Non-alcoholic Fatty Liver Disease (NAFLD): A Systematic Review and Meta-analysis from an Omics Perspective

Luxita Sharma1*, Manoj Kumar Kashyap2,3 and Dhananjay Sharma4

1Department of Dietetics and Applied Nutrition, Amity Medical School, Amity University, Haryana, India; 2Amity Stem Cell Institute, Amity Medical School, Amity University Haryana, Panchgaon, Manesar (Gurugram), India; 3Clinical Biosamples & Research Services (CBRS), Noida, India; 4Department of Dietetics and Applied Nutrition, Amity Medical School, Amity University Haryana, India

Received: January 28, 2023 | Revised: March 24, 2023 | Accepted: May 23, 2023 | Published online: Month 00, 2023

Abstract

Background and objectives: Non-alcoholic fatty liver disease (NAFLD) is the most common chronic liver disease in developed countries, contributing to ~24% of cases worldwide and includes non-alcoholic steatohepatitis. NAFLD is characterized by lipid accumulation rather than alcohol consumption. There are several diagnostic/prognostic biomarkers for NAFLD including CK-18, ALT, AST, GGT, and haptoglobin, but with limited sensitivity and specificity. Therefore, high-throughput OMICS approaches have been used to characterize NAFLD conditions for the identification of potential molecular signatures or differentially regulated molecules (DEMs) and early detection of NAFLD.

Methods: We analyzed the publically available data set (accession number: GSE63067) from the Gene Expression Omnibus (GEO) using the GEO2R program. The differentially expressed genes (DEGs) were filtered using the criteria where genes with p-value ≤0.05 and fold-change ≥2.0-fold (upregulated), and fold-change ≤0.5-fold (downregulated).

Results: We identified 264 differentially expressed genes (DEGs) between NAFLD and normal liver tissue samples, where 211 were upregulated and 53 were downregulated in NAFLD. Additionally, we identified novel genes SGMS2, and WNK3 that were not well understood in the molecular pathophysiology of NAFLD. Further gene ontology-based analysis revealed that among biological processes, cellular components, and molecular functions were also dysregulated in NAFLD. Collectively, this integration of a systemic-cum-meta-analysis approach suggests that an OMICS-based analysis may provide better solutions if microarray and other high-throughput study-based DEMs are to be cataloged systematically for sharing with the scientific community. This will allow potential candidates (DEMs) to be interrogated either alone or in combination with existing biomarkers for effective early detection and or diagnosis of NAFLD.

Conclusions: Our study shows that meta-analysis of publicly available data could be a good source for identification of DEGs in NAFLD, and it can be easily extrapolated in other disease conditions.

Keywords: Non-alcoholic fatty liver disease; NASH; Microarray; GEO; GEO2R; hepatocellular carcinoma.

Abbreviations: ALF, acute liver failure; BPs, biological process; CLD, chronic liver disease; CCs, cellular components; DEMs, differentially expressed genes; DEGs, differentially expressed genes; GO, gene expression omnibus; GEO, gene ontology; ICC, hepatocellular carcinoma; MFs, molecular functions; NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis; PNPLA3, patatin-like phospholipase domain-containing 3; PD-1, programmed death 1 ligand-1; TM6SF2, transmembrane 6 superfamily member 2; WNK3, with-no-lysine kinase 3.

*Correspondence to: Luxita Sharma, Amity Medical School, Department of Dietetics and Applied Nutrition, Amity University Haryana, Manesar (Gurugram), Haryana, India. ORCID: https://orcid.org/0000-0002-4700-4792. Tel: 91-0124-2337016, E-mail: lshrama@ggn.amity.edu

How to cite this article: Sharma L, Kashyap MK, Sharma D. Non-alcoholic Fatty Liver Disease (NAFLD): A Systematic Review and Meta-analysis from an Omics Perspective. Gene Expr 2023:000(000):000–000. doi: 10.14218/GE.2023.00005.

