733/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

The prevalence of type 2 diabetes mellitus (T2DM) with liver cirrhosis continues to increase globally. T2DM is identified as an independent risk factor for liver cirrhosis and an important prognostic factor for clinical outcomes in patients with liver cirrhosis. However, this co-occurring mechanism has not yet been elucidated. Therefore, this study aims to investigate the mechanisms underlying the co-pathogenesis of liver cirrhosis and T2DM and to provide reference information for future diagnoses and treatment of patients with liver cirrhosis associated with T2DM.

Methods

RNA-seq profile of liver cirrhosis and T2DM was downloaded from Gene Expression Omnibus (GEO) database and analyzed. Differentially expressed genes (DEGs) associated with liver cirrhosis and T2DM were identified using GEO2R. Thereafter, the co-differentially expressed genes (co-DEGs) associated with liver cirrhosis and T2DM were obtained from the intersection of the datasets on the DEGs. Subsequently, 175 overlapping DEGs were identified and further analyzed using a bioinformatic approach, which included Gene Ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses, protein–protein interaction (PPI) network analysis, transcription factors (TFs)–gene interaction network analysis, and drug candidate prediction analysis.

Results

The intersection of datasets on DEGs associated with liver cirrhosis and T2DM enabled the selection of 175 co-DEGs for subsequent analyses. Functional enrichment analyses showed that these co-DEGs are associated with inflammatory cytokine responses and positive regulation of transforming growth factor-β1 (TGF-β1). The KEGG analysis showed that advanced glycation end products–receptor for advanced glycation end products signaling pathway was markedly involved in liver cirrhosis associated with T2DM. Thereafter, a total of eight hub genes: SPARC, COL4A2, THBS1, LUM, TIMP3, COL3A1, IGFBP7, and FSTL1, associated with the diseases were identified using five algorithms from Cytoscape app for network centrality analysis and CytoHubba (a plug-in in the Cytoscape software). In total, 29 TFs of the hub genes were detected by NetworkAnalyst and Drug SIGnatureS DataBase, which predicted that retinoic acid is one of the promising agents that may be used for the treatment of liver cirrhosis associated with T2DM.

Conclusions

This study elucidated the common pathogenesis of liver cirrhosis and T2DM and predicted a potential clinical therapeutic drug. Therefore, these novel findings may contribute to the literature on the
pathogenesis of liver cirrhosis associated with T2DM.

**Introduction**

Liver cirrhosis is an end-stage liver disease caused by multiple hepatotoxic factors and biliary tract diseases. Chronic liver injury damages liver epithelial cells and triggers fibrogenesis, thereby inducing the activation of hepatic stellate cells (HSCs) and producing a large amount of extracellular matrix (ECM), eventually leading to tissue structure disorder and the formation of diffuse hepatic damage, ultimately causing liver dysfunction [1]. Notably, the progression of liver cirrhosis and the development of type 2 diabetes mellitus (T2DM) are connected. Grancini et al. showed that > 80% of patients with liver cirrhosis had impaired glucose tolerance and 48.6% of those patients had T2DM as a comorbidity [2].

T2DM is a metabolic disorder characterized by hyperglycemia due to insulin resistance [3]. The progression of T2DM affects the prognosis of liver cirrhosis. T2DM is an important risk factor that promotes the development of liver cirrhosis. According to six cohort studies conducted by Dyal et al., patients with liver cirrhosis associated with T2DM have an increased risk of developing advanced liver cirrhosis [odds ratio (OR) value, 2.25–9.24] [4]. Furthermore, T2DM may also decrease patient survival rate by increasing the risk of cancer in patients with liver cirrhosis associated with T2DM. Nishida found that the 5-year survival rate of patients with liver cirrhosis who exhibited normoglycemia was 94.7%, whereas the 5-year survival rate of patients with liver cirrhosis associated with T2DM declined to 56.6% [5]. Marengo et al. demonstrated that compared with normal-glucose-metabolism cirrhotic group, in T2DM cirrhotic group, the incidence of liver cancer was three-fold [6].

Liver cirrhosis and T2DM are common chronic diseases that are interrelated. The development of liver cirrhosis causes dysfunction in the receptors of the liver and pancreas, which leads to decreased insulin binding, resulting in insulin resistance and therefore the development of T2DM [7]. Cytokines may act as an important factor affecting the co-occurrence of the two diseases. An increase in the level of tumor necrosis factor α (TNF-α) may further aggravate the inflammation that contributes to liver damage [8]. Insulin resistance induces the production and activation of TNF-α, resulting in the massive release of insulin and insulin-like growth factors (IGFs). Increased production and secretion of IGF-1 may promote the proliferation of liver cancer cells and reduce their apoptosis, thereby accelerating the carcinogenesis of hepatic cells in liver cirrhosis [9].

