Key Genes Associated with Non-Alcoholic Fatty Liver Disease and Acute Myocardial Infarction

Background: With increasing research on non-alcoholic fatty liver (NAFLD) and acute myocardial infarction (AMI), many studies show a tight correlation between NAFLD and AMI, but the underlying pathophysiology is still not clear. This study was performed to identify the potential hub genes and pathways related to these 2 diseases by using the bioinformatics method.

Material/Methods: The Gene Expression Omnibus (GEO) dataset GSE63067 of NAFLD patients and normal controls was downloaded from the GEO database. The GSE60993 and GSE66360 datasets for AMI patients and healthy controls were also obtained. Differentially expressed genes (DEGs) of NAFLD and AMI datasets and the common genes between them were obtained. Further GO and KEGG enrichment analyses for common genes were performed. To define the pathogenesis associated with both NAFLD and AMI, a protein–protein interaction (PPI) network was constructed. Finally, SPSS software was utilized to analyze the diagnostic value of hub genes in the NAFLD and AMI datasets, respectively.

Results: Seventy-eight common genes were obtained in NAFLD and AMI with the threshold of P-value <0.05. Thirty-one GO terms and 10 KEGG pathways were obtained. Also, the top 10 hub genes (TLR2, LILRB2, CXCL1, FPR1, TLR4, TYROBP, MMP9, FCER1G, CLEC4D, and CCR2) were selected with P<0.05.

Conclusions: The results of this study suggest that some novel genes play an important role in the occurrence and progression NAFLD and AMI. More experimental research and clinical trials are needed to verify our results.

MeSH Keywords: Biological Markers • Liver Diseases • Myocardial Infarction

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/922492
Background

Non-alcoholic fatty liver disease (NAFLD) is a metabolic syndrome characterized by multiple definite causes, including cardiovascular and cerebrovascular diseases, DM, dietary rhythm, and environmental and genetic factors [1]. It was once considered as a benign compensatory disease; however, with the deepening of understanding of the disease, it is now known that it can progress from simple steatohepatitis to non-alcoholic steatohepatitis (NASH), and even cirrhosis and liver cancer [2]. Although its pathogenesis is unclear, inflammatory factors play an important role [3]. With the high prevalence of NAFLD, the number of patients and financial expenditure are increasing steadily, which imposes a heavy economic burden on society [4,5].

Acute myocardial infarction (AMI) is one of the main causes of death in developed countries [6]. Although therapeutic techniques have rapidly progressed, the morbidity and mortality rates remain at nearly 5% [7]. To date, the pathogenesis of AMI has not yet been fully studied; however, it is clear that inflammatory factors are related to the injury of cardiomyocytes and AMI [8].

In recent studies, AMI has been proved to be the most common cause of death among NAFLD patients [9,10]. A large-cohort study was performed on NAFLD patients in Korea, showing that AMI was more likely to occur in NAFLD patients compared with those without NAFLD after adjusting for potential risk factors of AMI and NAFLD [11]. Also, a meta-analysis including over 34,000 participants revealed the odds ratio was 1.64 (95% confidence interval 1.26–2.13) for combined fatal and non-fatal AMI [12]. Many NAFLD patients have abnormally high glucose and lipid levels, as well as high BMI, which are both risk factors for AMI [13]. Although research on the relationship between NAFLD and AMI has progressed steadily, the related genetics research is still limited and needs further exploration.

In the present study, 1 gene expression profile for NAFLD and 2 profiles for AMI were downloaded from the Gene Expression Omnibus (GEO) database, which is an open-access database that provides genetics information and can be used to identify new targets of disease by bioinformatics analysis [14]. Differentially expressed genes (DEGs) between disease samples and normal controls in NAFLD and AMI were found, and the common expressed genes of these 2 diseases were obtained as well. Cluster analysis was conducted to search for hub genes with multiple functions, including Gene Ontology (GO) term enrichment, Kyoto Encyclopedia of Genes and Genomes (KEGG), and protein-protein interaction (PPI) network analyses. Consequently, the identified hub genes may become a novel research focus, and the obtained molecular mechanisms and signal pathways may also help to explain the relationships between NAFLD and AMI.

