Western Diet-Induced Nonalcoholic Fatty Liver Disease Mice Mimic the Key Transcriptomic Signatures Observed in Humans

Tatsuya ISHIGURE[1,2], Tomohiko SASASE[1,2], Marika TOHMA[3], Kinuko UNO[2,3], Yasufumi TORINIWA[1], Tomoyuki SAITO[1], Yasuka SAIGO[1,2], Koji EDAMURA[4], Katsuhiro MIYAJIMA[3], Takeshi OHTA[2]

[1]Biological/Pharmacological Research Laboratories, Takatsuki Research Center, Central Pharmaceutical Research Institute, Japan Tobacco Inc., Osaka, Japan, [2]Laboratory of Animal Physiology and Functional Anatomy, Graduate School of Agriculture, Kyoto University, Kyoto, Japan, [3]Department of Food and Nutritional Science, Graduate School of Agriculture, Tokyo University of Agriculture, Tokyo, Japan, [4]i2i Labo, Yokohama Research Center, Central Pharmaceutical Research Institute, Japan Tobacco Inc., Kanagawa, Japan

Received September 22, 2023
Accepted February 29, 2024

Summary

Nonalcoholic fatty liver disease (NAFLD) is a chronic liver disease characterized by the accumulation of fat in the liver in the absence of excessive alcohol consumption or a secondary cause of hepatic steatosis. The prevalence of NAFLD is increasing worldwide and its management has become a public health concern. Animal models are traditionally used to elucidate disease mechanisms and identify potential drug targets; however, their translational aspects in human diseases have not been fully established. This study aimed to clarify the utility of animal models for translational research by assessing their relevance to human diseases using gene expression analysis. Weighted gene co-expression network analysis of liver tissues from Western diet (WD)-induced NAFLD mice was performed to identify the modules associated with disease progression. Moreover, the similarity of the gene co-expression network across species was evaluated using module preservation analysis. Nineteen disease-associated modules were identified. The brown module was positively associated with disease severity, and functional analyses indicated that it may be involved in inflammatory responses in immune cells. Moreover, the gene co-expression network of the brown module was highly preserved in human NAFLD liver gene expression datasets. These results indicate that WD-induced NAFLD mice have similar gene co-expression networks (especially genes associated with inflammatory responses) to humans and are thought to be a useful experimental tool for preclinical research on NAFLD.

Keywords

Nonalcoholic fatty liver disease (NAFLD) • Weighted gene co-expression network analysis (WGCNA) • Western diet (WD)

Corresponding author

T. Sasase, Biological/Pharmacological Research Laboratories, Takatsuki Research Center, Central Pharmaceutical Research Institute, Japan Tobacco Inc., Takatsuki, Osaka 569-1125, Japan, E-mail: tomohiko.sasase@jt.com

Introduction

Nonalcoholic fatty liver disease (NAFLD) is a chronic liver disease characterized by the ectopic accumulation of fat in the liver in the absence of excessive alcohol consumption or a secondary cause of hepatic steatosis (such as viral infections, medications, and genetic liver diseases). It consists of two distinct conditions; non-progressive simple steatosis, and a more advanced progressive form, nonalcoholic steatohepatitis (NASH). Among patients with NAFLD, a small subset develop NASH, and some of these may progress to cirrhosis and hepatocellular carcinoma [1]. It is often associated with metabolic disorders such as obesity, dyslipidemia, hypertension, and type 2 diabetes. With the rapid increase in such comorbidities, the prevalence of NAFLD is increasing, and it is estimated that approximately 40 % of the population worldwide has NAFLD [2,3]. Currently, NAFLD is the second leading cause of liver transplantation and is rapidly becoming the leading indication in the United States [4]. This increasing prevalence affects not only patients' quality of

life and life expectancy but also healthcare and socioeconomic costs. Therefore, management of NAFLD has become a major public health concern [5]. To date, much efforts have been devoted to the development of effective therapies. Recently, promising results have been reported, but there are still no approved drugs to prevent or regress these conditions [6]. One of the hurdles in the development of effective therapies is the lack of adequate animal models. Several types of animal models (diet-induced, chemically induced, genetically modified, and combinations thereof) have been developed as NAFLD models. However, owing to the complexity of the disease etiology, an ideal animal model that fully recaptures human pathophysiology has not been established [7,8]. The diet-induced model is the most common type of animal model for NAFLD. Typically, a high-fat diet (HFD; 60 % kcal from fat) is used to induce metabolic dysfunctions (such as obesity, hyperlipidemia, and insulin resistance) and hepatic steatosis. Despite hepatic steatosis, inflammatory cell infiltration is absent or weak, even after long-term feeding, and severe NASH and fibrosis are not observed. Hence, dietary macronutrients have been modified to accelerate the disease onset and progression. One of the modified HFDs is a Western diet (WD) containing 40 % kcal fat and varying percentages of sucrose/fructose and cholesterol. Mice treated with the WD develop steatohepatitis and mild fibrosis in a relatively short time (up to 24 weeks) [9,10]. Although, this animal model shows similar comorbidities and disease progression to human NAFLD, its relevance to human disease (which aspects of disease etiology are recaptured) is not fully understood. Transcriptome analysis is an important systems biology approach for characterizing the molecular mechanisms of disease development and progression. Traditionally, differential gene expression analysis has been used in transcriptome analysis; however, it has some limitations. It tends to focus on genes with large differences and does not include genes with little or no difference that may contribute to pathogenesis for analysis. In addition, it cannot consider the relationships between genes or the association with phenotype. Weighted gene co-expression network analysis (WGCNA) is a correlation-based network analysis method that detects clusters (modules) of highly correlated genes [11]. The modules are then characterized by assessing the association between phenotypic traits or by performing functional enrichment analysis. Moreover, based on the correlation between genes, we can identify functionally important genes (hub genes) in the module. In this study, we applied WGCNA to WD-induced NAFLD mice to further characterize their pathophysiology and evaluate its relevance to human NAFLD.

Methods

Animals

Male C57BL/6J mice were purchased from Charles River Laboratories Japan, Inc. (Kanagawa, Japan). Animals were housed under a 12-hr light/dark cycle at 23±3 °C and 55±15 % humidity. All the animals were fed standard rodent chow (CRF-1; normal diet, Oriental Yeast Co., Ltd., Tokyo, Japan) and tap water *ad libitum*.

All the experimental protocols for the use of laboratory animals followed the guidelines of the Animal Care Committee of the Central Pharmaceutical Research Institute of Japan Tobacco, Inc. (protocol No. 01125).

Dietary interventions and sample collection

At 8 weeks of age, the animals were divided into two groups and treated with either a Western diet (WD; high-fat, high-sucrose, high-cholesterol diet (TD.88137, Envigo, Hackensack, NJ) and fructose/glucose solution (23.1 g/l D-fructose + 18.9 g/l D-glucose)) or a normal diet and tap water for 8, 12, 16, and 24 weeks (n=8 per group for WD, n=4 per group for normal diet). During the course of the dietary interventions, animals whose conditions deteriorated (such as significant weight loss, low activity, and low body temperature) were excluded from subsequent analyses. After the dietary intervention, the animals were anesthetized and sacrificed for blood and liver sampling. Liver samples for hepatic lipid content and gene expression analyses were snap-frozen in liquid nitrogen and stored at -80°C until use. For histological analysis, liver samples were fixed in 10 % neutral-buffered formalin immediately after collection.

Blood biochemical analysis

The plasma levels of each biochemical parameter (glucose, triglycerides, total cholesterol, aspartate aminotransferase (AST), and alanine aminotransferase (ALT)) were measured using the corresponding product kits (Roche Diagnostics, Tokyo, Japan) and an automated analyzer (Hitachi, Tokyo, Japan).

Analysis of hepatic lipid content

The liver tissue (approximately 100 mg) was homogenized in methanol (0.5 ml) using a mixer mill (MM300; Retsch, Haan, Germany). For lipid extraction from the homogenized solution, 1 ml of chloroform was added and vigorously mixed. The samples were then centrifuged (10,000 × *g*, 5 min, 4°C), and a portion of the supernatant was dried with nitrogen gas. The lipid extract was dissolved in 2-propanol, and the triglyceride concentrations were measured using the corresponding product kits (Roche Diagnostics) and an automated analyzer (Hitachi).

