BJMG 25 (1), 2022 • 25-34
DOI: 10.2478/bjmg-2022-0006
sciendo
ORIGINAL ARTICLE

IDENTIFICATION OF KEY TARGET GENES AND PATHWAY ANALYSIS IN NONALCOHOLIC FATTY LIVER DISEASE VIA INTEGRATED BIOINFORMATICS ANALYSIS

Chen X.[1], Zhang L.[2], Wang Y.[1], Li R.[1], Yang M.[1], Gao L.[3*]

[*]**Corresponding Author:** Lei Gao, MD, College of Basic Medicine, Changchun University of Chinese Medicine, 1035 Boshuo, Road, Jingyue District, Changchun City, Jilin Province, 130117, China; Tel:+ 86-431-8604 5309, Email: gaolei790708@163.com

ABSTRACT

Purpose: This study aimed at exploring the mechanisms underlying nonalcoholic fatty liver disease (NAFLD) and developing new diagnostic biomarkers for nonalcoholic steatohepatitis (NASH). **Methods**: The microarray dataset GES83452 was downloaded from the NCBI-GEO database, and the differentially expressed RNAs (DERs) were screened between the NAFLD and non-NAFLD samples of the baseline and 1-year follow-up time point group based on the Limma package. **Results**: A total of 561 DERs (268 downregulated and 293 upregulated) were screened in the baseline time point group, and 1163 DERs (522 downregulated and 641 upregulated) were screened in the 1-year follow-up time point group. A total of 74 lncRNA–miRNA pairs and 523 miRNA–mRNA pairs were obtained in order to construct a lncRNA–miRNA–mRNA regulatory network. Subsequently, functional enrichment analysis revealed 28 GO and 9 KEGG pathways in the ceRNA regulatory network. *LEPR* and *CXCL10* are involved in the Cytokine–cytokine receptor interaction (P = 1.86E-02), and the *FOXO1* is involved in both the insulin signaling pathway (P = 1.79E-02) and the pathways in cancer (P = 2.87E-02). **Conclusion:** *LEPR*, *CXCL10*, and *FOXO1* were the characteristic target genes for NAFLD.

Key words: NAFLD; ceRNA regulation network; PharmGKB

[1] Endocrine Metabolic Disease Section, The Affiliated Hospital to Changchun University of Chinese Medicine, Changchun City, Jilin Province, 130021, China

[2] Department of Dermatology, The Affiliated Hospital to Changchun University of Chinese Medicine, Changchun City, Jilin Province, 130021, China

[3] College of Basic Medicine, Changchun University of Chinese Medicine, 1035 Boshuo Road, Jingyue District, Changchun City, Jilin Province, 130117, China

INTRODUCTION

Nonalcoholic fatty liver disease (NAFLD) is the most common type of chronic liver disease with a prevalence rate of 25% worldwide (1). Several lifestyle-related factors are associated with incident fatty liver such as alcohol intake, lower physical activity, smoking, and shift work. Poor lifestyle choices are often the main cause of fatty liver, these include smoking, drinking, lack of physical activity, and shift work, etc. In addition, high triglycerides, type 2 diabetes mellitus, obesity, and hypertension are associated with incident fatty liver. Therefore, lifestyle modification is strongly recommended to prevent fatty liver (2, 3). It is difficult to detect this problem in the earlier stages of the disease, and may thus further develop into advanced liver diseases, such as cirrhosis and hepatocellular carcinoma, bringing forth clinical challenges to the treatment of NAFLD (4). In the literature, the severity of NAFLD in patients with type 2 diabetes and obesity will be significantly affected, increasing the degree of deterioration of liver fibrosis and the possibility of further development of end-stage liver disease (5-7). Likewise, studies have shown that when NAFLD patients suffer from cardiovascular diseases and dyslipidemia, these factors have a negative impact on the natural progression of NAFLD (8-10). Nearly 40% of patients with NAFLD die of complications, as previously reported (1). However, the detailed mechanisms under which NAFLD develops remain largely unknown. Diet adjustment and weight loss can improve NAFLD, but it is difficult to maintain. Moreover, the theory of insulin resistance has been widely accepted clinically. Insulin sensitizers have a certain therapeutic effect, but they can cause adverse reactions such as increased body weight and its therapeutic target is too limited. Therefore, this study aimed at finding new molecular targets to provide a theoretical basis for new and effective treatment methods of NAFLD.

Long non-coding RNA (lncRNA) is the main component of the human transcriptome. Long non-coding RNA plays an important role in regulating cell migration, proliferation, invasion, and metastasis. It can also be used as a diagnostic marker or therapeutic target for malignant tumors and other diseases. Competitive endogenous RNA (ceRNA) is a transcript with the same microRNA (miRNA) response element, which binds to miRNA to compete and regulate its target gene, thereby affecting the biological behavior of the disease. Studies have confirmed that the mutual regulation between lncRNA and miRNA and their downstream target genes plays an important role in the occurrence and development of diseases (11).

The inflammatory component of nonalcoholic steatohepatitis (NASH) is more difficult to capture with ultrasound-assisted techniques. Although more and more technologies are applied in clinical practice, such as quantitative and contrast-enhanced ultrasound, there are still many technical barriers to be broken; and not all technologies have been successful in clinical and research practice (12). Due to the limitations of liver biopsy, searching for non-invasive and reliable diagnostic biomarkers for NAFLD is a priority for current research. Bioinformatics has been widely used to explore biomarkers of different diseases, but NAFLD-related biomarkers need to be further explored to help the early diagnosis and prognosis evaluation of NAFLD (13).

In this study, human samples from the Gene Expression Omnibus (GEO) database were used to identify key genes related to NAFLD and non-NASH samples during the baseline and 1-year follow-up time point, and to explore the underlying mechanism of NAFLD and develop new NAFLD diagnostic biomarkers. Then, the lncRNA–miRNA–mRNA network related to NAFLD was constructed by mapping the differentially expressed RNAs (DERs) into a global triple network via starBase and miRcode databases. This was done to identify which RNAs can be used as sensitive and specific markers for NAFLD. Furthermore, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed to explore the potential regulatory functions of RNAs. Finally, the PharmGKB database was used to search and obtain gene-related drug molecules in the ceRNA regulatory network and then build a gene–drug connection network to screen out important gene molecules and KEGG signaling pathways involved in genes.

