Published online 2022 July 27.

Research Article

Identification of Shared Gene Signatures in Different Stages of Nonalcoholic Fatty Liver Disease Using Integrated Microarray Datasets

B. Fatemeh Nobakht M. Gh¹, Yazdan Hasani Nourian¹ and Masoud Arabfard ¹, *

Corresponding author: Chemical Injuries Research Center, Systems Biology and Poisonings Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran. Email: arabfard@gmail.com

Received 2022 January 09; Revised 2022 May 22; Accepted 2022 June 08.

Abstract

Background: Nonalcoholic fatty liver disease (NAFLD) is the most common type of chronic liver disease worldwide. Left untreated, it can be a risk factor for developing cirrhosis or hepatocellular carcinoma (HCC). Although experts have made many efforts to find the underlying mechanisms of NAFLD, they remain a mystery.

Objectives: This study aimed to distinguish common gene signatures and pathways in the human liver during NAFLD progression through systems biology.

Methods: In this study, the researchers selected three microarray datasets, GSE48452, GSE63067, and GSE89632, from the NCBI GEO database to explore differentially expressed genes (DEGs) among healthy controls, simple steatosis, and nonalcoholic steatohepatitis (NASH) patients. Furthermore, protein-protein interaction (PPI) networks and pathway enrichment analyses were used to detect common genes and biological pathways in different stages of NAFLD.

Results: The current study included 45 healthy participants, 36 simple steatosis patients, and 46 NASH patients. Common genes for NAFLD progression were Chi3L1, ICAM1, MT1A, MT1H, ABCB11, ACOT1, CYP2C9, HSP90B1, and CPB2, which are involved in inflammation and oxidative stress pathways.

Conclusions: The present study investigated the shared vital genes and pathways between different stages of NAFLD, which may facilitate understanding NAFLD mechanisms and identifying potential therapeutic targets in this disease.

Keywords: Microarray, NASH, NAFLD, Protein-protein Interaction Network, Systems Biology

1. Background

Nonalcoholic fatty liver disease (NAFLD) is the most prevalent liver disease in the world, associated with lifestyle and obesity. It may progress to severe hepatic disorders, including simple steatosis, nonalcoholic steatohepatitis (NASH), fibrosis, cirrhosis, and hepatocellular carcinoma (HCC) (1). Nonalcoholic fatty liver disease affects around 90% of obese patients and 25% of people worldwide (2). Nonalcoholic steatohepatitis might lead to liver cirrhosis and, subsequently, HCC. It is the second most common risk factor for HCC in patients undergoing liver transplantation in the USA (3, 4). Moreover, NAFLD is associated with other comorbidities, including metabolic syndrome, type 2 diabetes, cardiovascular disorders, and extrahepatic cancers (5).

The NAFLD pathogenesis is associated with genetic and environmental factors. Multiple tissues, including the adipose tissue, hypothalamus, intestine, and liver, are involved in the pathogenesis of NAFLD. This multifactorial disease is related to the altered bowel microbiota, increased oxidative stress (OS), and dysregulated adipokines and hormones (2, 6). Given the complexities of this disease and limited data on the underlying molecular mechanisms involved in its progression, there are no approved remedies for NAFLD treatment despite lots of efforts to detect pharmacological therapies (7, 8). Since this disorder results from lifestyle, diet, and physical activity, alterations in this regard may be helpful but not easy to achieve. Consequently, novel and effective treatments are controversial. The NAFLD diagnosis and staging are typically established by hepatic biopsy. Hence, noninvasive biomarkers that might recognize patients at risk of progression to the advanced stage are immediately necessary and helpful for drug development.

2. Objectives

Identifying gene-specific expression panels could improve our understanding of the pathogenesis and pro-

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.

gression mechanisms of the disease or drug assessment. To the best of our knowledge, two studies have performed protein-protein interaction (PPI) networks and pathway enrichment analyses to screen differentially expressed genes (DEGs) in different stages of NAFLD compared with controls (9, 10). Wang et al. used one microarray dataset, GSE48452, to detect DEGs among control, healthy obese, steatosis, and NASH people (9). Another study used six datasets (GSE48452, GSE66676, GSE72756, GSE63067, GSE89632, and GSE107231) to find DEGs in NAFLD patients and controls. However, the GSE72756 dataset used in their study (in which normal liver tissues from NAFLD patients were used as controls) and the GSE107231 dataset do not contain three stages of NAFLD. Besides, the GSE66676 dataset contains liver tissue samples from adolescents (13.4 - 19.8 years) (10), while NAFLD in adolescents is different from that in adults (11). Thus, the present study further analyzed three microarray datasets, including GSE48452, GSE63067, and GSE89632, to explore DEGs among healthy controls, simple steatosis, and NASH patients. Furthermore, PPI networks and pathway enrichment analyses were conducted to find genes/proteins and biological pathways shared between different stages of NAFLD.

3. Methods

3.1. Data Processing

A literature search was carried out in gene expression omnibus (GEO) electronic database up to March 2021. The chosen keywords included "NAFLD," "nonalcoholic fatty liver disease," "simple steatosis," "NASH," and "nonalcoholic steatohepatitis."

Microarray studies were conducted on adults of any gender or ethnicity who met the inclusion criteria. The exclusion criteria included the studies that were randomized controlled trials, had been published before 2013, lacked a control group, reported gene expression from tissues other than the liver, and were animal, in vitro, or ex vivo research. Finally, three datasets, GSE63067, GSE89632, and GSE48452, were selected from the NCBI GEO database for the differential detection of genes (12-14). These three datasets followed the same grouping of samples.