Histological analysis

Liver samples were paraffin-embedded and thin-sectioned (3–5 μm) using standard techniques. The sections were stained with hematoxylin and eosin (H&E) and Sirius red. Hepatic steatosis, hepatocyte degeneration, and inflammatory cell infiltration were evaluated in H&E-stained tissue sections. Hepatic fibrosis was quantified in Sirius red-stained tissue sections.

RNA extraction and microarray analysis

Total RNA was extracted from liver tissue using the RNeasy Mini Kit (Qiagen, Hilden, Germany). Microarray analysis was performed using an Agilent SurePrint G3 Mouse Gene Expression 8 × 60 K Ver. 2 (Agilent Technologies, Santa Clara, CA, USA) according to the manufacturer's protocols.

Data preprocessing

Background correction and quantile normalization were performed using the limma package in R (version 3.6.3). The data were then log2 transformed to obtain standardized expression values. To reduce the potential for false discoveries, low and non-expressed genes were filtered out, and the remaining genes were used for subsequent analysis.

Weighted gene co-expression network construction and module detection

A weighted gene co-expression network was constructed following the protocols of WGCNA [11]. A pairwise Pearson correlation coefficient matrix was computed and transformed into an adjacency matrix with an appropriate soft-thresholding power (β). β was selected based on a scale-free fit index and mean connectivity values. A topological overlap matrix (TOM) was then constructed, a TOM-based dissimilarity was used as input, modules were identified in the dendrogram using the dynamic tree-cutting algorithm [12]. The minimum module size was defined as that of 30 genes.

Identification of the modules associated with disease severity

The module eigengene (ME) is defined as the first principal component of a given module, which represents the overall gene expression values in the module. The correlation between MEs and hepatic fibrosis was analyzed to identify modules associated with disease severity. Statistical significance was set at $P < 0.05$.

Estimation of the cell types of each module

Human liver-derived cell type marker genes (genes highly expressed in a specific cell type) [13] were used to estimate the cell type in each module. The human Entrez identifier was converted to the corresponding mouse Entrez identifier using the homologene package. Cell type deconvolution was performed by measuring the overlap between module genes and cell type marker genes. Statistical significance was set at $P < 0.05$.

Functional analysis of the modules of interest

To assess the biological function of the module, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed using clusterProfiler package [14]. Based on the overlapped gene counts, the top 10 GO terms and KEGG pathways were visualized. Statistical significance was set at $P < 0.05$.

Module preservation analysis

To evaluate the similarity of gene co-expression networks across species, human NAFLD liver gene expression datasets (GSE48452, GSE49541, and GSE130970) were downloaded from the NCBI Gene Expression Omnibus database (Table 1). After data preprocessing, the human Entrez identifier was converted to the corresponding mouse Entrez identifier using the homologene package. Module preservation calculations were performed using the modulePreservation function in WGCNA with 200 permutations [15]. Two composite statistics, Zsummary and Median rank, were used to evaluate network similarity. Zsummary is a summary of the permutation-based preservation Z-statistics. Zsummary > 10 implies strong evidence of module preservation, whereas 2 < Zsummary < 10 implies

weak-to-moderate evidence of preservation. If Zsummary is < 2, there is no evidence of module preservation. Median rank is another statistic used to test the level of preservation, and a lower value implies strong evidence of module preservation.

Table 1. Datasets used in the study (human NAFLD liver gene expression datasets).

Accession	Number of samples	Sample information
GSE48452	54†	control = 12 healthy obese = 16 steatosis = 9 NASH = 17
GSE49541	72	mild NAFLD (fibrosis stage 0-1) = 40 advanced NAFLD (fibrosis stage 3-4) = 32
GSE130970	78	fibrosis stage 0 (F0) = 25 fibrosis stage 1 (F1) = 28 fibrosis stage 2 (F2) = 9 fibrosis stage 3 (F3) = 14 fibrosis stage 4 (F4) = 2

† Samples acquired before bariatric surgery were used.

Table 2. Laboratory and biochemical parameters of each treatment group.

	8w		12w		16w		24w	
	WD	**Normal**	**WD**	**Normal**	**WD**	**Normal**	**WD**	**Normal**
Body weight (g)	42.8±1.7**	29.1±1.1	46.6±3.0**	29.9±2.5	46.7±2.7**	30.5±1.6	51.1±3.5**	33.7±3.3
Liver weight (g)	2.9±0.3**	1.5±0.1	4.0±0.6**	1.5±0.2	4.1±0.8**	1.5±0.1	4.8±1.0**	1.5±0.3
Liver weight (g/Body weight)	0.068±0.005**	0.051±0.002	0.085±0.007**	0.052±0.005	0.088±0.015**	0.050±0.001	0.093±0.019**	0.043±0.008
Glucose (mg/dl)	217.9±31.6	187.5±18.1	230.1±29.7	206.3±25.1	210.0±17.6	198.8±15.0	231.5±38.1	229.5±16.5
Triglycerides (mg/dl)	88.7±36.4	105.2±15.2	54.2±10.6	129.8±79.1	43.4±12.0*	96.8±29.5	81.5±41.9	106.0±47.6
Total Cholesterol (mg/dl)	245.6±15.4**	108.5±4.8	269.8±14.9**	114.8±15.7	286.9±38.6**	101.6±8.2	335.8±56.6**	105.8±13.0
AST (IU/l)	85.8±22.1*	49.9±14.6	160.2±27.8**	50.4±18.9	246.5±96.1**	43.0±9.1	327.1±81.8**	66.6±13.1
ALT (IU/l)	93.5±35.8**	30.0±3.8	200.4±32.4**	24.1±3.2	274.6±107.4**	23.7±5.3	373.9±99.3**	29.9±5.6
Liver triglycerides (mg/g tissue)	287.7±56.5**	12.8±3.8	330.4±28.6**	12.2±5.1	377.9±63.1**	13.5±1.2	202.8±43.4**	16.7±6.4

Data are shown as mean ± SD (Western diet (WD)-treated mice; n=6-8, Normal mice; n=4). *P<0.05 and **P<0.01, WD versus Normal. AST; aspartate aminotransferase, ALT; alanine aminotransferase.

Identification of hub genes in the module

Intramodular connectivity (K.in) was calculated as the sum of the adjacency over all genes in the module. Hence, genes with high K.in values were assumed to be highly connected to other genes in the module. We extracted genes with high K.in values (top 30 %) and imported them into the STRING database [16] to construct the protein-protein interaction (PPI) network. We used the Cytoscape software [17] and its plug-in cytohubba to identify the hub genes in the module. After calculating the degree centrality of each gene, genes with high degree centrality value (≥20) were defined as hub genes.

Table 3. List of hub genes in the brown module.

Gene symbol	Degree centrality
Tyrobp	41
Fcer1g	34
Ctss	34
Fcgr3	31
Emr1	30
Cd68	28
Hck	28
Syk	27
Ncf4	26
Rac2	26
Hcls1	25
Ly86	25
Lyz2	24
Fcgr4	24
Clec4n	24
C1qc	23
Laptm5	22
Vav1	22
Slc11a1	21
AF251705	21

Validation of hub genes in the human datasets

To confirm the relevance of hub genes in human diseases, we analyzed their expression profiles in human NAFLD liver gene expression datasets (Table 1). After data preprocessing, differentially expressed gene analysis was performed. Statistical significance was set at $P < 0.05$, and the expression values are shown as box and whisker plots.

Statistical analysis

In animal experiments, all values are expressed as mean ± standard deviation (SD). To confirm the effect of dietary intervention on biochemical and histological parameters, we evaluated the difference between the WD-treated mice and normal mice at each time point. Statistical significance was determined by F-test followed by Student's *t*-test or Aspin-Welch's *t*-test with $P < 0.05$ as the significance threshold. Statistical analyses for bioinformatics were performed according to standard protocols.

Results

Western diet-induced metabolic abnormalities and histologic changes in the liver

Compared to normal mice, WD-treated mice rapidly gained body weight and developed marked obesity. Feeding a WD also led to a significant increase in plasma total cholesterol levels. On the other hand, plasma glucose and triglyceride levels did not show consistent change between the treatment groups. Plasma levels of ALT and AST, a marker of liver damage, significantly increased in the WD-treated groups and showed an increasing trend throughout the treatment period. Liver weight (adjusted for body weight) increased in the WD-treated mice, which was thought to reflect lipid accumulation in the liver (Table 2). Liver histology (H&E staining) showed that WD-treated mice developed marked lipid accumulation in hepatocytes and hepatocyte degeneration as early as 8 weeks of dietary intervention. An increase in the number of inflammatory cells was also observed. At week 12, hepatic fibrosis, as assessed by Sirius red staining, was observed in the WD-treated groups, which markedly progressed at weeks 16 and 24 (Fig. 1A–C). These results indicate that WD induces metabolic abnormalities and histological changes in the liver (such as NASH and fibrosis).

