CrossMark

ORIGINAL PAPER

Transcriptional networks implicated in human nonalcoholic fatty liver disease

Hua Ye[1] · Wei Liu[2,3]

Abstract The transcriptome of nonalcoholic fatty liver disease (NAFLD) was investigated in several studies. However, the implications of transcriptional networks in progressive NAFLD are not clear and mechanisms inducing transition from nonalcoholic simple fatty liver (NAFL) to nonalcoholic steatohepatitis (NASH) are still elusive. The aims of this study were to (1) construct networks for progressive NAFLD, (2) identify hub genes and functional modules in these networks and (3) infer potential linkages among hub genes, transcription factors and microRNAs (miRNA) for NAFLD progression. A systems biology approach by combining differential expression analysis and weighted gene co-expression network analysis (WGCNA) was utilized to dissect transcriptional profiles in 19 normal, 10 NAFL and 16 NASH patients. Based on this framework, 3 modules related to chromosome organization, proteasomal ubiquitin-dependent protein degradation and immune response were identified in NASH network. Furthermore, 9 modules of co-expressed genes associated with NAFL/NASH transition were found. Further characterization of these modules defined 13 highly connected hub genes in NAFLD progression network. Interestingly, 11 significantly changed miRNAs were predicted to target 10 of the 13 hub genes. Characterization of modules and hub genes that may be regulated by miRNAs could facilitate the identification of candidate genes and pathways responsible for NAFL/NASH transition and lead to a better understanding of NAFLD pathogenesis. The identified modules and hub genes may point to potential targets for therapeutic interventions.

Introduction

Nonalcoholic fatty liver disease (NAFLD) represents a wide spectrum of disorders, the prevalence of which may be as high as 24 % of the population of the United States (Parekh and Anania 2007). NAFLD encompasses not only simple steatosis (also called nonalcoholic fatty liver or NAFL) but also includes nonalcoholic steatohepatitis (NASH), advanced fibrosis, or cirrhosis. Hepatic steatosis can progress to NASH, which can be distinguished from simple steatosis by the presence of ballooning hepatocytes, apoptotic cells, inflammatory infiltrate and/or collagen deposition (fibrosis) (Cohen et al. 2011). Currently, oxidative stress and inflammation are recognized as important contributors to NASH (Fujimoto et al. 2008; Podrini et al. 2013). However, mechanisms of the transition from NAFL to NASH are still largely unknown (Cohen et al. 2011). Mouse models of NAFLD have provided insights into

Communicated by I. Thiele.

✉ Wei Liu
antibodyliu@gmail.com

[1] Department of Gastroenterology, Ningbo Medical Treatment Center Lihuili Hospital, Ningbo 315040, China

[2] School of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China

[3] Department of Pathology, Human Centrifuge Medical Training Center, Institute of Aviation Medicine of Chinese PLA Air Force, Beijing 100089, China

possible pathological mechanisms contributing to its development, but the relative roles of these pathways in humans have not been conclusively determined due to species differences (Yu et al. 2010; Hebbard and George 2011). Some human-based studies integrate metabolism network and gene expression profiling to find potential metabolite biomarkers (Sookoian and Pirola 2012, 2013). Some use the protein interaction network or gene expression to find potential NAFLD drugs and mechanisms (Sookoian et al. 2011; Sookoian and Pirola 2012). However, a gene coexpression network-based analysis is still vacuum. Therefore, network approaches for investigation of human NAFLD pathogenesis and progression are required. Better understanding of the pathogenesis of liver steatosis and its progression to NASH will have a major impact on the prognosis and therapy of NAFLD patients. This is particularly important for individuals affected by NASH given the greater association of NASH than simple steatosis with cirrhosis and hepatocellular carcinoma (Koutsari and Lazaridis 2010).

Microarray gene expression profiling is a high-throughput method for genome-scale, quantitative examination of liver disease by simultaneously measuring expression levels for thousands of genes (Shackel et al. 2002). The technology has been applied to NAFLD gene expression (Younossi et al. 2005a; Yoneda et al. 2007; Greco et al. 2008) and provides insights into molecular aspects of NAFLD progression (Younossi et al. 2005b; Yoneda et al. 2008; Bertola et al. 2010; Gawrieh et al. 2010; Ahrens et al. 2013; Lopez-Vicario et al. 2014; Moylan et al. 2014). However, traditional differential expression analysis methods are biased against genes with large changes in expression, results of which are varied across different analytical methods and threshold selection without considering the relationship between changing genes, which, in turn, cause much useful information lost. The use of co-expression networks can surmount this problem because it allows for the examination of gene expression from a systematic perspective (Lee et al. 2004). Weighted gene co-expression network analysis (WGCNA) groups functionally related genes into modules based on the variability in microarray data among biological samples. The modularity of the biological system allows us to look into its components independently, and the relationships between genes within modules can be delineated. It is believed that modules are stable units in systems biology because the overall function of a module can remain the same while individual gene expression can be changed or replaced by other genes with similar redundant functions. Therefore, functional modules can more effectively reveal consistent differences during NAFLD progression.

Herein, we first applied WGCNA to publicly available microarray data covering normal control, NAFL and NASH. Genome-scale modules of co-expressed genes with clear functional annotations were identified. The analysis revealed the commonplace and characteristic between transcriptional networks of NAFL and NASH. Our results also suggested a functional link between differential microRNAs (miRNA) and hub genes during NAFL/NASH transition.

Materials and methods

Data acquisition

Microarray dataset from a previous study was used in this analysis: microarrays representing three clinically defined pathological groups; normal, NAFL and NASH. The raw data (.cel files) were obtained from the ArrayExpress database with accession number E-MEXP-3291. Briefly, 19 normal, 10 steatotic and 16 NASH livers were used for microarray analysis. All these samples were clinically characterized as previously described (Lake et al. 2011) (Supporting Table 1).

