MEDICAL SCIENCE MONITOR

MOLECULAR BIOLOGY

e-ISSN 1643-3750
© Med Sci Monit, 2019; 25: 2993-3004
DOI: 10.12659/MSM.914044

Received: 2018.11.09
Accepted: 2019.01.07
Published: 2019.04.23

Transcriptional Regulation of Early Growth Response Gene-1 (EGR1) is Associated with Progression of Nonalcoholic Fatty Liver Disease (NAFLD) in Patients with Insulin Resistance

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

C 1 Zedong Li
E 1 Peng Yu
E 2 Jiajia Wu
B 3 Fang Tao
A 1 Jun Zhou

1 Department of Minimally Invasive Surgery, The Second Xiangya Hospital, Central South University, Changsha, Hunan, P.R. China
2 Department of General Surgery, The Second Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, P.R. China
3 Department of General Surgery, The First Affiliated Hospital of Gannan Medical University, Ganzhou, Jiangxi, P.R. China

Corresponding Author: Jun Zhou, e-mail: lizd941217@csu.edu.cn
Source of support: Departmental sources

Background: The occurrence of nonalcoholic fatty liver disease (NAFLD) is closely related to type 2 diabetes, especially in patients with insulin resistance. The purpose of this research was to elucidate the major genes and transcriptional regulation of insulin resistance in the progression of NAFLD.

Material/Methods: We downloaded the gene expression matrix of GSE89632 from Gene Expression Omnibus. Then the principal component analysis was used to identify whether the samples were clustered. Differentially expressed genes were identified by limma R package. Enrichment analysis and protein-protein interaction network was used to find potential function and screening hub genes. We further used ChIP-seq data from ENCODE to predict the transcriptional regulation of hub genes. Finally, we verified the functions of hub genes with clinical information.

Results: These hub genes were significantly enriched in "response to insulin", "response to glucose", and "fat cell differentiation". ChIP-seq data showed that EGR1 (early growth response gene-1) may play an important role in the transcriptional regulation of SOCS1 (suppressor of cytokine signaling 1), SOCS3 (suppressor of cytokine signaling 3), and Fos gene family in the liver, as the low expression of EGR1 in patients with insulin resistance may promote the occurrence and development of NAFLD. Similarly, correlation analysis showed that EGR1 was positively correlated with the expression of SOCS1, SOCS3, and the genes of Fos gene family, and EGR1 was negatively correlated with the degree of steatosis.

Conclusions: Newly identified hub genes and their transcriptional regulation may promote understanding of the molecular mechanisms underlying insulin resistance related to the progression of NAFLD and provide a new therapy target and biomarkers.

MeSH Keywords: **Diabetes Mellitus, Type 2 • Early Growth Response Protein 1 • Fatty Liver • Insulin Resistance**

Abbreviations: **NAFLD** – nonalcoholic fatty liver disease; **DEGs** – differentially expressed genes; **PPI** – protein-protein interaction; **SS** – simple steatosis; **NASH** – nonalcoholic steatohepatitis; **HC** – healthy controls; **IR** – insulin resistance; **PCA** – principal component analysis; **GEO** – Gene Expression Omnibus; **GO** – Gene Ontology; **KEGG** – Kyoto Encyclopedia of Genes and Genomes; **ENCODG** – Encyclopedia of DNA Elements; **EGR1** – early growth response gene-1

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/914044

3031 1 10 25

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Background

Fatty liver is the most common liver disease in the world. Fatty liver without overdrinking is called nonalcoholic fatty liver disease (NAFLD). NAFLD is characterized by the accumulation of lipids in more than 5.5% of liver cells and results in liver sensitivity to inflammation and injury [1], including simple steatosis (no inflammation) and nonalcoholic steatohepatitis (NASH) (with necroinflammatory lesions) [2]. Persistent liver inflammation can cause the development of hepatic fibrosis and cirrhosis, as well as increase the incidence and prevalence of liver cancer [3].

One study found that 80% of patients with NAFLD were overweight or obesity [4], and about 70% of patients with type 2 diabetes and almost all patients with type 2 diabetes with obesity have liver steatosis or NAFLD. The occurrence of NAFLD is closely related to insulin resistance (IR). Hepatic inflammation is considered to be the result of IR, and the accumulation of hepatic lipids is also closely related to hepatic IR, but the specific mechanism has not yet been clarified [5,6]. The "two-hit hypothesis" was proposed to elucidate the relationship between NAFLD and IR, where "first hit" is the result of IR, which increases the accumulation of triglycerides and the transport of free fatty acids to the liver [7]. Many studies have shown that IR is one of the key factors in the development of NAFLD. However, the molecular mechanism of IR that promotes NAFLD development remains unclear. Therefore, researching and identifying key genes that affect the occurrence and development of NAFLD by IR can help elucidate the underlying molecular mechanisms, identify new diagnostic markers and therapeutic targets, and provide new ideas for the diagnosis and treatment of diabetic patients with fatty liver.

In this study, we downloaded the microarray dataset GSE89632 [8] from the Gene Expression Omnibus (GEO) (*http://www.ncbi.nlm.nih.gov/geo/*). According to the purpose of the study, the incomplete data samples were excluded, and a total of 46 samples met the inclusion criteria. The differentially expressed genes (DEGs) between fatty liver with IR or without IR and normal liver tissue without IR were determined by limma package of the R programming language. The protein-protein interaction network (PPI) of DEGs was constructed to identify the core genes in the process of fatty liver and the core genes of IR affecting the development and progression of fatty liver. Then, related pathways and functions of DEGs were analyzed using the Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG). The purpose of the present study was to investigate the hub genes and pathways of fatty liver development, and the ways IR works in the development of fatty liver. A new understanding of these mechanism will give us a solid base for future treatment of the fatty liver of patients with IR.

Table 1. Number of samples by insulin resistance and non-insulin resistance in NASH, SS, and HC groups.

Group	NASH	SS	HC
IR	11	9	–
NonIR	3	7	16

NASH – nonalcoholic steatohepatitis; SS – simple steatosis; HC – healthy control, IR – insulin resistance. A total of 46 samples were included in the study according to whether the patient had diabetes and insulin resistance and excluded healthy control samples with insulin resistance, including 11 cases of NASH with IR, 9 cases of SS with IR, 3 cases of NASH, 7 cases of SS, and 16 cases of HC.

Material and Methods

Microarray data analysis

GSE89632 is microarray data in the GEO database (*http://www.ncbi.nlm.nih.gov/geo/*), which is based on Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip. The gene expression matrix was downloaded from the Series Matrix File(s) directly. The data set included 63 samples: 20 patients with simple steatosis (SS), 9 patients with nonalcoholic steatohepatitis (NASH), and 24 healthy controls (HC). To determine whether a patient had diabetes and IR we used the IR index HOMA-IR >2.7 [9]. Healthy control with IR was excluded. A total of 46 samples were included in the study (Table 1).

Study design

We used prcomp function of R (version 3.5.0) as the principal component analysis based on the characteristics of IR and diabetes in the sample and the fatty liver type of the sample to determine whether there was a significant difference in gene expression between samples with or without IR and diabetes and different types of fatty liver.

