

RESEARCH

Open Access



Dynamic co-expression modular network analysis in nonalcoholic fatty liver disease

Jing Zheng¹, Huizhong Wu², Zhiying Zhang³ and Songqiang Yao^{1*}

Abstract

Background: Nonalcoholic fatty liver disease (NAFLD) is the most common chronic liver disease affecting people's health worldwide. Exploring the potential biomarkers and dynamic networks during NAFLD progression is urgently important.

Material and methods: Differentially expressed genes (DEGs) in obesity, NAFL and NASH were screened from GSE126848 and GSE130970, respectively. Gene set enrichment analysis of DEGs was conducted to reveal the Gene Ontology (GO) biological process in each period. Dynamic molecular networks were constructed by DyNet to illustrate the common and distinct progression of health- or obesity-derived NAFLD. The dynamic co-expression modular analysis was carried out by CEMiTool to elucidate the key modulators, networks, and enriched pathways during NAFLD.

Results: A total of 453 DEGs were filtered from obesity, NAFL and NASH periods. Function annotation showed that health-NAFLD sequence was mainly associated with dysfunction of metabolic syndrome pathways, while obesity-NAFLD sequence exhibited dysregulation of Cell cycle and Cellular senescence pathways. Nine nodes including COL3A1, CXCL9, CYCS, CXCL10, THY1, COL1A2, SAA1, CDKN1A, and JUN in the dynamic networks were commonly identified in health- and obesity-derived NAFLD. Moreover, CYCS, whose role is unknown in NAFLD, possessed the highest correlation with NAFLD activity score, lobular inflammation grade, and the cytological ballooning grade. Dynamic co-expression modular analysis showed that module 4 was activated in NAFL and NASH, while module 3 was inhibited at NAFLD stages. Module 3 was negatively correlated with CXCL10, and module 4 was positively correlated with COL1A2 and THY1.

Conclusion: Dynamic network analysis and dynamic gene co-expression modular analysis identified a nine-gene signature as the potential key regulator in NAFLD progression, which provided comprehensive regulatory mechanisms underlying NAFLD progression.

Keywords: Fatty liver, Obesity, Gene regulatory network

Background

Nonalcoholic fatty liver disease (NAFLD) is regarded as one of the most common chronic liver diseases. NAFLD is associated with obesity, and ranges from simple hepatic steatosis (nonalcoholic fatty liver [NAFL]) to

nonalcoholic steatohepatitis (NASH). Some patients with NAFLD eventually developed cirrhosis, fibrosis, or hepatocellular carcinoma [1]. Since the liver is required for the glucose metabolism process and energy homeostasis, NAFLD also acts as a high-risk factor for metabolic diseases such as type 2 diabetes. A recent study confirmed that NAFLD is closely related to high mortality, especially liver disease-specific and diabetes-specific deaths [2]. Due to the high prevalence, high mortality, and serious complications, exploring the potential biomarkers and

*Correspondence: songqiangyao888@163.com

¹ Department of Pharmacy, Zhejiang Medical & Health Group Hangzhou Hospital, No.1 Banshan Road, Kangjian nong, Hangzhou 310022, China
Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

elucidating the dynamic networks involved in NAFLD are becoming urgently important.

NAFLD is mainly caused by obesity. Lifestyle intervention such as weight loss with diet or surgery is a primary therapy for NAFLD [3]. Liver transplantation has been proven to be an efficient method to treat NAFLD especially at the end stage [4], but the shortage of the donor's liver and surgical complications limit the therapy. The treatment of NAFLD is to prevent NAFL from developing to NASH, which is highly associated with liver cirrhosis, fibrosis or serious hepatic complications. However, there is no specific drug approved for NAFLD until now. Moreover, NAFLD is increasingly being identified in non-obese patients. The common and distinct mechanisms involved in the non-obesity and obesity-derived NAFLD has not been fully elucidated. Expression profiling by high throughput sequencing has been a powerful tool to reveal key factors and pathways in NAFLD liver, and some studies have focused on the different stages of NAFLD [5, 6]. However, the dynamic alterations in key genes and co-expression networks were not fully elucidated during NAFLD progression. Searching key regulators, functional networks, and pathways in obesity, NAFL and NASH would be helpful to prevent and delay the development of NAFLD.

In this work, we focused on the dynamic gene landscape and pathways during NAFLD progression. Consistently changed genes were screened from GSE126848 [5] and GSE130970 [7]. Dynamic molecular networks and dynamic co-expression modular analysis were constructed to illustrate the key modulators, networks, and enriched pathways during NAFLD. The results would offer new clinical biomarkers for different stages of NAFLD, and provide potential drug targets for NAFLD treatment.

Material and methods

Microarray data

The flow diagram of this study was illustrated in supplementary figure 1. The search strategy is "(NAFLD) OR (nonalcoholic fatty liver disease)" in GEO (<http://www.ncbi.nlm.nih.gov/geo/>). The criteria are: 1. The dataset is an RNA sequencing dataset; 2. The dataset has more than 10 samples; 3. The dataset contains healthy, NAFL and NASH subjects. Finally, two datasets GSE126848 and GSE130970 meet the criteria. The RNA sequencing dataset GSE126848 was based on platform GPL18573 [5]. The liver biopsies comprised 14 healthy donors, 12 obese, 15 NAFL and 16 NASH patients. The dataset GSE130970 was based on platform GPL16791 [7]. According to the NIDDK NASH CRN criteria, 26 samples with the NAFLD activity score (NAS) ≥ 5 were regarded as NASH, 10 samples with steatosis and $NAS < 3$ were considered

as NAFL, 4 samples with $NAS = 0$ was chosen as normal tissues. Due to the limited histological information, the liver histology of the rest 36 samples could not be defined and these samples were not enrolled in our analysis.

Data processing and identification of DEGs

Basing on R (version 3.5.1), *rtracklayer* and *SummarizedExperiment* packages were applied to annotate the raw data using the human genome reference GRCh38 [8]. The *edgeR* package was applied to conduct the background correction, normalization and \log_2 -counts-per-million transformation. Empirical Bayes method depending on the *limma* package was performed to identify DEGs of GSE126848 and GSE130970. The thresholds for DEGs filtration were set as $|\log_2 \text{ fold change (FC)}| \geq 1$ and P value < 0.05 . Since only GSE126848 contained the obese samples, the DEGs in obesity were filtered from GSE126848. Common DEGs in NAFL and NASH were screened by the intersection of GSE126848 and GSE130970. GSE126848 contained all of the three disease types, therefore, the expression profiles of these DEGs at different stages were then extracted from GSE126848 for the dynamic expression pattern and gene set enrichment analysis.

Gene set enrichment analysis of DEGs

Gene set enrichment analysis (GSEA) of the DEGs was conducted using *gseaGO* function within the R package *clusterProfiler* [9]. The permutation number was set as 1500. The *org.Hs.eg.db* package was applied to map gene identifiers. Ridgeline plots were performed using *ggridges* and *ggplot2* packages. Gene set with a P value < 0.05 was considered to be significant.

Analysis of DEGs expression pattern during NAFLD progression

The union of DEGs in obesity, NAFL and NASH was input and normalized by Short Time-series Expression Miner (STEM) to explain the NAFLD progression. STEM is a software designed for clustering, comparing, and displaying gene expression data from time-course experiments [10]. The maximum number of expression models was set as 60, and P values were adjusted using FDR. The other parameters were set as default. Clusters with colors mean statistically significant numbers of genes enriched.

Analysis of dynamic molecular interaction networks during NAFLD progression

To analysis the key genes involved in the NAFLD progression, dynamic molecular interaction networks were constructed using *DyNet* (1.0.0) application in *Cytoscape* (3.6.1). DEGs were input into the *STRING* database to construct the protein-protein interaction (PPI) network,

and DyNet [11] was applied to analyze the most ‘rewired’ nodes across dynamic network states. To analyze the heterogeneous mechanism of NAFLD, the rewired nodes from obese and non-obese NAFLD progression sequences were intersected and subsequently enriched by KEGG within the R package clusterProfiler [9]. Node degree variance was calculated by DyNet and was displayed as Dn-Score.