Microarray data analysis

Microarray data analysis was performed using the R software and Bioconductor affy package. Raw expression data were log_2 transformed and normalized by quantile normalization. Differential expression was detected using the SAM package (significance analysis of microarrays, http://www-stat.stanford.edu/~tibs/SAM/) and the significance threshold was FDR <0.01 and fold changes >1.3. (Details of the gene expression analysis are given in Supporting Information)

Weighted gene co-expression network analysis (WGCNA)

Unsigned co-expression networks were constructed on the basis of 5000 genes with highest expression variance using the WGCNA package in R (Langfelder and Horvath 2008). For each set of genes a pairwise Pearson correlation coefficient matrix is computed, and an adjacency matrix is calculated by raising the correlation matrix to a power (Zhang and Horvath 2005). The power of 14 was chosen using the scale-free topology criterion and was used for all three networks: the network built using control versus NAFL samples, controls versus NASH samples or NAFL versus NASH samples. The weighted network was transformed into a network of topological overlap (TO)—an advanced co-expression measurement that considers not only the correlation of 2 genes with each other, but also the extent of their shared correlations across the weighted

network (Zhang and Horvath 2005). Genes were hierarchically clustered on the basis of their TO. Finally, modules were identified on the dendrogram using the Dynamic Tree Cut algorithm (Langfelder et al. 2008). Each module was summarized using singular value decomposition so that each module eigengene (ME) represented the first principal component of module expression profiles (Zhang and Horvath 2005). Thus, ME explains the maximum amount of variation of the module expression levels and is considered the most representative gene expression in a module. The resulting MEs were used to extend the modules' gene composition from 5000 network genes up to the genome scale (19,742 genes). For each module, kME is defined as the correlation between gene expression values and the module eigengene. Genes were assigned to a module according to the best matching ME according to its kME values. To avoid capturing weak associations, it was additionally required that the |kME| value is not less than 0.7. Genes weakly correlated with all of the MEs (|kME| <0.7) were assigned to none of the modules. To visualize the pairwise relationships between genes, VisAnt (Hu et al. 2009) was used. About 100 pairs of genes with the highest intramodular TO value were depicted (Supporting Information).

The disease status was correlated with module eigengene by Pearson correlation.

Functional annotation of the modules

Gene ontology (GO) enrichment and KEGG pathway analysis for network modules were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang da et al. 2009) with the background list of all genes on the array. Association of modules with genomic aberration was detected using DAVID on the basis of overrepresentation of genes encoded at neighboring chromosomal locations. To discover potential common transcription factors that may control transcription of genes in a module, transcription factor binding site (TFBS) information was retrieved. In DAVID, an overrepresentation of a term is defined as a modified Fisher's exact P value with an adjustment for multiple tests using Benjamini method. In addition, we related modules to biological processes in NAFLD on the basis of the literature data.

Identification of potential miRNA gene targets

Gene expression can be regulated by miRNA (Lim et al. 2005). Analysis of 3′-UTR sequences of co-expressed genes can provide evidence of miRNA regulation. Predicted gene targets of all differentially expressed miRNAs were identified using several databases, including micro-rna.org, MicroCosm and TargetScan. For genes that were not available from these resources, a PubMed literature search was performed. To associate miRNA with predicted target gene, correlation in expression level for each of these miRNA-gene pairs were calculated using SPSS 17.0.

Results

Neither age nor gender is a major driver of gene expression differences between the NAFL and NASH groups

Owing to the relatively small number of array samples used in this analysis ($n = 46$), we assessed if individual variation, age or gender contributes to expression differences. We found that the overall gene expression variation between individuals of normal, NAFL and NASH were 5.5 ± 3.1 %, 5.2 ± 3.1 % and 5.4 ± 3.2 %, respectively. We also applied linear regression of expression values against age and gender and then assessed differential expression between the NAFL and NASH groups using the residual values. We observed an over 95 % overlap between differentially expressed genes using either the residual values or the raw data, indicating that neither age nor gender was a major driver of gene expression differences between the control and NAFLD groups (Supporting Fig. 1). In addition, supervised hierarchical clustering based on the top 200 variant expressed genes showed distinct clustering of all NASH samples (Supporting Fig. 2). The inter-array correlation was also calculated and clustered to ensure that no aberrant gene expression existed (Supporting Fig. 3).

Network of NAFL versus control

Thirteen modules were identified from WGCNA of control and NAFL group. No modules were significantly associated with the disease status (Supporting Table 2). Interestingly, DAVID analysis revealed that module S1, which was enriched in proteins involved in sterol biosynthetic process ($P = 6.3E-29$), contained 4 genes bearing loci contribute to polygenic dyslipidemia ($P = 2.2E-2$). And module S7 was enriched in genes located on Y chromosome ($P = 2.9E-21$). Other eleven modules were also significantly associated with specific biological processes: oxidation reduction, fatty acid metabolic process, ion homeostasis, DNA replication, M phase, protein transport, RNA processing, immune response and negative regulation of ubiquitin-protein ligase activity during mitotic cell cycle. Although some modules are similar in function, they are from different cellular compartments. For example, microsomal module S2 and mitochondrial module S3 both perform oxidative reduction. It has been demonstrated that DNA methylation of liver mitochondrion is associated with histological severity of NAFLD and insulin

resistance (Sookoian et al. 2010; Pirola et al. 2013a). These results indicate that WGCNA can group genes by function in a robust manner.

Network of NASH versus control

WGCNA identified fourteen modules between NASH and control group, including three modules (N1, N8, N14) significantly correlated with disease status (Supporting Table 3). DAVID analysis indicated three disease-related modules are significantly enriched with genes in the following biological processes: proteasomal protein catabolic process, immune response and chromosome organization. Module N5 contained 6 genes bearing loci contribute to polygenic dyslipidemia ($P = 8.4E{-}4$). Also, two KEGG pathways were represented in modules N8 and N14, including proteasome ($P = 6.3E{-}16$) and cell adhesion molecules ($P = 1.1E{-}4$). Results of DAVID analysis for other modules are presented in Supporting Table 3. The heat-map of module genes illustrated these modules are differentially expressed and are correlated with disease status. These significant module networks were visualized (Supporting Fig. 4).

Network for NASH versus NAFL

To understand the molecular mechanisms responsible for NAFL to NASH transition, network of NAFLD progression was also analyzed. Fourteen modules were identified in progression network, including nine modules significantly associated with disease status. DAVID analysis results of these nine modules are shown in Table 1. The heat-map of module genes illustrated these modules are differentially expressed and are correlated with disease status. These significant module networks were visualized (Supporting Fig. 5).

It has been demonstrated that WGCNA modules generated from small sample sizes ($n < 30$) are still robust (Farber 2010). However, owing to the relatively small number of array samples used in this group ($n = 26$), we wanted to quantitatively assess the modules' stability. For each module, intramodular connectivity in 1000 module gene sets generated by sampling 13 of the 26 array samples were calculated. Then, correlations between the true module gene connectivity values and those from the 1000 randomly selected sets were calculated. The correlations of each module were higher than 0.7 and the average correlation was 0.8 (Supporting Fig. 6). The result proved that the modules are robust to a 50 % sample exclusion.

Furthermore, two recently published NAFLD datasets GSE48452 (Ahrens et al. 2013) and GSE61620 (Horvath et al. 2014) were used for validation. The module preservation analysis results suggested that all modules were preserved in these two datasets. All modules have a Zsummary statics greater than 2, which is the threshold for module preservation. The lowest preservation is the module 2. The detailed module preservation statistics is provided in Table 2.

