Identification of autophagy-related signatures in nonalcoholic fatty liver disease and correlation with non-parenchymal cells of the liver†

Kaiwei Chen, Ling Wei, Shengnan Yu, Ningning He* and Fengjuan Zhang*

Non-alcoholic fatty liver disease (NAFLD) is a chronic hepatic disease. The incidence and prevalence of NAFLD have increased greatly in recent years, and there is still a lack of effective drugs. Autophagy plays an important role in promoting liver metabolism and maintaining liver homeostasis, and defects in autophagy levels are considered to be related to the development of NAFLD. However, the molecular mechanisms of autophagy in NAFLD still remain unknown. In this study, we identified 6 autophagy-associated hub genes using gene expression profiles obtained from the GSE48452 and GSE89632 datasets. Biomarkers were screened according to gene significance (GS) and module membership (MM) using weighted gene co-expression network analysis (WGCNA), and the immune infiltration landscape of the liver in NAFLD patients was explored using the CIBERSORT algorithm. Subsequently, we analyzed the relationship between liver non-parenchymal cells and autophagy-related hub genes using scRNA-seq data (GSE129516). Finally, we separated the NAFLD patients into two groups based on 6 hub genes by consensus clustering and screened 10 potential autophagy-related small molecules based on the cMAP database.

1. Introduction

Non-alcoholic fatty liver disease (NAFLD) is a chronic hepatic disease characterized by fat deposition in more than 5% of hepatocytes and whose causative factors exclude long-term alcohol abuse and specific injuries.1–3 With the global trend of obesity and other metabolic syndromes in recent years, the prevalence of NAFLD has increased greatly and there is still a lack of effective drugs.4–7 Approximately 30% of patients with NAFLD may develop chronic hepatitis, known as non-alcoholic steatohepatitis (NASH), and risk developing cirrhosis and hepatocellular carcinoma.8 Meanwhile, the development of NAFLD also influences other metabolic diseases, such as diabetes9 and hypertension.10 Therefore, the quality of life of NAFLD patients is severely affected and how to treat NAFLD remains a global medical challenge.

Autophagy is a highly conserved cyclic process present in all eukaryotes that relies on the lysosomal pathway for the degradation of cytoplasmic proteins and organelles. There are three main types of autophagy, namely microautophagy, macroautophagy and chaperone-mediated autophagy.11,12 Autophagy plays an important role in cell survival and maintenance, and defects in this process are often associated with the pathology of many diseases.13,14 For example, it has been reported in recent years that autophagy can maintain the stability of the intestinal barrier by influencing tight junction regulation and protecting against cell death, while disruption of the intestinal barrier is an important etiology of chronic inflammatory bowel disease.15 Since the lysosome was discovered in the rat liver in 1963, the concept of autophagy has been rapidly established. Over the next 60 years, more and more studies have shown that autophagy plays an important role in promoting liver metabolism and maintaining liver homeostasis.16 Defective levels of autophagy were considered related to the development for a host of liver diseases, especially NAFLD.17 In genetic and high-fat diet-induced obese mice, significant defects in hepatic autophagy were observed, resulting in the promotion of endoplasmic reticulum stress and the development of insulin resistance.18 It has also been reported that impaired $\delta$-nitrosoglutathione reductase (GSNOR) in obese human liver causes autophagy-lysosomal degradation and that hepatic...
steatosis can be ameliorated by pharmacological activation of autophagy. However, detailed studies of autophagy-related genes in NAFLD are still lacking. Therefore, exploring its gene expression patterns and searching for key genes may provide a new perspective for the treatment of NAFLD.