Weighted gene co-expression network construction and module detection

After data preprocessing, 15228 genes were used to construct the gene co-expression network. Principal component analysis (PCA) score plots and hierarchical average linkage clustering showed a relatively clear separation by treatment group, suggesting that dietary intervention has a strong effect on gene expression profiles. We confirmed that there were no outlier samples that needed to be eliminated from subsequent analyses (Supplementary Fig. 1A, B). In accordance with the scale-free topology criterion, $\beta = 9$ was selected (Supplementary Fig. 2A). A dynamic tree-cutting algorithm was used to identify modules from the hierarchical clustering dendrogram, and 31 distinct co-expression modules were obtained (Supplementary Fig. 2B). Genes that did not fit into a distinct module were assigned to the grey module, which was ignored in this study.

Identification of the modules associated with disease severity

Hepatic fibrosis is an important feature of NAFLD severity, and its degree is associated with the incidence of liver-related morbidity and mortality

[18]. To identify the genes and pathways associated with disease severity, we evaluated the relationship between each module and disease severity by assessing the correlation between MEs and hepatic fibrosis. We detected 7 modules that were positively and 12 modules that were negatively correlated with hepatic fibrosis (Fig. 2A).

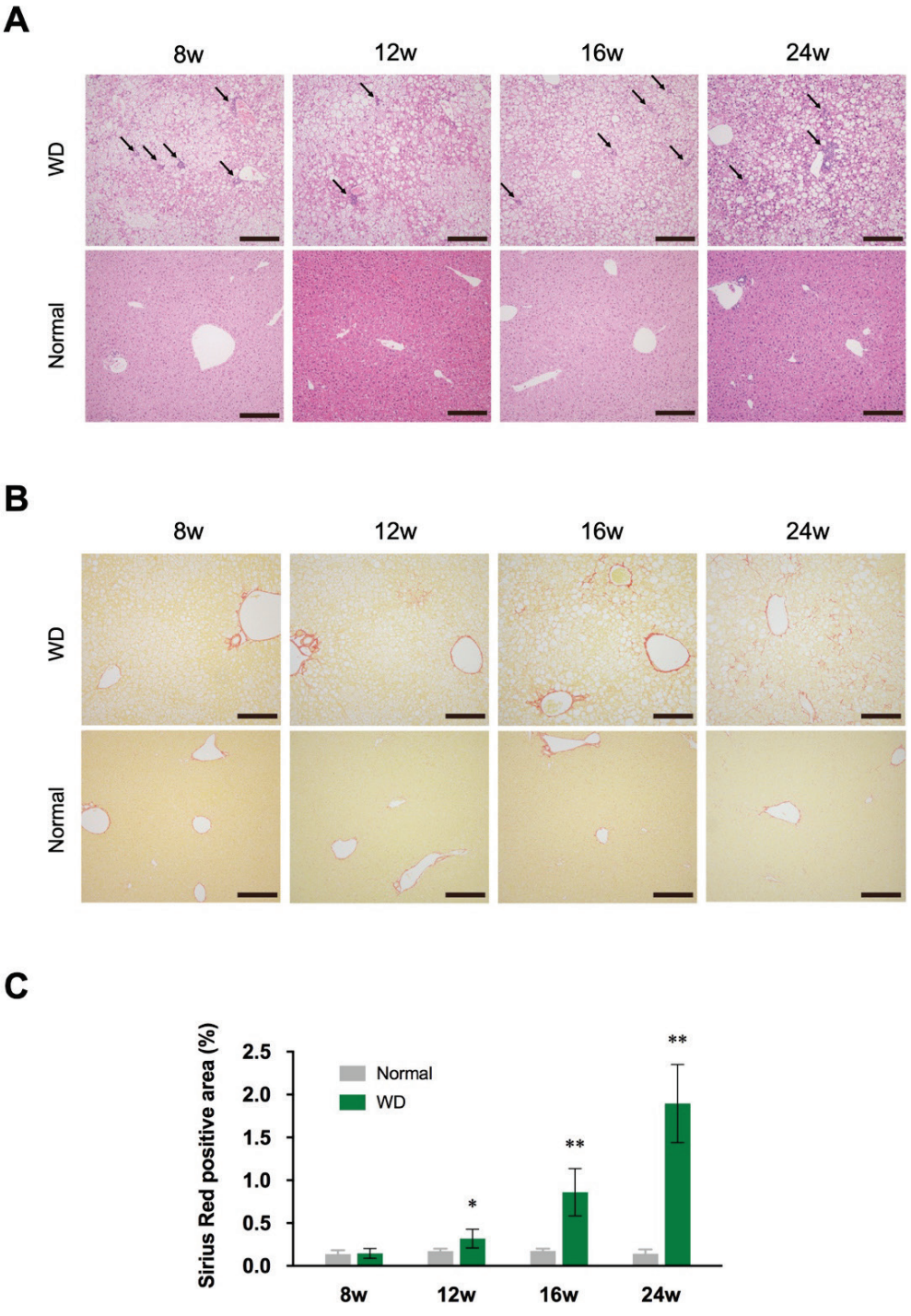
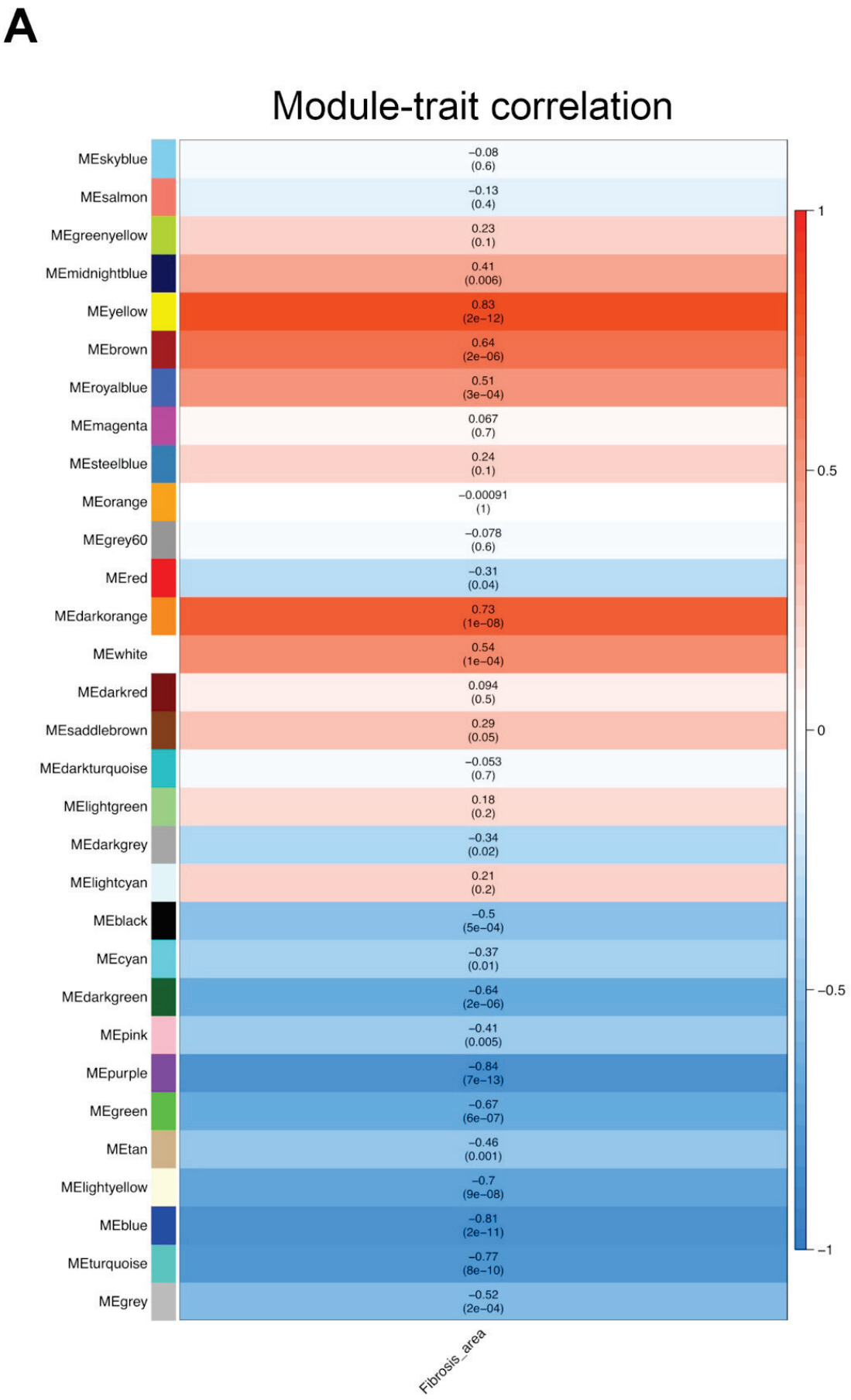