Two modules function in protein degradation and cell adhesion in progression network, respectively

After viewing the global properties of the three networks, we next examined details of particular modules. In progression network, the most significantly altered module P4 ($P = 9.73E{-}8$) was down-regulated in NASH and was also the second largest module in terms of gene number (Fig. 1a; Table 1). Module P4 was associated with protein degradation, and contained endoplasmic reticulum and proteasome (Fig. 1a). Genes in this module were enriched with target genes of transcription factor including MECOM ($P = 3.4E{-}5$), POU2F1 ($P = 9.5E{-}4$), MEF2 ($P = 1.2E{-}3$), NRF2 ($P = 1.4E{-}3$) and CEBP ($P = 3.9E{-}2$). The module was centered at SRPRB and was strongly connected with other components (Fig. 1a). SRPRB strongly interacts with ABCB6, a membrane transport protein important for cadmium ion transmembrane transport and heme transport. SRPRB is a subunit of the signal recognition particle receptor, involved in small GTPase-mediated signal transduction. The largest module P8 (Fig. 1b) includes mainly extracellular matrix component playing a role in cell adhesion ($P = 3.8E{-}21$). The module was up-regulated in NASH and is also the second most significant module.

We next focus on the three smallest modules (P1, P2 and P7, Supporting Table 4) in progression network. Cadmium ion binding proteins were significantly enriched in module P1. Hub genes in this module were from metallothionein family (Supporting Fig. 5A). Genes in this module were down-regulated, indicating ion dysregulation during transition of NAFL to NASH. Modules P2 and P7 were participating in ribosome biogenesis and immune response. The ribosome biogenesis complex contains 58 genes centered around SNORD44 and SNORD75 (Supporting Fig. 5B), while immune response complex consists of 55 genes centered around IGKC and IGHG1 (Supporting Fig. 5D). Interestingly, module P7 also enriched ribosome proteins indicating elevated protein synthesis. Together, our module data suggest the differences in protein degradation, ribosomal biogenesis and immune response between NAFL and NASH can be used as molecular features distinguishing these two disease stages.

To decide which of these nine modules in the progression network is associated with NAFLD phenotype,

Table 1 Functional annotation of the modules in progression network

Module (no. genes)	GO[a]	Overrepresented KEGG pathway[a]	Overrepresented TFBS[a]	Hubs[b]
P1 (22)	Cadmium ion binding (4.0E−17) Copper ion binding (1.4E−10)		MYB (3.1E−2) USF (2.1E−2)	MT1IP, MT1DP, MT1X
P2 (58)	Ribosome biogenesis (1.5E−6) Nucleolus (7.1E−10) RNA binding (3.4E−4)			SNORD44, SNORD75, SNORD30
P4 (711)	Proteasomal protein catabolic process (6.2E−19) Endoplasmic reticulum (2.5E−19) Proteasome complex (8.4E−19) Oligosaccharyl transferase activity (1.6E−7) Unfolded protein binding (7.0E−5)	Proteasome (1.0E−20) N-Glycan biosynthesis (6.1E−8) Alanine, aspartate and glutamate metabolism (1.8E−2)	MECOM (3.4E−5) OCT1 (9.5E−4) MEF2 (1.2E−3) NRF2 (1.4E−3) CEBP (3.9E−2)	SRPRB, NANS, PDIA6
P5 (104)	Translational elongation (4.8E−10) Cytosolic ribosome (6.8E−9) Structural constituent of ribosome (7.2E−8)	Ribosome (5.0E−9)		RPL18, RPL28, RPL8
P7 (55)	Immune response (4.0E−2) Antigen binding (4.7E−8)	Ribosome (3.4E−4)		IGKC, IGHG1, IGKV4-1
P8 (1435)	Cell adhesion (1.7E−8) Extracellular matrix (1.7E−7) Growth factor binding (4.0E−3)	Focal adhesion (6.8E−13) ECM-receptor interaction (1.1E−8)	FOXO1 (2.2E−32) SOX5 (2.6E−31) SRY (2.1E−30) STAT5B (3.0E−30) FOXO4 (1.2E−29) HFH3 (2.6E−29)	TCF4, ANTXR1, VIM
P9 (109)	Response to molecule of bacterial origin (1.0E−4) Cell surface (1.9E−2)	T cell receptor signaling pathway (4.1E−3) Toll-like receptor signaling pathway (1.2E−2)	NFKB (4.8E−4) CREL (6.2E−3) STAT (1.4E−2) FREAC7 (2.8E−2)	PTGS2, LCP2, DOCK10
P13 (416)	mRNA processing (2.2E−6) Ribonucleoprotein complex (1.4E−6) RNA binding (5.4E−4)	Spliceosome (6.0E−4) Oxidative phosphorylation (3.5E−2)	YY1 (2.1E−5) AML1 (1.2E−5) E2F (1.6E−5) ELK1 (1.6E−5)	STOML2, PAN3, YIF1A
P14 (251)	Protein transport (6.8E−5) Nuclear lumen (8.0E−3)		SP1 (5.7E−3)	SRRM2, CCT4, SNRNP70

[a] Functional terms overrepresented with highest statistical significance in the given module (Benjamini–Hochberg corrected P values are given in brackets)

[b] Genes that are the top 3 most connected in the given module

an independent genetic validation was performed. We directly used a list of genes with significant SNP loci from a recently published paper (Chalasani et al. 2010). These genes were identified to be associated with NAFLD phenotype as described in the original work. These genes were mapped to nine progression modules (Supporting Table 5). The mean intramodular connectivity of the associated versus non-associated genes were calculated in all modules. Module P8 had a maximum 15 genes bearing SNPs genotyped in the study. These 15 genes indeed had a higher connectivity than other genes in the module ($P = 0.009$). The results indicate that module P8 may be involved in NAFLD progression.

Identification of key hub genes and overrepresentation of TFBS in NAFLD progression network

Highly connected intramodular "hub" genes play pivotal role in maintaining the module structure (Barabasi et al. 2011). Evidence suggests that a gene's network position has significant functional implications, with more centralized genes in the network more likely to be vital to proper cellular function than peripheral genes (nodes). It has previously been shown that hubs play important roles in yeast protein networks (Jeong et al. 2001) and in glioblastoma gene networks (Ivliev et al. 2010), where hubs have been shown to be therapeutic targets or signatures for

Table 2 Gene network modules from E-MEXP-3291 are well preserved in GSE48452 and GSE61620

Module	Zsummary in GSE48452	Zsummary in GSE61620
1	17.0	13.0
2	2.3	4.5
3	25.0	22.0
4	4.2	10.0
5	16.0	20.0
6	58.0	49.0
7	6.4	4.7
8	26.0	24.0
9	15.0	14.0
10	11.0	8.7
11	11.0	5.5
12	16.0	11.0
13	15.0	14.0
14	23.0	19.0

Zsummary<2 implies no evidence for module preservation, 2<Zsummary<10 implies weak to moderate evidence, and Zsummary>10 implies strong evidence for module preservation

prognosis. Also, co-expressed genes in some modules may be co-regulated by one or more common transcription factor (TF) (Segal et al. 2003). In progression network, top three hub genes within each module and overrepresentation of TFBS for TFs of the module are listed in Table 1. For example, MT1IP, MT1DP and MT1X are hub genes in module P1, and all belong to metallothionein family. Small nucleolar RNAs SNORD44, SNORD75 and SNORD30 are hub genes in module P2 involved in ribosome biogenesis. Common TFs overrepresented in in modules are also summarized in Table 1.