MATERIALS AND METHODS

Microarray data and data preprocessing

GES83452 (14) in the NCBI-GEO (https://www.ncbi.nlm.nih.gov/) (15) database were downloaded on April 10, 2020, which included a total of 231 samples, including 159 patients at baseline (44 no NASH, 104 NASH, and 4 undefined) and 79 patients at 1-year follow-up (54 no NASH, 22 NASH, and 3 undefined) based on platforms GPL16686 [HuGene-2_0-st] Affymetrix Human Gene 2.0 ST Array [transcript (gene) version].

Screening significantly differentially expressed RNAs and functional enrichment analyses

The mRNA and lncRNA in the GES83452 datasets were reannotated using the HUGO Gene Nomenclature Committee (http://www.genenames.org/) (16) based on information of Transcript ID, RefSeq ID, etc., which contain 4600 lncRNAs and 19195 protein coding genes. The Limma package (version 3.34.0, https://bioconductor.org/packages/release/bioc/html/limma.html) (17) in R was used to identify DERs between the NAFLD and non-NAFLD samples of the baseline and 1-year follow-up time point group. False discovery rate (FDR) < 0.05 and $|\log_2$ fold change (FC)| > 0.5 were used as the cutoff criteria to define DERs, and the ggplot2packages in R was used to visualize the volcano plots. The heat map was plotted using the pheat map package (version 1.0.8, https://cran.r-project.org/package=pheat map) (18) in R and was presented by two-way hierarchical clustering heat maps (19) based on Euclidean distance (20). $P < .05$ was considered statistically significant. The Venn software online (http://bioinformatics.psb.ugent.be/webtools/Venn/) was used to detect overlapping DERs among the baseline and 1-year follow-up time point groups. Then, GO and KEGG enrichment analyses were performed on intersection mRNAs that commonly contained DERs using the online tool DAVID (version 6.8, https://david.ncifcrf.gov/) (21, 22) $P < .05$ was considered as significant enrichment.

Construction of ceRNA network

The miRNAs related to NAFLD included in the Human MicroRNA Disease Database (HMDD) database (http://www.cuilab.cn/hmdd) were downloaded (23). We constructed a ceRNA network based on NAFLD directly related to lncRNAs and miRNAs, as well as the miRNAs with significantly consistent expression. Firstly, we downloaded the connection relationship pairs of lncRNA-miRNA in the DIANA–LncBase (version 2, http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php) (24). The regulatory relationship between significantly DElncRNA and NAFLD-related differentially expressed miRNA (DEmiRNA) was retained, with retention connection score (miRNA target gene score (miTG–score): the target gene score of DEmiRNA; the higher the value, the greater the probability of targeting) higher than 0.6, thereby the lncRNA–miRNA connection network was constructed.

Then, the starBase database (version 2.0, http://starbase.sysu.edu.cn/) (25) was used to predict target genes regulated by miRNA linked to lncRNA, and the comprehensive target gene prediction information from five databases (targetScan, picTar, RNA22, PITA and miRanda) was provided in the StarBase database. The target miRNA regulatory target gene relationship pair was selected in at least one of the databases, and the miRNA–mRNA pairs with the opposite significant differential expression direction was retained to construct the miRNA–mRNA connection network. Finally, a ceRNA regulation network composed of lncRNA–miRNA–mRNA was constructed by combining lncRNA–miRNA and miRNA–mRNA, and the ceRNA network was visualizatied by using the cytoscape (version 3.6.1, http://www.cytoscape.org/).

The screened target genes in the ceRNA regulatory network were submitted to DAVID 6.8 online tool to perform functional annotation based on GO biological processes and KEGG pathway enrichment analysis, the P value < 0.05 as the significance threshold.

Construction of drug–gene regulation network

The pharmacogenetics and pharmacogenomics knowledge base (PharmGKB) (https://www.pharmgkb.org/) (26) collected the most complete genotype and phenotype information related to the drug genome and was classified systematically, which contained 27,007 genes related to 3579 drugs and 3410 diseases. In this study, the PharmGKB database was used to search for and obtain the gene-related drug molecules in the regulated ceRNA network; then the gene–drug connection network was constructed, the important gene molecules were screened out, and the KEGG signaling pathway of those genes participated in in-depth analysis.

RESULTS

Data preprocessing and DERs screening

A total of 9698 mRNAs and 1116 lncRNAs were detected after data preprocessing, and a total of 561 DERs (48 lncRNA and 513 mRNA; 268 downregulated and 293 upregulated) were screened in the baseline time point group, and 1163 DERs (114 lncRNA and 1049 mRNA; 522 downregulated and 641 upregulated) were screened in the 1-year follow-up time point group, with FDR < 0.05 and $|\log_2 FC| > 0.5$ as the cutoff criteria. We identified all DERs shown in the volcano map according to the value of $|\log_2 FC|$ and displayed DERs on a heat map (Figure 1A, B). The expression values of the DERs were 2-way hierarchically clustered, and the color contrast indicated that there was a significant difference in expression levels between the NAFLD and non-NAFLD samples (Figure1 C, D). Subsequently, a total of 220 overlapping DERs were identified between the baseline and 1-year follow-up time points, which were used by the Venn diagram software (Figure 2).

In addition, the functional enrichment analysis of the overlapping DERs based on online DAVID analyses revealed 22 significantly related GO biological processes and 9 KEGG pathways, with $P < .05$ as the cutoff criteria (Table 1). We found that chemotaxis (GO, 0006935; P = 3.110E-04), unsaturated fatty acid biosynthetic process (GO, 0006636; P = 4.770E-04), and cell-cell signaling (GO, 0007267; P = 1.513E-03) were the three most significant pathways in GO biological processes. Meanwhile, fatty acid metabolism (hsa01212, P = 2.300E-04), PPAR signaling pathway (hsa03320, P = 1.090E-03), and Toll-like receptor signaling pathway (hsa04620, P = 1.479E-03) were the three most significant pathways in KEGG signaling pathways.