In this study, we identified hub genes related to autophagy by weighted gene co-expression network analysis and differential analysis in NAFLD based on the GSE48452 and GSE89632 datasets as discovery cohorts. The immune infiltration landscape in the liver was explored using the CIBERSORT algorithm. Subsequently, we verified the expression of hub genes using an animal model and patient liver tissue sections at the gene level and protein level to reveal the effect of NAFLD on autophagy-related genes. We then explored the relationship between hub genes and liver non-parenchymal cells using the scRNA-seq dataset GSE129516, providing new perspectives and therapeutic targets for NAFLD and autophagy. In addition, via the CMAP database, we identified 10 small molecule drugs that may affect autophagy using differential genes after consensus clustering of 6 hub genes.

2. Materials and methods

2.1. Datasets and preprocessing

Gene expression profiles for NAFLD were sourced from the gene expression omnibus (GEO) database. A comprehensive list of 232 autophagy-related genes (ARGs) was extracted from the human autophagy database (Table S1, ESI).

2.2. Differential expression analysis

We used the R package “sva” to remove batch effects between datasets and obtained a discovery cohort containing 34 patients with liver steatosis and 51 healthy individuals by merging the datasets GSE48452 and GSE89632. Differentially expressed genes (DEGs) between individuals with steatosis and healthy controls were determined using the “limma” package, applying a significance threshold of P-value < 0.05 and |log 2FC| > 0.5.

2.3. Functional annotation and pathway enrichment analysis

The R package “org.Hs.eg.db” was employed for gene ontology (GO) annotation of genes as the background reference. The identified genes were subsequently mapped to the background dataset and subjected to enrichment analysis using the R package “clusterProfiler”.

2.4. Construction of a co-expression network and identification of modules

The GSE89632 dataset served as the discovery cohort for the weighted gene co-expression network analysis (WGCNA). To maintain the scale-free topological network structure characteristics and reduce the computational workload, a subset of 7127 genes with a variation coefficient greater than 0.05 was selected for network construction. First, outlier samples were filtered using hierarchical clustering and the appropriate soft threshold β was evaluated using the “pickSoftThreshold” function. Modules were delineated by clustering genes exhibiting comparable expression patterns, employing a minimum module size cutoff of 30 and merging identical modules based on a threshold value of 0.2. Subsequently, module eigengene (ME) was calculated to identify modules significantly associated with the disease. Furthermore, the relationships between modules and clinical characteristics were evaluated by grouping genes displaying akin expression patterns within each module.

2.5. Identification of autophagy-related hub genes and biomarkers

To identify hub genes, we extracted genes that were common to ARGs, DEGs, and the modules that had the strongest correlation with NAFLD features in WGCNA. From these hub genes, potential biomarkers were selected based on a module membership (MM) > 0.8 and a gene significance (GS) > 0.2 in the WGCNA. We evaluated the performance of the selected biomarkers using receiver operating characteristic (ROC) analysis utilizing the R package “pROC.”

2.6. Estimation of immune cell infiltration

The infiltration landscape of immune cells in the liver of healthy individuals and NAFLD patients was evaluated using the CIBERSORT algorithm. The results were visualized with the R packages “ggboxplot” and “corrplot”.

2.7. Single-cell analysis

Dataset GSE129516 contains single-cell sequencing data of liver non-parenchymal cells from 3 NASH mice. The “Seurat” package was utilized for subsequent data processing. Cells and genes with low expression levels were excluded from the analysis, with additional quality control criteria applied to ensure that no more than 5% of mitochondria per cell were present, and that genes had a feature count between 500 and 6000. Following data preprocessing, a total of 15,380 cells were selected for subsequent analysis. A total of 3000 high variant genes (HVGs) were identified using the “vst” method following normalization performed with the “LogNormalize” method. Twenty principal components (PCs) were selected to conduct t-SNE analysis and the resolution parameter was set to “0.2” to identify different clusters. The ‘BlueprintEncodeData’ dataset from the R package ‘cellIndex’ was used as a reference to annotate each cluster. The R package ‘AUCell’ was used to assess the degree of cell correspondence to ARGs.