GSE63067 was determined in samples from three groups:

(1) Nine patients diagnosed with early-stage NASH (NA),

(2) Two patients diagnosed with nonalcoholic hepatic steatosis (NA), and

(3) Seven controls (NA) with normal liver function and histology.

GSE89632 was determined in samples from three groups:

(1) Nineteen patients diagnosed with early-stage NASH (10 females and nine males with a mean age of 43.5 years and BMI of 31.9),

(2) Twenty patients diagnosed with nonalcoholic hepatic steatosis (six females and 14 males with a mean age of 44.7 years and BMI of 28.6), and

(3) Twenty-four controls (13 females and 11 males with a mean age of 37.2 years and BMI of 26.1) with normal liver function and histology.

GSE48452 was determined in samples from three groups:

(1) Eighteen patients diagnosed with early-stage NASH (14 females and four males with a mean age of 45.4 years and BMI of 45.9),

(2) Fourteen patients diagnosed with nonalcoholic hepatic steatosis (10 females and four males with a mean age of 41.6 years and BMI of 48.2), and

(3) Fourteen controls (nine females and five males with a mean age of 51.8 years and BMI of 25.14) with normal liver function and histology.

The selected datasets included microarray data processed by the LIMMA package in R, followed by performing quantile normalization and quality control and identifying and removing the noisy data and outliers by hierarchical clustering. The calculation of DEGs was done by the eBays Method. Among different data analysis methods, such as integration and fusion, the integration method was selected, involving a separate analysis of each dataset. In this method, each dataset is analyzed without any specific dependency, and finally, the results are integrated at the decision level. The DEGs with $|FC| \ge 2$ and P-value \le 0.05 were selected as the final DEG list for all three data groups. Data were collected from three groups of healthy, steatosis, and NASH participants (Table 1).

For both groups in each dataset, a differential list was obtained. The lists were DEGs related to two healthy and steatosis groups named SH, healthy and NASH named NH, and steatosis and NASH named NS, respectively.

3.2. Physical Protein Interaction Network

To plot the physical interaction network, the researchers combined all groups (i.e., SH, NS, and NH) in the dataset, and then the tool https://string-db.org was used. Each list's physical and experimental interactions were examined, and the interaction network was extracted. Moreover, to study the deeper connections between gene interactions in liver tissue, the researchers entered all gene lists as input to the HIPPIE tool (http://cbdm-01.zdv.unimainz.de). The interaction data were extracted by filtering interactions related to liver tissue. Finally, all the extracted networks were entered into the Cytoscape

Table 1. Characteristics of the Included Gene Expression Omnibus Datasets				
GSE_ID	Participants	Tissues	Platform	Year
GSE63067	9 NASH,2 Steatosis, 7 Healthy	Liver	GPL570	2014
GSE89632	19 NASH, 20 Steatosis, 24 Healthy	Liver	GPL14951	2016
GSE48452	18 NASH, 14 Steatosis, 14 Healthy	Liver	GPL11532	2013

(https://cytoscape.org), and Network Analyzer identified the hub genes (with the highest degree).

3.3. Similarity Analysis

In order to investigate the similarities between the three processing sets, the study selected identical groups in pairs and overlapping genes. Also, since the enrichment analysis was performed for each group, the overlaps in each analysis were examined.

3.4. Enrichment Analysis

The present study applied Enrichr (maayanlab.cloud) to find the biological pathways and the gene ontology involved in the gene list. The data were used to investigate the biological similarities of the gene sets in the gene list.

4. Results

4.1. Identification of DEGs in NAFLD

All three microarray datasets were first analyzed separately. The results of data normalization and clustering can be found in Supplementary File 1. The DEGs were screened using the "LIMMA" package in R software according to the data processing section criteria. The DEG lists of the three microarrays are shown in Supplementary Files 2, 3, and 4 for GSE48452, GSE63067, and GSE89632, respectively. The results of the three microarrays are presented in Table 2. Besides, the overlapping genes in this step are shown in Figure 1. No genes were shared between the three groups in the GSE48452 dataset (A). In the GSE63067 dataset (B), two genes were found to be shared between the three groups, Intercellular Adhesion Molecule 1 (ICAM1) and Paternally Expressed Gene-10 (PEG10). In the GSE89632 dataset (C), three genes were found to be shared between the three groups, Flavin-Containing Monooxygenases 1 (FMO1), PEG10, and Metallothionein 1A (MT1A).

By examining three datasets to find the list of shared genes between all the three sets, the researchers came across five genes, CHI3L1 (Chitinase 3 Like 1), GFBP2 (Insulinlike growth factor (IGF) binding protein 2), NRG1 (Neuregulin 1), PEG10, and FADS2 (Fatty Acid Desaturase 2) (Figure 2).

able 2. Results of Differentially Expressed Genes Calculation						
GSE_ID	GS	GSE63067 GSE89632		GSE48452		
	UPs	DOWNs	UPs	DOWNs	UPs	DOWNs
SH	283	66	186	366	11	5
NS	229	59	17	10	16	5
NH	120	19	244	321	24	10

4.2. Overlapping Genes in Matched Groups

The DEGs of each processed group are used to identify the similarities between the two sets. We plotted the Venn diagram for the overlapping genes in the three groups (Figure 3). IGFBP2 (insulin-like growth factor-binding proteins) and PEG10 were the shared genes in the list of NH (Figure 3A). There were no shared genes in the list of NS (Figure 3B). There was only the PEG10 gene as the shared gene in the list of SH (Figure 3C).