Construction of ceRNA regulation network

A total of 77 miRNAs directly related to NAFLD were downloaded from the HMDD database. After the lncRNA and miRNA connection relationship pairs were downloaded and the regulation relationship between significantly DElncRNA and NAFLD-related DEmiRNAs were selected, a total of 74 connection pairs were retained to construct an lncRNA–miRNA connection network with a connection coefficient higher than 0.6. In addition, after the target genes of the miRNA were screened, we compared the regulated target genes with the significant DEmRNAs in target modules and retained the opposite relationship pairs of the expressions of significant differential direction. The miRNA–mRNA regulation network was constructed by using a total of 523 pairs of regulation relationships. Finally, as shown Figure 3, a ceRNA regulation network was constructed.

A total of 28 GO biological processes and 9 KEGG pathways of the mRNAs in the ceRNA regulatory network were obtained, with $P < .05$ as the significance threshold (Table 2). We found that a lipid biosynthetic process (GO, 0008610; P = 1.15E-06), a steroid metabolic process (GO, 0008202; P = 1.26E-05), and a steroid biosynthetic process (GO, 0006694; P = 9.34E-05) were the three most significant pathways in GO biological processes. Meanwhile, the biosynthesis of unsaturated fatty acids (hsa01040, P = 2.39E-05), terpenoid backbone biosynthesis (hsa00900, P = 1.090E-03), and heparan sulfate biosynthesis (hsa00534, P = 1.38E-02) were the three most significant pathways in KEGG signaling pathways.

Construction of drug regulation gene network

The gene-related drug molecule connection pairs were downloaded from the PharmGKB database. A total of

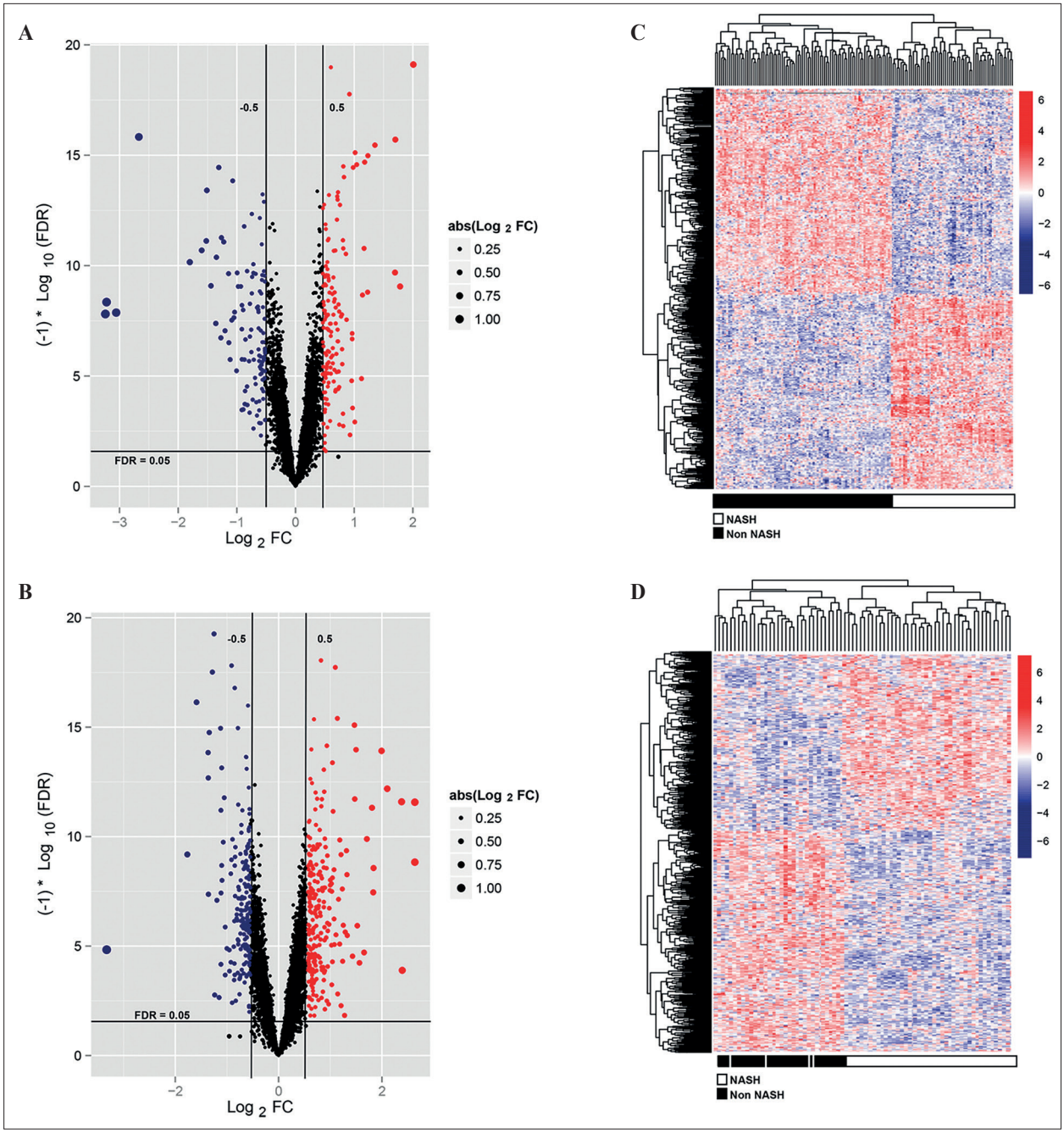

Figure 1. The hierarchically clustering analysis of screened differentially expressed RNAs (DERs). Left: log_2FC-log_{10} (FDR) volcano map for GSE83452 using the significant DERs. Blue and red dots indicate significant DERs. The horizontally dashed line indicates FDR < 0.05. Two vertical lines indicate |Log2FC| > 0.5. A: Baseline time points; B: 1-year follow-up time points. Right: Two-way hierarchically clustered heat map for GSE83452 using the DERs. Red: upregulated DERs. Blue: downregulated DERs. C: Baseline time points; D: 1-year follow-up time points.