2.8. Identification of small molecule drugs

The GSE213621 dataset was used as a new discovery cohort. The R package “ConsensusClusterPlus” was used for consensus clustering. The broad institutes connectivity map (CMAP) database was used for identifying small molecule drug candidates associated with autophagy.

2.9. Animal experiments

Male C57BL/6j mice (18–20 g) were obtained from Jinan Pengyue Laboratory Animal Breeding Company (Jinan, China). After one week of adaptation, the mice were randomly divided...
into two groups (8 mice per group): (1) NC group: fed a normal diet (13.5% of energy from fat) and (2) HFD group: fed a high fat diet (HFD) (60% of energy from fat), both for 13 weeks. Food/water intake and body weight were monitored on a weekly basis throughout the duration of the experiment. At the end of the experiment, the liver and white adipose tissues (WATs) were collected and stored at −80 °C for subsequent analysis. The animal procedures were conducted following the Guidelines for Care and Use of Laboratory Animals of Qingdao University and approved by the Animal Ethics Committee of the Medical College of Qingdao University (QDU-AEC-2022361).

2.10. Collection of human liver samples
Paraffin sections for normal liver tissues and nonalcoholic fatty liver disease tissues were collected from the Affiliated Hospital of Qingdao University. The study protocol received ethical approval from the Ethics Committee of the Affiliated Hospital of Medical College of Qingdao University (QYPYWZL27720). Informed consent was obtained for all clinical samples.

2.11. Histological analysis
The liver tissue and WAT were collected and fixed in 10% formalin for 24 hours. Subsequently, they were embedded in paraffin, dehydrated, and cut into 5 µm-thick sections with a gradient concentration of alcohol. Then, the sliced sections were stained with hematoxylin and eosin (H&E).

2.12. Immunohistochemistry (IHC) and immunofluorescence (IF)
After pretreatment with 3% hydrogen peroxide, sections of liver samples were blocked with 10% goat serum (Gibco, USA) and then incubated overnight at 4 °C with primary antibody (diluted at 1:400). After primary antibody incubation, the unbound antibodies were washed away with PBS, and then the sections were incubated with the secondary antibody (IgG) at a dilution of 1:300 for 2 hours at room temperature. The positively stained areas were then developed using diaminobenzidine (DAB) substrate. The sections were subsequently counter-stained with hematoxylin. P62 for IF (#GB11531), MYC for IHC (GB13076), and CCL2 for IHC (GB11199) were purchased from Servicebio Technology Co., Ltd (Wuhan, China).

2.13. RNA extraction and quantitative real-time PCR (RT-qPCR) analysis
A SparkJade kit (Jinan, China) was used to extract the total RNA and reverse transcription, GAPDH was used as an internal control for control, and the 2^ΔΔCt method was used to quantify gene expression. The primers for the RT-qPCR are shown in Table S2 (ESI†).

2.14. Western blot
The total protein of liver tissue samples was extracted using a cold RIPA lysis buffer (Solarbio, Beijing, China). The proteins were quantified using the Bradford assay (Bio-Rad, Hercules, CA, USA). Antibodies against Beclin-1 (#AF5128) and P62 (#AF5384) were purchased from Affinity Biosciences (Jiangsu, China). The antibody against β-actin (#AC026) was purchased from ABelonal (Wuhan, China). Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) was performed at 100 V for approximately 2 h. Antibody treatment was performed with primary antibodies at a rate of 1:1000 overnight at 4 °C. The membrane was washed three times with tris-buffered saline solution containing Tween 20 and secondary antibodies were added at a ratio of 1:5000 for 2 h at room temperature. The transferred protein band on the polyvinylidene difluoride membrane was visualized by inducing an enhanced chemiluminescence reaction. ImageJ software was used for quantitative analysis.

2.15. Statistical analysis
Statistical analyses were performed using the “ggpubr” and “stats” R packages with the R software version 4.1.1. Unpaired Student’s t-test was used to compare two groups. One-way ANOVA was used to compare three or more groups; subsequently, Tukey HSD was used to compare multiple groups with each other. p < 0.05 was considered statistically significant.