Introduction

The liver is the largest internal organ of the body, providing important metabolic, exocrine, and endocrine functions. These include lipids lipogenesis, carbohydrates (gluconeogenesis, regulation of glucose level through glycogen storage, and synthesis of albumin, haptoglobin, and blood homeostasis by secreting clotting factors, as well as the production of bile salt, hemoglobin, iron, vitamin, ammonia, copper, and drugs. The liver regulates the number of chemicals, nutrients, microbial pathogens, and endotoxin components in the blood that exit the gastrointestinal tract through portal veins. It also participates in immune surveillance against pathogens and overcoming the effects of benign antigens through cytokine signaling and acute phase protein production. Hepatocytes...
are the main parenchymal cells in the liver, accounting for ~70% of its volume. These cells form an extracellular matrix that is primarily composed of glycoproteins, collagens, proteoglycans, and hyaluronic acid. Hepatic stellate cells are another important cell present in the liver and represent 5–8% of the volume of the liver. They are located between sinusoidal endothelial and liver epithelial cells and act as a reservoir of vitamin A. Kupffer cells are immune cells of the liver that clear endotoxins and provide a rapid response to hepatic damage and aging red blood cells.

Various physiological changes affect the pathogenesis of liver diseases and may lead to conditions such as acute and chronic liver diseases, Hepatitis A, B, C, D, and E, autoimmune liver diseases including autoimmune hepatitis and primary biliary cirrhosis, alcoholic liver disease, non-alcoholic liver disease, metabolic associated fatty liver diseases, drug-induced liver diseases, and genetic conditions such as hemochromatosis, Wilson’s disease, alpha-1 antitrypsin deficiency, hepatocellular carcinoma, and liver transplant condition.

Liver functions are insulted due to either acute or chronic liver failure. Acute liver failure (ALF) is defined by sudden hepatocellular necrosis that leads to the drastic deterioration of hepatic function, which begins within an hour and lasts up to six months after being jaundiced or a pre-existing liver disease. Recurrent cases of ALF are often due to drug addiction and acetaminophen overdose (hepatotoxicity). Other causes include autoimmune hepatitis and acute viral hepatitis. Chronic liver disease (CLD) is a condition that involves a long-term decline in liver function due to chronic damage to the liver. After 1–2 decades with CLD, the individual suffers from liver cirrhosis (a condition in which liver cells are replaced by fibrosis, scar tissues, and regenerative nodules).

CLD is a combination of abnormalities characterized by a progressive decline in hepatic functions over six or more years, resulting in chronic inflammation. Recent studies suggest that ~5.5 million individuals worldwide are affected with CLD and for ~40,000 it will be fatal. In the United States, CLD is the 12th leading cause of death with 4,000–5,000 deaths and 11,000–17,000 hospitalizations annually. Several factors such as excess alcohol consumption, poisoning, autoimmune disease, viral or pathogenic infections, metabolic imbalance, and genetic disorders can trigger the destruction of the parenchymal layer of the liver which is capable of releasing important liver injury biomarkers such as alanine and aspartate transaminase.

Materials and methods

Non-alcoholic fatty liver disease/non-alcoholic steatohepatitis

While non-alcoholic steatohepatitis (NASH) and non-alcoholic fatty liver disease (NAFLD) are both liver conditions related to the accumulation of fat in the liver, NAFLD is a broader term that refers to the presence of fat in the liver without any evidence of inflammation or liver damage. NASH, however, is a more severe form of NAFLD that is characterized by liver inflammation and damage.

Mortality rate and cause of death in NAFLD and other liver-related disease

The most common causes of liver cirrhosis are alcoholism, chronic viral hepatitis, and accumulation of excess fat in liver cells, causing NAFLD, NASH, and hepatocyte hyperplasia, which could eventually lead to end-stage liver disease namely hepatocellular carcinoma (Fig. 1). The mortality rate in different liver-associated abnormalities is summarized in Table 1.

NAFLD is considered the most common liver disease in developed countries and is prevalent in ~25% of the world population. The disease is characterized by hepatic lipid accumulation rather than alcohol consumption. Presently in Western and Asian countries, the prevalence of NAFLD varies at 25–35% and 5–15% respectively, and in India, it ranges from 9–53%. Excess calorie and fat intakes leading to obesity and related comorbidities are the leading risk factor for NAFLD. Conversely, the intake of low-calorie foods and good fat such as omega-3 polyunsaturated fats reduces insulin resistance, intrahepatic triglycerides content, and risk of steatohepatitis. The incidence of NAFLD rapidly rises along with an incidence increase of diabetes, obesity, and dyslipidemia. The prevalence is considerably higher in people with diabetes (60–70%) and in morbidly obese (75–92%) than in the general population. NAFLD includes non-alcoholic fatty liver and NASH. Due to higher exposure to free fatty acids, diacylglycerol, and oxidized cholesterol, lipotoxicity is a common pathophysiological condition seen in NAFLD patients.