154 connection pairs were obtained by selecting the parts related to the genes in the ceRNA regulatory network, and a gene–drug connection network was constructed (Figure 4). Compared with the pathway in which the RNAs significantly participated in the ceRNA regulatory network constructed in the previous step, the leptin receptor (LEPR) and *CXCL10* are involved in the Cytokine–cytokine receptor interaction (P = 1.86E-02), and *FOXO1* is involved in both the Insulin signaling pathway (P = 1.79E-02) and the pathways in cancer (P = 2.87E-02).

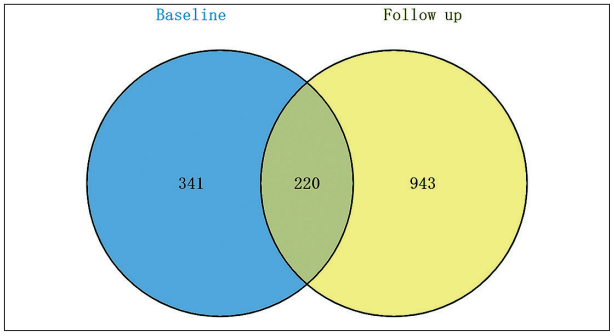

Figure 2. Authentication of overlapping DERs in the GSE83452 datasets via Venn diagrams software. Blue represents DERs of the baseline time points group. Yellow represents DERs of the 1-year follow-up time points group.

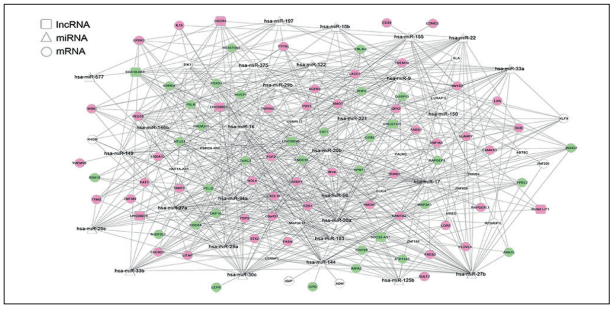

Figure 3. The lncRNA–miRNA–mRNA ceRNA network. Squares, triangles, and circles represent lncRNA, miRNA, and mRNA, respectively. Green and red dots indicate the significantly downregulated DERs at both baseline and 1-year follow-up time points, and the white dots indicate DERs whose expression difference direction has changed.

Table 1. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis for the differentially expressed RNAs (DERs).

Category	Term	Count	P-Value	Genes
Biology Process	GO:0006935~chemotaxis	8	3.110E-04	*MAP2K1, CXCL9, AMOT,* etc
	GO:0006636~unsaturated fatty acid biosynthetic process	4	4.770E-04	*FADS1, SCD, FADS2,* etc
	GO:0007267~cell-cell signaling	10	1.513E-03	*SH2D1A, ADM, FADS1,* etc
	GO:0007568~aging	8	1.848E-03	*TYMS, ADM, APOD,* etc
	GO:0007166~cell surface receptor signaling pathway	10	2.538E-03	*MARCO, PRLR, TSPAN3,* etc
	GO:0055114~oxidation-reduction process	15	4.070E-03	*KDM6A, HSD17B2, PYROXD2,* etc
	GO:0016337~single organismal cell-cell adhesion	6	4.312E-03	*MPZL2, PKHD1, FAT1,* etc
	GO:0006629~lipid metabolic process	7	6.425E-03	*APOD, PLIN1, APOF,* etc
	GO:0070098~chemokine-mediated signaling pathway	5	6.698E-03	*TFF2, CXCL9, ACKR3,* etc
	GO:0032496~response to lipopolysaccharide	7	7.895E-03	*ADM, DUSP10, CXCL9,* etc
	GO:0010508~positive regulation of autophagy	4	8.546E-03	*RNF152, FOXO1, TRIM22,* etc
	GO:0035338~long-chain fatty-acyl-CoA biosynthetic process	4	9.779E-03	*SCD, FASN, ELOVL6, ACSL5*
	GO:0006915~apoptotic process	13	1.716E-02	*PEG10, RASSF6, LITAF,* etc
	GO:0002250~adaptive immune response	6	2.033E-02	*SH2D1A, EOMES, CD1C,* etc
	GO:0006959~humoral immune response	4	2.223E-02	*SH2D1A, BST1, LTF, CD28*
	GO:0006968~cellular defense response	4	2.767E-02	*SH2D1A, CXCL9, LBP,* etc
	GO:0060326~cell chemotaxis	4	3.124E-02	*CXCL9, CCL5, DOCK4,* etc
	GO:0045766~positive regulation of angiogenesis	5	3.346E-02	*ADM, LRG1, RHOB,* etc
	GO:0032868~response to insulin	4	3.374E-02	*ADM, INSIG2, FADS1, PCK1*
	GO:0008203~cholesterol metabolic process	4	3.504E-02	*INSIG2, APOF, LEPR, ERLIN1*
	GO:0001889~liver development	4	4.331E-02	*COBL, ONECUT1, DBP, RPGRIP1L*
	GO:0006954~inflammatory response	9	4.788E-02	*LXN, AOX1, LYZ,* etc
KEGG Pathway	hsa01212:Fatty acid metabolism	6	2.300E-04	*FADS1, SCD, FASN,* etc
	hsa03320:PPAR signaling pathway	6	1.090E-03	*PLIN1, SCD, FADS2,* etc
	hsa04620:Toll-like receptor signaling pathway	7	1.479E-03	*CTSK, MAP2K1, CXCL9,* etc
	hsa01040:Biosynthesis of unsaturated fatty acids	4	2.350E-03	*FADS1, SCD, FADS2, ELOVL6*
	hsa04640:Hematopoietic cell lineage	6	3.475E-03	*CR1, CD2, CD1C,* etc
	hsa04152:AMPK signaling pathway	6	1.465E-02	*LEPR, SCD, FASN,* etc
	hsa04668:TNF signaling pathway	5	3.695E-02	*MAP2K1, IL15, CREB3L3,* etc
	hsa00760:Nicotinate and nicotinamide metabolism	3	4.529E-02	*BST1, ENPP3, AOX1*
	hsa04060:Cytokine-cytokine receptor interaction	7	4.659E-02	*PRLR, LEPR, CXCL9,* etc

Table 2. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis for the mRNA in the ceRNA regulatory network.