3. Results
3.1. Identification of DEGs and functional annotation and pathway enrichment of DEGs
Fig. 1 shows the flowchart of this study. We preprocessed and evaluated the GSE4845224 and GSE8963225,26 datasets obtained from the GEO database. 51 healthy control samples and 34 NAFLD samples were collected for further analysis. The PCA results showed that batch effects between the two datasets were eliminated (Fig. 2A). The P values < 0.05 and |log 2FC| > 0.5 as thresholds were used to identify DEGs, and 659 DEGs were identified, including 275 up-regulated and 384 down-regulated genes (Fig. 2B) (Table S3, ESI†). Subsequently, we performed GO and KEGG enrichment analysis to better understand the biological significance underlying these DEGs. The GO terms for biological processes (BPs) show that DEGs are mainly enriched in response to the external environment and regulate cell death, such as apoptosis and programmed cell death (Fig. 2C). Notably, autophagy has been demonstrated to regulate cell fate through various crosstalk signals,27 and the intricate interplay between hepatic autophagy and apoptosis ultimately dictates the progression of liver diseases.28 These results also predicted an important link between autophagy and DEGs. Meanwhile, the enrichment results of KEGG indicated that a series of pathways associated with inflammatory responses were activated, including the IL-17 signaling pathway, TNF signaling pathway, and PI3K-Akt signaling pathway (Fig. 2D). This also predicts that NAFLD patients are in a state of inflammatory activation in vivo and have the potential to progress to NASH.

3.2. WGCNA construction and key module identification
A co-expression network was constructed to identify the modules most relevant to NAFLD using the expression data of 7127 CV genes in the GSE89632 dataset. A scale-independent topological network (a scale-free R^2 value of 0.86 and a soft
threshold of 15) and the mean connectivity network were established (Fig. 3A). By hierarchical clustering and module merging, 10 gene modules were obtained (Fig. 3B). Overall, the blue module showed the highest negative correlation with the NAFLD activity score ($r = -0.45$, $P = 3 \times 10^{-4}$), steatosis ($r = -0.58$, $P = 8 \times 10^{-4}$), aspartate transaminase ($r = -0.42$, $P = 8 \times 10^{-4}$), and alanine transaminase ($r = -0.60$, $P = 3 \times 10^{-7}$), while the brown module showed the highest positive correlation (Fig. 3C). Based on the correlation coefficient and $P$ value, we selected the blue module containing 543 genes as the characteristic module of NAFLD (Table S4, ESI†). Meanwhile, 172 genes were selected from the blue module as the module hub genes based on MM $> 0.8$ and GS $> 0.2$ (Fig. 3D).

3.3. Identification of autophagy-related hub genes and biomarkers

To identify hub genes associated with autophagy in NAFLD, we obtained genes shared by DEGs, blue module genes, and ARGs (Fig. 4A). VENN analysis resulted in 6 hub genes, namely BAG1, FOS, MYC, CCL2, CXCR4, and PPP1R15A. Fig. 4B shows the correlation between 6 hub genes, BAG1 was weakly correlated with other hub genes, and PPP1R15A was the most correlated with MYC and FOS.

The overlap of blue module hub genes with DEGs and ARGs as a new NAFLD biomarker was associated with autophagy (Fig. 4C). We identified 4 genes (FOS, MYC, CCL2, and PPP1R15A). The diagnostic significance of 4 biomarkers was confirmed by ROC analysis (Fig. 4D). All the genes showed AUC values greater than 0.75. MYC had the highest AUC value of 0.91 (95% CI), while FOS had the lowest AUC value of 0.80 (95% CI). This result suggests that FOS, MYC, CCL2, and PPP1R15A may be new NAFLD biomarkers associated with autophagy.