Expression analysis of DEGs

The gene expression matrix of the samples included in the analysis was extracted from the total gene expression matrix, and DEGs were identified by the limma R package [10] between: NASH with IR tissue samples and the HC tissue samples, SS with IR tissue samples and HC tissue samples, NASH without IR tissue samples and HC tissue samples, and SS without IR tissue samples and HC tissue samples. Genes with |log2Fold Change| >1.5 and $P<0.05$ were considered as potential meaningful DEGs. Subsequently, common DEGs of 4 groups and common DEGs of NASH with IR to HC and SS with IR to HC were screened for further analysis.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

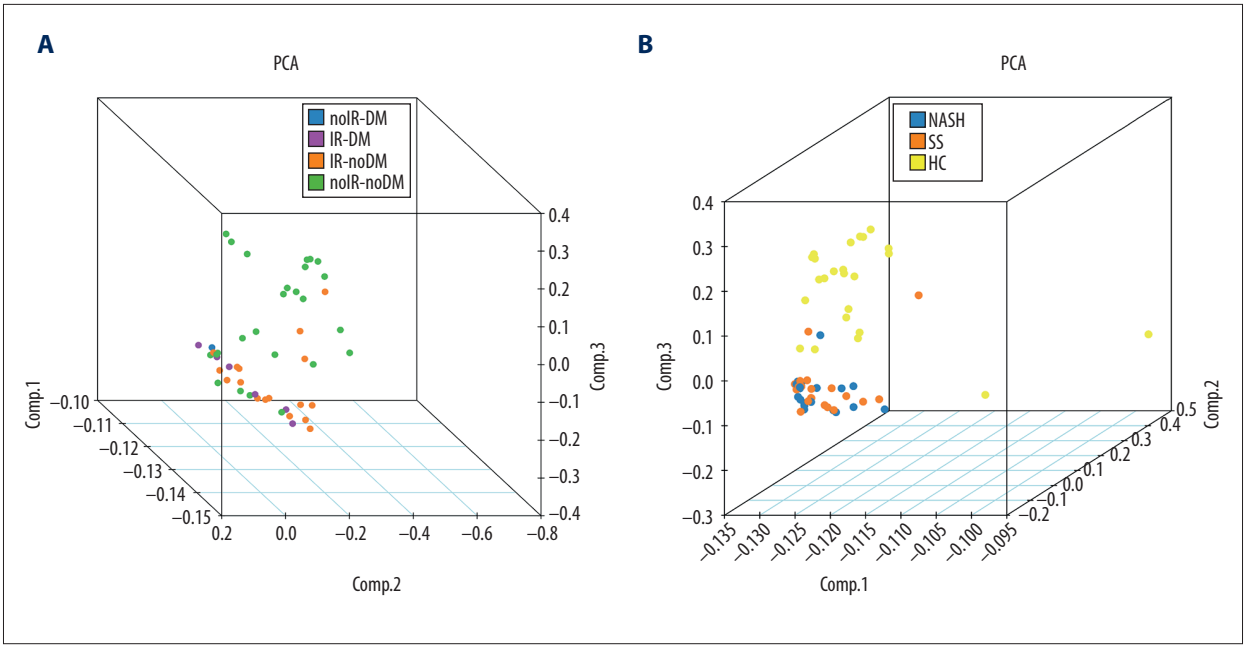


Figure 1. Different groups were set for principal component analysis. (**A**) Analysis was based on whether the samples had diabetes or insulin resistance. The results showed that the samples with insulin resistance exhibited clustering, and the samples with diabetes showed clustering too. (**B**) According to the sample tissue type grouping, the analysis results show that SS and NASH are clustered, and HC samples are clustered separately. SS – simple steatosis; NASH – nonalcoholic steatohepatitis; HC – healthy controls.

GO and KEGG enrichment analysis

The clusterProfiler [11] is a method used to analyze and visualize functional profiles of genomic coordinates, genes, and gene clusters. To identified DEGs we used clusterProfiler (version 3.8.0) for GO [12] and KEGG [13] enrichment analyses, and $P<0.05$ was considered a statistically significant difference.

PPI network construction and module analysis

We used the online Search Tool for the Retrieval of Interacting Genes (STRING; 2017 release) to evaluate the interactive relationships among the differentially expressed genes. K-core decomposition was used to help screen out stable structures in the network, narrow the scope of research, and identify hub-gene. Cytoscape (version 3.6.1) was used for PPI network visualization, where the Molecular Complex Detection plugin (MCODE; version 1.31) completes k-core resolution. In addition, the MCODE was used to screen out significant modules based on the constructed PPI networks with the criteria of degree cutoff=2, node density cutoff=0.1, node score cutoff=0.2, and k-core=2.

Regulation of transcription factors on hub genes

ChIP-seq was used as a classical experiment for finding transcription factor binding sites (TFBS). ChIP-seq experiments can find a target gene regulated by transcription factors and check the ChIP-seq signal of the transcription factor near the gene. Encyclopedia of DNA Elements (ENCODE, *https://www.encodeproject.org*) has a lot of ChIP-seq data with high quality, based on a large amount of ChIP-seq data, transcriptional regulation of transcription factors can be predicted. According to the peak of the transcription factor near the transcription start site (TSS), it can be speculated that the transcription factor binds to the promoter, thus participating in the transcriptional regulation of the gene.

Spearman's correlation coefficients calculation

The direction of association between mRNA and clinical information were calculated by Spearman's rank correlation coefficient from 43 samples that were selected from GSE89632 with full clinical information available. If a biochemical index tends to increase when mRNA increases, the Spearman's correlation coefficient is positive. If a biochemical index tends to decrease when mRNA increases, the Spearman's correlation coefficient is negative. R (version 3.5.0) was used to calculate the Spearman's rank correlation coefficient between each of the 2 biochemical indexes or mRNAs. By checking the threshold table, at n=43, we had 99% confidence that the 2 random variables were related when the Spearman's rank correlation coefficient was greater than 0.391.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