Material and Methods

Microarray data collection

The microarray datasets of NAFLD, AMI, and normal controls were downloaded from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo). The GSE63067 dataset contained the gene expression profiles of 11 NAFLD patients and 7 non-NAFLD controls. In the GSE60993 dataset, 7 AMI patients and 7 healthy controls were involved in analysis, and 49 AMI patients and 50 non-AMI controls were also selected from GSE66360. Because these gene expression profiles all originated from a free open-access database on the internet, our research did not require Ethics Committee approval.

Identification of differentially expressed genes (DEGs) in NAFLD and AMI

DEGs between NAFLD and normal controls, AMI patients, and corresponding controls were identified using the limma R package, which is an efficient analysis method in bioinformatics [15]. The selected criteria in NAFLD datasets were set as P-value <0.05 and \(|\log_{2} FC| >0.5\). In AMI section, the cut-off values were adjusted-P<0.05 and \(|\log_{2} FC| >1\). After we utilized these screening conditions, 2 sets of DEGs were identified, then we put these DEGs that came from the 2 diseases into an online analysis tool Venn (http://bioinformatics.psb.ugent.be/webtools/Venn/) to obtain their intersection genes. These intersecting common genes were used for subsequent analysis.

Functional enrichment analyses for common DEGs

The Gene Ontology (GO) [16] classification, which contains molecular functions, biological processes, and cellular component, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) [17] pathway enrichment analyses were performed on the intersecting common DEGs using the R package. P-values 

<0.05 were defined as statistically significant.

Construction of protein–protein interaction (PPI) network and identification of hub genes

To further explore the interaction among the common genes obtained above, we used the Search Tool for the Retrieval of Interacting Genes (STRING) 11.0 (http://string-db.org/) [18] to construct a PPI network. The minimum required interaction score was considered high confidence (0.700) as the criteria of statistical significance. Cytoscape 3.6.1 (https://cytoscape.org) [19] was utilized to present the results. In the network outcome, the nodes represented the proteins, while the lines represented the interactions between proteins [20]. After we identified the hub genes among these common genes, we installed the
cytoHubba plugin (http://hub.iis.sinica.edu.tw/cytohubba/), which can be downloaded for free via Cytoscape software [21].

Statistical analysis

Using SPSS 22.0 (SPSS, Inc, Chicago, IL, USA), we constructed receiver operating characteristic (ROC) curves and calculated the area under the curve (AUC) of the hub genes to compare the AUC as the index of the model. These results showed the diagnostic efficiency of genes. P-values <0.05 were regarded as statistically significant.

Results

Differential expression analysis of DEGs in NAFLD and AMI

In the NAFLD dataset GSE63067, 823 DEGs were filtered when we compared the 11 NAFLD samples with 7 normal controls. Two AMI-related datasets (GSE60993 and GSE66360) were enrolled in the study, the merged dataset contained 56 AMI patient samples and 57 of their control samples. Heatmaps showing the gene expression profiles of NAFLD and AMI are presented in Figures 1 and 2, after homogenization, there were 823 and 200 DEGs achieved with adjusted-P<0.05, respectively. Using the Venn Diagram online tool, 78 intersecting common genes of 2 diseases were obtained and are shown in Figure 3. The details of these common genes are also shown in the Supplementary Table 1. The information on the whole study process is presented in Figure 4.

GO and KEGG enrichment pathway analysis

Functional enrichment and KEGG pathway analyses of 78 common genes were performed in NAFLD and AMI at the threshold of P-value <0.05. Changes in GO biological processes (BP) mainly included inflammatory reactions (e.g., inflammatory response, innate immune response, immune response, and neutrophil chemotaxis) and signal transduction, shown in Figure 5. In Figure 6, changes in cellular component (CC) were notably focused on enrichment of cell outer membranes, such as plasma membrane, integral component of plasma membrane, extracellular space, cell surface, and anchored component of membranes. Moreover, in the molecular function (MF) section (Figure 7), changes were significant in receptor activity (RAGE receptor binding and signaling pattern recognition receptor activity), binding-related function (carbohydrate binding and IgG binding), and pantetheine hydrolase activity. In particular, changes in the KEGG pathway were mostly enriched in several immune system diseases (leishmaniasis, tuberculosis, malaria, and rheumatoid arthritis), immune-related pathways (IL-17 signaling pathway and NF-kappa B signaling pathway), and transcriptional mis-regulation in cancer, which are illustrated by a dot plot in Figure 8.