Dynamic co-expression modular analysis

The CEMiTool is a comprehensive R package that allowed the users to identify co-expressed gene modules, hubs, and determine significant module functions [12]. The R packages WGCNA and CEMiTool were combined to calculate the correlation between key nodes identified above and gene co-expression modules. A $P < 0.05$ was defined as significant correlation. Then, files including GO gene sets (C5) from MSigDB 7.0 and human protein–protein interaction (PPI) from the mentha database

were regarded as enrichment and PPI background. The association of module activity was determined using fgsea package. The biological functions related to modules were annotated using the over representation analysis (ORA).

Results

Function annotation of DEGs in obese, NAFL and NASH patients

GEO datasets GSE130970 and GSE126848 were enrolled in our study, and the clinicopathological features of patients in two datasets were presented in Table 1. DEGs in NAFL and NASH were firstly screened by the intersection of GSE126848 and GSE130970, respectively. As a result, 31 upregulated DEGs and 8 downregulated DEGs were filtered from NAFL samples, while 99 upregulated DEGs and 23 downregulated DEGs were screened from NASH subjects (Fig. 1A). Moreover, 159 upregulated DEGs

Table 1 Clinicopathological features of patients in different datasets

Variables	GSE130970			GSE126848			
	NT	NAFL	NASH	NT	Obesity	NAFL	NASH
Age (years)							
Mean	53	50.5	53.9	/	/	/	/
Gender							
Male	1	5	9	14	12	9	12
Female	3	5	17	0	0	6	4
Lobular inflammation grade							
0	4	3	0	/	/	/	/
1	0	7	18	/	/	/	/
2	0	0	8	/	/	/	/
Cytological ballooning grade							
0	4	10	1	/	/	/	/
1	0	0	7	/	/	/	/
2	0	0	18	/	/	/	/
Steatosis grade							
0	4	0	0	/	/	/	/
1	0	9	2	/	/	/	/
2	0	1	13	/	/	/	/
3	0	0	11	/	/	/	/
Fibrosis stage							
0	4	7	1	/	/	/	/
1	0	3	12	/	/	/	/
2	0	0	4	/	/	/	/
3	0	0	8	/	/	/	/
4	0	0	1	/	/	/	/
NAFLD activity score							
NAS=0	4	0	0	/	/	/	/
0<NAS<3	0	10	0	/	/	/	/
NAS≥5	0	0	26	/	/	/	/

and 191 downregulated DEGs were found in obese patients. The Venn diagram showed that 8 upregulated DEGs (including TP53I3, FNDC5, INHBE, SERPINE1, PRKCE, ACSL4, IP6K3 and CXCL10), and 4 downregulated DEGs (including B3GAT1, P4HA1, IGFBP2 and GPR88) were shared by all three groups (Fig. 1B).

In order to elucidate the functional gene sets involved in the process of obesity, NAFL and NASH, we conducted the GO biological process enrichment using DEGs from each period, respectively. In the obesity group, the top 2 upregulated gene sets ranked by normalized enrichment score (NES) were blood coagulation, hemostasis, and the top 2 downregulated gene sets were epidermis development, cornification (Fig. 1C). In the NAFL group, the top 2 upregulated gene sets were cell cycle phase transition, mitotic cell cycle process, and the top 2 downregulated gene sets were actin-myosin filament sliding, digestion (Fig. 1D). For DEGs in NASH, the top 2 upregulated gene sets were DNA replication, cell cycle process, and the top 2 downregulated gene sets were antimicrobial humoral response, digestion (Fig. 1E).

Dynamic landscape of DEGs during NAFLD progression

To further explore the markers during NAFLD progress, dynamic profiles of all DEGs in Fig. 1B were analyzed. Altogether 60 expression trends were determined in our study. Here we summarized four featured expression trends among these profiles, and the top five DEGs ranked by expression fold changes in each profile were presented in the line charts. The first trend was that the DEGs specifically changed in obesity group, including profile 41 and profile 20 (Fig. 2B). The second trend was that the DEGs particularly changed in NAFL group. Profile 36 and profile 50 were upregulated, and profile 11 was downregulated in NAFL compared with other stages (Fig. 2C). The third trend was DEGs only changed in NASH group. HIST1H1D and THBS2 in profile 31 was in line with this trend (Fig. 2D). The fourth trend was that DEGs changed in all three stages. We found that DEGs in profile 46, 49, 59, 57, 55, 54, 58, 47, 56 were upregulated, and profile 15, 0, 3, 12, 6, 1, 14 were downregulated (Supplementary table 1). For example, EME1, OLFM2, PNMT, PRKCE, SPTBN5 in profile 46 were upregulated, and DCD, KRT14, SCGB2A2, PIP, SCGB1D2 in profile 15 were downregulated in all three periods (Fig. 2E).

Dynamic network analysis in health- and obesity-NAFL-NASH sequences

To explore the common and distinct mechanisms of healthy or obese subjects-derived NAFLD, dynamic networks were constructed by Dynet in the Cytoscape software. As a result, 20 key nodes in the health-NAFL-NASH sequence and 178 key nodes in the obesity-NAFL-NASH sequence were identified by DyNet (Fig. 3A). Distinct pathways of these rewired nodes in different sequences were analyzed by KEGG (Fig. 3B). Only eight pathways including p53 signaling pathway, AGE – RAGE signaling pathway in diabetic complications, PPAR signaling pathway, Endocrine resistance, Toll – like receptor signaling pathway, HIF-1 signaling pathway, Relaxin signaling pathway, and Apoptosis were significantly enriched in the health-NAFL-NASH sequence. The top 10 KEGG pathways enriched in obesity-NAFL-NASH sequence included cell cycle, Chemokine signaling pathway, Progesterone-mediated oocyte maturation, Oocyte meiosis, p53 signaling pathway, Viral protein interaction with cytokine and cytokine receptor, Toll – like receptor signaling pathway, Cellular senescence, Pertussis and Neuroactive ligand – receptor interaction (Fig. 3B). Nine rewired nodes COL3A1, CXCL9, CYCS, CXCL10, THY1, COL1A2, SAA1, CDKN1A, and JUN were commonly included in both health- and obesity-NAFL-NASH sequences (Fig. 3C). The expression levels of these nine node genes were all up-regulated during NAFLD progression derived from obese or non-obese individuals (Fig. 3D), suggesting these nine genes may play fundamental roles in NAFLD development. The PPI network of these nine key nodes was displayed in Fig. 3E.

We next analyzed the correlation between these nine genes and the clinical parameters (Fig. 3F). We demonstrated that all of the nine genes were positively correlated with the NAFLD activity score, and CYCS has the highest correlation ($r=0.7$, $P=4.96 \times 10^{-7}$). CDKN1A ($r=0.53$, $p=4.28 \times 10^{-4}$) was the most relevant gene with age. Moreover, CYCS was the most relevant gene with the lobular inflammation grade ($r=0.59$, $P=6.95 \times 10^{-5}$) and the cytological ballooning grade ($r=0.62$, $P=1.66 \times 10^{-5}$). CXCL10 ($r=0.55$, $P=2.37 \times 10^{-4}$) was the most relevant gene with the steatosis grade, and THY1 ($r=0.67$, $P=2.04 \times 10^{-6}$) was the most relevant gene with the fibrosis stage.

(See figure on next page.)