3.4. Immune infiltration analysis

The infiltration of immune cells in NAFLD was determined using the CIBERSORT algorithm. We evaluated differences in the abundance of immune cells between NAFLD and healthy controls using the GSE89632 dataset as the discovery cohort. The NAFLD group had significantly higher levels of CD8 T cells ($P < 0.01$), gamma delta T cells ($P < 0.001$), activated NK cells ($P < 0.05$), M2 macrophages ($P < 0.001$), and resting mast cells ($P < 0.001$) than healthy controls. In the latter group, the levels of B naive cells ($P < 0.001$), plasma cells ($P < 0.01$), monocytes ($P < 0.001$), activated mast cells ($P < 0.001$) and neutrophils were higher ($P < 0.01$) (Fig. 4E). Fig. 4F demonstrates the
A significant correlation between the hub gene and immune cells, in which activated mast cells and B naïve cells show a positive correlation with the hub gene, while M2 macrophages show a strong negative correlation with the hub gene. These results suggest that some degree of immune dysregulation occurs in the liver of NAFLD patients.

3.5. Validation of autophagy in HFD-induced NAFLD mice

Validation of hub gene expression levels with a high-fat diet (HFD)-induced mouse obesity model. Compared to the control group, the mice showed a significant increase in bodyweight after 14 weeks of HFD feeding and showed significant body-weight differences from the fifth week (Fig. 5A). Meanwhile, there was a significant increase in liver weight \( P < 0.01 \) and white adipose tissues \( P < 0.01 \), demonstrating the successful establishment of an obesity model (Fig. 5B and C). H&E staining of adipose tissue showed that the HFD mice had a significantly greater fat volume than NC mice (Fig. 5D). H&E staining of the liver tissue showed that the hepatocytes of the HFD mice were enlarged and accompanied with large fat vacuoles, indicating that they had suffered from severe steatosis (Fig. 5E). To further explore impaired liver autophagy in HFD-induced NAFLD mice, we measured the levels of autophagy-associated proteins by immunofluorescence and western blotting. The immunofluorescence results showed a significant increase in P62 levels in the livers of HFD-induced NAFLD mice compared to normal mice (Fig. 5F and G). Additionally, the western blotting results indicated that the expression of Beclin-1, a key protein that initiates autophagy, was decreased, while the autophagy substrate P62 was elevated in HFD mice (Fig. 5H). These findings suggest that steatosis leads to severe autophagy deficiency.

3.6. Validation of hub gene expression levels in HFD-induced NAFLD mice

We further detected the mRNA expression levels of hub genes in liver tissues using RT-qPCR. The results showed that their expression levels were significantly down-regulated in HFD-induced NAFLD mice compared with NC mice, which was consistent with our prediction (Fig. 6A–F). The expression of
MYC and CCL2 in protein levels was also analyzed using immunohistochemistry in liver tissues for normal mice and HFD-induced NAFLD mice. The results showed that the expression of MYC and CCL2 was significantly downregulated in the HFD-induced NAFLD mice (Fig. 6G and H). Meanwhile, the expression of MYC and CCL2 was significantly decreased in the samples from NAFLD patients (Fig. 6I and J). These results indicate that autophagy related hub genes were significantly reduced in NAFLD mice and patients, further suggesting that autophagy plays an important role in the occurrence and development of NAFLD.