4.3. PPI Networks

Gene interactions for each group indicated a significant number of interactions; therefore, for each group, the genes with the highest degree of interaction can be considered significant points in the network. Three global interaction networks were extracted from three gene lists that can be seen in Figures 4A - C (Red color indicates the nodes with the highest degree in this network, and as the color tends to yellow, the degree of nodes decreases).

Besides, to study the genetic connections in a more detailed way in the liver tissue, we examined these connections specifically. These liver tissue connections suggest that hub genes can play a crucial role in disease due to their location and role in the interaction network. The Signal Transducer and Activator of Transcription 3 (STAT3) and JUN genes, as the gene with the highest degree, play a crucial role in liver tissue.

Using the network analyzer plugin in the Cytoscape, we analyzed the hub genes topologically, and the top hub genes with the highest degrees were extracted using the criteria. In the interaction networks, the gene with the highest degree has the most connections and plays an intermediate role between other genes; it is known as a bottleneck. In case of any problem with the network, this in-



Figure 1. The Venn diagram for overlapping genes in each dataset (each color represents a group)



Figure 2. The Venn diagram for overlapping datasets (each color represents a dataset)

termediate gene can help understand the network's structure.

4.4. Enrichment Analysis

The biological pathways and processes shared between datasets are listed in Figure 5, which is the output of UpSet (https://asntech.shinyapps.io). We obtained three enrichment lists by examining the overlap between the ontologies of genes and the shared biological pathways in each group. Tables 3 - 5 show all enrichment list subscriptions.

The most important terms of gene ontology in the biological process are signal transduction, immune response, and anti-apoptosis in the NH group, metabolism and energy pathways in the NS group, and cell communication, signal transduction, and immune response in the SH group. They are also parts of the pathways that significantly overlap in the study groups.

Although the results obtained from data fusion are consistent with individual data, it is noteworthy that reviewing the results in different datasets can be very helpful in confirming the results.

5. Discussion

In the present study, PPI networks were produced based on differential gene expression provided from liver



Figure 3. The Venn diagram for overlapping genes in each combined group (each color represents a dataset)

Table 3. Top 10 Biological Pathways for the NH Group		
Biological Pathways	P-Value (Hypergeometric Test)	Bonferroni Method
TNF signaling pathway	1.18E-09	3.10E-07
Viral protein interaction with cytokine and cytokine receptor	4.44E-08	3.875E-06
AGE-RAGE signaling pathway in diabetic complications	2.63E-07	0.00001722
Cytokine-cytokine receptor interaction	8.41E-09	1.102E-06
IL-17 signaling pathway	0.00001786	0.0007798
JAK-STAT signaling pathway	2.835E-06	0.0001485
Malaria	0.0002654	0.005349
Chagas disease	0.00004321	0.001258
C-type lectin receptor signaling pathway	0.00005311	0.001391
Osteoclast differentiation	0.00002855	0.001068

Table 4. Top 10 Biological Pathways for the NS Group				
Biological Pathways	P-Value (Hypergeometric Test)	Bonferroni Method		
Fatty acid elongation	0.0009678	0.01901		
Lipid and atherosclerosis	3.302E-06	0.000908		
Proteoglycans in cancer	8.957E-06	0.001232		
Fructose and mannose metabolism	0.002087	0.02733		
Bacterial invasion of epithelial cells	0.0002898	0.01139		
Leishmaniasis	0.0002898	0.01139		
IL-17 signaling pathway	0.0001709	0.01139		
Tight junction	0.0001256	0.01139		
Th17 cell differentiation	0.0004155	0.01217		
Neomycin, kanamycin, and gentamicin biosynthesis	0.0803	0.2389		



Figure 4. The interaction network of genes shared in the groups



B, Biological Process



Figure 5. The common biological pathways and biological processes used for enrichment analysis

Table 5. Top 10 Biological Pathways for the SH Group				
Biological Pathways	P-Value (Hypergeometric Test)	Bonferroni Method		
IL-17 signaling pathway	4.33E-11	1.18E-08		
Leishmaniasis	2.58E-09	1.17E-07		
TNF signaling pathway	2.64E-10	3.58E-08		
Osteoclast differentiation	6.54E-10	4.45E-08		
Malaria	8.67E-07	0.00002621		
Rheumatoid arthritis	6.07E-08	0.000002359		
Transcriptional misregulation in cancer	6.43E-10	4.45E-08		
Viral protein interaction with cytokine and cytokine receptor	1.93E-07	0.00006571		
Legionellosis	0.000003845	0.0000747		
Inflammatory bowel disease	0.000002811	0.00005882		

tissue of simple steatosis and NASH patients compared with healthy controls. Centrality and modularity analyses of PPI networks helped us detect common genes and pathways involved in these diseases. In order to evaluate the results based on the at-hand evidence, we used the Human Protein Atlas Database, and the results were matched with the list of approved proteins. The results revealed that many proteins had been examined in the experiments.