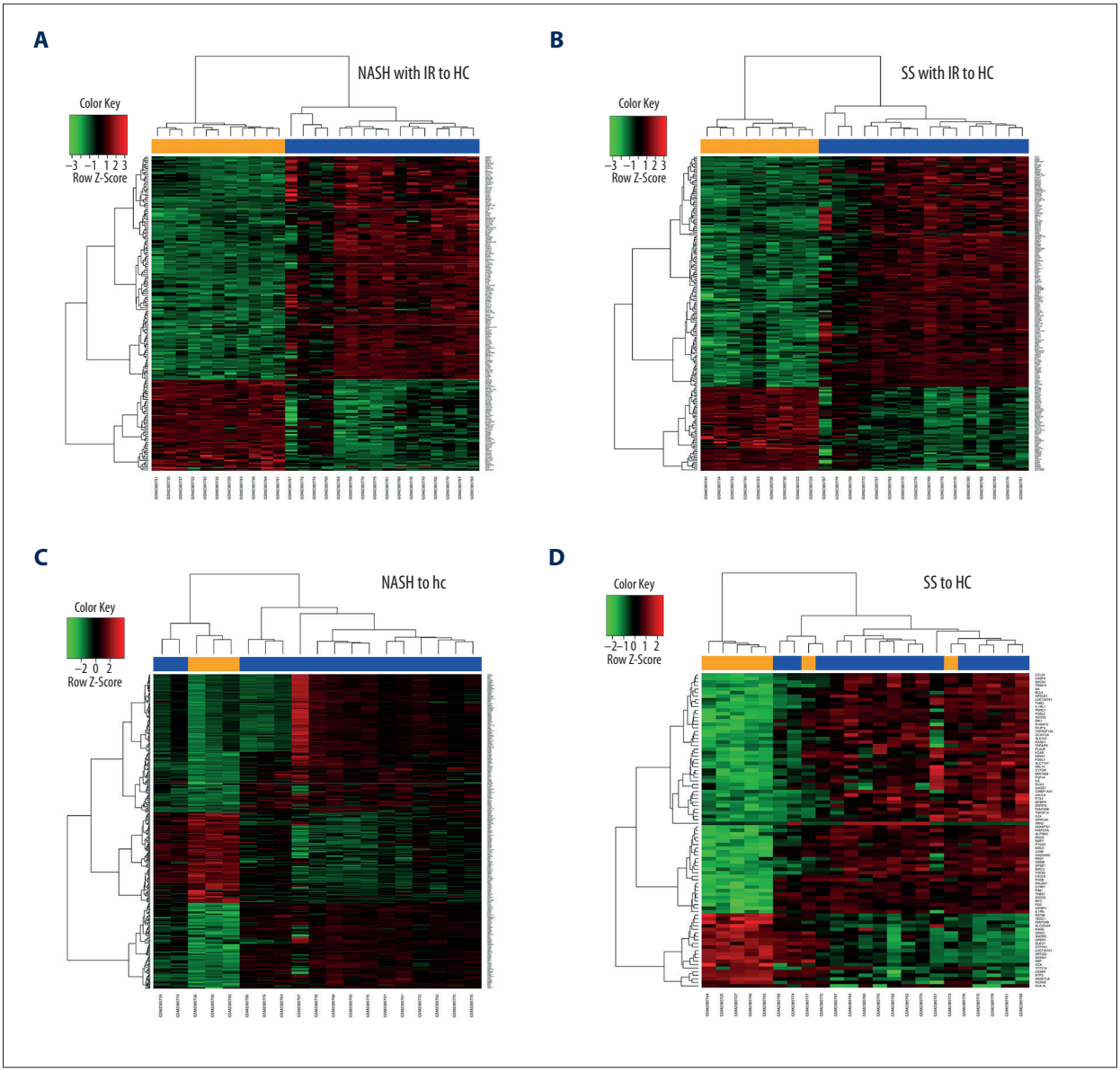


Figure 2. (**A–D**) Heat map of the differentially expressed genes of 4 groups that |log2Fold Change| >1.5 and *P*<0.05. Yellow represent NASH with IR, SS with IR, NASH and SS; blue represent HC. SS – simple steatosis; NASH – nonalcoholic steatohepatitis; HC – healthy controls; IR – insulin resistance.

Results

Study design

Grouped according to whether the samples had IR and diabetes, the results showed that the samples were clustered according to whether they had IR characteristics (Figure1A). According to the fatty liver type of the samples, SS and NASH were clustered, and HC samples were clustered (Figure 1B). We speculated that there were differences in gene expression between IR and non-IR samples, but no significant differences in gene expression between NASH and SS samples. Accordingly, we designed the groupings (Table 1) to compare the DEGs between fatty liver and normal liver to try to find the genes that play a major role in patients with IR.

The identification of DEGs

DEGs were identified by limma R package. The group NASH with IR to HC identified a total of 196 DEGs: 57 genes were upregulated, and 139 genes were downregulated. The group SS with IR to HC identified a total of 206 DEGs: 55 genes were upregulated, and 151 genes were downregulated. The group NASH to HC identified a total of 352 DEGs: 101 genes were

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

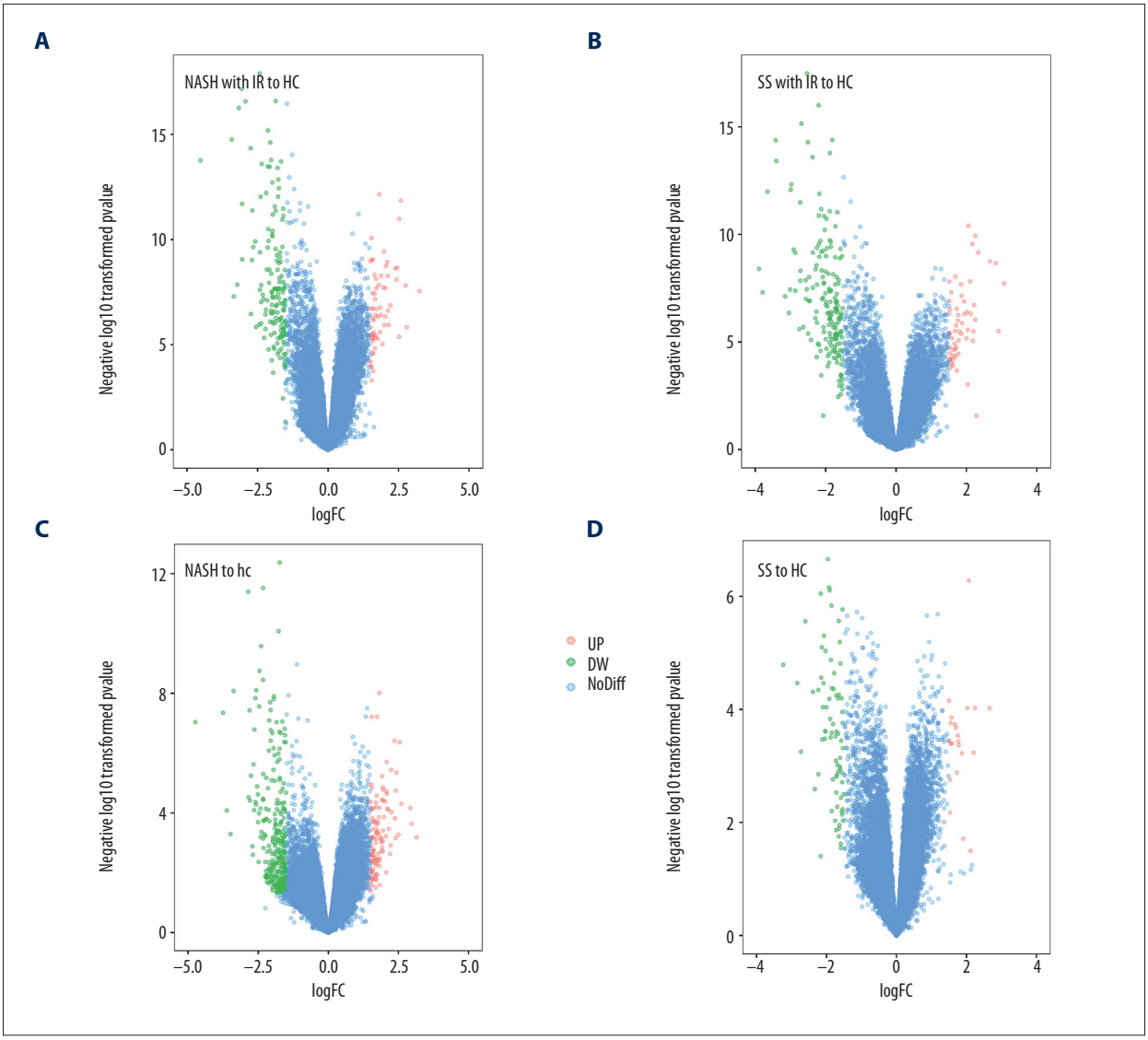
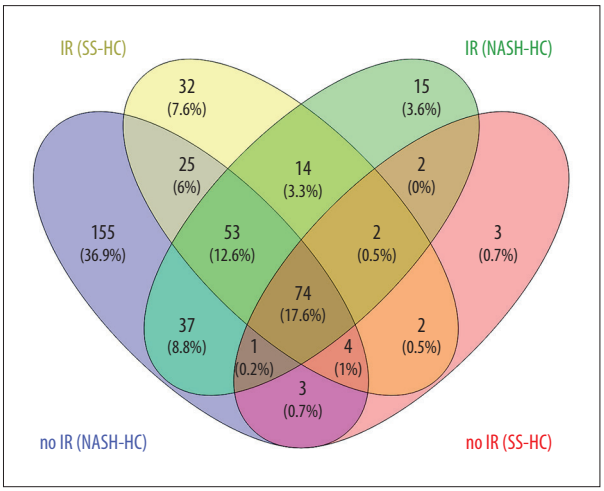