Category	Term	Count	P-Value	Genes
Biology Process	GO:0008610~lipid biosynthetic process	12	1.15E-06	*PRLR, ADM, ISPD,* etc
	GO:0008202~steroid metabolic process	9	1.26E-05	*PRLR, INSIG2, ADM,* etc
	GO:0006694~steroid biosynthetic process	6	9.34E-05	*PRLR, ADM, FDPS,* etc
	GO:0006633~fatty acid biosynthetic process	5	8.64E-04	*FADS1, SCD, FASN,* etc
	GO:0046649~lymphocyte activation	6	4.38E-03	*PRLR, EOMES, IL15,* etc
	GO:0016053~organic acid biosynthetic process	5	9.79E-03	*FADS1, SCD, FASN,* etc
	GO:0046394~carboxylic acid biosynthetic process	5	9.79E-03	*FADS1, SCD, FASN,* etc
	GO:0045321~leukocyte activation	6	9.84E-03	*PRLR, EOMES, IL15,* etc
	GO:0006355~regulation of transcription, DNA-dependent	18	1.19E-02	*CAMTA2, MAP2K1, LITAF,* etc
	GO:0030334~regulation of cell migration	5	1.31E-02	*MAP2K1, CLIC4, AMOT,* etc
	GO:0051252~regulation of RNA metabolic process	18	1.47E-02	*CAMTA2, MAP2K1,* etc
	GO:0007267~cell-cell signaling	9	1.52E-02	*ADM, FADS1, FAT1,* etc
	GO:0051094~positive regulation of developmental process	6	1.71E-02	*MAP2K1, EOMES, AMOT,* etc
	GO:0001775~cell activation	6	1.93E-02	*PRLR, EOMES, IL15,* etc
	GO:0040012~regulation of locomotion	5	2.00E-02	*MAP2K1, CLIC4, AMOT,* etc
	GO:0051270~regulation of cell motion	5	2.04E-02	*MAP2K1, CLIC4, AMOT,* etc
	GO:0009719~response to endogenous stimulus	7	2.19E-02	*MAP2K1, ADM, FADS1,* etc
	GO:0006631~fatty acid metabolic process	5	2.21E-02	*FADS1, SCD, FASN,* etc
	GO:0045860~positive regulation of protein kinase activity	5	3.23E-02	*PRLR, MAP2K1, CD24,* etc
	GO:0060429~epithelium development	5	3.42E-02	*STX2, MAP2K1, RPGRIP1L,* etc
	GO:0033674~positive regulation of kinase activity	5	3.61E-02	*PRLR, MAP2K1, CD24,* etc
	GO:0051347~positive regulation of transferase activity	5	4.06E-02	*PRLR, MAP2K1, CD24,* etc
	GO:0001568~blood vessel development	5	4.33E-02	*LEPR, FOXO1, AMOT,* etc
	GO:0051254~positive regulation of RNA metabolic process	7	4.51E-02	*CAMTA2, MAP2K1, CSRNP1,* etc
	GO:0001944~vasculature development	5	4.66E-02	*LEPR, FOXO1, AMOT,* etc
	GO:0009725~response to hormone stimulus	6	4.80E-02	*MAP2K1, ADM, FADS1,* etc
	GO:0030030~cell projection organization	6	4.84E-02	*STX2, MAP2K1, ADM,* etc
	GO:0007243~protein kinase cascade	6	4.94E-02	*PRLR, MAP2K1, DUSP10,* etc
KEGG Pathway	hsa01040:Biosynthesis of unsaturated fatty acids	4	2.39E-05	*FADS1, SCD, FADS2, ELOVL6*
	hsa00900:Terpenoid backbone biosynthesis	2	8.23E-03	*FDPS, MVK*
	hsa00534:Heparan sulfate biosynthesis	2	1.38E-02	*EXT1, HS3ST3B1*
	hsa04910:Insulin signaling pathway	3	1.79E-02	*MAP2K1, FASN, FOXO1*
	hsa04060:Cytokine-cytokine receptor interaction	4	1.86E-02	*PRLR, LEPR, IL15, CXCL10*
	hsa04010:MAPK signaling pathway	4	1.93E-02	*MAP2K1, DUSP10, MAP3K13, FGF2*
	hsa04630:Jak-STAT signaling pathway	3	2.21E-02	*PRLR, LEPR, IL15*
	hsa05200:Pathways in cancer	4	2.87E-02	*MAP2K1, FZD1, FOXO1, FGF2*
	hsa03320:PPAR signaling pathway	2	3.28E-02	*SCD, FADS2*