Our findings suggested three overlapping genes in three groups (Chi3L1, ICAM1, and MT1A), whereas two groups shared several genes. A biological pathway was also shared by all groups (IL-7). Herein, these genes' roles in various disease phases are explained. Intercellular Adhesion Molecule 1 (ICAM1) is a glycoprotein located on the cell surface and expressed in some cells such as liver, endothelial, epithelial, and hematopoietic cells. The overexpression of ICAM1 may occur by some inflammatory cytokines, including IL-1 and TNF- α (15). Ito et al. suggested that serum ICAM1 might be a diagnostic marker for NASH as its concentration increases in NASH patients compared with healthy individuals, which aligns with our finding. Moreover, there was a positive correlation between serum ICAM1 levels and the severity of the liver injury and inflammation (15). Another study showed that hepatic expression of ICAM1 is significantly higher in NASH patients than in simple steatosis and normal controls. Moreover, a significant correlation was observed between the steatosis degree and hepatic ICAM1 expression (16). It has been hypothesized that overexpression of hepatic ICAM1 enhances leukocyte adherence and increases Reactive Oxygen Species (ROS) production, indicating the severity of necroinflammatory activity (17).

The molecular processes linked with steatosis development and disease progression were investigated in this study. Our findings presented the most critical genes in the liver tissue associated with the progression of different stages of NAFLD. Bile acids may play a role in the development and treatment of NAFLD/NASH by signaling through their designated nuclear receptor (i.e., farnesoid X receptor)(FXR; NR1H4) as a critical regulator of glucose and lipid metabolism, as well as inflammation (18, 19). The bile salt export pump (BSEP and ABCB11) excretes bile acids from hepatocytes into the gallbladder. In humans, there has been evidence of links between BSEP polymorphisms and elevated serum triglycerides (TG), cholesterol, and obesity (20). When mice overexpressing BSEP are fed a lithogenic diet or a methionine choline-deficient (MCD) diet, they develop moderate hepatic steatosis (21). Diminished expression of ATP Binding Cassette Subfamily B Member 11 (ABCB11), and therefore reduced BSEP pump, were also found to be one of the causes of steatosis progression to NASH, according to the findings of Okushin et al. (22). In fact, lowering the expression of this pump causes an increase in bile acids in liver cells, resulting in cell damage (22). Our findings also revealed that the expression of the ABCB11 gene in simple steatosis was decreased compared to healthy persons, suggesting that this drop in expression could be one of the crucial elements in the development and progression of fatty liver disease. These findings were also confirmed in transthyretin-ABCB11 (TTR-ABCB11) mice, which have much greater levels of ABCB11 compared to the wild ones. The findings of investigations suggest that this type of mouse develops hepatic steatosis less than a regular one (23, 24).

At complexes I and III of the electron transport chain, mitochondrial Fatty Acid (FA) oxidation creates a small amount of ROS (25). Antioxidant activity counteracts this ROS generation, protecting the mitochondria from OS (26). On the other hand, high membrane potential might cause a substantial amount of ROS by increasing FA oxidation, enhancing antioxidant capacity, and resulting in oxidative stress (27). Controlling the rate at which FAs are oxidized in a cell is essential for reducing ROS and oxidative stress that might arise with an increase in oxidation. Acyl-CoA thioesterase 1 (ACOT1) catalyzes the conversion of acyl-CoA to FAs and CoA, which subsequently limits OS (28). In mice with diabetic cardiomyopathy, overexpression of ACOT1 in cardiomyocytes decreases FA oxidation and ROS generation (29). These findings imply that ACOT1 controls FAs destined for oxidative pathways, presumably by eliminating acyl-CoAs as substrates. Also, ACOT1 may protect the liver from the harmful consequences of excessive FA oxidation by delaying FA oxidation (28). Moreover, PPAR α has been shown to have anti-inflammatory and protective effects in inflammatory disorders of the liver, and ACOT1 leads to the supply of ligands required for its activation. Consequently, enhanced ACOT1 expression may help prevent steatosis and the development of steatosis to other inflammatory liver disorders (28, 30). Our study also indicated that the expression of the ACOT1 gene is lower in steatosis than in healthy conditions, suggesting that this decrease in expression might be one of the causes of FA metabolism disturbance and the beginning of the fatty liver disease.

Chitinase 3-like 1 (Chi3L1), a newly identified human glycoprotein, has a role in inflammation, tissue remodeling, and visceral fat storage (31). Increased Chi3L1 gene expression is a factor in increasing insulin resistance and, thus, fat accumulation in the liver. In fact, mice with the Chi3L1 gene knockout are more insulin sensitive, which slows the progression of the disease to NASH (32). Chi3L1 is generated by various inflammatory cells and can enhance the production of inflammatory cytokines such as IL-13, IL-6, IL1, and IFN- γ (33). Chi3L1 serum levels have been linked to different phases of liver fibrosis in recent clinical trials (34, 35). According to preliminary research, Chi3L1 gene expression can discriminate between isolated simple steatosis and NASH by whole genome RNA sequencing compared to existing biomarker scoring systems (RNA-seq) (32). Our results also showed that the expression of the Chi3L1 gene is higher in patients in the NASH stage of fatty liver disease than in simple steatosis. Therefore, this gene can be used as a biomarker to detect the progression of NAFLD.

As known, CYP1, CYP2, and CYP3 are members of the cytochrome P450 (P450) enzyme family that are responsible for the metabolism of roughly 75% of all therapeutically relevant medications. With the rising incidence of NAFLD, individuals with this condition are expected to be a new group at risk for changes in these critical drugmetabolizing enzymes (36). One of the CYPs important in the metabolism of endogenous and exogenous substances is CYP2C9. Fisher et al. (36) reported that hepatic mRNA and enzymatic activity of CYP2C9 increased with progressive stages of NAFLD, per our findings. CYP2C9 activity in-

Hepat Mon. 2022; 22(1):e122362.

creases in hypoxic conditions (37). Studies have shown that hypoxia is associated with NAFLD severity and provides a plausible reason for the enhanced CYP2C9 activity in human and animal models (36, 38). Hypoxia can also cause inflammation, a feature of NAFLD (39).