The condition leads to hepatic steatosis contributing to the production of fats from non-fat sources by de novo lipogenesis process. Incidences of liver cancer, i.e., hepatocellular carcinoma (HCC) have been observed more frequently in NAFLD patients, compared with cirrhosis patients. Some of the important predictors for NAFLD-linked HCC that are commonly associated with HCC are old age, gender, and genetic factors. In virus-associated cirrhosis, Hepatitis-B/Hepatitis-C and Hepatitis-B/Hepatitis-D viruses co-existingly increase the possibility of HCC. A comparative analysis showed that individuals infected with the Hepatitis-B virus have a 5–100-fold likelihood to get HCC compared to individuals with the Hepatitis-C virus who have a 15–20-fold likelihood to get HCC.

Genomics aspect of NAFLD

Excess fatty acid metabolism increases the generation of reactive oxygen species (ROS)—which are capable of damaging the liver cells—such as superoxide radical (O2−), hydrogen peroxide (H2O2), and hydroxyl radicals (·OH). Kupffer cells of the liver are relatively more sensitive to oxidative stress than hepatocytes, and produce tumor necrosis factor-α that may cause inflammation and apoptosis. ROS initiates lipid peroxidation in hepatic stellate cells, which causes cell proliferation and collagen synthesis leading to hepatic fibrosis. Epigenetic changes in DNA methylation, histone modification, and non-coding RNA-mediated gene silencing have been reported in different liver aberrancies.

We know that to delineate the disease biology, a number of techniques have been used including conventional molecular biology techniques such as southern blot, northern blot, and conventional and real-time PCR. However, as low throughput techniques, they have limitations. A handful of genes such as patatin-like phospholipase domain-containing 3 (PNPLA3), transmembrane 6 superfamily member 2 (TM6SF2), and glucokinase regulatory protein (GCKR) have been reported in the context of NAFLD. PNPLA3 has acyltransferase and lipase activity, which encodes adiponutrin and is located on the 22q13 locus. Adiponutrin variant p.I148M alters fat composition by reducing polyunsaturated fatty acids from diacylglycerol into phosphatidylcholine which causes an increase in triglycerides and diacylglycerol. Genotype rs738409 is an indicator of the risk of NAFLD that histologically confirms steatosis. TM6SF2 encrypts a regulatory protein in very low-density lipoprotein secretion. TM6SF2 variant p.E167K (rs58542926) mainly affects the biosynthesis of PUFA and evacuates PUFA from hepatic polyunsaturated phosphatidylcholines.
Sharma L. et al: Differential regulated genes in NAFLD

GCKR is pointed at chromosome 2, which inhibits glucokinase enzyme in hepatocyte nuclei. GCKR p.(rs780094) variant increases glycogen and glycolysis simultaneously increases de novo lipogenesis, which is linked to NAFLD risk. Some additional genes are also linked with NAFLD such as membrane-bound O-acyltransferase domain-containing 7 (MBOAT7), APOB, MERTK, SERPINA1, IL28B, and HFE. MBOAT (rs641738) is mildly linked with liver fibrosis and inflammation but not with steatosis and increases NAFLD risk.

Proteomics studies for NAFLD profiling

A number of proteomics studies have been carried out to profile NAFLD for the identification of markers toward early diagnosis, or progression. Differential Gel Electrophoresis (DIGE) was used in combination with matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) to profile NASH vs. control liver tissue samples. A total of 43 differentially regulated proteins were identified. Among those 22 and 21 were significantly higher (ratio ≥1.5, \( p < 0.05 \)) in the steatosis and NASH samples compared to the control sample, respectively. Among these GRP78 (78 kDa glucose-regulated protein) was upregulated in the steatosis sample, but downregulated in the NASH sample. The appearance of the spots for the same protein in the upregulated versus downregulated stage suggests that there may be involvement of post-translational modifications of these proteins which may be crucial in the pathophysiology of NAFLD.