Fig. 1 Function annotation of DEGs in obese, NAFL and NASH patients by GSEA. **A** Consistently changed DEGs between health, NAFL and NASH in GSE126848 and GSE130970. **B** DEGs in obesity NAFL and NASH. **C–E** The top 5 significant upregulated and downregulated gene sets based on normalized enrichment score (NES) were presented as ridgeline plots (**C** obese; **D** NAFL; **E** NASH). The top 2 upregulated and downregulated gene sets in disease groups (**C** obesity; **D** NAFL; **E** NASH) were ranked by NES on the right

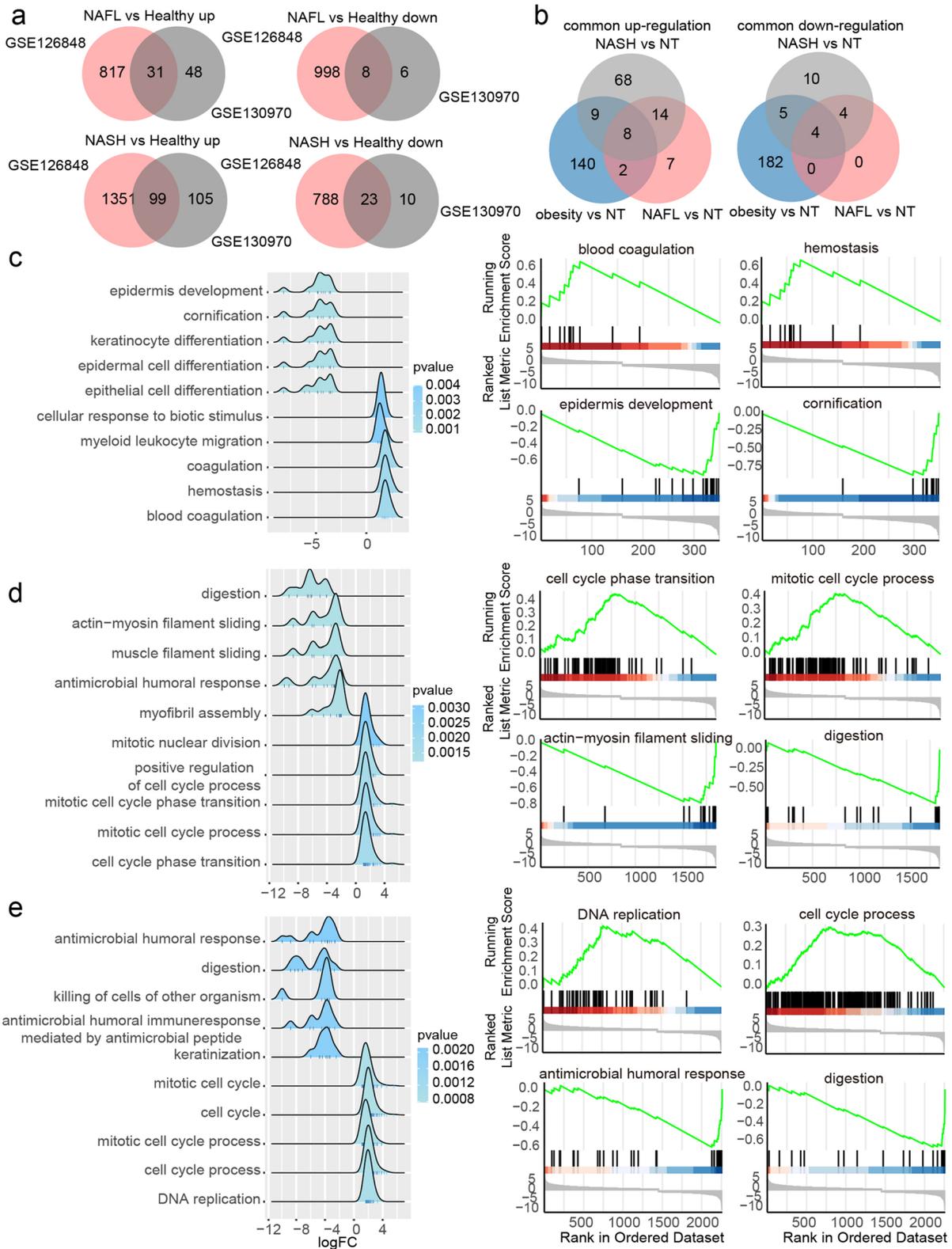
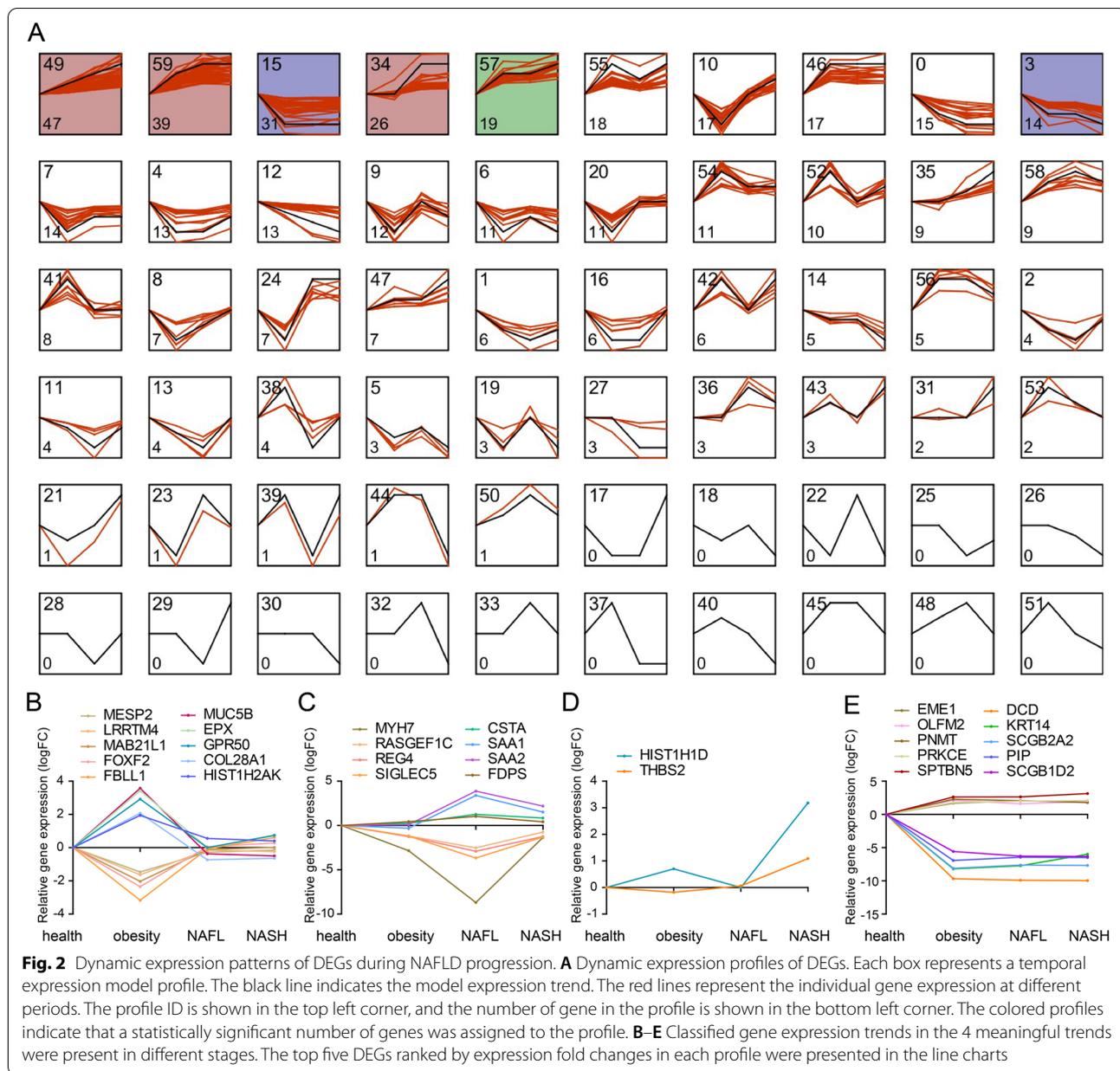


Fig. 1 (See legend on previous page.)



Identification of dynamic co-expression module during NAFLD progression

To investigate the dynamic co-expression networks during NAFLD progression, CEMiTool was carried out to disclose highly correlated gene modules. The sample dendrogram and clinical trait heatmap were displayed in Fig. 4A. A power of $\beta=3$ was selected by scale independence and mean connectivity (Fig. 4B and C). Five modules were enriched by co-expression network analysis. Module 4 (M4) was inhibited in health and obesity groups, and activated in NAFL and NASH. M2 was inhibited in health and obesity groups, and only activated

in NASH. M5 was activated in health and obesity groups, and only inhibited in NAFL. M3 was activated in health and obesity groups, and inhibited in NAFL and NASH (Fig. 4D). The relationships between these modules and nine fundamental nodes from dynamic networks were analyzed by WGCNA (Fig. 4E). JUN, CYCS, COL1A2, THY1, CXCL10, CXCL9 and COL3A1 were positively correlated with M2. CDKN1A, SAA1, CYCS, CXCL10 were negatively correlated with M3. CDKN1A, CYCS, COL1A2, THY1, CXCL10, and COL3A1 were positively correlated with M4. None of the nine rewired nodes were significantly correlated with M5.