Sterol regulatory element-binding proteins (SREBPs) are critical regulators of lipogenesis, regulating the expression of genes involved in fatty acid production and uptake, as well as cholesterol and phospholipid uptake (40). Heat shock protein 90 (HSP90) is a protein family that has a role in various pathophysiological events. These proteins, particularly HSP90B1, are also involved in lipid homeostasis via the SREBP pathway (41, 42). Previous studies indicated that HSP90 gene expression increases in the inflammatory stages of NAFLD disease, and inhibition of this protein by factors such as 17-AAG could reduce disease progression (43). These findings were also confirmed in our research, suggesting that in the NASH stage of the disease, the expression of HSP90 family proteins, particularly HSP90B1 and HSP90B1, increases compared to simple steatosis.

Another hub gene, MT1A, is a family of Metallothioneins with a high affinity for binding essential and toxic metals such as zinc and cadmium, respectively (44, 45). The biological functions of MT1A are protection against toxic metals, defense against OS, and protection against cytotoxicity and genotoxicity (45-47). Additionally, MT1A has an indispensable role in the regeneration of hepatocytes (48). In line with our results, Arendt et al. demonstrated the downregulation of MT1A gene expression in patients with simple steatosis compared to healthy controls; they observed a further decrease in NASH patients compared to patients with simple steatosis (13).

Carboxypeptidase B2 (CPB2) is another protein discovered in both databases. The liver produces plasma carboxypeptidase, which is transformed into the active enzyme by thrombin (49). To the best of our knowledge, there is not enough data on its link to human liver diseases. As a result, further research on this protein in different stages of liver illnesses is recommended.

One biological pathway, IL-7, was shared by all groups, as tabulated in Tables 3-5. T-cell development has been demonstrated to be dependent on IL-7 (50). It has been suggested that IL-7 and its receptor (CD127) have a role in NAFLD development. In NAFLD, serum IL-7 levels were inversely linked with fibrosis (51). In addition, the liver is a crucial producer of IL-7 in response to TLR stimulation. The TLR signaling in Kupffer cells drives the transition of simple steatosis to NASH (52). The gut microbiome can produce TLR ligands. TLR ligands can induce the production of proinflammatory cytokines in liver cells. Using probiotics or prebiotics to normalize the gut flora is a viable therapy option for NAFLD (53).

The strength of our study is that our findings are based on a combination of three datasets, and the genes discussed are those found to be shared in all three datasets. However, this study also has some limitations, including lacking an experimental section. For future studies, we recommend that the researchers carefully evaluate the expression of the proposed genes and proteins involved in developing NAFLD by experimental methods such as qRT-PCR.

This study revealed unique vital genes found in both simple steatosis and NASH. After further verification, the shared genes (Chi3L1, ICAM1, MT1A, MT1H, ABCB11, ACOT1, CYP2C9, HSP90B1, and CPB2) could be used as biomarkers of NAFLD progression and a putative candidate for therapeutic targets in NAFLD. Moreover, pathway analysis found that DEGs significantly enriched in several pathways are involved in inflammation and oxidative stress pathways.

Supplementary Material

Supplementary material(s) is available here [To read supplementary materials, please refer to the journal website and open PDF/HTML].

Footnotes

Authors' Contribution: MA designed the project, conducted bioinformatics studies, and performed data analysis. BFNMG and YHN participated in designing the project, contributed to supervision and data analysis, and wrote the manuscript. All authors read and approved the final manuscript.

Conflict of Interests: All the authors declare no conflicts of interest.

Data Reproducibility: The datasets analyzed in the current study are available in the Figshare data repository (figshare.com/articles/dataset/NAFLD/14387666).

Funding/Support: The authors declare no sources of support or funding.

References

- 1. Polyzos SA, Mantzoros CS. Nonalcoholic fatty future disease. *Metabolism.* 2016;**65**(8):1007–16. doi: 10.1016/j.metabol.2015.12.009. [PubMed: 26805015].
- Polyzos SA, Kountouras J, Mantzoros CS. Obesity and nonalcoholic fatty liver disease: From pathophysiology to therapeutics. *Metabolism*. 2019;92:82–97. doi: 10.1016/j.metabol.2018.11.014. [PubMed: 30502373].
- Wree A, Broderick L, Canbay A, Hoffman HM, Feldstein AE. From NAFLD to NASH to cirrhosis-new insights into disease mechanisms. *Nat Rev Gastroenterol Hepatol*. 2013;10(11):627-36. doi: 10.1038/nrgastro.2013.149. [PubMed: 23958599].