Other proteomics studies have reported a large number of molecules but few have made it into the clinical setting. Several diagnostic markers have been documented for NAFLD, but their efficiency has been compromised due to their limited sensitivity and specificity. The partial list of different diagnostic markers for NAFLD has been summarized in Table 2.
Methods

To overcome the problem of low throughput, scientists have begun using high-throughput techniques such as microarray or RNAseq. These techniques can generate a high volume of data providing a perspective at the whole transcriptome level. Keeping this in mind, we searched the microarray repository popularly known as the gene expression omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). The workflow used in this study is shown in Figure 2. We used Boolean operators such as AND, OR, and NOT to limit our search to specific studies.\(^6\) We selected studies where NAFLD tissue samples of human origin were compared with the normal liver tissue samples. To analyze the microarray, we used the GEO2R program developed by NCBI and freely available on the NCBI website (Link: https://www.ncbi.nlm.nih.gov/geo2r/)—GEO has been integrated with the NCBI and freely available on the NCBI website (Link: https://www.ncbi.nlm.nih.gov/geo2r/)—GEO has been integrated with the NCBI and freely available on the NCBI website (Link: https://www.ncbi.nlm.nih.gov/geo2r/). The package facilitates the transformation of the GEO data, which can be exported to .txt files. The genes are presented in order of significance (based on \(p\)-value), but one can sort the data based on Log2-fold change as well. While using the GEO2R program, the assignment of the samples must be performed in order to differentiate between NAFLD and normal liver samples. As such, the NAFLD samples were assigned first followed by the normal liver samples. To identify differentially regulated genes (DEGs), genes with a change \(\geq 2\)-fold and \(p\)-value \(< 0.05\) were considered as upregulated, and genes with a change of \(\leq 0.5\)-fold and \(p\)-value \(< 0.05\) were considered as downregulated or unexpressed genes. We then downloaded the NAFLD data set (GSE accession No. GSE63067) from the NCBI database. For this study, we used the GEO2R tools, the dataset was analyzed to identify the DEGs between the NAFLD and normal samples.

Results and discussion

This analysis identified 264 DEGs between the NAFLD and normal samples. Among these, 211 genes were upregulated in the NAFLD samples compared to the normal samples with at least \(p < 0.05\), and a change of \(\geq 2\)-fold (Table 3).\(^7\) Similarly, 53 genes were downregulated in the NAFLD samples compared to the normal samples with at least \(p < 0.05\), and \(\leq 0.5\)-fold change. In addition, we used the ShinyGO enrichment tool to identify the involvement of a variety of pathways as seen in the analysis of the DEGs identified from the microarray dataset. This allowed us to identify different pathways that were upregulated or downregulated in the case of the NAFLD samples. After identifying the DEGs from different studies using GEO2R, we subjected them to the ShinyGO enrichment tool\(^8\) for the analysis of gene ontology (GO) biological processes (BPs), molecular functions (MFs), and cellular components (CCs) in the NAFLD samples compared to the normal samples. Using the ShinyGO enrichment tool, the DEGs were mapped to the respective chromosomes and it was found that the dysregulated DEGs in the NAFLD samples were from all the chromosomes (Fig. 3).

Known and differentially regulated genes in NAFLD

**C-C motif chemokine ligand 20 (CCL20)**

CCL20 is located on 2q33-q37. It is an extracellular protein that contains a signal peptide motif. It is secretory in nature and has been reported in biological fluids such as plasma. CCL20 is an...
Sharma L. et al: Differential regulated genes in NAFLD

anti-microbial protein that is also secretory in nature. Stellate cells produce CCL20 in fibrotic conditions. However, CCL20 is a poor prognostic marker for renal cell carcinoma and pancreatic cancer. 86,87 CCL20 is secreted by tumor-associated macrophages, where cancer cells are activated by CCL20 through activation of AKT. CCL20 was upregulated 5.96-fold in the NAFLD samples compared to the normal samples in the present study, which is in concordance with a previously published study on NAFLD. 72