Fig. 1. Liver histology in Western diet (WD)-treated mice and normal mice. Representative images of hematoxylin and eosin (H&E)-stained (**A**) and Sirius red-stained (**B**) liver sections from WD-treated and normal mice. Black arrow indicates infiltered inflammatory cells. Scale bars: 200 μm (H&E, Sirius red). (**C**) Fibrous area expressed as the percentage positive area for Sirius red. Data are expressed as mean ± SD (WD-treated mice; n=6-8, Normal mice; n=4). **P*<0.05 and ***P*<0.01, WD versus Normal.


A
Module-trait correlation
MEskyblue
MEsalmon
MEgreenyellow
MEmidnightblue
MEyellow
MEbrown
MEroyalblue
MEmagenta
MEsteelblue
MEorange
MEgrey60
MEred
MEdarkorange
MEwhite
MEdarkred
MEsaddlebrown
MEdarkturquoise
MElightgreen
MEdarkgrey
MElightcyan
MEblack
MEcyan
MEdarkgreen
MEpink
MEpurple
MEgreen
MEtan
MElightyellow
MEblue
MEturquoise
MEgrey
−0.08 (0.6)
−0.13 (0.4)
0.23 (0.1)
0.41 (0.006)
0.83 (2e−12)
0.64 (2e−06)
0.51 (3e−04)
0.067 (0.7)
0.24 (0.1)
−0.00091 (1)
−0.078 (0.6)
−0.31 (0.04)
0.73 (1e−08)
0.54 (1e−04)
0.094 (0.5)
0.29 (0.05)
−0.053 (0.7)
0.18 (0.2)
−0.34 (0.02)
0.21 (0.2)
−0.5 (5e−04)
−0.37 (0.01)
−0.64 (2e−06)
−0.41 (0.005)
−0.84 (7e−13)
−0.67 (6e−07)
−0.46 (0.001)
−0.7 (9e−08)
−0.81 (2e−11)
−0.77 (8e−10)
−0.52 (2e−04)
Fibrosis_area
1
0.5
0
−0.5
−1

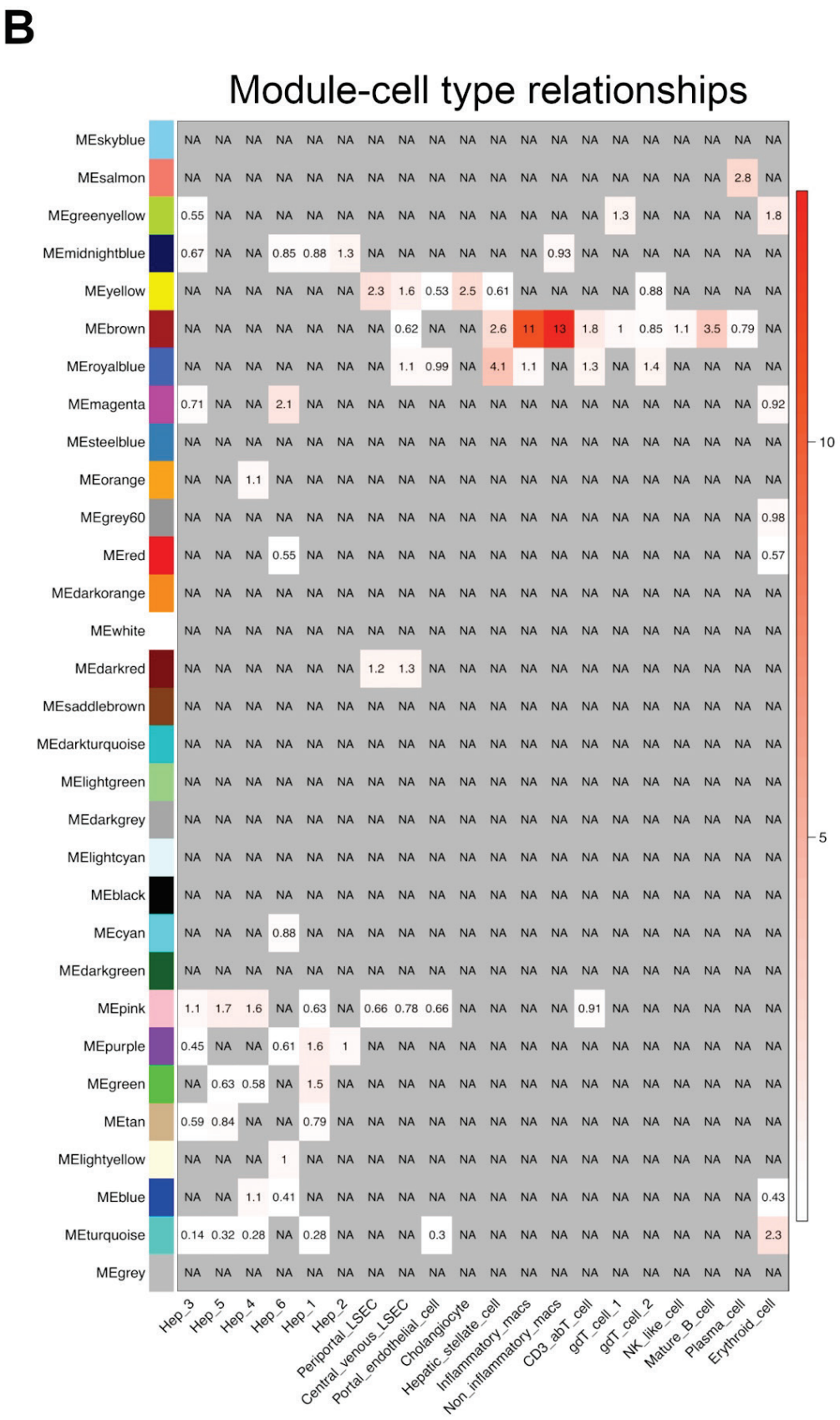


Fig. 2. Module-trait correlation and module-cell type relationships. (**A**) Each row corresponds to a module, and each column corresponds to a trait. Each cell contains the correlation coefficient and *P* value. The table was color-coded based on the correlation coefficient according to the color legend. (**B**) Each row corresponds to a module, and each column corresponds to a cell type. Each cell contains the -log transformed *P* value. The table was color-coded based on the *P* value according to the color legend. If no overlap was detected between module genes and cell type marker genes, the cell was noted as NA and colored grey.