3.7. Single-cell data analysis of liver non-parenchymal cells

Dataset GSE129516 was used for single-cell analysis. After filtering, 15,380 liver non-parenchymal cells were obtained. These cells were obtained from the livers of mice with NASH. Liver non-parenchymal cells are highly associated with disease progression in NAFLD and may also play an important role in hepatic autophagy. The expression profiles of the samples are shown in Fig. 7A and B. We normalised the data and identified 3000 highly variable genes (HVGs) using “VST”, of which the top 10 HVGs are shown in Fig. 7C. Eighteen cell types were identified using the t-SNE method (Fig. 7D). Eight liver non-parenchymal cells were identified by annotation, including skeletal muscle cells, endothelial cells, smooth muscle cells, adipocytes, macrophages and melanocytes. Skeletal muscle cells and fibroblasts were the most abundant (Fig. 8A). First, we analyzed the distribution of expression in various liver non-parenchymal cells for 6 hub genes (Fig. 8B). The results showed that CXCR4 was more expressed in adipocytes; meanwhile, MYC and PPP1R15A were more expressed in fibroblasts, whereas FOS and BAG1 were more widely expressed. Then, we explored the relationship between liver non-parenchymal cells and ARGs and found that fibroblasts and adipocytes had the highest AUC values (Fig. 8C and D).

**Fig. 3** Identification of the key modules and genes associated with NAFLD by WGCNA. (A) Estimating the scale independence index of the 1–20 soft threshold power and determining the mean connectivity of the 1–20 soft threshold power. (B) Module clustering dendrogram derived from the 1-tomatrix. The different color bands represent different modules. (C) Correlations between the various modules and traits. The number of each module depicts the correlation coefficient with the trait, and the color of the module ranges from red to green, showing high to low correlation. (D) Scatter plot of the blue module genes. The vertical coordinate represents the GS score, and the horizontal coordinate represents the MM score for each gene.
Fig. 4  Identification of hub genes and biomarkers related to autophagy in NAFLD. (A) Genes common to the blue module genes, ARGs and DEGs shown by the Venn diagram. (B) Correlation of autophagy-related hub genes in NAFLD. (C) Genes common to the blue hub module genes, ARGs and DEGs shown by the Venn diagram. (D) ROC curves for FOS, MYC, CCL2, and PPP1R15A. (E) The boxplot shows the immune cell infiltration difference between the NAFLD and control groups. (F) Correlation between the immune cell and the hub gene. *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001.
Fig. 5  Decreased levels of liver autophagy in HFD mice. The body weight change (A), liver weight change (B) and WAT weight change (C). H&E staining for WATs (D) and liver tissue (E). (F) Immunohistochemistry results of P62. (G) Quantitative immunohistochemistry results of P62. (H) Western blot and quantitative analysis of proteins Beclin-1 and P62. *P < 0.05, **P < 0.01 and ***P < 0.001.
Therefore, we hypothesised that these two types of cells might be most closely associated with autophagy.

3.8. Identification of small molecule drugs

The NAFLD samples in the GSE213621 dataset were classified into two different clusters with \( K = 2 \) by autophagy-related hub gene consensus clustering (Fig. 9A). NAFLD patients were divided into high and low autophagy groups. t-Distributed stochastic neighbourhood embedding (t-SNE) dimensionality reduction analysis showed that these two clusters were significantly different in their distribution patterns (Fig. 9B). We obtained DEGs via differential analysis. Based on 150 upregulated and downregulated genes, we identified potential therapeutic small molecules related to autophagy with the Broad Institutes Connectivity Map (cMAP) database. Ten potential small molecule drugs were selected based on transcriptional activity score (Tas) values as follows: ligustilide (Tas = 1.9437), AH-11110 (Tas = 1.9418), gatifloxacin (Tas = 1.8789), diclofensine (Tas = 1.8676), vandetanib (Tas = 1.8386), BGT-226 (Tas = –1.8561), atipamezole (Tas = –1.8578), penfluridol (Tas = –1.8659), withaferin-A (Tas = –1.9066), and GSK-1059615 (Tas = –1.8281). Fig. 9C and D show 3D conformations of the 10 chemicals mentioned above, which may reverse or induce specific gene expression and thus have an effect on autophagy.