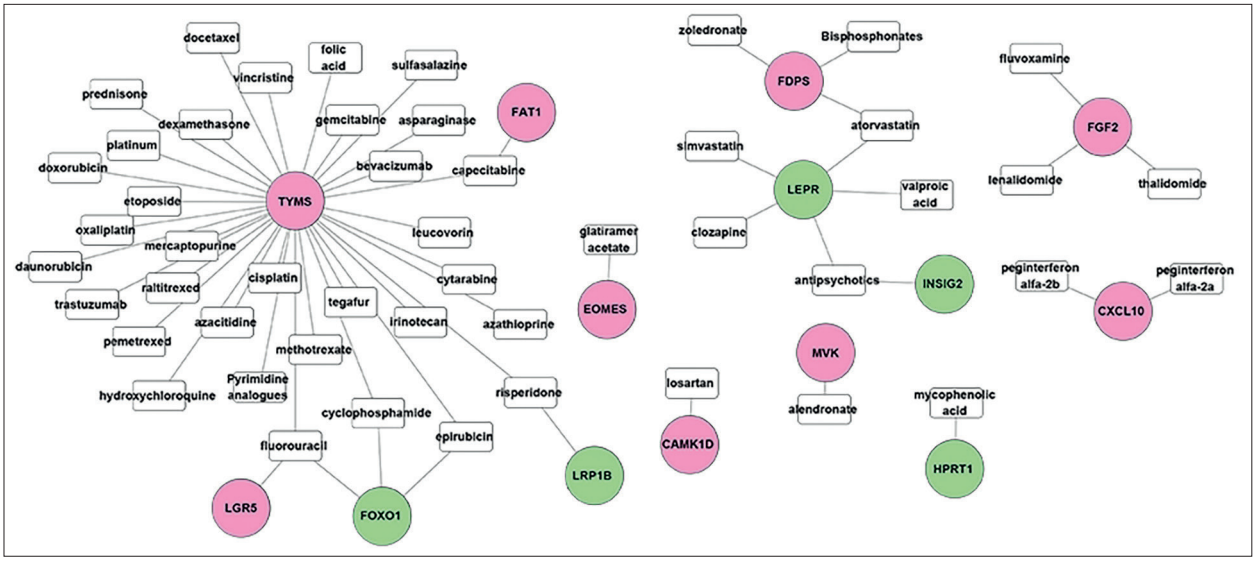

Figure 4. Gene–drug connection network. Squares represent drug molecules, circles represent genes, and green and red dots represent the significantly downregulated DERs at both baseline and 1-year follow-up time points.

DISCUSSION

The characteristics of NAFLD include necrotizing inflammation and lipid accumulation in the liver, as well as continuous improvement of living standards leading to overnutrition. In addition, bad living habits lead to the incidence of NAFLD on a global scale (27). The specific mechanism of the transition from benign steatosis to steatohepatitis in NAFLD is not fully understood, and there are currently no pharmacological options for the treatment of NAFLD. Therefore, the treatment of NASH mainly depends on lifestyle changes, such as strengthening exercises, reducing weight, and a light diet (28). Although current studies have shown that weight loss improves the histological characteristics of NAFLD, most patients have not however achieved the goal of curing NAFLD. There are some potentially valuable molecules, nevertheless, which are currently being clinically evaluated (29). For example, *PNPLA3* (30), *TM6SF2* (31), *MBOAT7* (32), and *HSD17B13* (33), molecules that predispose an individual to the spectrum of NAFLD-related disease, have been found to play a role in macrophage phagocytosis, immune response, oxidative stress, and inflammation, insulin signaling, and lipid metabolism in NAFLD susceptibility and progression (34). But there is no unmet clinical need for drug discovery and development for patients with NAFLD. Increased levels of toxic lipids (free fatty acids or free cholesterol) can lead to liver cell damage and trigger inflammation is the pathogenesis of NAFLD as is currently understood. In addition, oxidative stress, proinflammatory chemokines and cytokines have been proven to lead to liver inflammation, which, in turn, leads to damage and fibrosis of the liver. Therefore, the identification of proinflammatory cytokines related to lipotoxicity may improve our understanding of the pathogenesis of NAFLD, helping to develop new pharmacological methods.

In this study, a total of 220 overlapping DERs were identified between the baseline and 1-year follow-up time points. In addition, functional enrichment analysis of overlapping DERs, based on online DAVID analyses, revealed 22 significantly related GO biological processes and 9 KEGG pathways, with $P < .05$ as the cutoff criteria. We found that chemotaxis ($P = 3.110E-04$), unsaturated fatty acid biosynthetic process ($P = 4.770E-04$), and cell-cell signaling ($P = 1.513E-03$) were the three most significant pathways in GO biological processes. Meanwhile, fatty acid metabolism ($P = 2.300E-04$), PPAR signaling pathway ($P = 1.090E-03$), and Toll-like receptor signaling pathway ($P = 1.479E-03$) were the three most significant pathways in KEGG signaling pathways. Afterwards, a ceRNA regulatory network was constructed. The GO and pathway enrichment analyses indicated that the mRNAs of the ceRNA regulatory network were involved in various important biological functions and metabolic pathways associated with NAFLD, including lipid biosynthetic process, steroid metabolic process, steroid biosynthetic process, biosynthesis of unsaturated fatty acids, terpenoid backbone biosynthesis, heparan sulfate biosynthesis, Cytokine–cytokine receptor interaction, Insulin signaling pathway, and the pathways in cancer. To further understand the functional mechanism of the ceRNA network, a drug regulation gene network was constructed which included 154 gene–drug connection pairs. Subsequently, *LEPR*, *CXCL10*, and *FOXO1* were investigated using the PharmGKB database. It was revealed that the therapeutic

effect of antipsychotics, atorvastatin, valproic acid, risperidone, clozapine, olanzapine, simvastatin, and quetiapine were produced, thus possibly targeting to *LEPR* through the Cytokine–cytokine receptor interaction pathway. The therapeutic effect of Peginterferon alfa-2a and peginterferon alfa-2b were produced by targeting *CXCL10* through the Cytokine–cytokine receptor interaction pathway. The therapeutic effect of Epirubicin, cyclophosphamide, and fluorouracil were produced by targeting *FOXO1* through the Insulin signaling pathway or the pathways in cancer.

LEPR is responsible for encoding the leptin receptor that binds to leptin in target tissues. Due to its role in regulating lipid metabolism and insulin resistance, it is considered to be a candidate gene for NAFLD and coronary atherosclerosis (35). Simultaneously, An et al. (36) found that *LEPR* Q223R polymorphism may lead to a significant risk of NAFLD and coronary atherosclerosis, which is consistent with the results of this study. The CXC motif chemokine ligand 10 (*CXCL10*) is a particularly important pro-inflammatory cytokine related to lipotoxicity, which can recruit inflammatory cells to the site of tissue injury (37, 38). Studies have shown that *CXCL10* is upregulated in NAFLD patients (39), and revealed that *CXCL10* may be a key molecule that contributes to the transition from benign steatosis to steatohepatitis, promoting liver cell damage and inflammation (40).