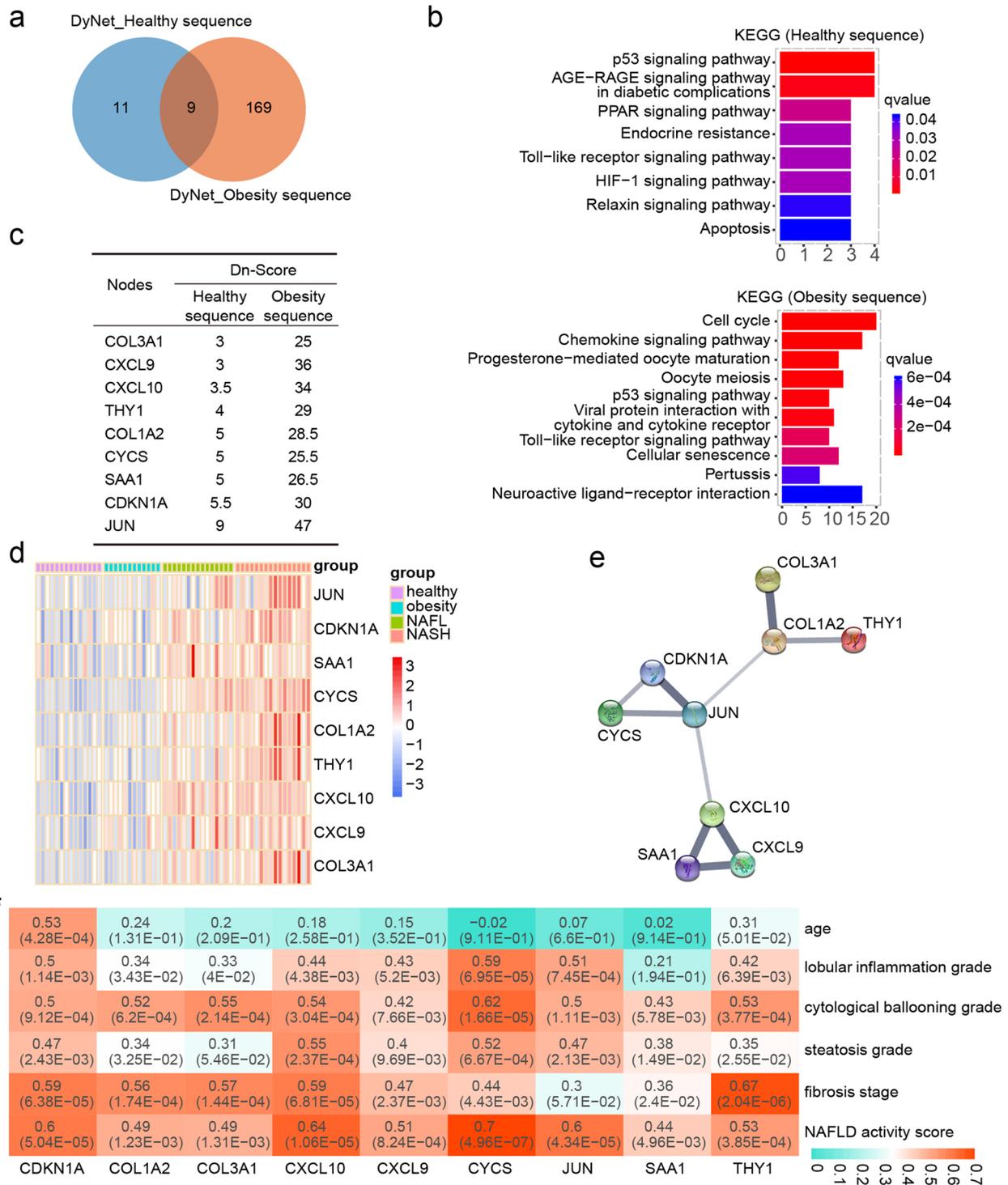


Fig. 3 Dynamic molecular networks along the NAFLD progression sequence. **A** The node genes involved in dynamic networks of obesity and non-obesity derived NAFLD. Dynamic networks were created and Dn-Scores of node genes were calculated by DyNet. **B** Pathway enrichment of node genes in healthy- or obesity-NAFLD sequences, respectively. **C** The Dn-Score of common nodes shared by healthy- and obesity-NAFLD sequences. **D** Expression heatmap of nine common nodes at different NAFLD stages. **E** The protein-protein interactions of nine common nodes were retrieved by the STRING database. **F** Pearson correlation analysis of the nine nodes and clinical parameters of NAFLD