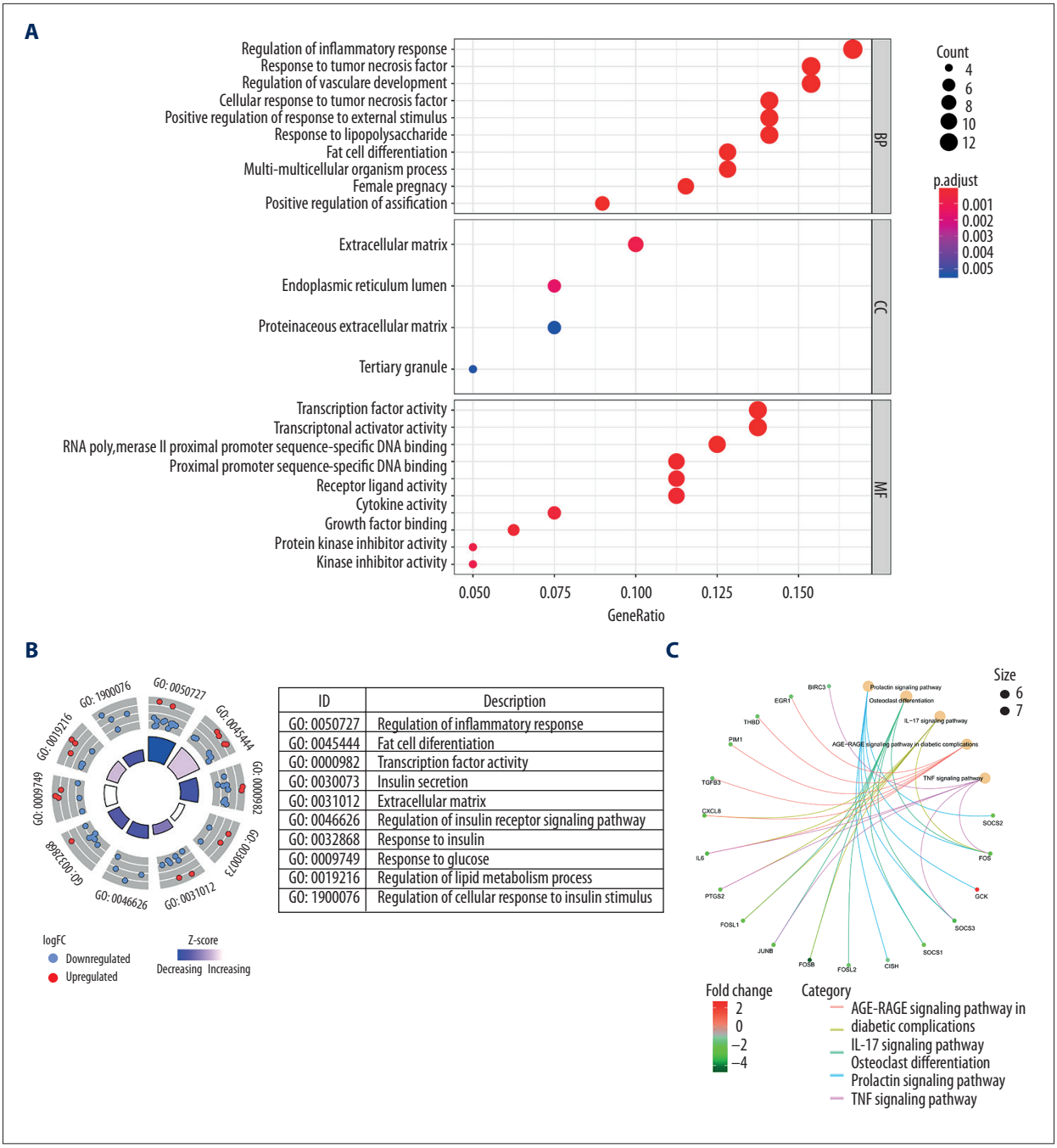
Figure 3. (**A–D**) Volcano plot of the differentially expressed genes of 4 groups. Red indicates DEGs with a log2FC >1.5 and *P*<0.05, green indicates DEGs with a log2FC >1.5 and *P*<0.05, FC, fold change. DEGs – differentially expressed genes.



upregulated, and 251 genes were downregulated. The group SS to HC identified a total of 89 DEGs: 21 genes were upregulated, and 68 genes were downregulated. Four expression heat maps (Figure 2) and 4 volcano plots (Figure 3) for the identified DEGs were constructed.

Figure 4. Using Venn diagram to display differentially expressed genes: microarray analysis of liver tissue showed that in liver with IR, comparing SS-to-HC and NASH-to-HC, and that in liver without IR, comparing SS-to-HC and NASH-to-HC, that 74 genes were found in DEGs of 4 groups; 14 genes were found in the groups with IR individually. SS – simple steatosis; NASH – nonalcoholic steatohepatitis; HC – healthy controls; IR – insulin resistance; DEGs – differentially expressed genes.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]



ID	Description
GO: 0050727	Regulation of inflammatory response
GO: 0045444	Fat cell diferentiation
GO: 0000982	Transcription factor activity
GO: 0030073	Insulin secretion
GO: 0031012	Extracellular matrix
GO: 0046626	Regulation of insulin receptor signaling pathway
GO: 0032868	Response to insulin
GO: 0009749	Response to glucose
GO: 0019216	Regulation of lipid metabolism process
GO: 1900076	Regulation of cellular response to insulin stimulus

Figure 5. Functional enrichment analysis of 74 DEGs that associated with the development of fatty liver and the function of IR in fatty liver development. (**A**) GO analysis revealed that DEGs were significantly enriched in biological process terms (BP), cell component terms (CC), and molecular function terms (MF). (**B**) The GO terms included insulin, glucose, and lipid. (**C**) KEGG analysis were used to obtain significantly enriched KEGG terms. Fold change: LogFC of NASH with IR to HC group were used to show fold change. DEGs – differentially expressed genes; IR – insulin resistance; GO – Gene Ontology; KEEG – Kyoto Encyclopedia of Genes and Genomes; NASH – nonalcoholic steatohepatitis; HC – healthy controls.

Among the DEGs, 74 differential genes were found in all 4 groups, which suggested that these 74 DEGs play a major role in the development of fatty liver. There were 14 DEGs that were found only in 2 groups: the NASH with IR to HC group and the SS with IR to HC group, which suggests that these 14 DEGs could be key genes for IR in fatty liver formation (Figure 4).