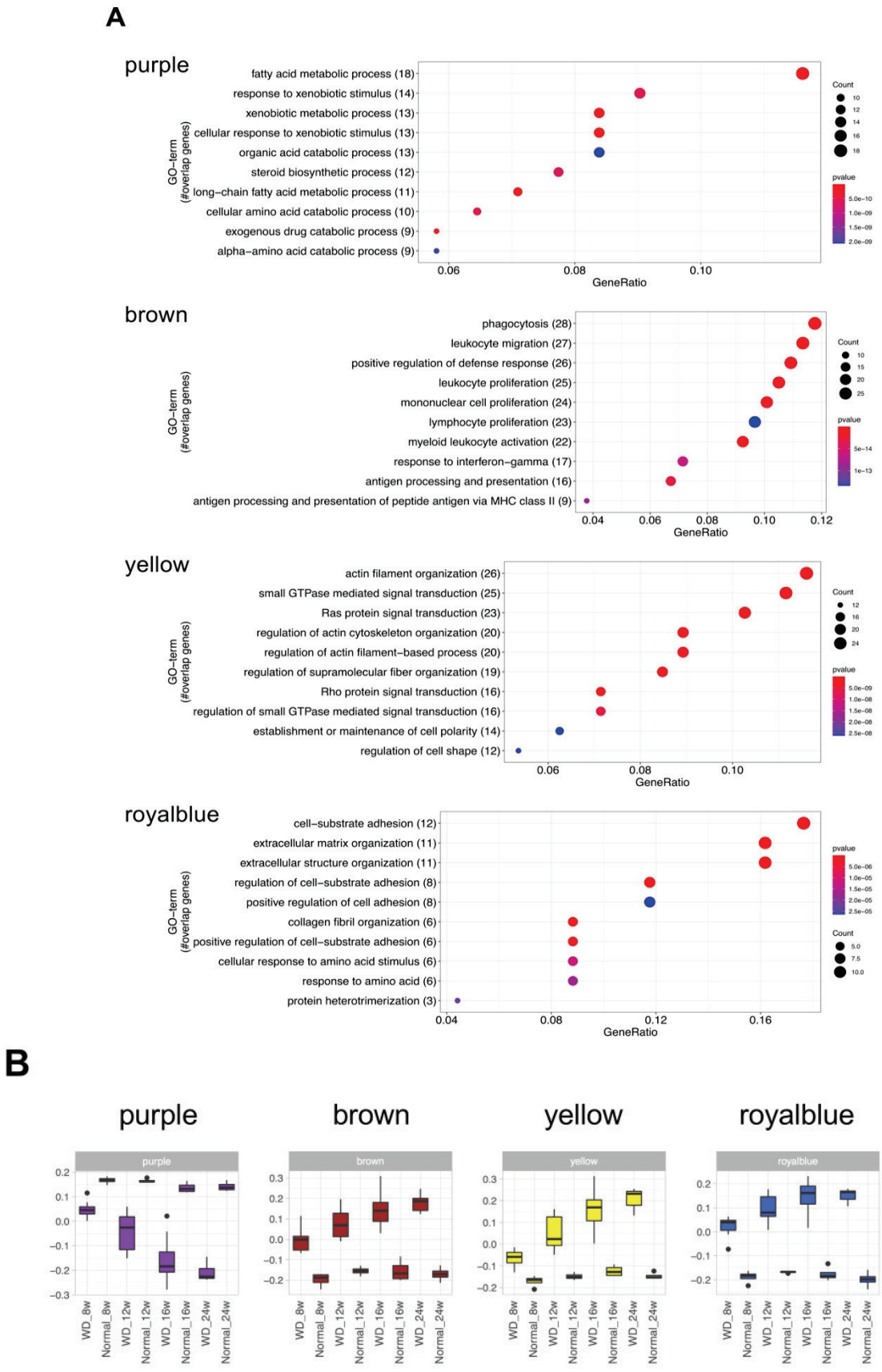


Fig. 3. Functional analysis of the modules of interest. (**A**) Gene Ontology (GO) enrichment analysis (Biological process) of each module. The node was color-coded based on the *P* value according to the color legend, and the size of the node refers to the number of enriched gene count. (**B**) Expression levels of the module eigengenes in each treatment group. Data are shown as box and whisker plots. WD; Western diet-treated mice.

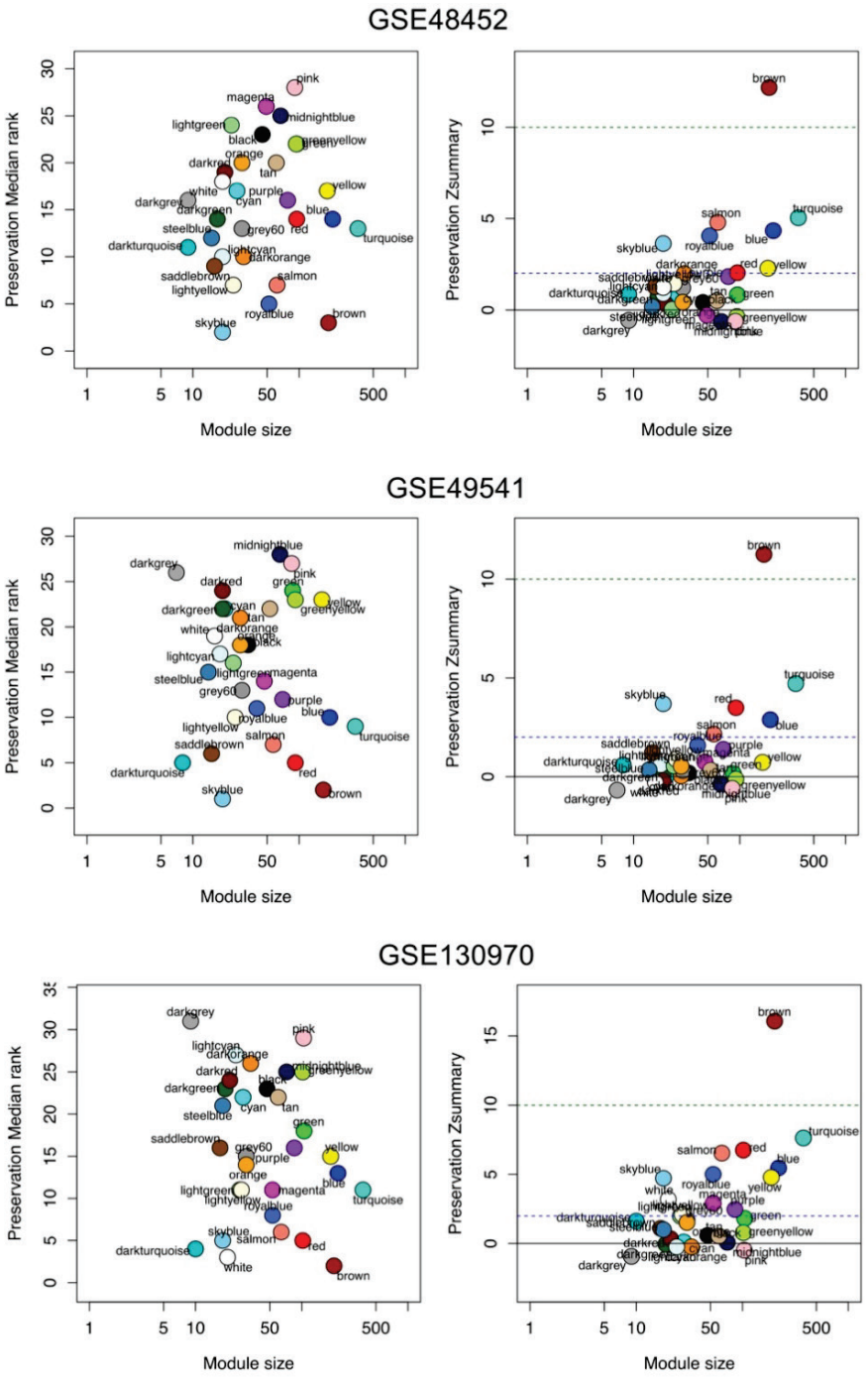


Fig. 4. Preservation of gene co-expression networks in human NAFLD datasets. Each module is represented by its color-code and name. The left figure shows the composite statistic, preservation Median rank. A lower value implies strong evidence of module preservation. The right figure shows the composite statistic, preservation Zsummary. The dashed blue and green lines indicate the thresholds Z=2 and Z=10, respectively. Zsummary > 10 implies strong evidence of module preservation, 2 < Zsummary < 10 implies weak-to-moderate, and Zsummary < 2 implies no evidence of module preservation.

Functional characterization of the modules associated with disease severity

To elucidate the functional role of each module, cell type deconvolution was performed by measuring the overlap between module genes and human liver-derived cell type marker genes [13]. Of the 30 modules identified, several modules showed a relatively clear association with specific cell types. For example, the brown module was associated with immune cells such as macrophages and mature B cells, and the royalblue module was associated with hepatic stellate cells. In addition, the yellow module was relevant to liver sinusoidal endothelial cells and cholangiocytes. Hepatocytes were enriched in 5 modules (midnightblue, magenta, pink, purple, and green) (Fig. 2B). As the purple, brown, yellow, and royalblue modules showed a strong correlation with disease severity and a relatively clear cell type, we focused on these modules. GO and KEGG pathway enrichment analyses indicated that the purple module was associated with fatty acid and xenobiotic metabolism and the brown module was involved in inflammatory responses. The yellow and royalblue

modules were associated with actin cytoskeleton and extracellular matrix organization, respectively (Fig. 3A, Supplementary Fig. 3). The boxplot of the MEs indicated that these pathways were already altered after 8 weeks of dietary intervention and progressed throughout the treatment period (Fig. 3B).