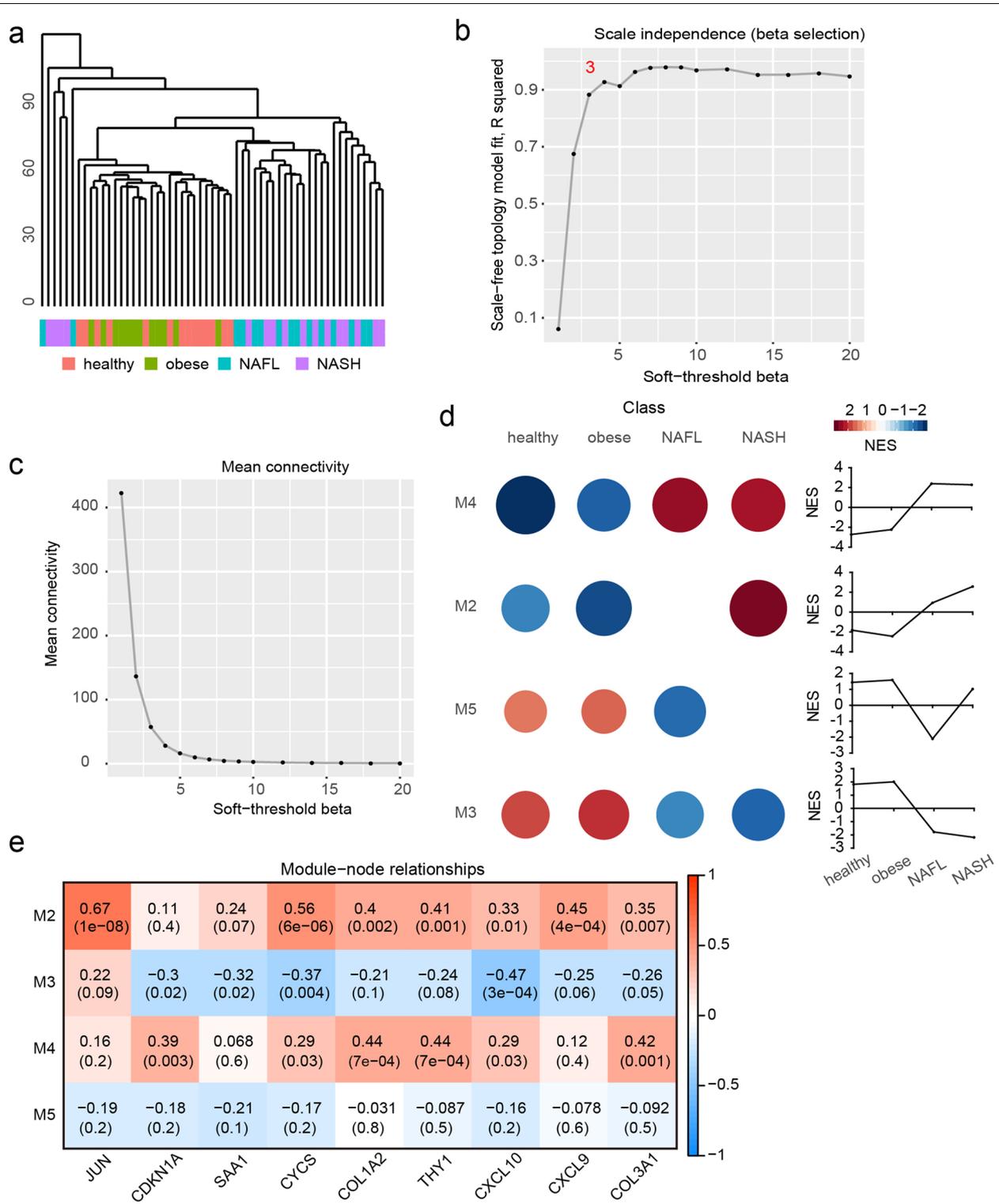
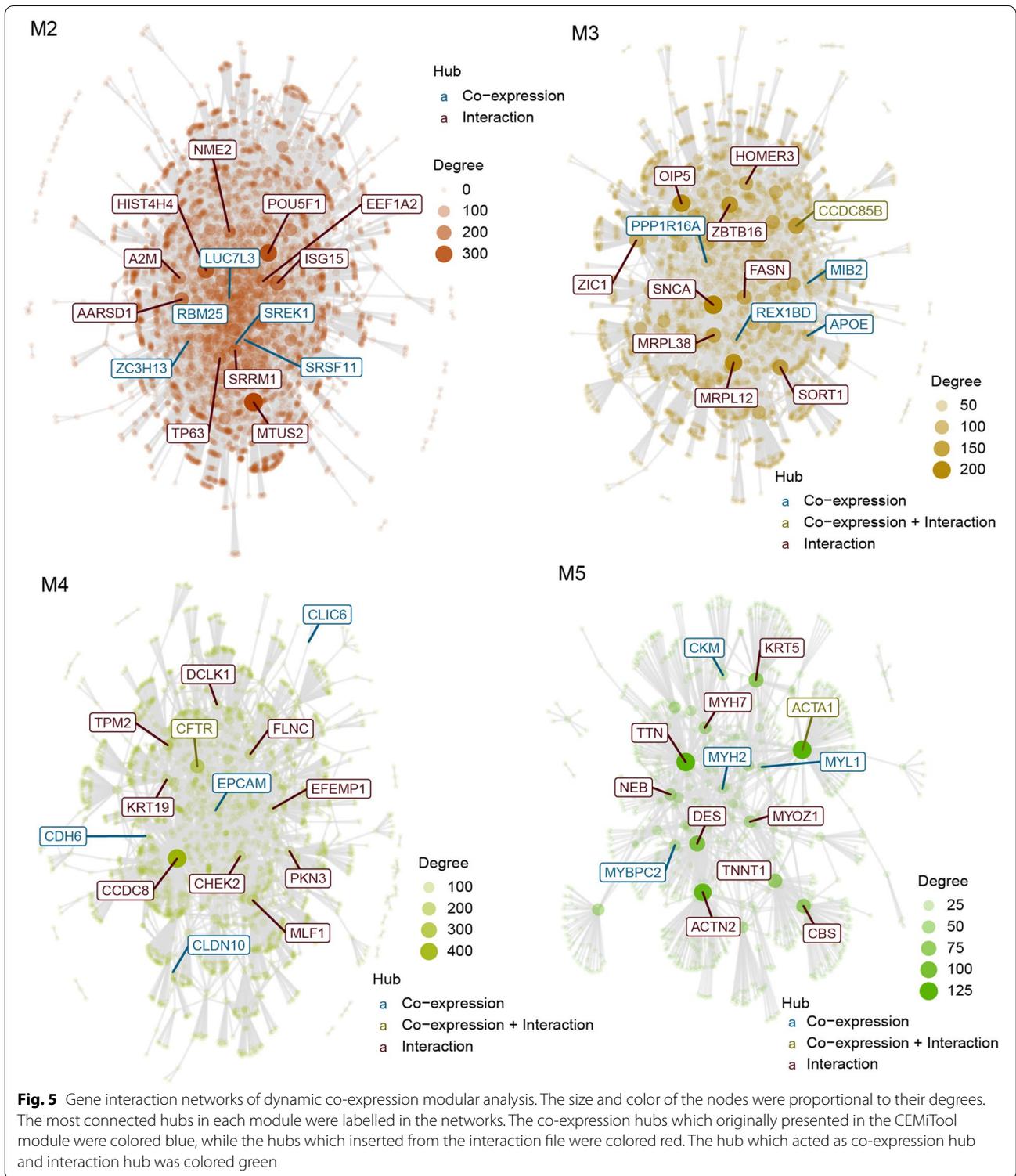
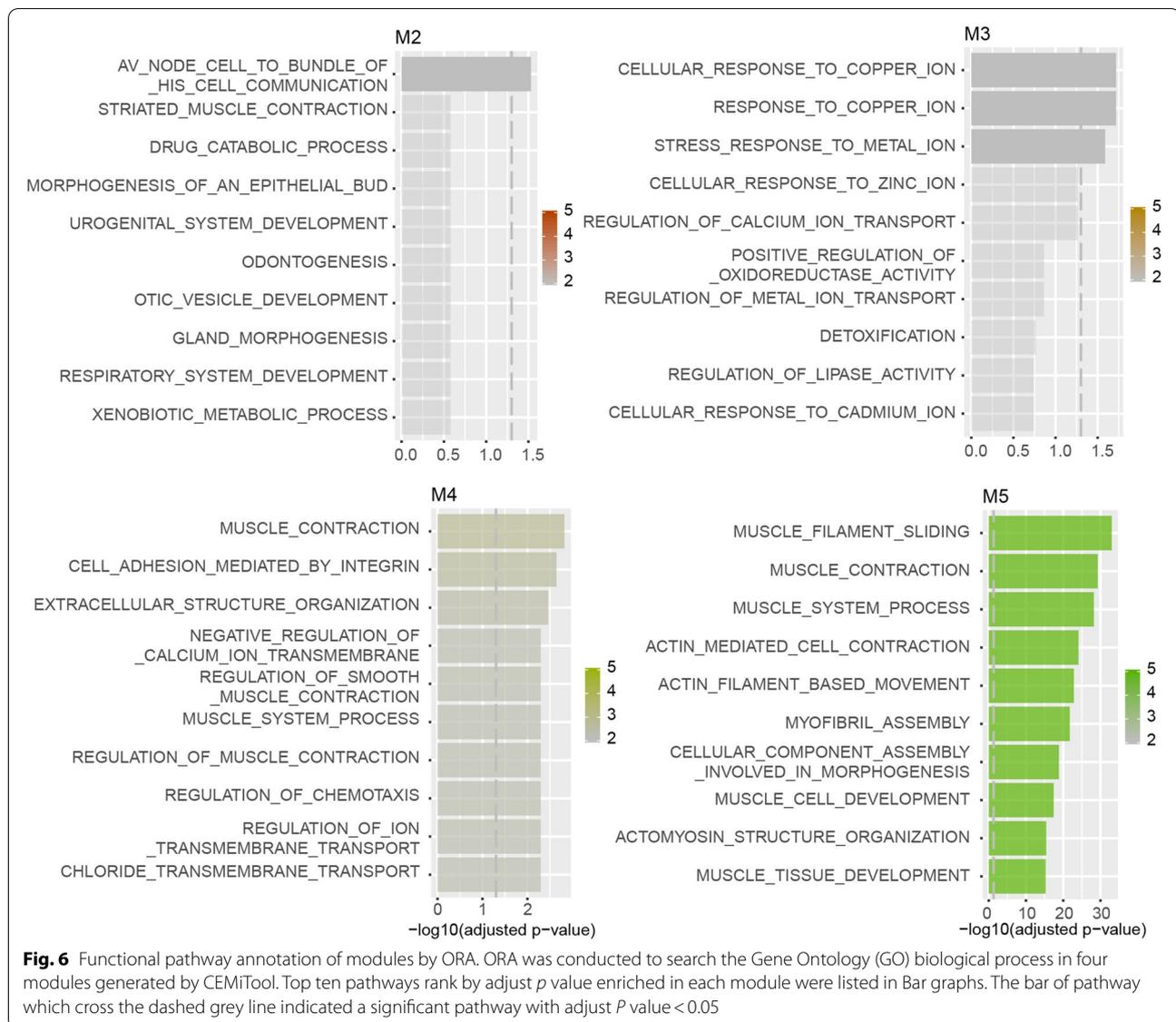