PPI network analysis and hub gene selection

To distinguish the hub genes from the common genes, a PPI network was constructed. As seen in Figure 9, toll-like receptor 2 (TLR2), leukocyte immunoglobulin-like receptor subfamily B2 (LILRB2), C-X-C motif chemokine ligand 1 (CXCL1), formyl
peptide receptor 1 (FPR1), toll-like receptor 4 (TLR4), tyrosine kinase binding protein (TYROBP), matrix metalloproteinase 9 (MMP9), Fc-receptor common gamma chain (FCER1G), C-type lectin domain family 4 member D (CLEC4D), and chemokine receptor 2 (CCR2) proteins interact with other proteins by >5, which was the central node of the protein interaction network. Figure 10 shows the concrete scores of these hub genes. Finally, we chose the top 6 genes for use in further research.

Validation of diagnostic value of hub genes

To validate the diagnostic value of the top 6 hub genes obtained from the above analysis, we constructed ROC curves and calculated the corresponding area under the curve (AUC) of these gene expression levels in the NAFLD and AMI datasets. Figure 11 shows the results of NAFLD. The AUC for TLR2, LILRB2, CXCL1, FPR1, TLR4, and TYROBP in NAFLD patients and normal controls were 0.818 [95% confidence interval (CI), 0.619–1.000; P<0.05], 0.798 (95% CI, 0.584–1.000; P<0.05), 0.805 (95% CI, 0.558–1.000; P<0.05), 0.818 (95% CI, 0.610–1.000; P<0.05), 0.792 (95% CI, 0.563–1.000; P<0.05), and 0.831 (95% CI, 0.631–1.000; P<0.05). Figure 12 shows the ROC curve in AMI patients and non-AMI controls. The AUC for 6 hub genes in AMI were 0.762 [95% CI, 0.672–0.851; P<0.05], 0.798 (95% CI, 0.715–0.881; P<0.05), 0.742 (95% CI, 0.651–0.832; P<0.05), 0.748 (95% CI, 0.656–0.839; P<0.05), 0.758 (95% CI, 0.666–0.849; P<0.05), and 0.677 (95% CI, 0.577–0.777; P<0.05) (Figure 12B).

Discussion

NAFLD and AMI are highly prevalent diseases around the world. A previous study showed that NAFLD patients are at increased risk of both hepatic complications and cardiac metabolic complications [22]. With longer course of disease, the risk increases gradually, so it is necessary to explore the molecular mechanisms in these 2 diseases and discover the early targets to prevent disease development [23].

In this study, through searching the datasets of NAFLD and AMI from GEO, we found 78 common DEGs between these diseases. We also performed GO enrichment and KEGG pathway enrichment analyses, and a PPI network was constructed to identify the top 10 hub genes from among the common DEGs. We chose 6 hub genes to verify diagnostic value in NAFLD and
Figure 4. Flow diagram of study.

Figure 5. GO analysis of biological processes (BP).

Figure 6. GO analysis of cellular component (CC).
AMi patients (P<0.05). These genes may have an important ability to predict risk of AMI and T2DM.

Toll-like receptors (TLRs) belong to the family of pattern recognition receptors and play a vital role in recognizing bacterial and viral components, like lipopolysaccharides, bacterial DNA, and peptidoglycan [24,25]. Notably, among the TLRs family, toll-like receptor 2 (TLR2) and TLR4 have been reported to have a close relationship with pathogenesis of NAFLD [26]. A previous study indicated that the innate immune system plays an important role in the progression of NAFLD, and in this process, TLR4 is activated and participates in the disease sequence of steatosis, steatohepatitis, liver fibrosis, cirrhosis, and liver cancer [27]. In addition, Chiu et al. showed that, in a model of C57BL/6J mice fed with fresh fecal mixture obtained from NAFLD donors, the mice had higher TLR2 and TLR4 mRNA levels compared to control animals, which indicated these 2 factors participate in the gene regulation of disease [28]. A similar

Figure 7. GO analysis of molecular function (MF).