## Hub genes or TFs as miRNA targets

A cell's activity is organized as a network of functionally related modules: sets of genes co-regulated to respond to different conditions. Genes that are highly connected are referred to as hubs and have been proven to be important in disease and in controlling module behavior (Miller et al. 2008). NAFLD is not a consequence of an abnormality in a specific gene, but reflects the perturbations of the complex network formed by connected genes or modules. We have characterized NAFLD progression network mentioned above. So, we next want to explore which differentially expressed miRNAs may regulate NAFLD progression through module hub genes or modular regulatory TFs. Differential expression of 193 miRNA probe sets present on the chip was calculated using SAM package (Tusher et al. 2001). We focused our search on the 27

hub genes and the 15 differential miRNAs. We found that 12 of the 20 hub genes were the predicted targets for 13 differentially expressed miRNAs (Table 3). 6 of the 12 hub genes MT1DP, MT1X, SRPRB, PDIA6, NANS and YIF1A were significantly down-regulated, and the other 6 hub genes IGHG1, TCF4, VIM, PTGS2, PAN3 and SRRM2 were significantly up-regulated in NAFL/NASH transition (Table 4). Also, 7 of the module regulatory TFs are predicted to be targets of 9 differential miRNAs (Table 3). To provide further evidence that these genes are the targets of differential miRNAs, significant correlation in expression level for each of these miRNA-gene pairs was shown (Supporting Table 6). The most significant of miRNAs:hub gene pairs is miR-30e:NANS ($r = 0.69$, $P = 9.20E-5$). miR-30e also negatively regulates VIM ($r = -0.62$, $P = 5.88E-4$), which is important in lymphocyte adhesion and transmigration (Nieminen et al. 2006) during inflammation. miR-192, significantly down regulated in NAFLD, was predicted to positively regulate MT1X ($r = 0.53$, $P = 4.85E-3$), which is a hub gene involved in cadmium ion homeostasis. The significantly down regulated miR-15a was predicted to positively regulate PDIA6 ($r = 0.53$, $P = 5.34E-3$), which is a hub gene involved in unfolded protein binding within endoplasmic reticulum lumen.

Because miRNAs can target hub genes, we hypothesized that modules may be controlled by miRNAs. To this end, we correlated modular gene expression with those significantly altered miRNAs to test the hypothesis. Interestingly, several modules showed a significant correlation with expression levels of miRNAs (Supporting Fig. 7). For example, the highest correlation was between miR-192 and module P8 ($r = 0.82$, $P = 3.13E-7$). Module P8 was enriched with genes involved in cell adhesion ($1.7E-8$) and inflammatory response ($1.2E-2$). Module P8 expression was also associated with miR-122 expression ($r = 0.64$, $P = 4.62E-4$) and miR-24-2 expression ($r = 0.65$, $P = 3.41E-4$).

Candidate key target genes selection for NAFL/NASH pathogenesis

As nine modules correlated with NAFLD disease status, a panel of candidate genes based on modules P1, P2, P4, P5, P7, P8, P9, P13 and P14 for distinguishing NASH from NAFL subjects are listed in Table 4. These genes showed significant differential expression as indicated by the *P* values, and they were important components of the NASH-related modules as reflected by their strongly positive or negative values of module membership (kME). The expression of these genes in individual patient was shown (Supporting Fig. 8).
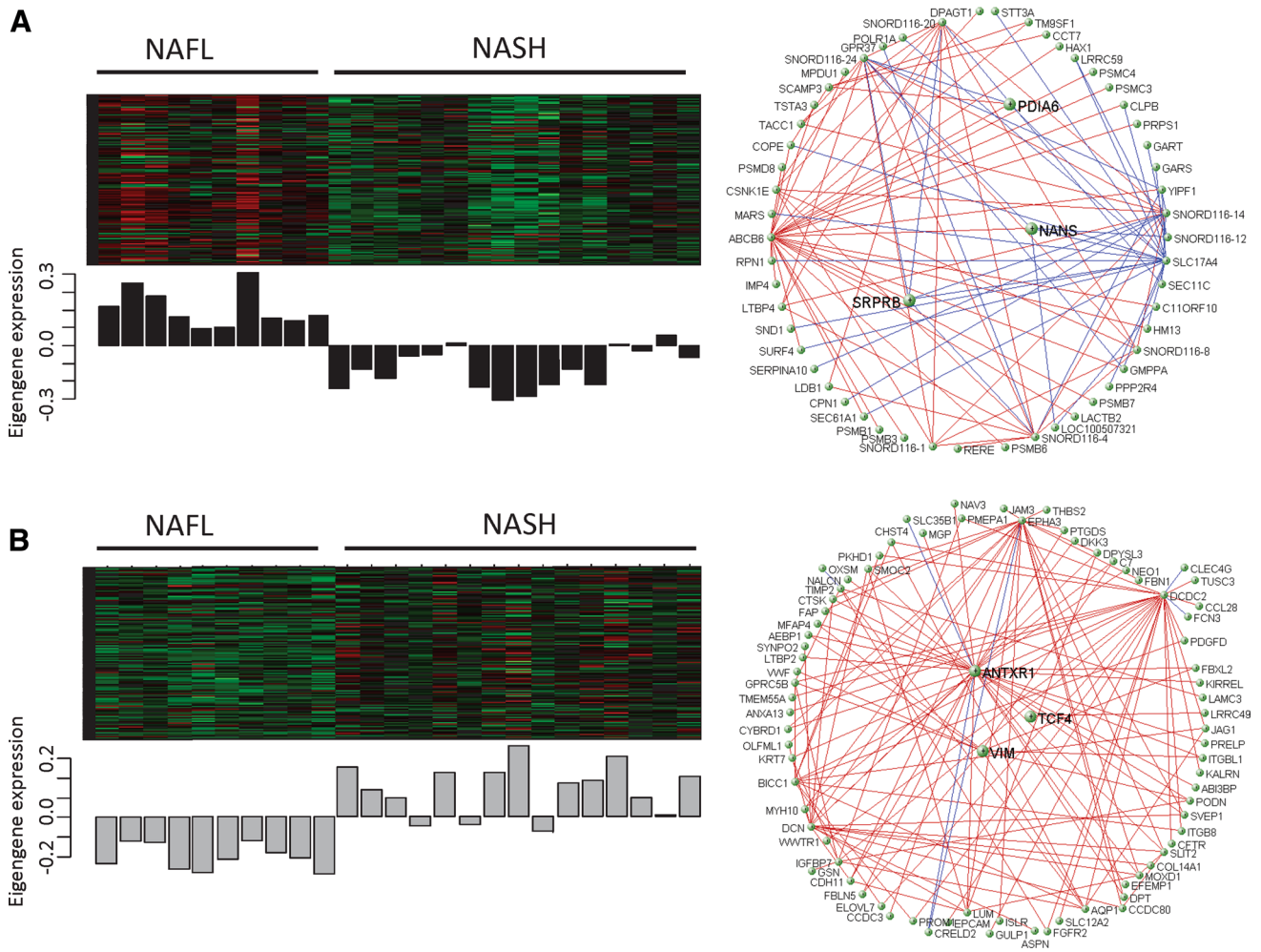
Fig. 1 Modules of P4 and P8 in progression network are the two most significant ones. Expression within both modules is shown in the heat-map and summarized with the module eigengene (*below columns*). The two modules were visualized using VisANT to plot the 100 strongest connections within each module. *Red lines* denote positive correlation between two genes, while *blue* denotes negative correlation. **a** The P4 module contains genes co-regulated in proteasome pathway. **b** The P8 module is enriched for genes involved in focal adhesion pathway. Both pathways are related to inflammatory responses (color figure online)