Fig. 4 Dynamic co-expression module identification during NAFLD progression. **A** Sample dendrogram and clinical trait heatmap. **B** Analysis of scale independence for various soft-thresholding powers. **C** Analysis of mean connectivity for various soft-thresholding powers. **D** Modules were enriched by co-expression network analysis. The activities of modules at different stage during NAFLD progression were calculated by NES from GSEA, and the modules enriched in each stage with adjust P value < 0.05 were shown in heatmap. **E** Module-node relationships were assessed by WGCNA. Nine common node genes identified above were selected to evaluate the correlation with Modules





Gene interaction networks of dynamic co-expression modules

Then, we analyzed the gene networks of the dynamic modules (Fig. 5). SREK1, SRSF11, ZC3H13, RBM25, LUC7L3 were co-expression hubs of M2. NME2, HIST4H4, POU5F1, EEF1A2, A2M, ISG15, AARSD1, SRRM1, TP63, and MTUS2 were interaction hubs of M2. MIB2, REX1BD, PPP1R16A, APOE, CCDC85B were co-expression hubs of M3. HOMER3, OIP5, ZBTB16, ZIC1, SNCA, FASN, CCDC85B, MRPL38, MRPL12, and SORT1 were interaction hubs of M3. CLDN10, EPCAM, CFTR, CDH6, CLIC6 were co-expression hubs of M4. DCLK1, TPM2, FLNC, KRT19, EFEMP1, CCDC8, CHEK2, PKN3, CFTR and MLF1 were interaction hubs of M4. MYL1, ACTA1, CKM, MYH2, MYBPC2 were

co-expression hubs of M5. KRT5, MYH7, TTN, NEB, DES, MYL1, MYOZ1, TNNT1, CBS and ACTN2 were interaction hubs of M5.

Functional pathway annotation of modules by ORA

ORA was conducted to search the GO biological process in modules (Fig. 6). Av node cell to bundle of his cell communication was enriched in M2. Response to copper ion, cellular response to copper ion, and stress response to metal ion were enriched in M3. Muscle contraction, cell adhesion mediated by integrin, extracellular structure organization, negative regulation of calcium ion transmembrane transport, regulation of chemotaxis, regulation of muscle contraction, muscle system process, regulation of smooth muscle contraction, and

negative regulation of ion transmembrane transport were enriched in M4. Muscle filament sliding, muscle contraction, muscle system process, actin mediated cell contraction, actin filament based movement, myofibril assembly, cellular component assembly involved in morphogenesis, muscle cell development, actomyosin structure organization, and muscle tissue development were enriched in M5.

Discussion

NAFLD represents a spectrum of liver disorders including NAFL and NASH, while NASH is tightly associated with the end stage of liver disease [13]. Obesity is one of the main causes of NAFLD. Thus, it is urgent to prevent the progressions from obesity to NAFL and from NAFL to NASH. Here, we attempted to explore the key molecular pathways and dynamic co-expression networks along the NAFLD progression course.

Functional pathway annotation of DEGs showed that cell cycle related pathways were upregulated in NAFL and NASH. It is consistent with previous research that canagliflozin could attenuate the development of NASH partly by the induction of cell cycle arrest [14]. Our data indicated that cell cycle related pathways also participated in NAFL. Aravinthan A found the permanent cell cycle arrest in NAFLD [15], which indicated an opposite role in NAFLD. Therefore, the roles of cell cycle in NAFLD need further exploration.

To investigate the gene alteration trends from obesity to NASH, the dynamic profiles of DEGs were conducted. In the current study, we screened several expression models that specifically changed at particular disease stages, which could be novel markers to identifying obesity, NAFL and NASH, respectively. SAA1 and SAA2 are serum amyloid A family members, and only upregulated in NAFL. They are highly expressed in response to liver inflammation [16], and play important roles in lipid metabolism [17]. Moreover, SAA1 was one of the 9 rewired nodes in healthy and obesity-NAFL-NASH sequences. SAA1 acted as a hub gene of PPI network constructed by DEGs from NAFLD dataset GSE106737 and GSE83452 [18]. Increased hepatocyte SAA1 aggravated liver inflammation in NAFLD [19]. Our findings suggested that SAA1 and SAA2 may play key roles in early NAFL stage. THBS2 only upregulated in NASH. It has been proved that the gene expression of THBS2 increased in fatty liver of NAFLD [20]. THBS2 increased during diet-induced mice hepatic fibrosis progression [21] and was up-regulated in the fibrosis stage 3–4 state of NAFLD patients [22]. We speculated that THBS2 might be related to the hepatic fibrosis in the late NASH period. Besides, we also found DEGs that were

obviously changed in all periods. PRKCE is a member of the Protein kinase C family. Protein kinase C members, such as PKC δ , are closely related to insulin sensitivity and body glucose tolerance [23]. Activation of PRKCE links the NAFLD to hepatic insulin resistance [24, 25]. PRKCE has been proven to be a critical DEG in NAFLD, and the expression level and DNA methylation of PRKCE and IGFBP2 were altered in obesity-derived NAFLD [26]. Hepatic steatosis leads to hepatic insulin resistance by activating PRKCE [27]. These results were consistent with our data, PRKCE kept upregulating at obesity, NAFL, and NASH stages. We considered that the PRKCE regulated the lipid abundance, steatosis, insulin resistance, thereby modulated the pathogenetic process of NAFLD. These results indicated the important roles of these DEGs in obese and NAFLD development.

Although NAFLD is commonly associated with obesity, it is increasingly being identified in non-obese patients. Therefore, we further investigate the common and distinct mechanisms between the healthy-NAFL-NASH sequence and the obesity-NAFL-NASH sequence. The pathophysiology of non-obese NAFLD is still not clear [28]. It has been reported that metabolic syndrome promotes the progression to NASH in non-obese patients with NAFLD [29]. Our result indicated that AGE – RAGE signaling pathway in diabetic complications, PPAR signaling pathway, Endocrine resistance were enriched in healthy-NAFL-NASH sequence, which are all associated with metabolic syndrome. On the other hand, the obesity-NAFLD sequence was mainly enriched in cell cycle, chemokine signaling pathway, cellular senescence.

By combination analysis of dynamic networks and dynamic co-expression modules, we also identified nine genes that essential for the progression of NAFLD derived from obese and non-obese individuals. These genes mainly contained collagens, chemokines and oncogenes. CYCS has the highest correlation with NAFLD activity score, lobular inflammation grade and the cytological ballooning grade. CYCS functions as a central component of the electron transport chain in mitochondria. Although the role and effect of CYCS during NAFLD progression have not been explored, a recent study demonstrated that CYCS was associated with hepatic lipid metabolic misalignment [30]. These findings suggested that CYCS might be a novel regulator in promoting the progression of NAFLD. The PPI network showed that CYCS interacted with CDKN1A and JUN. JUN was most positively correlated with M2. JUN and EGR were reported to drive the reprogramming of the Kupffer cell to a scar-associated macrophage phenotype by changing the liver X receptor

functions during diet-induced NASH [31]. We also demonstrated that M2 was only activated in NASH, and we surmised that this module may contribute to liver fibrosis to some extent. M3 was inhibited in NAFL and NASH, and was most negatively correlated with CXCL10. CXCL10 has been reported to inhibit autophagic protein degradation and the accumulation of ubiquitinated proteins, thereby promoted the development of steatohepatitis [32]. Blockade of CXCL10 protected against steatohepatitis development in mice, and CXCL10 levels were significantly higher in human NASH [33–35]. We found that CXCL10 upregulated in NAFL and NASH samples, and was indeed the most relevant gene with the steatosis grade. Our result demonstrated that M4 was activated in NAFL and NASH, and may relate to the development of NAFLD as well. M4 was most positively correlated with COL1A2 and THY1. It has been reported that in severe NAFLD patients the levels of pro-IL1 β mRNA correlate with the expression of COL1A1 [36]. Similar to COL1A1, COL1A2 is also a fibrosis-related gene [37], however, the role of COL1A2 in NAFLD has not been fully elucidated. We found that THY1 was the most relevant gene with the fibrosis stage. THY1 encodes a cell surface glycoprotein and member of the immunoglobulin superfamily of proteins, while little is known about its effects on NAFLD.