Figure 8. KEGG pathway enrichment analysis of common genes.
Figure 9. PPI network construction and the hub genes in different color.

Figure 10. The accordance percentage of hub genes.
phenomenon can be seen in AMI pathogenesis, in which it was reported that TLR2/1 and TLR4 agonists mediate substantial direct platelet activation in AMI patients [29,30]. Furthermore, another study showed that TLR2 or TLR4 gene defects in mice can prevent cardiac remodeling, resulting in unchanged cardiac function and geometry after AMI [31].

Leukocyte immunoglobulin-like receptors subfamily B (LILRB) members are inhibitory receptors that can negatively regulate immune cell activation spontaneous by signaling intracellular immunoreceptor tyrosine-based inhibitory motifs (ITIMs) [32]. The LILRB2 gene locates in the region of human chromosome 19q13.4, and some immune-related receptors and genes have been reported [33]. Yan et al. studied expression of B cell-associated genes in people with AMI, stable angina, and healthy controls (n=20 per group), finding that LILRB2 was significantly overexpressed in the AMI group compared with the other 2 groups [34].

CXCL1 is a neutrophil chemokine that, when overexpressed in rat liver, results in hepatic dysfunction with neutrophil infiltration [35,36]. In a recent study, the expression of CXCL1 was significantly increased in the livers of NASH mice compared with mice that had simple hepatic steatosis [37]. Saeed et al. conducted an experiment on wild-type (WT) mice fed high-fat diets to form a NAFLD model, and showed that the gene expression of CXCL1 was higher in liver tissue than in normal controls [38]. Similarly, it has been reported that the level of CXCL1 is evaluated in AMI patients [39,40]. Pordel et al. also...
reported that the plasma levels of CXCL1 chemokine in AMI patients were notably higher than in healthy participants, possibly due to AMI activating an intense inflammatory response, and release of related pro-inflammatory cytokines such as CXCL1 and CXCL5 [41,42].

Formyl peptide receptor 1 (FPR1) was first found in phagocytic leukocytes; it is a G protein-coupled 7 transmembrane cell surface receptor (GPCR). FPR1 has a variety of pathophysiologic functions such as inflammation, wound healing, glioblastoma progression, and defense against HCV infection [43,44]. The study of the relationship between NAFLD and FPR1 is incomplete, and previous research reported that liver cancer occurring at the end of NAFLD involves chronic inflammation, hepatocyte injury, and neovascularization [45]. Consist with other studies, Honda et al. induced liver necrosis in mice, finding that neutrophils quickly accumulated, but when FPR1 was inhibited, the neutrophil migration into the liver necrotic area was notably decreased [46]. Moreover, Zhou et al. studied the effect of FPR1 on ischemia/reperfusion injury in rats, showing that low expression of FPR1 depressed inflammation level, cardiomyocyte apoptosis, and ventricular remodeling [47]. Recent research also indicated that FPR1 thus might become a potential novel target in AMI treatment for improving the prognosis of MI patients [48,49].

Tyrosine kinase binding protein (TYROBP) is a regulatory protein of a variety of activated receptors in NK cells, which can bind to activated receptors in the form of noncovalent bonds, activate signal transduction, activate NK cells, and perform effective functions [50]. In a TYROBP gene knockout model, the level of pro-inflammatory cytokines was decreased, such as TNF-α, IL-10, IL-6, and MCP-1, and TYROBP was found to synthesize lipopolysaccharide and induce the production of pro-inflammatory cytokines mediated by toll-like receptor [51]. NK cells participate in the pathophysiological process of NAFLD and AMI [52,53]; suggesting that TYROBP may also play an important role through NK cells in these diseases, but further research on this is needed.