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

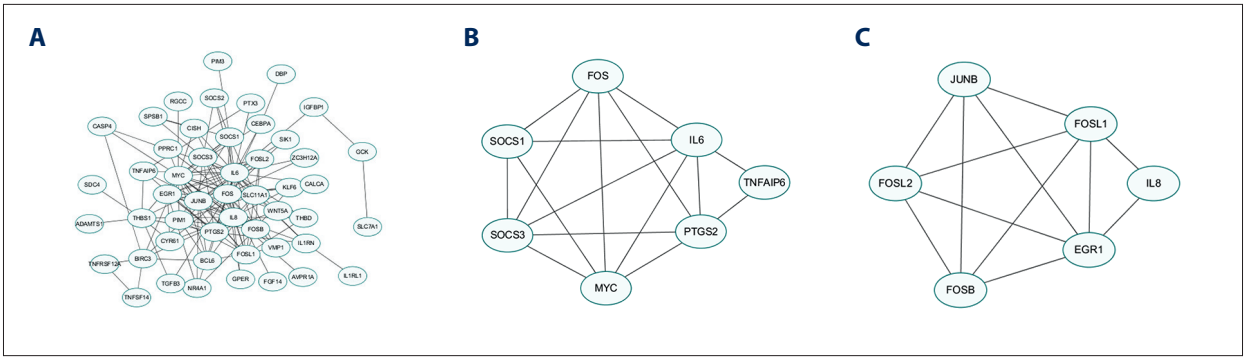


Figure 6. (**A**) Protein-protein interaction network of 74 DEGs was associated with the development of fatty liver and the function of IR in fatty liver development. (**B**) Module 1 identified by MCODE. (**C**) Module 2 identified by MCODE. DEGs – differentially expressed genes; IR – insulin resistance; MCODE – Molecular Complex Detection.

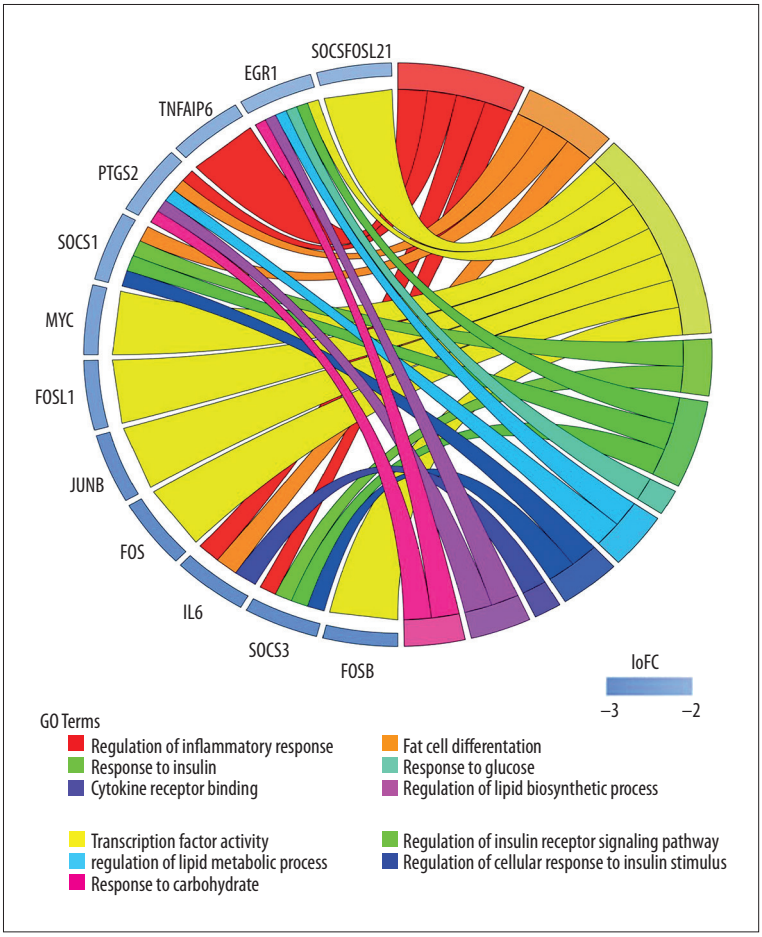


Figure 7. Using MCODE to screen out 2 significantly enriched modules, resulted in a total of 12 genes. These 12 hub genes were significantly enriched in the GO terms displayed. MCODE – Molecular Complex Detection; GO – Gene Ontology.

Functional enrichment analysis

To identify the GO and KEGG pathways that had the most significant involvement with fatty liver related differential genes and IR related differential genes, DEGs were analyzed by clusterProfiler R package for GO and KEGG pathway analysis. In biological process terms of GO analysis, the DEGs were mainly enriched in "positive regulation of ossification", "female pregnancy", "multi-multicellular organism process", "fat cell differentiation", and "response to lipopolysaccharide". In cell component terms, DEGs were mainly enriched in "tertiary granule", "proteinaceous extracellular matrix", "endoplasmic reticulum lumen", and "extracellular matrix". In molecular function terms, DEGs were mainly enriched in "transcription factor activity", "transcriptional activator activity", "RNA polymerase II in short promoter sequence-specific DNA binding",