Evaluation of the similarity to human disease by module preservation analysis

Although this animal model has comorbidities and a natural history similar to human NAFLD, its relevance to human diseases is not fully understood. Therefore, we performed a module preservation analysis to confirm the similarity across species. The gene co-expression network of the brown module (inflammatory responses in immune cells) was highly preserved in all three human NAFLD datasets. In addition, the purple (fatty acid and xenobiotic metabolism), yellow (actin cytoskeleton organization), and royalblue (extracellular matrix organization) modules showed moderate evidence of preservation in some datasets (Fig. 4).

Hub gene identification in the module

As the brown module showed the strongest evidence of module preservation in human datasets, we selected it for further analysis. To identify functionally important genes in the module, genes with high K.in values were extracted as candidate hub genes, and a protein-protein interaction (PPI) network was constructed. As shown in Fig. 5, some genes were highly interconnected in the PPI network, indicating that they play a central role in the module. Based on the criteria described in the Materials and Methods section, 20 genes were selected as hub genes (Table 3). Next, we examined the gene expression changes of hub genes in the human NAFLD liver gene expression datasets to assess their relevance to human diseases (Table 1). The expression of *TYROBP*, *CTSS*, and *SYK* tended to increase in obesity and NAFLD, suggesting that they may be involved in disease initiation and progression (Fig. 6).

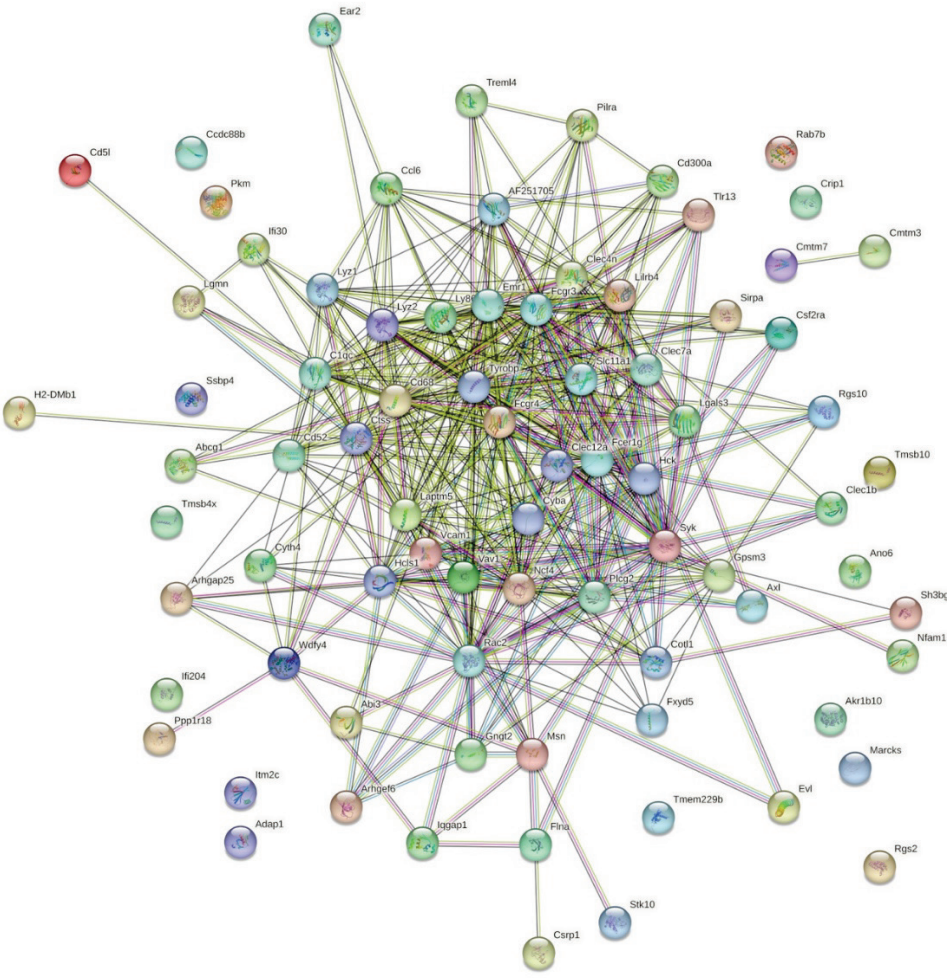


Fig. 5. Protein-protein interaction (PPI) network of the brown module. The PPI network was visualized using Cytoscape software.

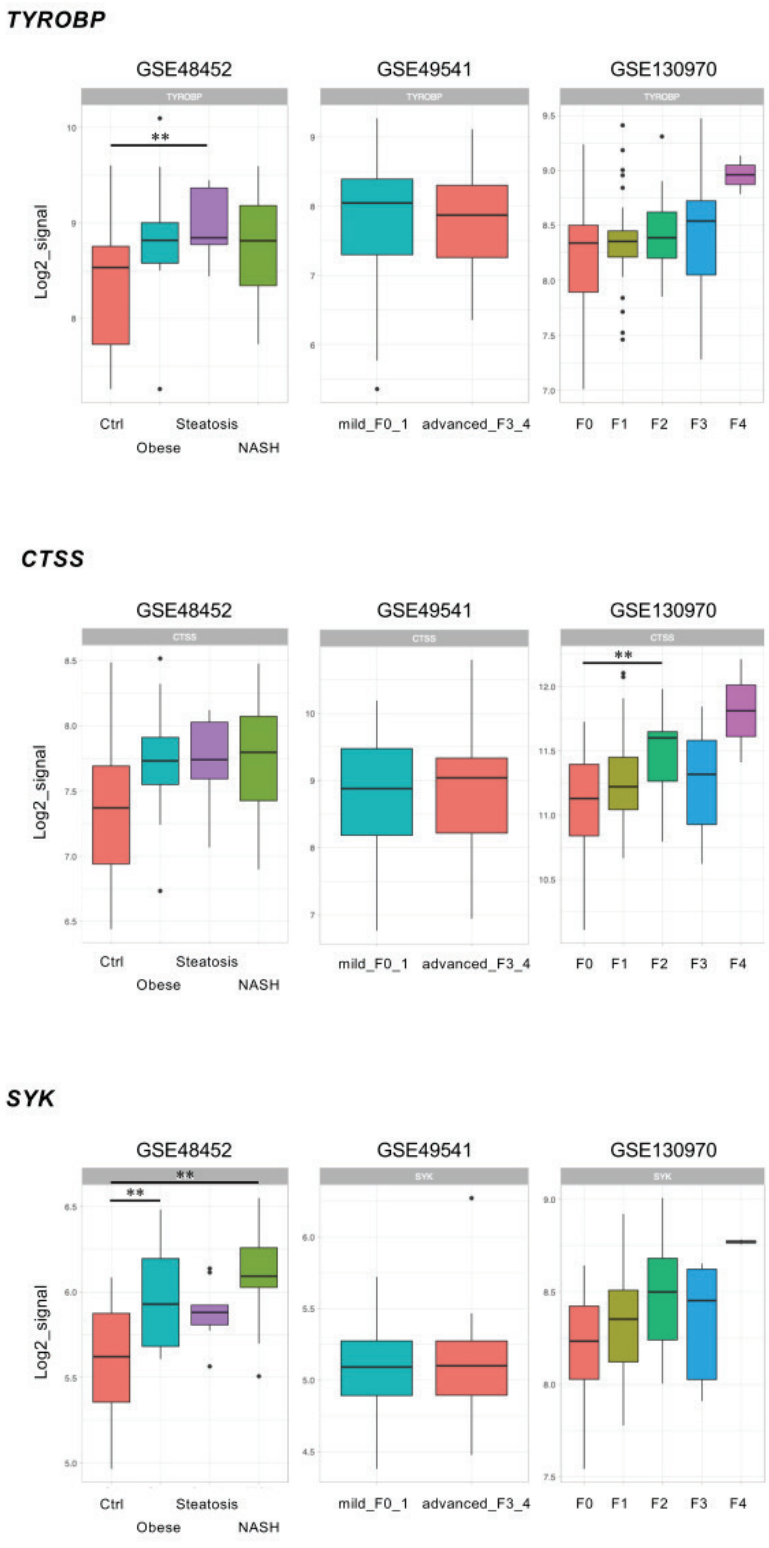


Fig. 6. Characterization of the brown module. Expression levels of the representative hub genes in human NAFLD datasets. Data are shown as box and whisker plots. *P < 0.05 and ** P < 0.01, versus Ctrl (GSE48452), mild_F0_F1 (GSE49541), and F0 (GSE130970).