Discussion

Currently, only a few human NAFLD transcriptome studies have been reported, due to clinic sample availability and microarray data accessibility. NAFLD animal model transcriptome (Hebbard and George 2011) results are not consistent with human due to species divergence (Odom et al. 2007). Furthermore, most of these studies are focused on just one stage, ignoring the property of disease progression. Here, we re-analyze this NAFLD progression microarray dataset. However, a big challenge in microarray studies is how to interpret gene lists into function, especially when the focus is on gene-by-gene analysis of differential expression. Because individual gene signatures distinguishing NAFL and NASH may not be stable across patients or studies, we turned to gene network analysis to examine consistent functional module differences between NAFL and NASH by employing WGCNA. WGCNA provides a framework based on the intrinsic transcriptome organization measured in a given study and allows identification of hub genes that play central roles in a specific context. As an alternative to traditional differential expression analyses which centered on most significant genes, WGCNA groups co-expressed genes that are biologically related on a genome-wide scale in an unbiased manner, which may shed light on key therapeutic targets discovery and mechanisms of disease.

Another advantage of WGCNA approach is that it facilitates annotation of disease-related genes by allowing assertion of a biological role for the module. The biological role of a module is based on the genes contained within the module that have known biological functions or

Table 3 miRNA-gene pairs as predicted from database

Module	miRNA:hub gene pairs	miRNA:TF pairs
P1	miR-29c:MT1DP, miR-192:MT1X[†]	
P4	miR-23b:PDIA6, miR-15a:SRPRB*, miR-15a:PDIA6[†], let-7f:PDIA6[†], miR-30e:NANS[‡]	miR-122:MECOM, let-7d:POU2F1, let-7f-1:POU2F1*, let-7g:POU2F1, miR-23b:POU2F1, miR-122:POU2F1
P5	miR-148b:RPL8*	
P7	miR-29c:IGHG1	
P8	miR-101-2:TCF4*, miR-105-2:TCF4*, miR-122:VIM[‡], miR-30e:VIM[‡]	miR-148b:SOX5*, miR-23b:STAT5B
P9	miR-101-2: PTGS2	let-7d:FOXL1, let-7f-1:FOXL1, miR-192:SOAT1
P13	miR-27b:PAN3, miR-105-2:PAN3*, miR-148b:YIF1A[†], miR-30e:YIF1A[†]	miR-29c:YY1
P14	let-7f-1:SRRM2*, let-7d:SRRM2, let-7g:SRRM2*	

The significance of miRNA-gene pairs was calculated by Pearson correlation

* $P < 0.05$

[†] $P < 0.01$

[‡] $P < 0.001$

Table 4 Module membership and Log_2 (NASH/NAFL) ratio of 12 selected hub genes

Symbol	Gene name	Module	kME	P value	Fold change	q value (%)
MT1DP	Metallothionein 1D, pseudogene	P1	0.95	9.29E−14	0.62	0.23
MT1X	Metallothionein 1X	P1	0.89	7.33E−10	0.37	0
PDIA6	Protein disulfide isomerase family A, member 6	P4	0.95	2.76E−14	0.46	0
SRPRB	Signal recognition particle receptor, B subunit	P4	0.93	2.14E−12	0.45	0
NANS	N-Acetylneuraminic acid synthase	P4	0.96	8.96E−15	0.40	0
RPL8	Ribosomal protein L8	P5	0.93	2.86E−12	NA	NA
IGHG1	Immunoglobulin heavy constant gamma 1	P7	0.94	7.02E−13	3.30	0.03
TCF4	Transcription factor 4	P8	0.97	1.34E−16	2.35	0
VIM	Vimentin	P8	0.96	5.46E−16	3.12	0
PTGS2	Prostaglandin-endoperoxide synthase 2	P9	0.95	1.79E−14	2.03	0.18
PAN3	Poly(A) specific ribonuclease subunit homolog	P13	0.88	2.86E−09	1.68	0
YIF1A	Yip1 interacting factor homolog A	P13	−0.89	4.54E−10	0.54	0
SRRM2	Serine/arginine repetitive matrix 2	P14	0.92	9.92E−12	1.35	0.45

Module membership (kME) was measured by correlating its gene expression profile with the module eigengene of a given module. P value was calculated by correlation test. Fold change and q value were calculated by SAM

are associated with a particular functional property. This approach can shed light on the function of novel genes or genetic findings on the basis of guilt-by-association. For example, Kruppel-like factor 6 (KLF6) belongs to the Kruppel-like family of transcription factors known to play diverse roles in differentiation, development, cell growth, apoptosis and angiogenesis. KLF6 expression was increased in association with increased steatosis, inflammation and fibrosis in NAFLD livers (Miele et al. 2008; Ray 2013). A functional polymorphism in the KLF6 (rs3750861) is associated with advanced NAFLD (Miele et al. 2008). In our analysis, KLF6 was up regulated 2.35-fold in NASH compared with NAFL. KLF6 resides in module P4 whose primary function is proteasome protein degradation. Expression of KLF6 (kME = −0.83) is negatively correlated with protein degradation. Interestingly, ubiquitin–proteasome degradation is involved in the development of inflammation through multiple pathways, including MHC-mediated antigen presentation, cytokine and cell cycle regulation and apoptosis (Wang and Maldonado 2006). The association between nonsynonymous gene variants and NAFLD histologic severity has been previously reported (Romeo et al. 2008). Recently identified genetic variants of MSRA (rs17151287), RBMS3 (rs9878305) and GCKR (rs780094), which are members of the module P8, are associated with histologic NAFLD

severity (Speliotes et al. 2011). Functions associated with module P8 include cell adhesion and growth factor binding (see Supplemental Table 4), thereby implicating these pathways in NAFLD. Moreover, genetic variants of PELI2 (rs9323291) and EFCAB4B (rs887304) in module P13 were associated with histologic NAFLD (Speliotes et al. 2011) and lobular inflammation of NAFLD (Chalasani et al. 2010), respectively.