Our study revealed that peginterferon alfa-2a and peginterferon alfa-2b can downregulate the expression of *CXCL10*, suggesting a potential role of *CXCL10* in the development of intrahepatic inflammation through the Cytokine–cytokine receptor interaction pathway, and demonstrated that *CXCL10* is an independent risk factor for patients with NAFLD. *FOXO1* is an important transcriptional effector. It is widely expressed in various types of tissues and plays an important role in the signaling pathway of insulin and insulin-like growth factor 1 (41). In addition, the expression levels of most genes related to adipocyte differentiation are affected by the coordination of *FOXO1* (42). Yue Li et al. (43) conducted a comprehensive analysis of the relevant information about the activity of *FOXO1* in lipid metabolism, and found that *FOXO1* has a significant inhibitory effect on the production of fibrotic effector cells, and pointed out that *FOXO1* has the potential to become a target for the treatment of NAFLD, but the related mechanism needs to be further verified by experiments (44). L. Valenti et al. (45) found that *FOXO1* may affect the susceptibility of NAFLD, and regulating the level of *FOXO1* mRNA in order to regulate the relevant cytokines in insulin signaling to promote the progression of liver injury. This study found that epirubicin, cyclophosphamide, and fluorouracil can downregulate the expression of *FOXO1*, suggesting that these drugs may produce therapeutic effect by targeting *FOXO1* through the insulin signaling pathway or the pathways in cancer. However, further study is necessary to validate this hypothesis.

CONCLUSION

In conclusion, this study constructed and analyzed a ceRNA network, a network which may provide some evidence for future studies focusing on the molecular mechanisms of NALFD. *LEPR*, *CXCL10*, and *FOXO1* may function as ceRNAs to serve critical roles in NALFD. In addition, antipsychotics, atorvastatin, valproic acid, risperidone, clozapine, olanzapine, simvastatin, quetiapine, peginterferon alfa-2a, peginterferon alfa-2b, epirubicin, cyclophosphamide, and fluorouracil may produce therapeutic effect for patients with NALFD.

ABBREVIATIONS

Nonalcoholic fatty liver disease (NAFLD)
Long noncoding RNAs (lncRNAs)
Competing endogenous RNAs (ceRNAs)
Nonalcoholic steatohepatitis (NASH)
Gene Expression Omnibus (GEO)
Differentially expressed RNAs (DERs)
Gene Ontology (GO)
Kyoto Encyclopedia of Genes and Genomes (KEGG)
False discovery rate (FDR)
Human MicroRNA Disease Database (HMDD)
Leptin receptor (LEPR)

Declaration of Interest: The authors declare no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Funding: This study was funded by Natural Science Foundation of Jilin Provincial Science and Technology Department (Grant Number:20200201569JC)

REFERENCES:

1. Erman H, Beydogan E, Cetin SI, Boyuk B. Endocan: A Biomarker for Hepatosteatosis in Patients with Metabolic Syndrome. Mediators Inflamm. 2020; 2020: 3534042.
2. Aizawa M, Inagaki S, Moriyama M, Asano K, Kakehashi M. Modeling the natural history of fatty liver using lifestyle-related risk factors: Effects of body mass index (BMI) on the life-course of fatty liver. PLoS One. 2019; 14(10):e0223683.
3. Zelber-Sagi S, Lotan R, Shlomai A, Webb M, Harrari G, Buch A, et al. Predictors for incidence and remission of NAFLD in the general population during a

seven-year prospective follow-up. Journal of hepatology. 2012; 56(5): 1145-1151.

4. Sookoian S, Pirola CJ. Review article: shared disease mechanisms between non-alcoholic fatty liver disease and metabolic syndrome - translating knowledge from systems biology to the bedside. Aliment Pharmacol Ther. 2019; 49(5): 516-527.
5. Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ. Mechanisms of NAFLD development and therapeutic strategies. Nat Med. 2018; 24(7):908-922.
6. Rinella ME. Nonalcoholic fatty liver disease: a systematic review. JAMA. 2015; 313(22):2263-73.
7. Noureddin M, Rinella ME. Nonalcoholic Fatty liver disease, diabetes, obesity, and hepatocellular carcinoma. Clin Liver Dis. 2015; 19(2):361-79.
8. Fabbrini E, Magkos F, Mohammed BS, Pietka T, Abumrad NA, Patterson BW, et al. Intrahepatic fat, not visceral fat, is linked with metabolic complications of obesity. Proc Natl Acad Sci USA. 2009;106(36):15430-5.
9. Konerman MA, Jones JC, Harrison SA. Pharmacotherapy for NASH: Current and emerging. J Hepatol. 2018; 68(2):362-375.
10. Sookoian S, Rosselli MS, Gemma C, Burgueño AL, Fernández Gianotti T, Castaño GO, Pirola CJ. Epigenetic regulation of insulin resistance in nonalcoholic fatty liver disease: impact of liver methylation of the peroxisome proliferator-activated receptor γ coactivator 1α promoter. Hepatology. 2010; 52(6):1992-2000
11. Liu NK, Xu XM. MicroRNA in central nervous system trauma and degenerative disorders. Physiol Genomics. 2011; 43(10):571-80
12. Piazzolla VA, Mangia A. Noninvasive Diagnosis of NAFLD and NASH. Cells. 2020; 9(4):1005
13. Zou B, Yeo YH, Nguyen VH, Cheung R, Ingelsson E, Nguyen MH. Prevalence, characteristics and mortality outcomes of obese, nonobese and lean NAFLD in the United States, 1999-2016. J Intern Med. 2020; 288(1):139-151
14. Lefebvre P, Lalloyer F, Baugé E, Pawlak M, Gheeraert C, Dehondt H, et al. Interspecies NASH disease activity whole-genome profiling identifies a fibrogenic role of PPARα-regulated dermatopontin. JCI Insight. 2017 Jul 6;2(13):e92264
15. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, et al. NCBI GEO: mining tens of millions of expression profiles--database and tools update. Nucleic Acids Res. 2007; 35(Database issue):D760-5
16. Braschi B, Denny P, Gray K, Jones T, Seal R, Tweedie S, et al. Genenames.org: the HGNC and VGNC resources in 2019. Nucleic Acids Res. 2019; 47(D1):D786-D792
17. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43(7):e47
18. Wang L, Cao C, Ma Q, Zeng Q, Wang H, Cheng Z, et al. RNA-seq analyses of multiple meristems of soybean: novel and alternative transcripts, evolutionary and functional implications. BMC Plant Biol. 2014; 14:169
19. Bonow RO. Braunwald`s Heart Disease: A Textbook of Cardiovascular Medicine, Single Volume, 9th Edition.
20. Szekely, G., Rizzo, M. Hierarchical Clustering via Joint Between-Within Distances: Extending Ward's Minimum Variance Method. Journal of Classification; 2005; 22, 151–183
21. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009; 4(1):44-57
22. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009; 37(1):1-13
23. Huang Z, Shi J, Gao Y, Cui C, Zhang S, Li J, et al. HMDD v3.0: a database for experimentally supported human microRNA-disease associations. Nucleic Acids Res. 2019; 47(D1):D1013-D1017
24. Paraskevopoulou MD, Vlachos IS, Karagkouni D, Georgakilas G, Kanellos I, Vergoulis T, et al. DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. Nucleic Acids Res. 2016; 44(D1):D231-8
25. Li JH, Liu S, Zhou H, Qu LH, Yang JH. StarBase v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res. 2014; 42(Database issue):D92-7
26. Thorn CF, Klein TE, Altman RB. PharmGKB: the pharmacogenetics and pharmacogenomics knowledge base. Methods Mol Biol. 2005; 311:179-91
27. Chen HJ, Liu J. Actein ameliorates hepatic steatosis and fibrosis in high fat diet-induced NAFLD by regulation of insulin and leptin resistant. Biomed Pharmacother. 2018; 97: 1386-1396.
28. Pan X, Zheng M, Zou T, Liu W, Gu X, Zhang X, et al. The LEPR K109R and Q223R Might Contribute