- Wong RJ, Aguilar M, Cheung R, Perumpail RB, Harrison SA, Younossi ZM, et al. Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. *Gastroenterology*. 2015;148(3):547-55. doi: 10.1053/j.gastro.2014.11.039. [PubMed: 25461851].
- Armstrong MJ, Adams LA, Canbay A, Syn WK. Extrahepatic complications of nonalcoholic fatty liver disease. *Hepatology*. 2014;59(3):1174– 97. doi: 10.1002/hep.26717. [PubMed: 24002776].
- Nobakht Motlagh Ghoochani BF, Ghafourpour M, Abdollahi F, Tavallaie S. Pro-oxidant antioxidant balance in patients with non-alcoholic fatty liver disease. *Gastroenterol Hepatol Bed Bench*. 2019;12(2):124–30. [PubMed: 31191836]. [PubMed Central: PMC6536017].
- Lavine JE, Schwimmer JB, Van Natta ML, Molleston JP, Murray KF, Rosenthal P, et al. Effect of vitamin E or metformin for treatment of nonalcoholic fatty liver disease in children and adolescents: the TONIC randomized controlled trial. *JAMA*. 2011;**305**(16):1659–68. doi: 10.1001/jama.2011.520. [PubMed: 21521847]. [PubMed Central: PMC3110082].
- Lindor KD, Kowdley KV, Heathcote EJ, Harrison ME, Jorgensen R, Angulo P, et al. Ursodeoxycholic acid for treatment of nonalcoholic steatohepatitis: results of a randomized trial. *Hepatology*. 2004;**39**(3):770-8. doi: 10.1002/hep.20092. [PubMed: 14999696].
- Wang R, Wang X, Zhuang L. Gene expression profiling reveals key genes and pathways related to the development of nonalcoholic fatty liver disease. *Ann Hepatol.* 2016;15(2):190–9. doi: 10.5604/16652681.1193709. [PubMed: 26845596].
- Jia X, Zhai T. Integrated Analysis of Multiple Microarray Studies to Identify Novel Gene Signatures in Non-alcoholic Fatty Liver Disease. Front Endocrinol (Lausanne). 2019;10:599. doi: 10.3389/fendo.2019.00599. [PubMed: 31551930]. [PubMed Central: PMC6736562].
- Fitzpatrick E, Dhawan A. Childhood and Adolescent Nonalcoholic Fatty Liver Disease: Is It Different from Adults? *J Clin Exp Hepatol.* 2019;9(6):716-22. doi: 10.1016/j.jceh.2019.05.005. [PubMed: 31889753]. [PubMed Central: PMC6926239].
- Frades I, Andreasson E, Mato JM, Alexandersson E, Matthiesen R, Martinez-Chantar ML. Integrative genomic signatures of hepatocellular carcinoma derived from nonalcoholic Fatty liver disease. *PLoS One*. 2015;**10**(5). e0124544. doi: 10.1371/journal.pone.0124544. [PubMed: 25993042]. [PubMed Central: PMC4439034].
- Arendt BM, Comelli EM, Ma DW, Lou W, Teterina A, Kim T, et al. Altered hepatic gene expression in nonalcoholic fatty liver disease is associated with lower hepatic n-3 and n-6 polyunsaturated fatty acids. *Hepa*tology. 2015;61(5):1565–78. doi: 10.1002/hep.27695. [PubMed: 25581263].
- Ahrens M, Ammerpohl O, von Schonfels W, Kolarova J, Bens S, Itzel T, et al. DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. *Cell Metab.* 2013;18(2):296–302. doi: 10.1016/j.cmet.2013.07.004. [PubMed: 23931760].
- Ito S, Yukawa T, Uetake S, Yamauchi M. Serum intercellular adhesion molecule-1 in patients with nonalcoholic steatohepatitis: comparison with alcoholic hepatitis. *Alcohol Clin Exp Res.* 2007;**31**(1 Suppl):S83– 7. doi: 10.1111/j.1530-0277.2006.00292.x. [PubMed: 17331172].
- Sookoian S, Castano GO, Burgueno AL, Rosselli MS, Gianotti TF, Mallardi P, et al. Circulating levels and hepatic expression of molecular mediators of atherosclerosis in nonalcoholic fatty liver disease. *Atherosclerosis*. 2010;209(2):585–91. doi: 10.1016/j.atherosclerosis.2009.10.011. [PubMed: 19896127].
- Theruvath TP, Ramshesh VK, Zhong Z, Currin RT, Karrasch T, Lemasters JJ. Icam-1 upregulation in ethanol-induced Fatty murine livers promotes injury and sinusoidal leukocyte adherence after transplantation. *HPB Surg.* 2012;2012:480893. doi: 10.1155/2012/480893. [PubMed: 22778492]. [PubMed Central: PMC3385666].