Furthermore, the synthesis and secretion of inflammatory mediators induced by hyperglycemia not only increases oxidative stress response in the liver, which results in mitochondrial damage, lipid peroxidation, and decreased density lipoprotein oxidation but also stimulates the release of inflammatory factors (IFs) and the activation of HSCs, further leading to the progression of liver cirrhosis [10]. Additionally, the advanced glycation end products–receptor for advanced glycation end products (AGE–RAGE) signaling pathway may be one of the critical pathways involved in liver cirrhosis associated with T2DM. Moreover, the accumulation of AGEs owing to glycosylation in T2DM is associated with persistent hyperglycemia in the disease. This accumulation may further activate nuclear factor kappa B (NF-κB) signaling and other
signaling pathways, thereby potentially influencing the redox balance in the hepatic cells and generating numerous IFs and reactive oxygen species (ROS), which accelerate the development of liver cirrhosis [11].

In summary, compared to patients with liver cirrhosis, the prognosis and treatment of patients with liver cirrhosis associated with T2DM require more attention. The association between liver cirrhosis and T2DM has become a matter of concern because of the increased prevalence of their co-occurrence, which has led to a few studies being conducted in this area. However, there are limited studies on the common pathogenesis of the two diseases. Therefore, the association between the two diseases and the identification of potential treatment options require further investigation.

In recent years, data integration using bioinformatic analysis has become an important method in the field of Life Sciences [12]. In the present study, the common hub genes associated with liver cirrhosis and T2DM were obtained via the Cytoscape software. Based on the information obtained regarding the hub genes, the common molecular pathogenesis and potential therapeutic drugs were elucidated. These results provide new insights into the common pathogenesis of liver cirrhosis and T2DM.

Materials And Methods

1. Collection of data

This study is based on Gene Expression Omnibus (GEO) datasets (http://www.ncbi.nlm.nih.gov/geo), which are used for relevant data collection [13]. Two datasets were obtained by using “liver cirrhosis” and “T2DM” as the Medical Subject Headings (MeSH) terms and “human beings” as the sample type. The two datasets, GSE14323 and GSE29221, were selected and analyzed for this study. GSE14323 included data on 12 normal liver tissue samples from healthy patients and 12 liver tissue samples from patients with liver cirrhosis. GSE29221 included data on 12 normal bone tissue samples from healthy patients and 12 skeletal muscle tissue samples from patients with T2DM.

2. Identification of differentially expressed genes (DEGs)

GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/) is an interactive analysis tool based on the GEOquery package and the Limma package of R language. It is used to identify DEGs in tissue samples obtained from normal and diseased patient groups [14]. In this study, p < 0.05 and |log FC| > 1 were used to screen the DEGs. Thereafter, the intersection between the two gene sets was formed via the online Venn drawing website (http://bioinformatics.psb.ugent.be/webtools/Venn/) to simultaneously obtain co-DEGs involved in liver cirrhosis and T2DM. Furthermore, a volcano map was created using the R statistical program (R Core Team, version 3.6.3), and the DEGs with the |logFC| > 1 were labeled as upregulated, whereas the DEGs with the |logFC| < −1 were labeled as downregulated.

3. Functional enrichment analyses of DEGs

Functional enrichment analyses were performed via Database for Annotation, Visualization, and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/home.jsp) website, with Gene Ontology (GO) and
Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses used in combination. GO is a database of biological process (BP), molecular function, and cellular component, among which BP is relevant for this study [15]. KEGG is a database dedicated to the analysis of gene-related pathways [16]. Finally, a visualization analysis was conducted using the GOplot package in the R language.

4. Construction of protein–protein interaction (PPI) network and analysis of gene modules

The PPI network is based on the search tool Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) (https://string-db.org/), which retrieves interacting genes. The STRING showed how intracellular proteins in the tissue samples interact with each other [17]. Subsequently, the PPI networks were established from the datasets on DEGs via Cytoscape app (http://www.cytoscape.org), an image software, with interaction scores greater than medium confidence (0.400) [18]. Meanwhile, the plug-in molecular complex detection technology (MCODE) of the Cytoscape app was used to analyze key functional gene modules, with all parameters set to default values. Thereafter, the gene modules were analyzed via GO and KEGG pathway enrichment analyses [19].