Discussion

Animal models are an important part of preclinical research. In recent years, new technologies, such as induced pluripotent stem cells, organ-on-a-chip, and gene editing, have led to dramatic advances in research on disease mechanisms and the identification of potential drug targets; however, animal models remain important research tools. In particular, because of the involvement of various factors (overnutrition, low activity, and genetics, *etc.*) [19,20] and inter-organ crosstalk [21] in the pathogenesis of the disease, it is difficult to reproduce the pathogenesis of NAFLD *in vitro*. To date, several types of animal models have

been reported as NAFLD models (diet-induced, chemically induced, genetically modified, and their combination; more than 30 models in total), and it is crucial to understand the advantages and disadvantages of each model and adapt them to human conditions. The WD-induced NAFLD mouse is one of the useful animal models that reproduces the key biochemical and histological features observed in human NAFLD. Compared with HFD, the WD contains several macronutrients that accelerate NAFLD pathology. Cholesterol is one of the macronutrients that contributes to disease development and progression. It affects various cellular functions and signaling pathways, such as mitochondrial function, hypoxic response, and inflammatory response [22]. Indeed, hepatic free cholesterol levels are elevated in NASH patients, and dietary cholesterol intake is independently associated with the presence of advanced fibrosis [23,24]. Fructose is another factor associated with NAFLD development and progression. Dietary fructose is rapidly absorbed and metabolized in the liver, which strongly induces hepatic lipid accumulation, insulin resistance, and mitochondrial dysfunction [25]. High dietary fructose intake also induces intestinal dysbiosis and increased intestinal permeability, leading to LPS-induced hepatic inflammation [26]. In the present study, we detected some time-dependent effects of dietary intervention on NAFLD-related biochemical and histological parameters. The liver damages and concomitant steatohepatitis reflected in plasma ALT levels and liver histology (hepatocyte degeneration and inflammatory cell infiltrations) were observed in the WD-treated groups as early as 8 weeks of dietary intervention (Table 2, Fig. 1A). In terms of liver fibrosis, a significant increase was observed after 12 weeks of dietary intervention (Fig. 1B, C). These results suggest that WD-induced NAFLD mice can develop steatohepatitis at 8 weeks or shorter of dietary intervention and progress to liver fibrosis after week 12.

We also conducted transcriptome analysis (WGCNA) for further characterizing their pathophysiology and evaluating its relevance to human NAFLD. The brown module was strongly correlated with disease severity, and its co-expression network was highly preserved in the human NAFLD datasets. It has been reported that co-expressed genes are likely to be co-regulated and functionally related [27]. Indeed, this module was estimated to be associated with immune cells (such as macrophages and mature B cells), and the module genes were enriched in pathways related to the inflammatory response. As shown in Fig. 3B, the MEs of the brown module were upregulated after 8 weeks of WD treatment, and this trend continued throughout the experimental period. Intriguingly, these gene expression changes were consistent with the liver histology. After 8 weeks of dietary intervention, inflammatory cell infiltration was detected in H&E-stained liver sections, indicating that inflammatory processes were already activated in the liver. Macrophages are a primary immune cell type that contribute to the development and progression of NAFLD [28,29]. Kupffer cells are embryo-derived tissue-resident macrophages that play a homeostatic role in the liver. They typically reside on the luminal side of the liver sinusoid and are involved in tissue remodeling, responses to gut-derived pathogens, and the regulation of immune responses. Bone marrow-derived monocytes are another source of liver macrophages that are implicated in the pathogenesis of NAFLD. In hepatocyte lipotoxicity, various stressors such as chemokines and damage-associated molecular patterns are secreted, which in turn recruit monocytes into the liver and induce their differentiation into macrophages. Although their contribution to the patho-genesis of NAFLD has not been fully elucidated, they are thought to exhibit proinflammatory phenotypes that exacerbate liver injury and accelerate liver fibrosis.

Most cellular networks (such as metabolism, gene expression, and protein-protein interactions) exhibit a scale-free topology, and a small number of molecules with many interactions are considered functionally important [30]. Therefore, to further characterize this animal model, we identified hub genes in the brown module. The expression of *TYROBP*, *CTSS* and *SYK* tended to increase in the livers of patients with NAFLD, suggesting that they may be implicated in disease pathogenesis. TYROBP is a transmembrane adaptor polypeptide that forms a signaling complex with cell surface innate immunoreceptors and thus plays a key role in immune responses [31]. In a sepsis model, *Tyrobp* knockout mice produced fewer proinflammatory cytokines, suggesting that TYROBP may be involved in the progression of NAFLD [32]. CTSS is a lysosomal cysteine protease that participates in the degradation of antigens for presentation on MHC class II molecules. CTSS is also involved in the regulation of SIRT1 activity, and its inhibition reduces NFκB-dependent hepatic inflammation [33]. In addition, the number of CTSS-positive immune cells is higher in the liver of NAFLD patients, suggesting that this protein is

associated with NAFLD pathology [34]. Finally, SYK is a non-receptor protein tyrosine kinase mainly expressed in hematopoietic cells and is involved in various biological processes such as innate and adaptive immunity [35]. Hepatic SYK expression is increased in NAFLD patients and positively correlated with NAFLD activity score, ALT activity, and hepatic expression of the macrophage marker CD68. Moreover, in mice fed a methionine- and choline-deficient diets, *Syk* knockdown in myeloid cells prevented liver injury and inflammatory cell infiltration in the liver [36]. These results suggest that SYK inhibition could be a potential therapeutic strategy for the treatment of NAFLD.

In conclusion, transcriptome analysis provides molecular insights into the pathophysiology of NAFLD. WD-induced NAFLD mice showed biochemical and histological changes in the liver similar to human NAFLD. This animal model also showed a similar gene co-expression network to humans especially genes related to immune cell functions and some hub genes are identified in this study. Moderate preservation evidence was also found for the fatty acid, xenobiotic metabolism, actin cytoskeleton organization, and extracellular matrix organization. These results suggest that the animal models coupled with advanced molecular analyses provide valuable insights into the pathogenesis of NAFLD and therapeutic avenues.

Authors' Contributions

TI and TO conceptualized the project. TI, T Saito, MT, KU, YT, YS, KE and KM performed the experiments. TI, T Saito, and KE analyzed the data. TI, T Sasase, and TO drafted the manuscript.

Conflict of Interest

Tatsuya Ishigure, Tomohiko Sasase, Yasufumi Toriniwa, Tomoyuki Saito, Yasuka Saigo, and Koji Edamura are employees of Japan Tobacco Inc. Marika Tohma, Kinuko Uno, Katsuhiro Miyajima, and Takeshi Ohta have no conflict of interest in relation to this study.

Acknowledgements

This research received no specific grants from any funding agency in the public, commercial, or not-for-profit sectors. The authors thank Dr. Miki Sugimoto for her general instructions and JT Creative Service for long-term animal care.