- Wang H, Chen J, Hollister K, Sowers LC, Forman BM. Endogenous bile acids are ligands for the nuclear receptor FXR/BAR. *Mol Cell*. 1999;3(5):543–53. doi: 10.1016/s1097-2765(00)80348-2. [PubMed: 10360171].
- Sinal CJ, Tohkin M, Miyata M, Ward JM, Lambert G, Gonzalez FJ. Targeted disruption of the nuclear receptor FXR/BAR impairs bile acid and lipid homeostasis. *Cell*. 2000;**102**(6):731–44. doi: 10.1016/s0092-8674(00)00062-3. [PubMed: 11030617].
- Gerloff T, Stieger B, Hagenbuch B, Madon J, Landmann L, Roth J, et al. The sister of P-glycoprotein represents the canalicular bile salt export pump of mammalian liver. J Biol Chem. 1998;273(16):10046–50. doi: 10.1074/jbc.273.16.10046. [PubMed: 9545351].
- Fuchs CD, Krivanec S, Steinacher D, Mlitz V, Wahlstrom A, Stahlman M, et al. Absence of Bsep/Abcb11 attenuates MCD diet-induced hepatic steatosis but aggravates inflammation in mice. *Liver Int.* 2020;**40**(6):1366–77. doi: 10.1111/liv.14423. [PubMed: 32141703]. [PubMed Central: PMC7317533].
- 22. Okushin K, Tsutsumi T, Enooku K, Fujinaga H, Kado A, Shibahara J, et al. The intrahepatic expression levels of bile acid transporters are inversely correlated with the histological progression of non-alcoholic fatty liver disease. J Gastroenterol. 2016;51(8):808-18. doi: 10.1007/s00535-015-1148-y. [PubMed: 26601667].
- Sundaram SS, Whitington PF, Green RM. Steatohepatitis develops rapidly in transgenic mice overexpressing Abcb11 and fed a methionine-choline-deficient diet. *Am J Physiol Gastrointest Liver Physiol.* 2005;**288**(6):G1321-7. doi: 10.1152/ajpgi.00455.2004. [PubMed: 15650132].
- 24. Figge A, Lammert F, Paigen B, Henkel A, Matern S, Korstanje R, et al. Hepatic overexpression of murine Abcb11 increases hepatobiliary lipid secretion and reduces hepatic steatosis. J Biol Chem. 2004;279(4):2790–9. doi: 10.1074/jbc.M307363200. [PubMed: 14570929].
- 25. Garcia-Ruiz C, Colell A, Morales A, Kaplowitz N, Fernandez-Checa JC. Role of oxidative stress generated from the mitochondrial electron transport chain and mitochondrial glutathione status in loss of mitochondrial function and activation of transcription factor nuclear factor-kappa B: studies with isolated mitochondria and rat hepatocytes. *Mol Pharmacol.* 1995;**48**(5):825–34. [PubMed: 7476912].
- Slimen IB, Najar T, Ghram A, Dabbebi H, Ben Mrad M, Abdrabbah M. Reactive oxygen species, heat stress and oxidative-induced mitochondrial damage. A review. Int J Hyperthermia. 2014;30(7):513–23. doi: 10.3109/02656736.2014.971446. [PubMed: 25354680].
- Brookes PS. Mitochondrial H(+) leak and ROS generation: an odd couple. *Free Radic Biol Med.* 2005;**38**(1):12-23. doi: 10.1016/j.freeradbiomed.2004.10.016. [PubMed: 15589367].
- Franklin MP, Sathyanarayan A, Mashek DG. Acyl-CoA Thioesterase 1 (ACOT1) Regulates PPARalpha to Couple Fatty Acid Flux With Oxidative Capacity During Fasting. *Diabetes*. 2017;66(8):2112-23. doi: 10.2337/db16-1519. [PubMed: 28607105]. [PubMed Central: PMC5521868].
- Yang S, Chen C, Wang H, Rao X, Wang F, Duan Q, et al. Protective effects of Acyl-coA thioesterase 1 on diabetic heart via PPARalpha/PGCtalpha signaling. *PLoS One*. 2012;7(11). e50376. doi: 10.1371/journal.pone.0050376. [PubMed: 23226270]. [PubMed Central: PMC3511550].
- Nakamura MT, Yudell BE, Loor JJ. Regulation of energy metabolism by long-chain fatty acids. *Prog Lipid Res.* 2014;53:124–44. doi: 10.1016/j.plipres.2013.12.001. [PubMed: 24362249].
- Kzhyshkowska J, Gratchev A, Goerdt S. Human chitinases and chitinase-like proteins as indicators for inflammation and cancer. *Biomark Insights*. 2007;2:128–46. [PubMed: 19662198]. [PubMed Central: PMC2717817].
- Zhang S, Sousa A, Lin M, Iwano A, Jain R, Ma B, et al. Role of Chitinase 3-Like 1 Protein in the Pathogenesis of Hepatic Insulin Resistance in Nonalcoholic Fatty Liver Disease. *Cells*. 2021;10(2). doi: 10.3390/cells10020201. [PubMed: 33498326]. [PubMed Central:

PMC7909438].