5. Selection and validation of hub genes

The Cytoscape app was used to screen and select 30 hub genes that have the same expression levels in each of the six algorithms: the betweenness centrality (BC) algorithm of the Cytoscape plug-in, the Cytoscape app for network centrality analysis (CytoNCA), and the EPC, MNC, MCC, Degree, and DMNC algorithms of the Cytoscape plug-in, CytoHubba, respectively. Thereafter, the intersection was formed using the Venn diagram to select final hub genes, which were analyzed using the GO and KEGG analyses [20, 21]. Finally, the T2DM dataset GSE7014 and the self-tested liver cirrhosis dataset were used as the verification datasets to verify the final hub genes, which were then visualized as a histogram.

6. Prediction of transcription factors (TFs)

TFs may be useful to investigate the effects of the expression of hub genes in vivo, either physiologically or pathologically. NetworkAnalyst (https://www.networkanalyst.ca/) was used to predict the TFs capable of regulating the hub genes and to construct an interaction network for the hub genes [22].

7. Prediction of drugs

By using Drug SIGnatures DataBase (DSigDB) (http://dsigdb.tanlab.org/DSigDBv1.0/), targeted drugs with high potential associated with the hub genes were predicted [23]. The DSigDB was accessed via the Enrichr platform, the data on hub genes were uploaded, and candidate drug data were obtained with p < 0.05 as a screening standard, and finally, the top 20 drug molecules were selected in sequence from lesser to greater P values.

Results

1. Identification of DEGs
Microarray gene expression datasets GSE14323 and GSE29221 include data on the samples from patients with liver cirrhosis and T2DM, respectively. Based on GSE14323, 1419 DEGs (1067 upregulated, 352 downregulated) were identified; based on GSE29221, 1332 DEGs (1035 upregulated, 296 downregulated) were identified (Fig. 1A, B). Thereafter, the Venn diagram crossover was obtained, which yielded 175 co-DEGs (Fig. 1C).

2. GO and KEGG pathway enrichment analyses

The DAVID was used to analyze 175 co-DEGs derived from the GO and KEGG pathway enrichment analyses. In total, 106 co-DEGs based on BP were screened and pathways associated with 16 co-DEGs, with \( p < 0.05 \) as the criterion, were identified using the KEGG analysis. The results of the GO analysis revealed that these genes were mainly involved in cell adhesion, ECM organization, and collagen fibril organization (Fig. 2A). The KEGG pathway showed that the genes were mainly involved in four main pathways: proteoglycan cancer pathway, protein digestion and absorption, focal adhesions, and P13K/Akt signaling pathway (Fig. 2B).

3. Construction of PPI network

Based on the 175 co-DEGs obtained from the GO and KEGG pathway enrichment analyses, by using STRING, the PPI visualization network was created, in the Cytoscape, with a combined score of \( >0.4 \). A total of 134 nodes were involved in the grid. Subsequently, the hub genes were screened based on the BC developed by the CytoNCA (Fig. 3A). The high correlation sub-network of MCODE was selected by the plug-in of the Cytoscape. Three closely-related gene modules were obtained, and the functional enrichment analyses of the gene modules were also conducted (Fig. 3B, C, D). The GO analysis showed that the gene modules were mainly enriched in ECM organization and cell adhesion (Fig. 4A). The KEGG analysis revealed that these genes were involved in cellular metabolism, cellular inflammation, and the regulation of the cell cycle (Fig. 4B).

4. Selection and functional enrichment analyses of hub genes

The top 30 gene candidates were screened by employing five methods of topological analysis from the CytoHubba plug-in (Table 1). Based on the intersection of the top 30 genes ranked by BC and from the aforementioned five algorithms, 8 hub genes were found to be crossed, which included \textit{SPARC}, \textit{COL4A2}, \textit{THBS1}, \textit{LUM}, \textit{TIMP3}, \textit{COL3A1}, \textit{IGFBP7}, and \textit{FSTL1} (Fig. 5B). The functional enrichment analyses showed that the hub genes were involved in the proteoglycan cancer pathway, protein digestion and absorption, ECM–receptor interaction pathway, AGE–RAGE signaling pathway, cytokine response pathway, and positive regulation of TGF-\( \beta \)1 (Fig. 5A).

5. Verification of expression of hub genes

The T2DM dataset GSE7014 from GEO and the self-tested liver cirrhosis dataset were selected as the validated datasets to verify the reliability of the expression levels of the hub genes. The results showed that compared with the normal tissue samples, the expression levels of the hub genes were upregulated.
in skeletal muscle tissue samples of patients with T2DM and liver tissue samples of patients with liver cirrhosis (Fig. 5C).