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

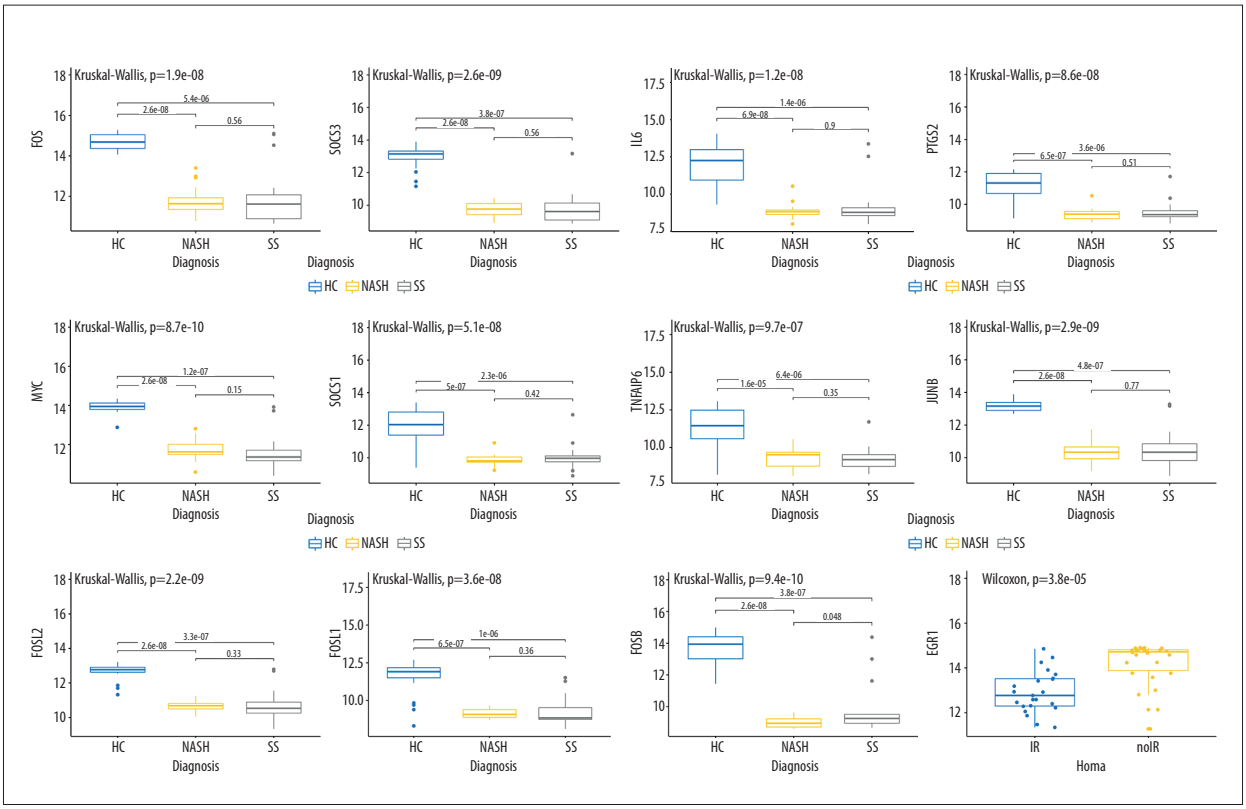


Figure 8. There was no significant difference in the expression of 11 fatty liver related hub genes between NASH and SS ($P>0.05$). There was a significant difference in the expression of HC compared with NASH and SS ($P<0.05$).There was a significant difference in the expression of EGR1 between IR and nonIR ($P<0.05$). SS – simple steatosis; NASH – nonalcoholic steatohepatitis; HC – healthy control, EGRI – early growth response gene-1; IR – insulin resistance.

and "proximal promoter sequence-specific DNA binding" (Figure 5A). Some of the GO terms include keywords such as insulin, glucose, and lipid, which may be related to the development of fatty liver, including "regulation of inflammatory response", "fat cell differentiation", "insulin secretion", "regulation of insulin receptor signaling pathway", "response to insulin", and "regulation of cellular response to insulin stimulus" (Figure 5B). KEGG pathway analysis demonstrated that DEGs were significantly enriched in "AGE-RAGE signaling pathway in diabetic complications", "IL-17 signaling pathway", "osteoclast differentiation", "prolactin signaling pathway", and "TNF signaling pathway"(Figure 5C).

PPI network construction and module analysis

To further understand the role of GEGs related to fatty liver and IR in the development of fatty liver, and the effect of IR on the development of fatty liver, we constructed a PPI network through String (Figure 6A) to reveal the interaction between DEGs. There were 82 nodes and 147 edges in the network, which was imported into Cytoscape and analyze using the plugin MCODE. Two significant modules which gained the highest MCODE score were selected (Figure 6B and Figure 6C), with 12 hub genes in total.

Notably, EGR1 (early growth response gene-1) belonged to the DEGs that were only found in the 2 groups: the NASH with IR to HC group and the SS with IR to HC group. The 12 hub genes in the modules were mainly enriched in "response to insulin", "response to glucose", and "fat cell differentiation" (Figure 7).

These results suggest that EGR1 might be the hub gene that plays a role in IR and fatty liver development, and the remaining 11 hub genes associated with fatty liver development. These results promoted us to hypothesize that these hub genes might be involved in the mechanism of IR promote fatty liver development. We included all NASH, SS, and HC samples in GSE89632, the differences in gene expression between the 11 fatty liver-related hub genes in NASH, SS, and HC were compared. The results showed that the expression of HC was significantly different from that of NASH and SS ($P<0.05$), and there was no significant difference between NASH and SS ($P>0.05$). When we included all the IR and non-IR samples in GSE89632, the expression levels of EGR1 in IR and nonIR groups were significantly different ($P<0.05$) (Figure 8).

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

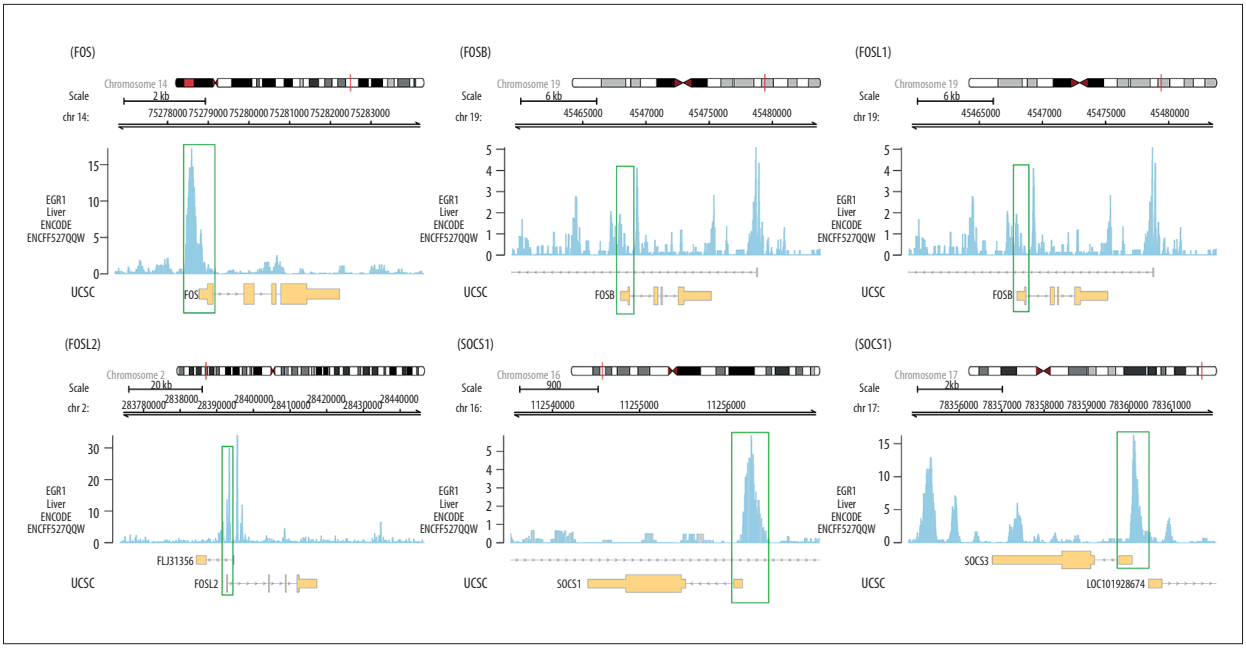


Figure 9. The transcription initiation site (TSS) is located in the green box, and there is a peak of the transcription factor EGR1 near the TSS. It is speculated that the transcription factor EGR1 binds to the promoter, thereby participating in the transcriptional regulation of SOCS1, SOCS3 and the genes of Fos gene family. EGR1 – early growth response gene-1; SOCS1 – suppressor of cytokine signaling 1; SOCS3 – suppressor of cytokine signaling 3.

Transcriptional regulation of hub genes by EGR1

EGR1 is a transcription factor. On the basis of the aforementioned analysis, we presumed that decreased EGR1 expression in IR patients promotes fatty liver formation, mainly through the transcriptional regulation of EGR1 on fatty liver-related hub genes. To further validate the conjecture, and to clarify the mechanism by which IR promotes fatty liver formation, we download ChIP-seq data of liver on EGR1-human (ENCFF527QQW) from ENCODE, to find out which genes the transcription factor EGR1 binds to. The peak of EGR1 near the TSS of the gene indicates that EGR1 has a binding signal at this position, and the transcription factor EGR1 is predicted to bind to the promoter, thereby participating in the transcriptional regulation of the genes. The results showed that EGR1 had a binding signal near the TSS of SOCS1 (suppressor of cytokine signaling 1), SOCS3 (suppressor of cytokine signaling 3), and the genes of Fos gene family (the Fos gene family consists of 4 members: FOS, FOSB, FOSL1, and FOSL2). The transcriptional regulation of EGR1 might be a key mechanism for fatty liver development of patients with IR (Figure 9).