Here, we present for the first time a systematic WGCNA of NAFL and NASH. Several lines of evidence suggest that the networks constructed are biologically significant. First, most of the identified modules are enriched for specific GO terms, KEGG terms or TFBS. Second, hub genes in several modules are key players in NAFLD as reported by published literature. Finally, multiple network modules execute similar functions between NAFL and NASH, including S1 and N5, S2 and N6, S4 and N7, S6 and N13, S9 and N11, S10 and N14, S13 and N8 as shown in Supporting Fig. 9. Besides, we identified nine modules that may play roles in NAFLD progression. These biological processes, including ion homeostasis, ribosome biogenesis, immune response, proteasomal protein degradation, translational elongation, cell adhesion, response to endotoxin, RNA splicing and protein transport, may possibly play in NAFL/NASH transition.

Recently, the NAFLD dataset we used has also been analyzed by other groups. The data was integrated with metabolic profile and network to identify metabolites for NASH diagnosis (Mardinoglu et al. 2014; Lake et al. 2015). These studies showed that non-essential amino acids serine, glycine, glutamate, glutamine, aspartate and alanine, and the essential amino acids valine and methionine seem to be involved in the appearance of NASH. In our analysis, we observed significant modules associated with amino acid metabolism and lipid metabolism in both steatosis (S1, S2, S3) and NASH (N5, N6) network. The result was represented in supplementary Tables 2 and 3. However, only amino acid metabolism was presented in module P4 of the progressive network, which suggests the continuous amino acid metabolic stress during NASH pathogenesis. One of the NASH associated modules, P4, was identified with enriched genes in alanine, aspartate and glutamate metabolism pathway. Module P4 is enriched with genes involved in proteasomal protein degradation, which suggests the potential origin of these up-regulated amino acids. Mardinoglu et al. also showed that blood concentrations of chondroitin and heparan sulfates are suitable for diagnosing NASH. In our analysis, we found that NASH associated module P8 (enriched with extracellular matrix genes) is involved in cell adhesion. The module P8 gene HSPG2 encodes a core protein of chondroitin and heparan sulfates, whose role in cell adhesion and inflammation has been reported (O'Boyle et al. 2009).

On the other hand, miRNAs are important regulators of gene expression and affect mRNA stability and function. Interestingly, 193 probe sets presented on the chip with miRNA gene annotation resulted in 15 differential miRNAs. Many of these miRNAs have been reported to play roles in NAFLD. For example, miR-122, the most abundant miRNA in the liver, was down regulated in NASH and has been shown to be involved in lipid and cholesterol metabolism, which are at the core of fatty liver disease (Cheung et al. 2008). It has been shown that miR-122 is decreased in the liver but increased in circulating forms and serves as potential biomarkers for predicting NAFLD progression (Pirola et al. 2013b). According to our analysis, miR-122 is 6.1-fold down-regulated, which is the most significantly changed miR. According to target gene prediction, miR-122 targets module P4 through several regulatory TFs of this module, and P8 through its hub gene VIM. Two recent studies show that miR-122 indeed can regulate cell adhesion (Tsai et al. 2009) and proteasomal ubiquitin-dependent protein degradation (Yang et al. 2011a) in hepatocellular carcinoma. In down-regulated module P4, two cellular components, endoplasmic reticulum ($P = 1.3E{-}27$) and proteasome complex ($P = 1.4E{-}22$), were significantly enriched, while this module was involved in proteasomal ubiquitin-dependent protein catabolic process ($P = 1.5E{-}23$) and unfolded protein binding ($P = 9.0E{-}9$). These results suggest that P4 may participate in endoplasmic reticulum-associated degradation (ERAD) pathway. Furthermore, miR-122 can target several module P4 regulatory TFs, including MECOM, POU2F1 and CUX1, which can regulate most of genes in this module (Table 3). One of P4 module hub genes is PDIA6, which is regulated by several other miRNAs such as miR-23b, miR-15a and let-7f. In case of module P8, its hub gene TCF4 is a transcription factor and is negatively regulated by both miR-105-2 (Yang et al. 2011b) and miR-101-2. Compared to NAFL, TCF4 overexpressed 2.3-fold in NASH. However, its role in NAFL/NASH transition has not yet been reported.

To demonstrate the usefulness of our modules in the development of efficient NAFLD treatment strategies, another large-scale analysis was performed for drug repositioning. The Connectivity Map (Lamb et al. 2006) microarray data were projected to the NASH network, and module eigengene was calculated for the 6,100 chemical molecules. According to the two most significant modules P4 and P8 in the network, we found the top 2 molecules (sulconazole and doxazosin) for these two modules. Interestingly, we found that the roles of doxazosin not only include treating hypertension but also include preventing hepatic steatosis (Menacho-Marquez et al. 2013), even alleviating insulin resistance (Shen et al. 2009). The drug is widely used to treat hypertension, but its roles in NASH are limited to

animal experiment. However, the validated pathogenetic link between the NAFLD and altered blood pressure status may suggest NAFLD as one of the possible pathogenesis mechanisms of hypertension (Latea et al. 2013).

However, our results are concluded from a limited number of samples, the power of the study should be improved by enrolling more samples and considering related clinic parameters such as BMI and serum enzymes in the future.

Although systems biology has been applied before to establish new gene networks implying even miRNAs in not only NAFLD but also AFLD (Sookoian and Pirola 2013), our comparative gene co-expression network analysis demonstrates that NAFL/NASH transition is associated with interaction between gene network and miRNA. miRNA may affect NAFLD transition through hub genes in modules, as well as transcription factors. The transcription factor TCF4 may be negatively regulated by miR-101-2 and miR-105-2, through which gene expression of cell adhesion module is affected during NAFLD progression. Finally, our approach can provide unique insights into the different regulatory mechanisms associated with NAFLD and will aid in identifying specific transcriptional networks and key genes that may involve in the development and progression of NASH. When validated, these candidate genes can provide molecular targets for diagnosis and therapy.

Conflict of interest Nothing to report.