CD274 is located at 9p24. It codes for the CD274 (cluster of differentiation 274) protein, which is also known as programmed death 1 ligand-1 (PD-L1) or B7 homolog 1 (B7-H1). It consists of a transmembrane domain and a signal peptide motif. It belongs to the molecular class ligand, and the primary localization is the plasma membrane. CD274 expresses on tumor cells and interacts with the PD1 receptor (which is usually present on the T-cells). This interaction leads to suppression of the immune response or so-called T-cell exhaustion. In the current study, PD-L1 was upregulated 4.78-fold in the NAFLD samples compared to the normal samples which is in agreement with a previous report on NAFLD, where CD274 was not only upregulated but was also part of the gene network associated with NAFLD. 73

S100 calcium-binding protein A8 (S100A8) S100A8 is located at 1q21.3 locus. It consists of EF-hand domain 1 (low Ca2+ affinity and binds to Zn2+) and EF-hand domain 2 (high Ca2+ affinity). Upregulation and overexpression of S100A8 have been reported in NASH with fibrosis samples compared with the control samples. In the current study, S100A8 was found to be directly correlated with the fibrotic status.

Moreover, S100A8 was found to be upregulated 3.34-fold in the NAFLD samples compared to the control samples.

With-no-lysine kinase 3 (WNK3) or WNK lysine deficient protein kinase 3 (WNK3) With-no-lysine kinase 3 or WNK lysine deficient protein kinase 3 (WNK3) is a serine/threonine protein kinase. WNK3 is located at the Xp11.22 locus. It belongs to the WNK family of kinases, which consists of four members including WNK3. Downregulation of WNK2 has been reported in HCC but no report exists for WNK3. WNK2 and WNK3 kinases are also known as “dark kinases” because little is known about them in signaling or tumor biology. In a recent study, inhibition of WNK3 also suppressed PD-L1 expression in cancer cells, which ultimately led to the activation of T-lymphocytes. 89 In our study, WNK3 was downregulated in the NAFLD samples compared to the normal samples. Furthermore, its role in the context of NAFLD requires further investigation.

Enrichment of biological process, cellular components, and molecular functions in NAFLD DEGs were further analyzed using the ShinyGO enrichment tool

Table 2. Different diagnostic markers for NAFLD along with their sensitivity and specificity