4. Discussion

NAFLD has become a predominant cause of chronic liver disease, as well as multisystem disease. Previous studies have reported that the clinical burden of NAFLD is not only limited to liver-related morbidity and mortality, but also exacerbates other chronic diseases such as chronic kidney disease, colorectal cancer, osteoporosis, psoriasis and various endocrine disorders. Despite some advances in NAFLD research over the last few decades, the effective treatment of NAFLD/NASH remains a challenge. There are still no drugs or procedures approved for the treatment of NAFLD/NASH, and lifestyle changes (diet, physical activity and exercise) remain the basis of its treatment. Autophagy is a cellular mechanism responsible for degrading proteins and organelles, which has been found to have a remarkable link to human physiology and disease. Previous reports have shown that pharmacological and genetic interventions to reduce autophagy can exacerbate diseases and mutations in autophagy-related processes can lead to severe human pathology. Dysfunctional autophagy in liver parenchymal and non-parenchymal cells leads to a variety of liver diseases, which include NAFLD. Targeting autophagy could be a potential strategy for the treatment of NAFLD; for example, Liu et al. demonstrated that scoparone...
restored autophagy levels in the liver while improving NAFLD mice and Chu et al. showed that cherry anthocyanins regulated NAFLD by promoting autophagy.

Differential analysis and WGCNA identified a total of 659 DEGs and 543 module genes, respectively. Enrichment analysis revealed that these genes are enriched in functional pathways highly relevant to autophagy, including cell death processes and inflammatory pathways. Subsequently, 6 hub genes related to autophagy were identified, namely BAG1, FOS, MYC, CXCR4, and PPP1R15A. BAG1 is a BCL2-associated athanogene containing a BAG protein structural domain that binds to heat shock protein 70 (Hsp70) thereby acting as a co-chaperone to regulate the activity of Hsp70. Up-regulation of autophagy associated with BAG-1 protein has been reported to have a protective impact on myocardium. FOS forms a tight but non-covalently linked complex with the JUN/AP-1 transcription factor for nuclear phosphoproteins. Indocyanine green photodynamic therapy upregulates FOS-mediated autophagy to treat melanoma. MYC is a target of the Wnt signalling pathway and has a critical function in the regulation of autophagy. CCL2 acts as a ligand for the C-C chemokine receptor and autophagy deficiency inhibits the transcription and production of CCL2. CXCR4 enhanced autophagy levels by stimulating ATG7 expression through ZEB1. PPP1R15A plays an important role in autophagy by regulating translation during starvation, thereby enabling lysosomal biogenesis and sustained autophagic flux. FOS, MYC, CCL2, and PPP1R15A were identified as diagnostic biomarkers and the ROC analysis showed that MYC showed the highest potential as a biomarker. The significant downregulation of these genes indicates a severe autophagy deficiency in the livers of NAFLD patients. Additionally, we calculated the correlations among these genes. Overall, except for BAG1, other hub genes exhibit high correlations, suggesting that they may have synergistic effects or common regulatory mechanisms in the process of autophagy in the livers of NAFLD patients.

We then validated our predictions with animal models and liver pathology slices from patients. Steatosis leads to autophagy deficiency, resulting in massive accumulation of the autophagy substrate P62. The 6 hub genes we identified were significantly downregulated in the liver of HFD mice. Also, liver non-parenchymal cells play a significant role in the progression of NAFLD. It has been reported that autophagy...
dysfunction in nonparenchymal cells can lead to various liver diseases, including NAFLD, viral hepatitis, and hepatocellular carcinoma.\textsuperscript{39} Moreover, Liu \textit{et al.} demonstrated that scoparone could improve NASH in mice by promoting macrophage autophagy but not hepatocyte autophagy.\textsuperscript{40} Our results suggest that adipocytes and fibroblasts are most closely related to autophagy, which may provide a new research direction for targeting autophagy in NAFLD.