References

Ahrens M, Ammerpohl O, von Schonfels W, Kolarova J, Bens S, Itzel T, Teufel A, Herrmann A, Brosch M, Hinrichsen H, Erhart W, Egberts J, Sipos B, Schreiber S, Hasler R, Stickel F, Becker T, Krawczak M, Rocken C, Siebert R, Schafmayer C, Hampe J (2013) DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. Cell Metab 18:296–302

Barabasi AL, Gulbahce N, Loscalzo J (2011) Network medicine: a network-based approach to human disease. Nat Rev Genet 12:56–68

Bertola A, Bonnafous S, Anty R, Patouraux S, Saint-Paul MC, Iannelli A, Gugenheim J, Barr J, Mato JM, Le Marchand-Brustel Y, Tran A, Gual P (2010) Hepatic expression patterns of inflammatory and immune response genes associated with obesity and NASH in morbidly obese patients. PLoS One 5:e13577

Chalasani N, Guo X, Loomba R, Goodarzi MO, Haritunians T, Kwon S, Cui J, Taylor KD, Wilson L, Cummings OW, Chen YD, Rotter JI (2010) Genome-wide association study identifies variants associated with histologic features of nonalcoholic fatty liver disease. Gastroenterology 139:1567–1576, 1576 e1561–1566

Cheung O, Puri P, Eicken C, Contos MJ, Mirshahi F, Maher JW, Kellum JM, Min H, Luketic VA, Sanyal AJ (2008) Nonalcoholic

steatohepatitis is associated with altered hepatic MicroRNA expression. Hepatology 48:1810–1820

Cohen JC, Horton JD, Hobbs HH (2011) Human fatty liver disease: old questions and new insights. Science 332:1519–1523

da Huang W, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1–13

Farber CR (2010) Identification of a gene module associated with BMD through the integration of network analysis and genome-wide association data. J Bone Miner Res 25:2359–2367

Fujimoto M, Tsuneyama K, Kainuma M, Sekiya N, Goto H, Takano Y, Terasawa K, Selmi C, Gershwin ME, Shimada Y (2008) Evidence-based efficacy of Kampo formulas in a model of non alcoholic fatty liver. Exp Biol Med (Maywood) 233:328–337

Gawrieh S, Baye TM, Carless M, Wallace J, Komorowski R, Kleiner DE, Andris D, Makladi B, Cole R, Charlton M, Curran J, Dyer TD, Charlesworth J, Wilke R, Blangero J, Kissebah AH, Olivier M (2010) Hepatic gene networks in morbidly obese patients with nonalcoholic fatty liver disease. Obes Surg 20:1698–1709

Greco D, Kotronen A, Westerbacka J, Puig O, Arkkila P, Kiviluoto T, Laitinen S, Kolak M, Fisher RM, Hamsten A, Auvinen P, Yki-Jarvinen H (2008) Gene expression in human NAFLD. Am J Physiol Gastrointest Liver Physiol 294:G1281–G1287

Hebbard L, George J (2011) Animal models of nonalcoholic fatty liver disease. Nat Rev Gastroenterol Hepatol 8:35–44

Horvath S, Erhart W, Brosch M, Ammerpohl O, von Schönfels W, Ahrens M, Heits N, Bell JT, Tsai P-C, Spector TD, Deloukas P, Siebert R, Sipos B, Becker T, Röcken C, Schafmayer C, Hampe J (2014) Obesity accelerates epigenetic aging of human liver. Proc Natl Acad Sci 111:15538–15543

Hu Z, Hung JH, Wang Y, Chang YC, Huang CL, Huyck M, DeLisi C (2009) VisANT 3.5: multi-scale network visualization, analysis and inference based on the gene ontology. Nucleic Acids Res 37:W115–W121

Ivliev AE, t Hoen PA, Sergeeva MG (2010) Coexpression network analysis identifies transcriptional modules related to proastrocytic differentiation and sprouty signaling in glioma. Cancer Res 70:10060–10070

Jeong H, Mason SP, Barabasi AL, Oltvai ZN (2001) Lethality and centrality in protein networks. Nature 411:41–42

Koutsari C, Lazaridis KN (2010) Emerging genes associated with the progression of nonalcoholic fatty liver disease. Hepatology 52:807–810

Lake AD, Novak P, Fisher CD, Jackson JP, Hardwick RN, Billheimer DD, Klimecki WT, Cherrington NJ (2011) Analysis of global and absorption, distribution, metabolism, and elimination gene expression in the progressive stages of human nonalcoholic fatty liver disease. Drug Metab Dispos 39:1954–1960

Lake AD, Novak P, Shipkova P, Aranibar N, Robertson DG, Reily MD, Lehman-McKeeman LD, Vaillancourt RR, Cherrington NJ (2015) Branched chain amino acid metabolism profiles in progressive human nonalcoholic fatty liver disease. Amino Acids 47:603–615

Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES, Golub TR (2006) The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science 313:1929–1935

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinform 9:559

Langfelder P, Zhang B, Horvath S (2008) Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. Bioinformatics 24:719–720

Latea L, Negrea S, Bolboaca S (2013) Primary non-alcoholic fatty liver disease in hypertensive patients. Australas Med J 6:325–330

Lee HK, Hsu AK, Sajdak J, Qin J, Pavlidis P (2004) Coexpression analysis of human genes across many microarray data sets. Genome Res 14:1085–1094

Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 433:769–773

Lopez-Vicario C, Gonzalez-Periz A, Rius B, Moran-Salvador E, Garcia-Alonso V, Lozano JJ, Bataller R, Cofan M, Kang JX, Arroyo V, Claria J, Titos E (2014) Molecular interplay between Delta5/Delta6 desaturases and long-chain fatty acids in the pathogenesis of non-alcoholic steatohepatitis. Gut 63:344–355

Mardinoglu A, Agren R, Kampf C, Asplund A, Uhlen M, Nielsen J (2014) Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. Nature Commun 5:3083

Menacho-Marquez M, Nogueiras R, Fabbiano S, Sauzeau V, Al-Massadi O, Dieguez C, Bustelo XR (2013) Chronic sympathoexcitation through loss of Vav3, a Rac1 activator, results in divergent effects on metabolic syndrome and obesity depending on diet. Cell Metab 18:199–211

Miele L, Beale G, Patman G, Nobili V, Leathart J, Grieco A, Abate M, Friedman SL, Narla G, Bugianesi E, Day CP, Reeves HL (2008) The Kruppel-like factor 6 genotype is associated with fibrosis in nonalcoholic fatty liver disease. Gastroenterology 135(282–291):e281

Miller JA, Oldham MC, Geschwind DH (2008) A systems level analysis of transcriptional changes in Alzheimer's disease and normal aging. J Neurosci 28:1410–1420

Moylan CA, Pang H, Dellinger A, Suzuki A, Garrett ME, Guy CD, Murphy SK, Ashley-Koch AE, Choi SS, Michelotti GA, Hampton DD, Chen Y, Tillmann HL, Hauser MA, Abdelmalek MF, Diehl AM (2014) Hepatic gene expression profiles differentiate presymptomatic patients with mild versus severe nonalcoholic fatty liver disease. Hepatology 59:471–482

Nieminen M, Henttinen T, Merinen M, Marttila-Ichihara F, Eriksson JE, Jalkanen S (2006) Vimentin function in lymphocyte adhesion and transcellular migration. Nat Cell Biol 8:156–162