Pathway enrichment showed that response to copper ion and cellular response to copper ion were enriched in M3. Copper has been proven to play a major role in NAFLD [38], and copper homeostasis could counteract the progression of NAFLD [39]. In the hub genes of M3, SNCA and APOE were reported to be associated with copper ion related pathways. SNCA interacts with copper in Parkinson's disease [40]. APOE deletion has no effect on copper-induced oxidative stress in the mice brain [41], but APOE might participate in protecting the liver from copper-induced damage [42]. The relationships between hub genes and copper ion related pathways remained unclear in NAFLD. FASN was an interaction hub gene in M3. FASN catalyzed the synthesis of palmitate and long-chain saturated fatty acids [43]. FASN expression was correlated significantly with the degree of hepatic steatosis, but not with inflammation or ballooning of hepatocytes [44]. The result was similar to another M3 correlated gene CXCL10, which was also relevant to the steatosis grade. Therefore, we speculated that M3 may closely related to hepatic steatosis. Cell adhesion mediated by integrin and regulation of chemotaxis were enriched in M4. Loss of cell adhesion molecule-1 had beneficial effects in NASH development by reducing inflammation, and β 7-integrin-deficiency results in increased steatohepatitis [45]. In the hub genes of M4,

EPCAM and EFEMP1 were associated with NAFLD. EPCAM, the epithelial cell adhesion molecule was upregulated in NASH [46]. EFEMP1 was identified as a hub gene in NAFLD fibrosis [22].

Conclusion

In summary, our study identified a nine-gene signature as the potential key regulator in NAFLD progression. The results provided potential pivotal genes and the clinical markers during NAFLD progression. Further experiments are still needed to explore the function of dynamic co-expression networks in NAFLD.

Abbreviations

NAFLD: Nonalcoholic fatty liver disease; DEGs: Differentially expressed genes; NASH: Nonalcoholic steatohepatitis; NAFL: Nonalcoholic fatty liver; GSEA: Gene set enrichment analysis; GO: Gene Ontology; NAS: NAFLD activity score; FDR: False Discovery Rate; FC: Fold change; STEM: Short Time-series Expression Miner; PPI: Protein–protein interaction; NES: Normalized enrichment score.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s41065-021-00196-8>.

Additional file 1: Supplementary figure 1. The flow diagram of dynamic network analysis.

Additional file 2: Supplementary table 1. DEGs in different profiles.

Acknowledgements

Not applicable.

Authors' contributions

Songqiang Yao designed the study. Jing Zheng, Huizhong Wu and Zhiying Zhang analyzed the data and prepared the figures. Jing Zheng wrote the manuscript, and Songqiang Yao revised the manuscript. The authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

The data are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Pharmacy, Zhejiang Medical & Health Group Hangzhou Hospital, No.1 Banshan Road, Kangjian nong, Hangzhou 310022, China.

²Department of Pharmacy, Zhejiang Quhua Hospital, Quzhou 324002, China.

³Department of Pharmacy, Hangzhou Jianggan District People's Hospital, Hangzhou 310016, China.

Received: 27 January 2021 Accepted: 4 August 2021
Published online: 21 August 2021