to the Risk of NAFLD: A Meta-Analysis. Curr Mol Med. 2018; 18(2): 91-99.

29. Zhang Y, Xiang D, Hu X, Ruan Q, Wang L, Bao Z. Identification and study of differentially expressed miRNAs in aged NAFLD rats based on high-throughput sequencing. Ann Hepatol. 2020; 19(3):302-312
30. Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. 2008; 40(12): 1461-1465.
31. Kozlitina J, Smagris E, Stender S, Nordestgaard BG, Cohen JCJNG. Exome-wide association study identifies a *TM6SF2* variant that confers susceptibility to nonalcoholic fatty liver disease. 2014; 46(4): 352-356.
32. Mancina RM, Dongiovanni P, Petta S, Pingitore P, Meroni M, Rametta R, et al. The MBOAT7-TMC4 Variant rs641738 Increases Risk of Nonalcoholic Fatty Liver Disease in Individuals of European Descent. 2016; 150(5): 1219-1230.e1216.
33. Eslam M, Hashem AM, Leung R, Romero-Gomez M, Berg T, Dore GJ, et al. ; International Hepatitis C Genetics Consortium (IHCGC). Interferon-λ rs12979860 genotype and liver fibrosis in viral and non-viral chronic liver disease. Nat Commun. 2015; 6:6422.
34. Petta S, Valenti L, Marra F, Grimaudo S, Tripodo C, Bugianesi E, et al. *MERTK* rs4374383 polymorphism affects the severity of fibrosis in non-alcoholic fatty liver disease. 2016; 64(3): 682-690.
35. Aller R, De Luis DA, Izaola O, González Sagrado M, Conde R, Pacheco D, et al. Lys656Asn polymorphism of leptin receptor, leptin levels and insulin resistance in patients with non alcoholic fatty liver disease. Eur Rev Med Pharmacol Sci. 2012;16(3):335-41.
36. An BQ, Lu LL, Yuan C, Xin YN, Xuan SY. Leptin Receptor Gene Polymorphisms and the Risk of Non-Alcoholic Fatty Liver Disease and Coronary Atherosclerosis in the Chinese Han Population. Hepat Mon. 2016; 16(4):e35055.
37. Neville LF, Mathiak G, Bagasra O. The immunobiology of interferon-gamma inducible protein 10 kD (IP-10): a novel, pleiotropic member of the C-X-C chemokine superfamily. Cytokine Growth Factor Rev. 1997; 8(3):207-19.
38. Luster AD, Unkeless JC, Ravetch JV. Gamma-interferon transcriptionally regulates an early-response gene containing homology to platelet proteins. Nature. 1985; 315(6021):672-6.
39. Bertola A, Bonnafous S, Anty R, Patouraux S, Saint-Paul MC, Iannelli A, et al. Hepatic expression patterns of inflammatory and immune response genes associated with obesity and NASH in morbidly obese patients. PLoS One. 2010; 5(10):e13577.
40. Zhang X, Shen J, Man K, Chu ES, Yau TO, Sung JC, et al. *CXCL10* plays a key role as an inflammatory mediator and a non-invasive biomarker of non-alcoholic steatohepatitis. Journal of hepatology. 2014; 61(6): 1365-1375.
41. Dong XC, Copps KD, Guo S, Li Y, Kollipara R, Depinho RA, et al. Inactivation of Hepatic Foxo1 by Insulin Signaling Is Required for Adaptive Nutrient Homeostasis and Endocrine Growth Regulation. 2008; 8(1): 0-76.
42. Munekata K, Sakamoto K. Forkhead transcription factor Foxo1 is essential for adipocyte differentiation. In Vitro Cell Dev Biol Anim. 2009; 45(10):642-51.
43. Li Y, Ma Z, Jiang S, Hu W, Li T, Di S, et al. A global perspective on FOXO1 in lipid metabolism and lipid-related diseases. Prog Lipid Res. 2017; 66: 42-49.
44. Xin Z, Ma Z, Hu W, Jiang S, Yang Z, Li T, et al. FOXO1/3: Potential suppressors of fibrosis. Ageing Res Rev. 2018; 41: 42-52.
45. Valenti L, Dongiovanni P, Rametta R, Fracanzani AL, Fargion SJD, Disease L. FOXO1 genotype influences the susceptibility to and severity of NAFLD by modulating FOXO1 expression. 2009; 41(3):A2–A3.