We also assessed associated gene biological functions and pathways, such as inflammatory reactions of GO terms and immune system diseases pathways, IL-17 signaling pathway, and NF-kappa B signaling pathway, which were reported to be closely involved in both diseases [54–56].

**Conclusions**

In this study, 2 AMI datasets and 1 NAFLD dataset were downloaded from GEO. After data selection, 78 intersection genes were obtained from these datasets when patient samples were compared to normal ones. Through GO and KEGG enrichment analyses, inflammation-related biological functions and pathways were obtained. Using PPI network construction, the hub genes TLR2, LILRB2, CXCL1, FPR1, TLR4, and TYROBP were selected and their diagnostic values were validated by SPSS data analysis. The outcomes of our study may provide potential targets for the prevention and treatment of NAFLD and AMI.

However, there are some limitations of our study. First, the DEGs screened by bioinformatics method can predict the occurrence of NAFLD and AMI, and our research team is working on *in vivo* and *in vitro* experiments and with clinical cases to confirm these results, which will be published in the near future. Second, for the included samples from different races and groups, the results do not necessarily apply to all populations, and research is needed in other populations. Finally, our sample sizes were relatively small, and larger-sample, multicenter research is needed.

**Acknowledgement**

The authors thank Dr Yuzhen Liang, the Second Affiliated Hospital of Guangxi Medical University, for editing the manuscript.

**Conflicts of interest**

None.
Supplementary Table 1. Seventy-eight common genes of NAFLD and AMI.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MGAM</td>
<td>FCGR2A</td>
<td>IL1R2</td>
<td>C1orf48</td>
<td>TYROBP</td>
<td>CLEC4D</td>
<td>TMCC3</td>
<td>FGR</td>
<td>SLC22A4</td>
</tr>
<tr>
<td>CSF3R</td>
<td>DUSP6</td>
<td>CSF1R1</td>
<td>CD55</td>
<td>CLEC4E</td>
<td>IL1R2</td>
<td>VNN1</td>
<td>CLEC7A</td>
<td>SLC2A3</td>
</tr>
<tr>
<td>CD36</td>
<td>FCN1</td>
<td>NCF2</td>
<td>GLUL</td>
<td>ICAM1</td>
<td>ACSL1</td>
<td>LRRK2</td>
<td>PILRA</td>
<td>CXCL1</td>
</tr>
<tr>
<td>CH3L1</td>
<td>GCA</td>
<td>TP53INP2</td>
<td>FCGR1B</td>
<td>MCEMP1</td>
<td>PDE4B</td>
<td>IL1RN</td>
<td>THBD</td>
<td>TLR4</td>
</tr>
<tr>
<td>IER3</td>
<td>CEBPD</td>
<td>IFNGR1</td>
<td>MAP3K8</td>
<td>CPVL</td>
<td>CCR2</td>
<td>CDKN1A</td>
<td>TLR8</td>
<td>MPEG1</td>
</tr>
<tr>
<td>MXD1</td>
<td>CD83</td>
<td>FCER1G</td>
<td>PFKFB3</td>
<td>S100A12</td>
<td>TREM1</td>
<td>BCL6</td>
<td>TLR2</td>
<td>G0S2</td>
</tr>
<tr>
<td>QPCT</td>
<td>PEL1</td>
<td>CXCL16</td>
<td>S100A8</td>
<td>PLAUR</td>
<td>CISH</td>
<td>SLC11A1</td>
<td>BCL2A1</td>
<td>GADD45B</td>
</tr>
<tr>
<td>MMP9</td>
<td>FPR1</td>
<td>VNN3</td>
<td>NRA2A</td>
<td>DOCK5</td>
<td>CCL20</td>
<td>VNN2</td>
<td>S100A9</td>
<td></td>
</tr>
</tbody>
</table>

References:


38. Saeed WK, Iun DW, Jang K et al: Mismatched effects of receptor interacting protein kinase-3 on hepatic steatosis and inflammation in non-alcoholic fatty liver disease. World J. Gastroenterol, 2018; 24: 5477–90.