References

- Samuel VT, Shulman GI. Nonalcoholic fatty liver disease as a nexus of metabolic and hepatic diseases. *Cell Metab.* 2018;27(1):22–41.
- Alvarez CS, Graubard BI, Thistle JE, Petrick JL, McGlynn KA. Attributable fractions of nonalcoholic fatty liver disease for mortality in the United States: results from the third National Health and Nutrition Examination Survey with 27 years of follow-up. *Hepatology.* 2020;72(2):430–40.
- Romero-Gomez M, Zelber-Sagi S, Trenell M. Treatment of NAFLD with diet, physical activity and exercise. *J Hepatol.* 2017;67(4):829–46.
- Pais R, Barritt AT, Calmus Y, Scatton O, Runge T, Lebray P, Poynard T, Ratzl V, Conti F. NAFLD and liver transplantation: current burden and expected challenges. *J Hepatol.* 2016;65(6):1245–57.
- Suppli MP, Rigbolt K, Veidal SS, Heeboll S, Eriksen PL, Demant M, Bagger JI, Nielsen JC, Oro D, Thrane SW, et al. Hepatic transcriptome signatures in patients with varying degrees of nonalcoholic fatty liver disease compared with healthy normal-weight individuals. *Am J Physiol Gastrointest Liver Physiol.* 2019;316(4):G462–72.
- Wruck W, Kashofer K, Rehman S, Daskalaki A, Berg D, Gralka E, Jozefczuk J, Drews K, Pandey V, Regenbrecht C, et al. Multi-omic profiles of human non-alcoholic fatty liver disease tissue highlight heterogenic phenotypes. *Sci Data.* 2015;2:150068.
- Hoang SA, Oseini A, Feaver RE, Cole BK, Asgharpour A, Vincent R, Siddiqui M, Lawson MJ, Day NC, Taylor JM, et al. Gene expression predicts histological severity and reveals distinct molecular profiles of nonalcoholic fatty liver disease. *Sci Rep.* 2019;9(1):12541.
- Lawrence M, Gentleman R, Carey V. rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics.* 2009;25(14):1841–2.
- Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic.* 2012;16(5):284–7.
- Li L, Pan Z, Yang X. Identification of dynamic molecular networks in peripheral blood mononuclear cells in type 1 diabetes mellitus. *Diabetes Metab Syndr Obes.* 2019;12:969–82.
- Goenawan IH, Bryan K, Lynn DJ. DyNet: visualization and analysis of dynamic molecular interaction networks. *Bioinformatics.* 2016;32(17):2713–5.
- Russo P, Ferreira GR, Cardozo LE, Burger MC, Arias-Carrasco R, Maruyama SR, Hirata T, Lima DS, Passos FM, Fukutani KF, et al. CEMiTool: a Bioconductor package for performing comprehensive modular co-expression analyses. *BMC Bioinformatics.* 2018;19(1):56.
- Schuppan D, Surabattula R, Wang XY. Determinants of fibrosis progression and regression in NASH. *J Hepatol.* 2018;68(2):238–50.
- Jojima T, Wakamatsu S, Kase M, Iijima T, Maejima Y, Shimomura K, Kogai T, Tomaru T, Usui I, Aso Y. The SGLT2 inhibitor canagliflozin prevents carcinogenesis in a mouse model of diabetes and non-alcoholic steatohepatitis-related hepatocarcinogenesis: association with SGLT2 expression in hepatocellular carcinoma. *Int J Mol Sci.* 2019;20(20):5237.
- Aravinthan A, Scarpini C, Tachtatzis P, Verma S, Penrhyn-Lowe S, Harvey R, Davies SE, Allison M, Coleman N, Alexander G. Hepatocyte senescence predicts progression in non-alcohol-related fatty liver disease. *J Hepatol.* 2013;58(3):549–56.
- Kaminsky-Kolesnikov Y, Raubach E, Abu-Halala D, Hahn M, Garcia-Ruiz C, Fernandez-Checa JC, Madar Z, Tirosh O. Cholesterol induces Nrf-2- and HIF-1alpha-dependent hepatocyte proliferation and liver regeneration to ameliorate bile acid toxicity in mouse models of NASH and fibrosis. *Oxid Med Cell Longev.* 2020;2020:5393761.
- Leclerc D, Christensen KE, Cauvi O, Yang E, Fournelle F, Bahous RH, Malyshova OV, Deng L, Wu Q, Zhou Z, et al. Mild methylenetetrahydrofolate reductase deficiency alters inflammatory and lipid pathways in liver. *Mol Nutr Food Res.* 2019;63(3):e1801001.
- Chen F, Zhou Y, Wu Z, Li Y, Zhou W, Wang Y. Integrated analysis of key genes and pathways involved in nonalcoholic steatohepatitis improvement after Roux-en-Y gastric bypass surgery. *Front Endocrinol (Lausanne).* 2020;11:611213.
- Li D, Xie P, Zhao S, Zhao J, Yao Y, Zhao Y, Ren G, Liu X. Hepatocytes derived increased SAA1 promotes intrahepatic platelet aggregation and aggravates liver inflammation in NAFLD. *Biochem Biophys Res Commun.* 2021;555:54–60.
- Huang S, Sun C, Hou Y, Tang Y, Zhu Z, Zhang Z, Zhang Y, Wang L, Zhao Q, Chen MG, et al. A comprehensive bioinformatics analysis on multiple Gene Expression Omnibus datasets of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. *Sci Rep.* 2018;8(1):7630.
- Lytle KA, Wong CP, Jump DB. Docosahexaenoic acid blocks progression of western diet-induced nonalcoholic steatohepatitis in obese Ldlr^{-/-} mice. *PLoS One.* 2017;12(4):e173376.
- Lou Y, Tian GY, Song Y, Liu YL, Chen YD, Shi JP, Yang J. Characterization of transcriptional modules related to fibrosing-NAFLD progression. *Sci Rep.* 2017;7(1):4748.
- Li M, Vienberg SG, Bezy O, O'Neill BT, Kahn CR. Role of PKCdelta in insulin sensitivity and skeletal muscle metabolism. *Diabetes.* 2015;64(12):4023–32.
- Petersen MC, Madiraju AK, Gassaway BM, Marcel M, Nasiri AR, Butrico G, Marcucci MJ, Zhang D, Abulizi A, Zhang XM, et al. Insulin receptor Thr1160 phosphorylation mediates lipid-induced hepatic insulin resistance. *J Clin Invest.* 2016;126(11):4361–71.
- Ter Horst KW, Gilljamse PW, Versteeg RI, Ackermans MT, Nederveen AJ, la Fleur SE, Romijn JA, Nieuwdorp M, Zhang D, Samuel VT, et al. Hepatic diacylglycerol-associated protein kinase Cepsilon translocation links hepatic steatosis to hepatic insulin resistance in humans. *Cell Rep.* 2017;19(10):1997–2004.
- Ahrens M, Ammerpohl O, von Schonfels W, Kolarova J, Bens S, Itzel T, Teufel A, Herrmann A, Brosch M, Hinrichsen H, et al. DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. *Cell Metab.* 2013;18(2):296–302.
- Samuel VT, Liu ZX, Qu X, Elder BD, Bilz S, Befroy D, Romanelli AJ, Shulman GI. Mechanism of hepatic insulin resistance in non-alcoholic fatty liver disease. *J Biol Chem.* 2004;279(31):32345–53.
- Eslam M, Fan JG, Mendez-Sanchez N. Non-alcoholic fatty liver disease in non-obese individuals: the impact of metabolic health. *Lancet Gastroenterol Hepatol.* 2020;5(8):713–5.
- Kawaguchi T, Torimura T. Is metabolic syndrome responsible for the progression from NAFLD to NASH in non-obese patients? *J Gastroenterol.* 2020;55(3):363–4.
- Wei L, Zhao C, Dong S, Yao S, Ji B, Zhao B, Liu Z, Liu X, Wang Y. Secoisolaricresinol diglucoside alleviates hepatic lipid metabolic misalignment involving the endoplasmic reticulum-mitochondrial axis. *Food Funct.* 2020;11(5):3952–63.
- Seidman JS, Troutman TD, Sakai M, Gola A, Spann NJ, Bennett H, Bruni CM, Ouyang Z, Li RZ, Sun X, et al. Niche-specific reprogramming of epigenetic landscapes drives myeloid cell diversity in nonalcoholic steatohepatitis. *Immunity.* 2020;52(6):1057–1074.e7.
- Zhang X, Wu WK, Xu W, Man K, Wang X, Han J, Leung WY, Wu R, Liu K, Yu J. C-X-C motif chemokine 10 impairs autophagy and autolysosome formation in non-alcoholic steatohepatitis. *Theranostics.* 2017;7(11):2822–36.
- Zhang X, Shen J, Man K, Chu ES, Yau TO, Sung JC, Go MY, Deng J, Lu L, Wong VW, et al. CXCL10 plays a key role as an inflammatory mediator and a non-invasive biomarker of non-alcoholic steatohepatitis. *J Hepatol.* 2014;61(6):1365–75.
- Pan X, Chiwanda KA, Liu A, Wen SW, Chen J, Luo J. Chemokines in non-alcoholic fatty liver disease: a systematic review and network meta-analysis. *Front Immunol.* 1802;2020:11.
- Tomita K, Freeman BL, Bronk SF, LeBrasseur NK, White TA, Hirsova P, Ibrahim SH. CXCL10-mediates macrophage, but not other innate immune cells-associated inflammation in murine nonalcoholic steatohepatitis. *Sci Rep.* 2016;6:28786.
- Wree A, McGeough MD, Peña CA, Schlattjan M, Li H, Inzaugarat ME, Messer K, Canbay A, Hoffman HM, Feldstein AE. NLRP3 inflammasome activation is required for fibrosis development in NAFLD. *J Mol Med (Berl).* 2014;92(10):1069–82.
- Ying L, Yan F, Zhao Y, Gao H, Williams BR, Hu Y, Li X, Tian R, Xu P, Wang Y. (-)-Epigallocatechin-3-gallate and atorvastatin treatment down-regulates liver fibrosis-related genes in non-alcoholic fatty liver disease. *Clin Exp Pharmacol Physiol.* 2017;44(12):1180–91.
- Heffernan MC, Park HM, Au-Yeung HY, Van de Bittner GC, Ackerman CM, Stahl A, Chang CJ. In vivo bioluminescence imaging reveals copper

- deficiency in a murine model of nonalcoholic fatty liver disease. *Proc Natl Acad Sci U S A*. 2016;113(50):14219–24.
39. Antonucci L, Porcu C, Iannucci G, Balsano C, Barbaro B. Non-alcoholic fatty liver disease and nutritional implications: special focus on copper. *Nutrients*. 2017;9(10):1137.
 40. Valensin D, Dell'Acqua S, Kozlowski H, Casella L. Coordination and redox properties of copper interaction with alpha-synuclein. *J Inorg Biochem*. 2016;163:292–300.
 41. Chen Y, Wang L, Geng JH, Zhang HF, Guo L. Apolipoprotein E deletion has no effect on copper-induced oxidative stress in the mice brain. *Biosci Rep*. 2018;38(5):BSR20180719.
 42. Chen Y, Li B, Zhao RR, Zhang HF, Zhen C, Guo L. Increased sensitivity of apolipoprotein E knockout mice to copper-induced oxidative injury to the liver. *Biochem Biophys Res Commun*. 2015;459(3):529–33.
 43. Wruck W, Graffmann N, Kawala MA, Adjaye J. Concise review: current status and future directions on research related to nonalcoholic fatty liver disease. *Stem Cells*. 2017;35(1):89–96.
 44. Dorn C, Riener MO, Kirovski G, Saugspier M, Steib K, Weiss TS, Gabel E, Kristiansen G, Hartmann A, Hellerbrand C. Expression of fatty acid synthase in nonalcoholic fatty liver disease. *Int J Clin Exp Pathol*. 2010;3(5):505–14.
 45. Drescher HK, Schippers A, Clahsen T, Sahin H, Noels H, Hornef M, Wagner N, Trautwein C, Streetz KL, Kroy DC. beta7-Integrin and MAdCAM-1 play opposing roles during the development of non-alcoholic steatohepatitis. *J Hepatol*. 2017;66(6):1251–64.
 46. Gonzalez-Rodriguez A, Valdecantos MP, Rada P, Addante A, Barahona I, Rey E, Pardo V, Ruiz L, Laiglesia LM, Moreno-Aliaga MJ, et al. Dual role of protein tyrosine phosphatase 1B in the progression and reversion of non-alcoholic steatohepatitis. *Mol Metab*. 2018;7:132–46.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