In recent years, with the continuous development of research, an increasing number of small molecule drugs have been utilized for the treatment of NAFLD.\textsuperscript{49–51} Based on the 6 hub genes we identified, NAFLD patients were categorized into two distinct groups. Potential new small molecule drugs were identified through the DEGs between these two groups. Among these, some small molecule drugs have already demonstrated considerable potential in the treatment of NAFLD. For example,
research by Li et al.\textsuperscript{52} has shown that ligustilide can inhibit lipid accumulation and regulate lipid metabolism, thus offering a potential treatment for NAFLD. Withaferin A, derived from Ayurvedic medicine, possesses extensive pharmacological activities. Studies have reported that withaferin A can significantly ameliorate liver damage, inflammation, and fibrosis in NAFLD mice, highlighting its potential as a therapeutic agent for NAFLD.\textsuperscript{53} The relationship between the other 8 small molecule drugs and NAFLD remains to be elucidated. Nonetheless, considering the findings related to ligustilide and withaferin A, these compounds exhibit substantial promise for the treatment of NAFLD.

Under normal physiological conditions, autophagy participates in basal lipid turnover by phagocytosis and degradation of lipid droplets, providing free fatty acids (FFAs) for ATP production. This ability of autophagy to degrade lipid droplets in hepatocytes is known as lipophagy.\textsuperscript{54} In contrast, the sustained lipid supply induced by a HFD inhibits the turnover of hepatic autophagy and lipolysis rates, leading to reduced fatty acid β-oxidation.\textsuperscript{55} The massive accumulation of lipids may have altered the membrane structure, resulting in a decrease in the efficiency of fusion between autophagosomes and lysosomes, which is a mechanism that may explain the inhibitory effect of HFD on autophagy. Starvation has an inducing effect on hepatic autophagy, and the livers of starved mice show increased numbers of lipid droplets, autophagosomes, lysosomes and autophagic lysosomes.\textsuperscript{54} Meanwhile, some enhancers of autophagy have been tested in HFD mice and have been shown to be protective in both reducing steatosis and improving insulin sensitivity.\textsuperscript{56–58} Therefore, targeting autophagy for treatment and research may be an effective approach for the future treatment of NAFLD.

5. Conclusions

In conclusion, our study identified 6 autophagy-related hub genes (\textit{BAG1}, \textit{FOS}, \textit{MYC}, \textit{CCL2}, \textit{CXCR4}, and \textit{PPP1R15A}) in NAFLD and validated them using animal models and patient tissue sections. \textit{FOS}, \textit{MYC}, \textit{CCL2}, and \textit{PPP1R15A} may be used as potential biomarkers associated with autophagy in NAFLD, which may also become a new target for future gene therapy. The link between autophagy defects in liver non-parenchymal cells and the progression of NAFLD is intriguing. The relationship we found between adipocytes and fibroblasts and
autophagy may provide a new perspective for the treatment of NAFLD. In addition, we identified 10 potential drugs that may target autophagy in NAFLD.

Author contributions

Conceptualization: Fengjuan Zhang and Ningning He; bioinformatics analysis: Kaiwei Chen and Ling Wei; animal experiments and related experiments: Shengnan Yu and Ningning He; writing – original draft preparation: Kaiwei Chen, Ling Wei and Ningning He; writing – review and editing: Ningning He and Fengjuan Zhang; supervision, project administration, and funding acquisition: Ningning He and Fengjuan Zhang. All authors have read and agreed to the published version of the manuscript.

Data availability

The original contributions presented in the study are publicly available. All data are available for download at https://www.ncbi.nlm.nih.gov/. The R code used in the study is available at https://github.com/illusion621/Identification-of-autophagy-related-signatures-in-nonalcoholic-fatty-liver-disease.

Conflicts of interest

No potential conflicts of interest are reported by the author(s).

Acknowledgements

This study was supported by the Natural Science Foundation of Shandong Province of China (ZR2022MH217). The authors thank all the researchers whose studies are described in this article.

Notes and references

18. L. Yang, P. Li, S. Fu, E. S. Calay and G. S. Hotamisligil, Cell Metab., 2010, 11, 467–478.