6. Prediction of TFs

The prediction of TFs of the hub genes was performed via NetworkAnalyst. The results of the prediction showed that \textit{THBS1}, \textit{COL4A2}, and \textit{TIMP3} were regulated by 29 TFs; \textit{SMC3} was capable of simultaneously regulating \textit{THBS1} and \textit{COL4A2}; and \textit{BHLHE40} was capable of simultaneously regulating \textit{THBS1}, \textit{COL4A2}, and \textit{TIMP3} (Fig. 6).

7. Prediction of drugs

To screen the potential intervention drugs, the DSigDB Enrichr platform was used to search the top 20 drugs, which were screened with \( p < 0.05 \) as the screening standard (Table 2). The results showed that retinoic acid (CTD 00006918), cyclosporin A (CTD 00007121), and cytarabine (CTD 00005743) were the potential targeted drugs that interacted with most of the hub genes.

**Discussion**

The incidence of liver cirrhosis associated with T2DM has been increasing. Compared to patients with liver cirrhosis, patients with T2DM as a comorbidity have more rapidly progressing liver cirrhosis, which may even result in liver cancer in severe cases \([9, 10]\). In the present study, compared to patients with liver cirrhosis, patients with liver cirrhosis associated with T2DM had considerably elevated basal C-polypeptide concentration and required higher insulin dosage. Therefore, insulin resistance is more severe in patients with liver cirrhosis associated with T2DM. Considerably high insulin resistance is the main factor affecting the development and prognosis of liver cirrhosis associated with T2DM \([24]\). However, there are no clinical treatments that may be considered the best therapy for patients with liver cirrhosis associated with T2DM. Therefore, the investigation of the common pathogenesis of liver cirrhosis and T2DM will help provide more information using which a potential targeted drug for the treatment of the co-occurring diseases may be obtained.

In this study, a combination of bioinformatic analyses was used to analyze eight common hub genes, including \textit{SPARC}, \textit{COL4A2}, \textit{THBS1}, \textit{LUM}, \textit{TIMP3}, \textit{COL3A1}, \textit{IGFBP7}, and \textit{FSTL1}, and associated TFs to form a complex interaction network to elucidate the pathogenesis of liver cirrhosis associated with T2DM. The etiology of both the diseases may be associated with the positive regulation of TNF-\(\beta\)1 and inflammatory cytokine response, according to the GO analysis wherein the genes were classified in terms of BP. IFs are one of the essential factors that induce T2DM by directly acting on pancreatic \(\beta\)-cells, thereby damaging cellular DNA. Its effects may trigger oxidative stress in pancreatic \(\beta\)-cells, which leads to apoptosis, resulting in the generation of numerous free radicals of oxygen that induce insulin resistance. As a result, the synthesis and secretion of insulin decrease \([25]\).
Inflammation is an important contributor to the development and progression of liver cirrhosis. An imbalance in the level of pro-inflammatory or anti-inflammatory factors may promote the transformation of HSCs and the activation of myofibroblasts, which contribute to accelerating the progression of liver cirrhosis [26].

TGF-β1 is involved in the regulation of various signaling pathways. It may trigger apoptosis of pancreatic β-cells and insulin resistance by the activation of P13K/Akt signaling pathway [27]. It may also participate in Smad3 signal transduction, thereby causing reduced insulin sensitivity and eventually affecting metabolic capacity [28]. Furthermore, TGF-β1 is a key factor in the development of liver cirrhosis; it stimulates the activation of HSCs and myofibroblasts and induces the production of ECM, thereby resulting in liver cirrhosis [29].

In the present study, the results of the analysis of molecular interactions suggested that the hub genes were predominantly enriched in the proteoglycan cancer pathway, proteolytic and absorbent protein pathway, ECM–receptor interaction pathway, and AGE–RAGE signaling pathway. Among these pathways, the AGE–RAGE signaling pathway is closely related to the occurrence of liver cirrhosis associated with T2DM. Dooley et al. suggested that the AGE–RAGE signaling pathway is associated with the myofibroblast transdifferentiation of rat HSCs and that RAGE may increase the level of TGF-β1 by activating protein kinase C [30].

In the present study, the bioinformatic analysis of TFs indicated that SMC3 and BHLHE40 were involved in simultaneously regulating THBS1 and COL4A2. THBS1 may lead to platelet aggregation, angiogenesis, and tumorigenesis. It may be an important risk factor for the occurrence of T2DM [31]. In addition, Kondou found that THBS1 is positively correlated with the occurrence of liver cirrhosis [32]. COL4A2 participates in the formation of ECM and type IV collagen. Dong demonstrated that COL4A2 is highly expressed in the subcutaneous adipose tissue of obese patients with T2DM [33]. However, the relationship between the expression of COL4A2 and the pathogenesis of T2DM is yet to be elucidated. Wei et al. analyzed data on multiple transcriptomes and revealed that COL4A2 is considerably associated with the progression of liver cirrhosis [34].