Spearman's correlation coefficients calculation

Our study found that IR may play a role in the progression of fatty liver through low expression of EGR1; the low expression of EGR1 leads to the low expression of SOCS1, SOCS3, and the genes of Fos gene family through transcriptional regulation. To further confirm these findings, the Spearman's correlation coefficients of each of the 2 indexes were calculated from 43 samples to identify negative reactions between EGR1 and fatty liver. We set a threshold of |0.391| to assess the correlation; any group with the correlation coefficient beyond |0.391| was considered relevant. According to the Spearman's correlation coefficients, the following conclusions were obtained. 1) The positive correlation between EGR1 expression and the gene expression of SOCS1, SOCS3, and the genes of Fos gene family was observed (Kappa coefficient >0.6, $P<0.05$). 2) Significant negative correlation was observed between gene expression of EGR1 and degree of steatosis (Kappa coefficient=–0.74, $P<0.05$) (Figure 10).

Discussion

Hepatic insulin resistance (IR) and development of nonalcoholic fatty liver disease (NAFLD) are closely related, and IR has been found to be an important driving force of type 2 diabetes [5]. IR could be a predictive factor for type 2 diabetes complicated with NAFLD [14,15], as 70% of patients with type 2 diabetes have been reported to have hepatic steatosis or NAFLD [16,17]. It is worth noting that one study reported that nearly all obese type 2 diabetes patients had hepatic steatosis or NAFLD [18]. High throughput sequencing has been applied widely, and may help the identification of target genes for understanding the mechanism of IR promoting NAFLD

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

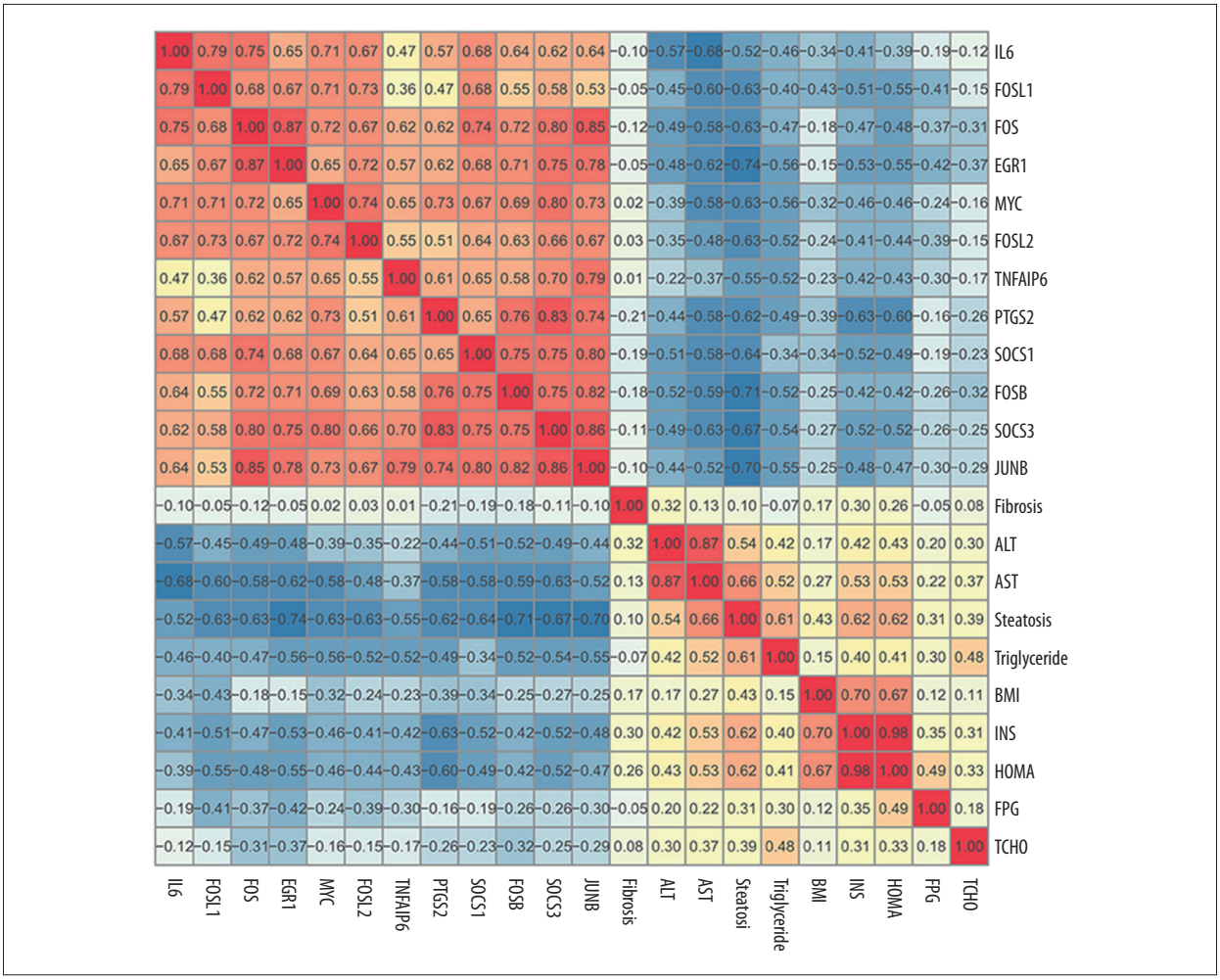


Figure 10. Incomplete samples were excluded. A total of 43 samples were selected. Correlation analysis was performed based on clinical information and gene expression levels. The correlation coefficient between each index was calculated by R (version 3.5.0). By querying the boundary value table, at n=43, when the Spearman's rank correlation coefficient ≥|0.391|, we had 99% confidence that 2 random variables were related. According to these results, EGR1 was positively correlated with SOCS1, SOCS3, and the genes of Fos gene family, and EGR1 was negatively correlated with the degree of steatosis. EGR1 – early growth response gene-1; SOCS1 – suppressor of cytokine signaling 1; SOCS3 – suppressor of cytokine signaling 3.

development and diagnosing or treating NAFLD in type 2 diabetes patients with IR.

In the present study, we downloaded gene profile datasets from GEO, divide the sample into 5 groups: healthy control (HC), simple steatosis (SS), nonalcoholic steatohepatitis (NASH), SS with IR, and NASH with IR). Bioinformatics analysis was performed, resulting in identification of 14 IR-related DEGs and 74 NASH-related DEGs. Functional enrichment analysis showed that DEGs were mainly enriched in "regulation of inflammatory response", "fat cell differentiation", "response to lipopolysaccharide", and "transcription factor activity". The pathways mainly enriched were "AGE-RAGE signaling pathway in diabetic complications", and "IL-17 signaling pathway"; EGR1 and FOSL1, the hub genes were enriched in these pathways. To further study the DEGs, PPI was constructed, and then the hub genes are screened from the PPI network by the plugin MCODE in which IR is considered to play a role through the key differential gene EGR1.