<table>
<thead>
<tr>
<th>Diagnostic Markers for NAFLD</th>
<th>NAFLD/Normal</th>
<th>Unit</th>
<th>Sensitivity/Specificity</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytokeratin (CK)-18</td>
<td>NAFLD</td>
<td>≥225 U/L</td>
<td>70%/82%</td>
<td>Papatheodoridis et al. 201054</td>
</tr>
<tr>
<td>HCV</td>
<td>≥225 U/L</td>
<td>67%/77%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibrosis</td>
<td>≥157.5 U/L</td>
<td>64%/61%</td>
<td></td>
<td>Joka et al. 201255</td>
</tr>
<tr>
<td>Alanine aminotransferase (ALT)</td>
<td>NASH</td>
<td>&gt;70 U/L</td>
<td>50%/60.7%</td>
<td>Verma et al. 201356</td>
</tr>
<tr>
<td>Advanced fibrosis</td>
<td>&gt;70 U/L</td>
<td>40%/57.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NASH</td>
<td>&gt;50 IU/L</td>
<td>72%/62%</td>
<td></td>
<td>Poynard et al. 200557</td>
</tr>
<tr>
<td>NASH</td>
<td>&gt;30 U/L</td>
<td>42%/80%</td>
<td></td>
<td>Kunde et al. 200558</td>
</tr>
<tr>
<td>Aspartate aminotransferase (AST)</td>
<td>Significant Fibrosis</td>
<td>42 U/L</td>
<td>57%/68%</td>
<td>Harrison et al. 202059</td>
</tr>
<tr>
<td>Advanced Fibrosis</td>
<td>40 U/L</td>
<td>77%/81%</td>
<td></td>
<td>Bril et al. 202060</td>
</tr>
<tr>
<td>γ-glutamyl transpeptidase (GGT)</td>
<td>NAFLD (Advanced fibrosis)</td>
<td>72.9 IU/L</td>
<td>64.2%/52.6%</td>
<td>Ha et al. 202261</td>
</tr>
<tr>
<td>Fibrosis</td>
<td>70 U/L</td>
<td>40%/72%</td>
<td></td>
<td>Harrison et al. 202059</td>
</tr>
<tr>
<td>Serum haptoglobin</td>
<td>≤67.13 ng/ml</td>
<td>80%/93.33%</td>
<td></td>
<td>Dawood et al. 202162</td>
</tr>
<tr>
<td>Adiponectin</td>
<td>NASH &amp; Non-NASH</td>
<td>13.5 µg/mL</td>
<td>92.3%/86.7%</td>
<td>Pirvulescu et al. 201263</td>
</tr>
<tr>
<td>Leptin</td>
<td>NASH &amp; Non-NASH</td>
<td>40 ng/mL</td>
<td>61.5%/65.9%</td>
<td></td>
</tr>
<tr>
<td>Hyaluronic acid (HA)</td>
<td>Hepatic fibrosis</td>
<td>46.1 µg/mL</td>
<td>85%/80%</td>
<td>Suzuki et al. 200564</td>
</tr>
<tr>
<td>Advanced fibrosis</td>
<td>46 ng/ml</td>
<td>80%/69.6%</td>
<td></td>
<td>Kumagai et al. 201665</td>
</tr>
<tr>
<td>Tissue inhibitor of metalloproteinase -1 and 2 (TIMP-1 &amp; TIMP-2)</td>
<td>NASH</td>
<td>–</td>
<td>96.7%/100%</td>
<td>Abdelaziz et al. 201566</td>
</tr>
<tr>
<td>Laminin</td>
<td>NAFLD (Fibrosis)</td>
<td>&gt;282 ng/ml</td>
<td>82%/89%</td>
<td>Dos Santos et al. 200567</td>
</tr>
<tr>
<td>YKL-40 or chitinase 3-like-1</td>
<td>NAFLD (Fibrosis)</td>
<td>165 ng/ml</td>
<td>70%/76.8%</td>
<td>Kumagai et al. 201665</td>
</tr>
</tbody>
</table>

HCV, hepatitis C virus; NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis.
Sharma L. et al: Differential regulated genes in NAFLD

The workflow for the analysis of the microarray dataset using the GEO2R online program of NCBI. A step-by-step workflow was used for the meta-data analysis of the microarray data derived from NAFLD patients. To find the relevant study, keywords ‘NALFD’, Homo sapiens, and ‘microarray’ were used in combination with the Boolean operators (AND, OR, or NOT) when searching the microarray data repository Gene Expression Omnibus. The selected study was processed to identify differentially regulated genes between NAFLD and normal liver tissue samples using the GEO2R program of NCBI. DEGs were selected based on ≥2.0-fold (upregulated) or ≤0.5-fold (downregulated) along with a p-value <0.05. The DEGs were processed using the ShinyGO enrichment tool for gene ontology, biological processes (BPs), cellular components (CCs), and molecular functions (MFs). DEGs, differentially expressed genes; HCC, hepatocellular carcinoma; NAFLD, non-alcoholic fatty liver disease.

Fig. 2. The workflow for the analysis of the microarray dataset using the GEO2R online program of NCBI. A step-by-step workflow was used for the meta-data analysis of the microarray data derived from NAFLD patients. To find the relevant study, keywords ‘NALFD’, Homo sapiens, and ‘microarray’ were used in combination with the Boolean operators (AND, OR, or NOT) when searching the microarray data repository Gene Expression Omnibus. The selected study was processed to identify differentially regulated genes between NAFLD and normal liver tissue samples using the GEO2R program of NCBI. DEGs were selected based on ≥2.0-fold (upregulated) or ≤0.5-fold (downregulated) along with a p-value <0.05. The DEGs were processed using the ShinyGO enrichment tool for gene ontology, biological processes (BPs), cellular components (CCs), and molecular functions (MFs). DEGs, differentially expressed genes; HCC, hepatocellular carcinoma; NAFLD, non-alcoholic fatty liver disease.