- Lee CG, Da Silva CA, Dela Cruz CS, Ahangari F, Ma B, Kang MJ, et al. Role of chitin and chitinase/chitinase-like proteins in inflammation, tissue remodeling, and injury. *Annu Rev Physiol*. 2011;73:479–501. doi: 10.1146/annurev-physiol-012110-142250. [PubMed: 21054166]. [PubMed Central: PMC3864643].
- Huang H, Wu T, Mao J, Fang Y, Zhang J, Wu L, et al. CHI3L1 Is a Liver-Enriched, Noninvasive Biomarker That Can Be Used to Stage and Diagnose Substantial Hepatic Fibrosis. *OMICS*. 2015;**19**(6):339–45. doi: 10.1089/omi.2015.0037. [PubMed: 26415140]. [PubMed Central: PMC4486713].
- Kumagai E, Mano Y, Yoshio S, Shoji H, Sugiyama M, Korenaga M, et al. Serum YKL-40 as a marker of liver fibrosis in patients with non-alcoholic fatty liver disease. *Sci Rep.* 2016;6:35282. doi: 10.1038/srep35282. [PubMed: 27739482]. [PubMed Central: PMC5064386].
- Fisher CD, Lickteig AJ, Augustine LM, Ranger-Moore J, Jackson JP, Ferguson SS, et al. Hepatic cytochrome P450 enzyme alterations in humans with progressive stages of nonalcoholic fatty liver disease. *Drug Metab Dispos*. 2009;**37**(10):2087–94. doi: 10.1124/dmd.109.027466. [PubMed: 19651758]. [PubMed Central: PMC2769034].
- Michaelis UR, Fisslthaler B, Barbosa-Sicard E, Falck JR, Fleming I, Busse R. Cytochrome P450 epoxygenases 2C8 and 2C9 are implicated in hypoxia-induced endothelial cell migration and angiogenesis. *J Cell Sci.* 2005;**118**(Pt 23):5489–98. doi: 10.1242/jcs.02674. [PubMed: 16291720].
- Li L, Chen SH, Zhang Y, Yu CH, Li SD, Li YM. Is the hypoxia-inducible factor-1 alpha mRNA expression activated by ethanol-induced injury, the mechanism underlying alcoholic liver disease? *Hepatobiliary Pancreat Dis Int.* 2006;5(4):560–3. [PubMed: 17085342].
- Eltzschig HK, Carmeliet P. Hypoxia and inflammation. N Engl J Med. 2011;364(7):656–65. doi: 10.1056/NEJMra0910283. [PubMed: 21323543]. [PubMed Central: PMC3930928].
- Horton JD, Goldstein JL, Brown MS. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J Clin Invest*. 2002;**109**(9):1125–31. doi: 10.1172/JCl15593. [PubMed: 11994399]. [PubMed Central: PMC150968].
- Kuan YC, Hashidume T, Shibata T, Uchida K, Shimizu M, Inoue J, et al. Heat Shock Protein 90 Modulates Lipid Homeostasis by Regulating the Stability and Function of Sterol Regulatory Elementbinding Protein (SREBP) and SREBP Cleavage-activating Protein. J Biol Chem. 2017;292(7):3016–28. doi: 10.1074/jbc.M116.767277. [PubMed: 28003358]. [PubMed Central: PMC5314195].
- Jhaveri K, Taldone T, Modi S, Chiosis G. Advances in the clinical development of heat shock protein 90 (Hsp90) inhibitors in cancers. *Biochim Biophys Acta*. 2012;**1823**(3):742-55. doi: 10.1016/j.bbamcr.2011.10.008. [PubMed: 22062686]. [PubMed Central: PMC3288123].
- Zheng ZG, Zhang X, Liu XX, Jin XX, Dai L, Cheng HM, et al. Inhibition of HSP90beta Improves Lipid Disorders by Promoting Mature SREBPs Degradation via the Ubiquitin-proteasome System. *Theranostics.* 2019;9(20):5769–83. doi: 10.7150/thno.36505. [PubMed: 31534518]. [PubMed Central: PMC6735373].
- Kagi JH, Schaffer A. Biochemistry of metallothionein. *Biochemistry*. 1988;27(23):8509–15. doi: 10.1021/bi00423a001. [PubMed: 3064814].
- Cherian MG, Jayasurya A, Bay BH. Metallothioneins in human tumors and potential roles in carcinogenesis. *Mutat Res.* 2003;**533**(1-2):201–9. doi:10.1016/j.mrfmmm.2003.07.013. [PubMed: 14643421].
- Fan LZ, Cherian MG. Potential role of p53 on metallothionein induction in human epithelial breast cancer cells. *Br J Cancer*. 2002;87(9):1019–26. doi: 10.1038/sj.bjc.6600549. [PubMed: 12434295]. [PubMed Central: PMC2364318].
- Cherian MG, Apostolova MD. Nuclear localization of metallothionein during cell proliferation and differentiation. *Cell Mol Biol (Noisy-legrand)*. 2000;46(2):347–56. [PubMed: 10774924].

- Cherian MG, Kang YJ. Metallothionein and liver cell regeneration. *Exp Biol Med (Maywood)*. 2006;231(2):138–44. doi: 10.1177/153537020623100203. [PubMed: 16446489].
- 49. Mosnier LO, Bouma BN. Regulation of fibrinolysis by thrombin activatable fibrinolysis inhibitor, an unstable carboxypeptidase B that unites the pathways of coagulation and fibrinolysis. Arterioscler Thromb Vasc Biol. 2006;26(11):2445–53. doi: 10.1161/01.ATV.0000244680.14653.9a. [PubMed: 16960106].
- Diedrich T, Kummer S, Galante A, Drolz A, Schlicker V, Lohse AW, et al. Characterization of the immune cell landscape of patients with NAFLD. *PLoS One.* 2020;**15**(3). e0230307. doi: 10.1371/journal.pone.0230307. [PubMed: 32168345]. [PubMed Central: PMC7069622].
- Estep M, Abawi M, Jarrar M, Wang L, Stepanova M, Elariny H, et al. Association of obestatin, ghrelin, and inflammatory cytokines in obese patients with non-alcoholic fatty liver disease. *Obes Surg.* 2011;21(11):1750–7. doi: 10.1007/s11695-011-0475-1. [PubMed: 21744131].
- Roh YS, Seki E. Toll-like receptors in alcoholic liver disease, nonalcoholic steatohepatitis and carcinogenesis. J Gastroenterol Hepatol. 2013;28 Suppl 1:38-42. doi: 10.1111/jgh.12019. [PubMed: 23855294]. [PubMed Central: PMC3721430].
- Miura K, Ohnishi H. Role of gut microbiota and Toll-like receptors in nonalcoholic fatty liver disease. World J Gastroenterol. 2014;20(23):7381–91. doi: 10.3748/wjg.v20.i23.7381. [PubMed: 24966608]. [PubMed Central: PMC4064083].