O'Boyle G, Mellor P, Kirby JA, Ali S (2009) Anti-inflammatory therapy by intravenous delivery of non-heparan sulfate-binding CXCL12. FASEB J 23:3906–3916

Odom DT, Dowell RD, Jacobsen ES, Gordon W, Danford TW, MacIsaac KD, Rolfe PA, Conboy CM, Gifford DK, Fraenkel E (2007) Tissue-specific transcriptional regulation has diverged significantly between human and mouse. Nat Genet 39:730–732

Parekh S, Anania FA (2007) Abnormal lipid and glucose metabolism in obesity: implications for nonalcoholic fatty liver disease. Gastroenterology 132:2191–2207

Pirola CJ, Gianotti TF, Burgueno AL, Rey-Funes M, Loidl CF, Mallardi P, Martino JS, Castano GO, Sookoian S (2013a) Epigenetic modification of liver mitochondrial DNA is associated with histological severity of nonalcoholic fatty liver disease. Gut 62:1356–1363

Pirola CJ, Gianotti TF, Castano GO, Sookoian S (2013b) Circulating MicroRNA-122 signature in nonalcoholic fatty liver disease and cardiovascular disease: a new endocrine system in metabolic syndrome. Hepatology 57:2545–2547

Podrini C, Borghesan M, Greco A, Pazienza V, Mazzoccoli G, Vinciguerra M (2013) Redox homeostasis and epigenetics in non-alcoholic fatty liver disease (NAFLD). Curr Pharm Des 19:2737–2746

Ray K (2013) NASH: KLF6 activates PPARalpha signalling in hepatic steatosis. Nat Rev Gastroenterol Hepatol 10:128

Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, Boerwinkle E, Cohen JC, Hobbs HH (2008) Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 40:1461–1465

Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, Friedman N (2003) Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat Genet 34:166–176

Shackel NA, Gorrell MD, McCaughan GW (2002) Gene array analysis and the liver. Hepatology 36:1313–1325

Shen KP, Lin HL, Hsieh SL, Kwan AL, Chen IJ, Wu BN (2009) Eugenosedin-A prevents hyperglycaemia, hyperlipidaemia and lipid peroxidation in C57BL/6J mice fed a high-fat diet. J Pharm Pharmacol 61:517–525

Sookoian S, Pirola CJ (2012) The genetic epidemiology of nonalcoholic fatty liver disease: toward a personalized medicine. Clin Liver Dis 16:467–485

Sookoian S, Pirola CJ (2013) Systems biology elucidates common pathogenic mechanisms between nonalcoholic and alcoholic-fatty liver disease. PLoS One 8:e58895

Sookoian S, Rosselli MS, Gemma C, Burgueno AL, Fernandez Gianotti T, Castano GO, Pirola CJ (2010) Epigenetic regulation of insulin resistance in nonalcoholic fatty liver disease: impact of liver methylation of the peroxisome proliferator-activated receptor gamma coactivator 1alpha promoter. Hepatology 52:1992–2000

Sookoian S, Gianotti TF, Rosselli MS, Burgueno AL, Castano GO, Pirola CJ (2011) Liver transcriptional profile of atherosclerosis-related genes in human nonalcoholic fatty liver disease. Atherosclerosis 218:378–385

Speliotes EK, Yerges-Armstrong LM, Wu J, Hernaez R, Kim LJ, Palmer CD, Gudnason V, Eiriksdottir G, Garcia ME, Launer LJ, Nalls MA, Clark JM, Mitchell BD, Shuldiner AR, Butler JL, Tomas M, Hoffmann U, Hwang SJ, Massaro JM, O'Donnell CJ, Sahani DV, Salomaa V, Schadt EE, Schwartz SM, Siscovick DS, Voight BF, Carr JJ, Feitosa MF, Harris TB, Fox CS, Smith AV, Kao WH, Hirschhorn JN, Borecki IB (2011) Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. PLoS Genet 7:e1001324

Tsai WC, Hsu PW, Lai TC, Chau GY, Lin CW, Chen CM, Lin CD, Liao YL, Wang JL, Chau YP, Hsu MT, Hsiao M, Huang HD, Tsou AP (2009) MicroRNA-122, a tumor suppressor microRNA that regulates intrahepatic metastasis of hepatocellular carcinoma. Hepatology 49:1571–1582

Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 98:5116–5121

Wang J, Maldonado MA (2006) The ubiquitin-proteasome system and its role in inflammatory and autoimmune diseases. Cell Mol Immunol 3:255–261

Yang F, Zhang L, Wang F, Wang Y, Huo XS, Yin YX, Wang YQ, Sun SH (2011a) Modulation of the unfolded protein response is the core of MicroRNA-122-involved sensitivity to chemotherapy in hepatocellular carcinoma. Neoplasia 13:590–600

Yang JH, Li JH, Shao P, Zhou H, Chen YQ, Qu LH (2011b) starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. Nucleic Acids Res 39:D202–D209

Yoneda M, Endo H, Nozaki Y, Tomimoto A, Fujisawa T, Fujita K, Yoneda K, Takahashi H, Saito S, Iwasaki T, Yamamoto S, Tsutsumi S, Aburatani H, Wada K, Hotta K, Nakajima A (2007) Life style-related diseases of the digestive system: gene expression in nonalcoholic steatohepatitis patients and treatment strategies. J Pharmacol Sci 105:151–156

Yoneda M, Endo H, Mawatari H, Nozaki Y, Fujita K, Akiyama T, Higurashi T, Uchiyama T, Yoneda K, Takahashi H, Kirikoshi H,

Inamori M, Abe Y, Kubota K, Saito S, Kobayashi N, Yamaguchi N, Maeyama S, Yamamoto S, Tsutsumi S, Aburatani H, Wada K, Hotta K, Nakajima A (2008) Gene expression profiling of non-alcoholic steatohepatitis using gene set enrichment analysis. Hepatol Res 38:1204–1212

Younossi ZM, Baranova A, Ziegler K, Del Giacco L, Schlauch K, Born TL, Elariny H, Gorreta F, VanMeter A, Younoszai A, Ong JP, Goodman Z, Chandhoke V (2005a) A genomic and proteomic study of the spectrum of nonalcoholic fatty liver disease. Hepatology 42:665–674

Younossi ZM, Gorreta F, Ong JP, Schlauch K, Del Giacco L, Elariny H, Van Meter A, Younoszai A, Goodman Z, Baranova A, Christensen A, Grant G, Chandhoke V (2005b) Hepatic gene expression in patients with obesity-related non-alcoholic steato-hepatitis. Liver Int 25:760–771

Yu Y, Ping J, Chen H, Jiao L, Zheng S, Han ZG, Hao P, Huang J (2010) A comparative analysis of liver transcriptome suggests divergent liver function among human, mouse and rat. Genomics 96:281–289

Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 4:Article17