References

1. Calzadilla Bertot L, Adams LA. The natural course of non-alcoholic fatty liver disease. Int J Mol Sci 2016;17. https://doi.org/10.3390/ijms17050774
2. Younossi Z, Tacke F, Arrese M, Chander Sharma B, Mostafa I, Bugianesi E, Wai-Sun Wong V, Yilmaz Y, George J, Fan J, Vos MB. Global perspectives on nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. Hepatology 2019;69:2672-2682. https://doi.org/10.1002/hep.30251
3. Younossi ZM, Golabi P, Paik JM, Henry A, Van Dongen C, Henry L. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review. Hepatology 2023;77:1335-1347. https://doi.org/10.1097/HEP.0000000000000004
4. Younossi ZM, Stepanova M, Ong J, Trimble G, AlQahtani S, Younossi I, Ahmed A, Racila A, Henry L. Nonalcoholic steatohepatitis is the most rapidly increasing indication for liver transplantation in the United States. Clin Gastroenterol Hepatol 2021;19:580-589 e585. https://doi.org/10.1016/j.cgh.2020.05.064
5. Allen AM, Lazarus JV, Younossi ZM. Healthcare and socioeconomic costs of NAFLD: A global framework to navigate the uncertainties. J Hepatol 2023. https://doi.org/10.1016/j.jhep.2023.01.026
6. Harrison SA, Loomba R, Dubourg J, Ratziu V, Noureddin M. Clinical trial landscape in NASH. Clin Gastroenterol Hepatol 2023. https://doi.org/10.1016/j.cgh.2023.03.041
7. Oligschlaeger Y, Shiri-Sverdlov R. NAFLD preclinical models: More than a handful, less of a Concern? Biomedicines 2020;8. https://doi.org/10.3390/biomedicines8020028
8. Chua D, Low ZS, Cheam GX, Ng AS, Tan NS. Utility of human relevant preclinical animal models in navigating NAFLD to MAFLD paradigm. Int J Mol Sci 2022;23. https://doi.org/10.3390/ijms232314762
9. Asgharpour A, Cazanave SC, Pacana T, Seneshaw M, Vincent R, Banini BA, Kumar DP, Daita K, Min HK, Mirshahi F, Bedossa P, Sun X, Hoshida Y, Koduru SV, Contaifer D, Jr., Warncke UO, Wijesinghe DS, Sanyal AJ. A diet-induced animal model of non-alcoholic fatty liver disease and hepatocellular cancer. J Hepatol 2016;65:579-588. https://doi.org/10.1016/j.jhep.2016.05.005

10. Abe N, Kato S, Tsuchida T, Sugimoto K, Saito R, Verschuren L, Kleemann R, Oka K. Longitudinal characterization of diet-induced genetic murine models of non-alcoholic steatohepatitis with metabolic, histological, and transcriptomic hallmarks of human patients. Biol Open 2019;8. https://doi.org/10.1242/bio.041251
11. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008;9:559. https://doi.org/10.1186/1471-2105-9-559
12. Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the dynamic tree cut package for R. Bioinformatics 2008;24:719-720. https://doi.org/10.1093/bioinformatics/btm563
13. MacParland SA, Liu JC, Ma XZ, Innes BT, Bartczak AM, Gage BK, Manuel J, Khuu N, Echeverri J, Linares I, Gupta R, Cheng ML, Liu LY, Camat D, Chung SW, Seliga RK, Shao Z, Lee E, Ogawa S, Ogawa M, Wilson MD, Fish JE, Selzner M, Ghanekar A, Grant D, Greig P, Sapisochin G, Selzner N, Winegarden N, Adeyi O, Keller G, Bader GD, McGilvray ID. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat Commun 2018;9:4383. https://doi.org/10.1038/s41467-018-06318-7
14. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16:284-287. https://doi.org/10.1089/omi.2011.0118
15. Langfelder P, Luo R, Oldham MC, Horvath S. Is my network module preserved and reproducible? PLoS Comput Biol 2011;7:e1001057. https://doi.org/10.1371/journal.pcbi.1001057
16. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, Jensen LJ, von Mering C. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res 2021;49:D605-D612. https://doi.org/10.1093/nar/gkaa1074
17. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498-2504. https://doi.org/10.1101/gr.1239303
18. Hagstrom H, Nasr P, Ekstedt M, Hammar U, Stal P, Hultcrantz R, Kechagias S. Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. J Hepatol 2017;67:1265-1273. https://doi.org/10.1016/j.jhep.2017.07.027
19. Younossi Z, Anstee QM, Marietti M, Hardy T, Henry L, Eslam M, George J, Bugianesi E. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. Nat Rev Gastroenterol Hepatol 2018;15:11-20. https://doi.org/10.1038/nrgastro.2017.109
20. Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ. Mechanisms of NAFLD development and therapeutic strategies. Nat Med 2018;24:908-922. https://doi.org/10.1038/s41591-018-0104-9
21. Santos J, Maio MC, Lemes MA, Laurindo LF, Haber J, Bechara MD, Prado PSD, Jr., Rauen EC, Costa F, Pereira BCA, Flato UAP, Goulart RA, Chagas EFB, Barbalho SM. Non-alcoholic steatohepatitis (NASH) and organokines: What is now and what will be in the future. Int J Mol Sci 2022;23. https://doi.org/10.3390/ijms23010498
22. Vergara D, Casadei-Gardini A, Giudetti AM. Oxidative molecular mechanisms underlying liver diseases: From systems biology to the personalized medicine. Oxid Med Cell Longev 2019;2019:7864316. https://doi.org/10.1155/2019/7864316
23. Ioannou GN. The role of cholesterol in the pathogenesis of NASH. Trends Endocrinol Metab 2016;27:84-95. https://doi.org/10.1016/j.tem.2015.11.008
24. Guveli H, Kenger EB, Ozlu T, Kaya E, Yilmaz Y. Macro- and micronutrients in metabolic (dysfunction) associated fatty liver disease: association between advanced fibrosis and high dietary intake of cholesterol/saturated fatty acids. Eur J Gastroenterol Hepatol 2021;33:e390-e394. https://doi.org/10.1097/MEG.0000000000002110
25. Federico A, Rosato V, Masarone M, Torre P, Dallio M, Romeo M, Persico M. The role of fructose in non-alcoholic steatohepatitis: Old relationship and new insights. Nutrients 2021;13. https://doi.org/10.3390/nu13041314

26. Lambertz J, Weiskirchen S, Landert S, Weiskirchen R. Fructose: A dietary sugar in crosstalk with microbiota contributing to the development and progression of non-alcoholic liver disease. Front Immunol 2017;8:1159. https://doi.org/10.3389/fimmu.2017.01159
27. Stuart JM, Segal E, Koller D, Kim SK. A gene-coexpression network for global discovery of conserved genetic modules. Science 2003;302:249-255. https://doi.org/10.1126/science.1087447
28. Barreby E, Chen P, Aouadi M. Macrophage functional diversity in NAFLD - more than inflammation. Nat Rev Endocrinol 2022;18:461-472. https://doi.org/10.1038/s41574-022-00675-6
29. Vonderlin J, Chavakis T, Sieweke M, Tacke F. The multifaceted roles of macrophages in NAFLD pathogenesis. Cell Mol Gastroenterol Hepatol 2023;15:1311-1324. https://doi.org/10.1016/j.jcmgh.2023.03.002
30. Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101-113. https://doi.org/10.1038/nrg1272
31. Tomasello E, Vivier E. KARAP/DAP12/TYROBP: three names and a multiplicity of biological functions. Eur J Immunol 2005;35:1670-1677. https://doi.org/10.1002/eji.200425932
32. Turnbull IR, McDunn JE, Takai T, Townsend RR, Cobb JP, Colonna M. DAP12 (KARAP) amplifies inflammation and increases mortality from endotoxemia and septic peritonitis. J Exp Med 2005;202:363-369. https://doi.org/10.1084/jem.20050986
33. de Mingo A, de Gregorio E, Moles A, Tarrats N, Tutusaus A, Colell A, Fernandez-Checa JC, Morales A, Mari M. Cysteine cathepsins control hepatic NF-kappaB-dependent inflammation via sirtuin-1 regulation. Cell Death Dis 2016;7:e2464. https://doi.org/10.1038/cddis.2016.368
34. Liu XH, Zhou JT, Yan CX, Cheng C, Fan JN, Xu J, Zheng Q, Bai Q, Li Z, Li S, Li X. Single-cell RNA sequencing reveals a novel inhibitory effect of ApoA4 on NAFL mediated by liver-specific subsets of myeloid cells. Front Immunol 2022;13:1038401. https://doi.org/10.3389/fimmu.2022.1038401
35. Mocsai A, Ruland J, Tybulewicz VL. The SYK tyrosine kinase: a crucial player in diverse biological functions. Nat Rev Immunol 2010;10:387-402. https://doi.org/10.1038/nri2765
36. Luci C, Vieira E, Bourinet M, Rousseau D, Bonnafous S, Patouraux S, Lefevre L, Larbret F, Prod'homme V, Iannelli A, Tran A, Anty R, Bailly-Maitre B, Deckert M, Gual P. SYK-3BP2 pathway activity in parenchymal and myeloid cells is a key pathogenic factor in metabolic steatohepatitis. Cell Mol Gastroenterol Hepatol 2022;13:173-191. https://doi.org/10.1016/j.jcmgh.2021.08.004