In previous studies, mechanisms or key genes of NAFLD development or IR have always been studied separately. We combined the analysis of NAFLD and IR for the purpose of studying the role of IR in the formation and development of NAFLD. Bioinformatics analysis found that IR in the development of NAFLD worked by decreasing EGR1 expression. EGR1 is a member of the immediate early gene family, its main feature is the recognition of the zinc finger domain of the highly

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

conserved consensus GC-rich nucleotide sequence (GCG(G/T) GG GCG) [19,20]. Previous studies have reported that the deletion of EGR1 increases sensitivity to lipopolysaccharide and prevents ethanol-induced fatty liver formation [21,22]. Moreover, in insulin-resistant adipocytes, EGR1 is involved in the IR-associated Egr-1/GGPPS/Erk1/2 pathway [23].

The research screened 11 hub genes associated with fatty liver formation, including IL6, PTGS2, MYC, TNFAIP6, JUNB, SOCS1, SOCS3, and Fos gene family. It is worth noting that it has been experimentally confirmed that the overexpression of SOCS1 and SOCS3 in the liver causes an increase in IR and is a key regulator of fatty acid synthesis in the liver [24]. The binding site of EGR1 is present in the SOCS1 promoter and EGR1 is an important transcriptional regulator of SOCS1 [25]. To further investigate the mechanism by which IR promotes the formation of NAFLD, we downloaded the ChIP-seq data of liver from ENCODE. Bioinformatics analysis showed that the transcription factor EGR1 may play a role in the regulation of fatty liver formation by regulating the transcription of SOCS1, SOCS3, and the genes of Fos gene family. Our analysis results were further confirmed by correlation analysis: the correlation between EGR1 expression and the gene expression of SOCS1, SOCS3, and the genes of Fos gene family was positive, and it was negative between gene expression of EGR1 and degree of steatosis. In this study, we found that the expression levels of EGR1 and other hub genes in liver tissue of patients with IR and fatty liver were reduced. However, some previous studies have shown that the increased expression of EGR1, SOCS1, and SOCS3 may play a major role in IR and fatty liver formation. The reason for the low expression of EGR1, SOCS1, and SOCS3 in fatty liver samples in our study is worth exploring further.

Our study identified hub genes and transcriptional regulation which may be involved in IR promoting NAFLD progression. These results may help us better understand the function of IR in NAFLD development of type 2 diabetes patients and provide a series of potential biomarkers.

Conclusions

In this study, we found 12 hub genes that may play a key role in the formation of fatty liver. In particular, EGR1 may be a key gene in which IR promotes fatty liver formation. In diabetic patients with IR, the transcription factor EGR1 may affect the formation of fatty liver by regulating the transcription of SOCS1, SOCS3, and the genes of Fos gene family. However, since this study is based on bioinformatics analysis, further experimental studies are needed to validate our conclusions.

Ethical Statement

The data of this study are from GEO and ENCODE database, and do not involve animal experiments and human specimens, no ethics related issues.

Acknowledgements

We thank the researchers who gave their data for this analysis.

Conflicts of interest

None.

References:

1. Hotamisligil GS: Inflammation and metabolic disorders. Nature, 2006, 444(7121): 860–67
2. Andronescu CI, Purcarea MR, Babes PA: Nonalcoholic fatty liver disease: Epidemiology, pathogenesis and therapeutic implications. J Med Life, 2018; 11(1): 20–23
3. Angulo P, Lindor KD: Non-alcoholic fatty liver disease. J Gastroenterol Hepatol, 2002; 17Suppl: S186–90
4. Diehl AM, Day C: Cause, pathogenesis, and treatment of nonalcoholic steatohepatitis. N Engl J Med, 2017; 377(21): 2063–72
5. Petersen MC, Shulman GI: Roles of diacylglycerols and ceramides in hepatic insulin resistance. Trends Pharmacol Sci, 2017; 38(7): 649–65
6. Funke A, Schreurs M, Aparicio-Vergara M et al: Cholesterol-induced hepatic inflammation does not contribute to the development of insulin resistance in male LDL receptor knockout mice. Atherosclerosis, 2014; 232(2): 390–96
7. Sheng X, Che H, Ji Q et al: The relationship between liver enzymes and insulin resistance in type 2 diabetes patients with nonalcoholic fatty liver disease. Horm Metab Res, 2018; 50(5): 397–402
8. Arendt BM, Comelli EM, Ma DW et al: Altered hepatic gene expression in nonalcoholic fatty liver disease is associated with lower hepatic n-3 and n-6 polyunsaturated fatty acids. Hepatology, 2015; 61(5): 1565–78
9. Geloneze B, Vasques AC, Stabe CF et al: HOMA1-IR and HOMA2-IR indexes in identifying insulin resistance and metabolic syndrome: Brazilian Metabolic Syndrome Study (BRAMS). Arq Bras Endocrinol Metabol, 2009; 53(2): 281–87
10. Ritchie ME, Phipson B, Wu D et al: limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res, 2015; 43(7): e47
11. Yu G, Wang L-G, Han Y, He Q-Y: ClusterProfiler: An R package for comparing biological themes among gene clusters. OMICS, 2012; 16(5): 284–87
12. Gene Ontology Consortium: The Gene Ontology (GO) project in 2006. Nucleic Acids Res 2006; 34(Database issue): D322–26
13. Ogata H, Goto S, Sato K et al: KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res, 1999; 27(1): 29–34
14. Bugianesi E, McCullough AJ, Marchesini G: Insulin resistance: A metabolic pathway to chronic liver disease. Hepatology, 2005; 42(5): 987–1000
15. Tiikkainen M, Tamminen M, Hakkinen AM et al: Liver-fat accumulation and insulin resistance in obese women with previous gestational diabetes. Obes Res, 2002; 10(9): 859–67
16. Leite NC, Salles GF, Araujo AL et al: Prevalence and associated factors of non-alcoholic fatty liver disease in patients with type-2 diabetes mellitus. Liver Int, 2009; 29(1): 113–19
17. Schindhelm RK, Heine RJ, Diamant M: Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. Diabetes Care, 2007; 30(9): e94; author reply e95

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

18. Silverman JF, O'Brien KF, Long S et al: Liver pathology in morbidly obese patients with and without diabetes. Am J Gastroenterol, 1990; 85(10): 1349–55
19. Hamilton TB, Borel F, Romaniuk PJ: Comparison of the DNA binding characteristics of the related zinc finger proteins WT1 and EGR1. Biochemistry, 1998; 37(7): 2051–58
20. Silverman ES, Collins T: Pathways of Egr-1-mediated gene transcription in vascular biology. Am J Pathol, 1999; 154(3): 665–70
21. McMullen MR, Pritchard MT, Wang Q et al: Early growth response-1 transcription factor is essential for ethanol-induced fatty liver injury in mice. Gastroenterology, 2005; 128(7): 2066–76
22. Derdak Z, Villegas KA, Wands JR: Early growth response-1 transcription factor promotes hepatic fibrosis and steatosis in long-term ethanol-fed Long-Evans rats. Liver Int, 2012; 32(5): 761–70
23. Shen N, Yu X, Pan FY et al: An early response transcription factor, Egr-1, enhances insulin resistance in type 2 diabetes with chronic hyperinsulinism. J Biol Chem, 2011; 286(16): 14508–15
24. Ueki K, Kondo T, Tseng YH, Kahn CR: Central role of suppressors of cytokine signaling proteins in hepatic steatosis, insulin resistance, and the metabolic syndrome in the mouse. Proc Natl Acad Sci USA, 2004; 101(28): 10422–27
25. Mostecki J, Showalter BM, Rothman PB: Early growth response-1 regulates lipopolysaccharide-induced suppressor of cytokine signaling-1 transcription. J Biol Chem, 2005; 280(4): 2596–605

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]