to identify enriched GO terms for biological processes, cellular components, and molecular functions in NAFLD. The top BPs enriched were flavonoid glucuronidation, leukocyte migration, positive regulation of cell migration, positive regulation of cellular component movement, positive regulation of locomotion, inflammatory response, response to the bacterium, cytokine-mediated signaling pathway, cellular response to cytokine stimulus, leukocyte activation, cell activation, response to cytokine, regulation of programmed cell death, cell migration, regulation of the apoptotic process, locomotion, and cell motility and localization (Fig. 4a). The top CCs enriched were secretory granules, secretory vesicles, collagen-containing extracellular matrix, and secretory granule membrane (Fig. 4b). Top MFs enriched were arachidonic acid binding, icosanoid binding, eicosatetraenoic acid binding, RAGE receptor binding, IgG binding, macrolide binding, retinoic acid binding, long-chain fatty acid binding, glucuronosyl transferase
Differential regulated genes in NAFLD

The pro-inflammatory lipid derivative arachidonic acid plays a crucial role in the development and progression of NAFLD. Arachidonic acid acts as a precursor to eicosanoids, another pro-inflammatory molecule, whose elevation can further induce the progression of NAFLD.

Identification of dysregulated pathways in NAFLD

We identified a number of dysregulated biological pathways in NAFLD. These include ascorbate and aldarate metabolism, pen-
sets from the GEO database. GEO2R is limited to a single pairwise comparison and is only applicable to certain types of datasets with user-submitted expression tables of a limited size. Additionally, GEO2R does not provide quality control plots or clustered heatmaps, which are useful tools for visualizing and interpreting gene expression data. While GEO2R can be a useful tool for the preliminary analysis of gene expression data, it may not be suitable for more complex analyses or larger datasets. Future research may require the use of other tools or software packages for more comprehensive gene expression analysis. Furthermore, for some GEO sets, when analyzed using GEO2R, there are no gene symbols but only NM IDs (usually due to the information provided for the chip), which must be converted into gene symbols using other online programs. Regardless of the limitations, multiple GEO datasets can be analyzed using GEO2R to enrich the identification of DEGs in NAFLD/NASH or any other abnormalities associated with the liver.

**Conclusion**

NAFLD encompasses both nonalcoholic fatty liver and NASH and can progress to cirrhosis and HCC if left untreated. Additionally, it is known that traditional biomarkers such as CK-18, ALT, AST, GGT, and haptoglobin have limited sensitivity and specificity for the diagnosis and prognosis of NAFLD. The meta-analysis of one microarray study identified over 200 DEGs in NAFLD liver tissue samples compared to normal liver tissue samples. These DEGs may provide important insights into the molecular mechanisms underlying NAFLD pathogenesis and progression, as well as potential diagnostic and prognostic biomarkers. It is also interesting to note that dysregulated GO biological processes, cellular components, and molecular functions were observed in NAFLD, indicating the involvement of multiple molecular pathways and cellular processes in this complex disease. The dysregulation of the arachidonic acid metabolism pathway is also a noteworthy finding, as this pathway has been implicated in various liver diseases and may represent a potential target for therapeutic intervention in NAFLD.

**Acknowledgments**

None to declare.

**Funding**

MKK is the recipient of an extramural grant (Sanction #: 5/13/55/2020/NCD-III)) from the Indian Council of Medical Research (ICMR), New Delhi, India, and the Teachers Associateship for Research Excellence (TARE, Grant Ref #: TAR/2018/001054) from the Science and Engineering Board (SERB), New Delhi.

**Conflict of interest**

The authors declare no conflict of interests.
Fig. 4. Enrichment analysis using the ShinyGO enrichment tool to identify top GO biological processes, cellular components, and molecular functions in NAFLD. NAFLD, non-alcoholic fatty liver disease; GO, gene ontology.
Author contributions
Study concept and design (LS and MK), statistical Analysis (MK and DS), manuscript drafting (LS, MK and DS), critical revision of the manuscript (MK and DS), supervision (LS). All the authors have read, critically evaluated, and approved it for submission.

Ethical statement
Not applicable because the data included in the study was collected from the dataset on NAFLD deposited to the Gene Expression Omnibus (GEO, a microarray repository). The study does not involve any human subjects, tissue samples, or cell lines.

References
[14] Golabi P, Paik JM, Arshad T, Younossi Y, Mishra A, Younossi ZM. Mor-


Sharma L. et al: Differential regulated genes in NAFLD