In this study, by using DSigDB, the drugs in the database were tested on the eight hub genes to identify the drugs with the potential to affect these genes. The most common drug molecules identified were retinoic acid (derivative of vitamin A), cyclosporin A, and cytarabine. Kang et al. discovered that retinoic acid markedly attenuates the cell toxicity induced by interleukin-1 (IL-1), interferon-γ, and other cytokines, resulting in the apoptosis of pancreatic β-cells [35]. Retinoic acid may also regulate the expression of IL-17A, a vital inflammatory factor associated with liver cirrhosis [36]. However, studies have shown that long-term use of cyclosporine A and cytarabine may clinically damage pancreatic β-cells, thereby leading to the onset of clinical T2DM and drug-induced liver injury [37, 38]. In summary, retinoic acid may be the drug for the treatment of liver cirrhosis associated with T2DM.

Therefore, the present study explored and analyzed the common DEGs, hub genes, and TFs between liver cirrhosis and T2DM. This information may help to further illustrate the pathogenesis of liver cirrhosis
associated with T2DM. However, the limitations of this study include the small sample size of the validation datasets because of which the validation datasets showed an upward trend, but there was a considerable difference between the trained group and the validation group. Therefore, further research is required to investigate the function of the hub genes in vitro. Additionally, clinical studies on the safety and efficacy of the candidate drugs must be conducted for further investigation.

Conclusion

By using a series of bioinformatic analyses, differential gene expression analysis was conducted between data on liver cirrhosis and T2DM. Various common etiological mechanisms are mediated by specific hub genes associated with liver cirrhosis and T2DM. Using retinoic acid as the treatment for liver cirrhosis associated with T2DM should be further verified through various experiments. In conclusion, the study results provide important information for further investigation of the molecular mechanisms involved in liver cirrhosis associated with T2DM.

Declarations

DATA AVAILABILITY STATEMENT

In this study, all data comes from publicly available databases and references to available data are included in the methodology section.

ETHICS STATEMENT

Not applicable.

AUTHOR CONTRIBUTIONS

YA J and D K. presented the idea and designed the whole outline of this article. ZY X wrote the original draft. MQ Y and LC Y contributed to figure and table preparation. The final version was revised by ZY X, MQ Y, LC Y, W Z and PJ L. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We thank the GEO, STRING, miRWalk, TargetScan and miRDB databases for their generous sharing of large amounts of data. We apologize for any omission of citations and references due to space limitations.

The study was supported by the Anti-aging Research Center of Wuhan University Education Development Foundation (No. 2002330) and the National Natural Science Foundation of China (no. 81972673).

References


15. Reijnders M, Waterhouse RM. Summary Visualizations of Gene Ontology Terms With GO-Figure! Front Bioinform. 2021;1:638255.


Tables

Tables 1-2 is available in the Supplementary Files section.

Figures

Volcano and Venn diagrams: (A) The volcano map of GSE14323. (B) The volcano map of GSE29221. Upregulated genes are marked in light red and downregulated genes are marked in light blue. (C) The two datasets show an overlap of 175 DEGs.
Figure 2

DEGs enrichment analysis results: The enrichment analysis results of (A) the KEGG and (B) the GO analyses. Adjusted P value < 0.05 was considered significant.
Figure 3

PPI network and significant gene modules: (A) The PPI network diagram. Red indicates genes with a high betweenness value and yellow indicates genes with a low betweenness value. (B–D) Three significant clustering gene modules.
Figure 4

Enrichment analysis of gene modules: The enrichment analysis results of (A) the GO and (B) the KEGG analyses. Adjusted P value < 0.05 was considered significant.
Figure 5

Venn diagram, enrichment analysis of hub genes, and the expression level of hub genes: (A) Results of the KEGG and the GO pathway enrichment analyses. Red indicates GO. Blue indicates the results of the KEGG analysis. (B) The six calculated results show an overlap of eight hub genes. (C) The highlighted red node represents GSE29221, blue represents GSE7014, green represents GSE14323, and dark blue indicates the data of self-tested liver cirrhosis.
Figure 6

TF–gene interaction network for hub genes: The highlighted red node represents the hub genes and other nodes represent TF–gene interactions.

Supplementary Files

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

- Table1.jpg
- Table2.jpg