Bioinformatics analysis of the NAFLD Interactome: Revealing candidate biomarkers of Non-Alcoholic Fatty Liver Disease

Athina I. Amanatidou☯ and George V. Dedoussis☯

Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, El. Venizelou 70, 17671, Athens, Greece

☯ **Correspondence to:** G. V. Dedoussis, A. I. Amanatidou, Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, El. Venizelou 70, 17671, Athens, Greece

E-mail addresses: dedousi@hua.gr (G.V. Dedoussis), aamanatid@gmail.com (A. I. Amanatidou)

Telephone: +302109549179 (G.V. Dedoussis), +306949293472 (A. I. Amanatidou)

Abstract

Non-alcoholic fatty liver disease (NAFLD) is a disease with multidimensional complexities. Many attempts have been made over the years to treat this disease but its incidence is rising. For this reason, the need to identify and study new candidate NAFLD biomarkers is of utmost importance. Systems-based approaches such as the analysis of protein-protein interaction (PPI) network could lead to the discovery of new disease biomarkers that can then be translated into clinical practice. The aim of this study is to analyze the interaction network of human proteins associated with NAFLD as well as their experimentally verified interactors and to propose new candidate proteins that may be involved in this disease. Computational analysis made it feasible to detect 77 candidate proteins associated with NAFLD, having high network scores. Furthemore, clustering analysis was performed to identify densely connected regions with biological significance in this network. Additionally, gene expression analysis was conducted to validate part of the findings of this research work. We believe that our research will be helpful in extending experimental efforts to address the pathogenesis and progression of NAFLD.

Keywords: Non-alcoholic fatty liver disease; nonalcoholic steatohepatitis; protein-protein interaction (PPI); protein-disease association; bioinformatics; biomarkers

1. Introduction

The liver is a vital digestive organ which performs many essential body's metabolic functions involving metabolism of lipids, bile acids, glucose and cholesterol [1]. Metabolic pathways do not operate independently within the liver; one pathway can heavily affect other pathways. The dysfunctional crosstalk of the hepatic pathways is a widespread health problem, responsible for about 2 million deaths worldwide each year [2]. The most common chronic liver disease worldwide is known as non-alcoholic fatty liver disease (NAFLD). It is an umbrella term which encompasses a spectrum of pathological conditions ranging from simple hepatic steatosis (SS) or non-alcoholic fatty liver (NAFL) to a more severe form nonalcoholic steatohepatitis (NASH), and NASH cirrhosis [3]. Although in the last decade, research advances demonstrate that NAFLD is a multisystem disease in which many complex processes are involved in its manifestation and development. In addition, growing number of studies demonstrates that NAFLD affects a variety of extrahepatic organs and regulatory pathways [4].

With the passage of time, NAFLD's health and socio-economic influence is rising, and the annual health costs in the United States are greater than $103 billion [5]. Henceforth, its timely and precise diagnosis is very significant, considering that its prevalence has rapidly reached global epidemic proportions in both adults and children [6]. Most patients are asymptomatic and the diagnosis of the disease is random in most cases [7].

The medical community has centered on the causes of the disease over the past few decades, and the identification of new diagnostic markers (biomarkers). Nonetheless, the gold standard for NAFLD diagnosis remains the liver biopsy but this procedure is inefficient as a diagnostic tool due to its invasive, expensive and sometimes serious complications [8]. In the foreseeable future, the key to NAFLD diagnosis and treatment could be the "molecular signature" of each NAFLD patient [9].

The data that derived from omics technologies which feed precision medicine have a major contribution to this effort. An increasing number of technical advancements have, to date,

produced a collection of many unused data as a whole. Therefore, it is necessary to move from single omics to multi-omics analysis, providing a broader window of its pathophysiology that scans different perspectives [9]. Network-based approaches integrate omics data such as protein-protein interaction (PPI) networks which are gaining ground in the scientific community as they provide valuable, quick and inexpensive tools for clarifying disease mechanisms and detecting new candidate disease-related proteins (or genes) [10].

Disease is rarely the result of an abnormality in a single gene but represents disruptions in the complex interaction network. Key biological factors that control the pathobiology of the disease are almost always the result of several pathobiological pathways interacting through an interconnected network [11]. Conventional methods which evaluate one gene or factor at a time have become less effective in tackling NAFLD's multidimensional complexities [1]. Given the fact that NAFLD research mostly includes studies on human clinical and animal model trials [9], the analysis of PPI network could be an ally to uncover candidate biomarkers and pathological pathways, as well as potential therapeutic targets, contributing to the development of noninvasive diagnosis.

In the present study, a PPI network analysis was conducted to identify new candidate NAFLD biomarkers through performing topological analyses. Besides, clustering analysis of the PPI network was achieved to identify densely connected regions. In order to reveal insights into the molecular mechanisms of the network's proteins, an enrichment analysis was performed. Moreover, an analysis of gene expression microarray data set was achieved to detect differential expressed genes (DEGs) between NAFLD samples and controls, as well as a pathway analysis of DEGs.

2. Methods

The research methodology used in this study includes the stages stated below. **Fig. 1** outlines the basic steps involved in the methodology.

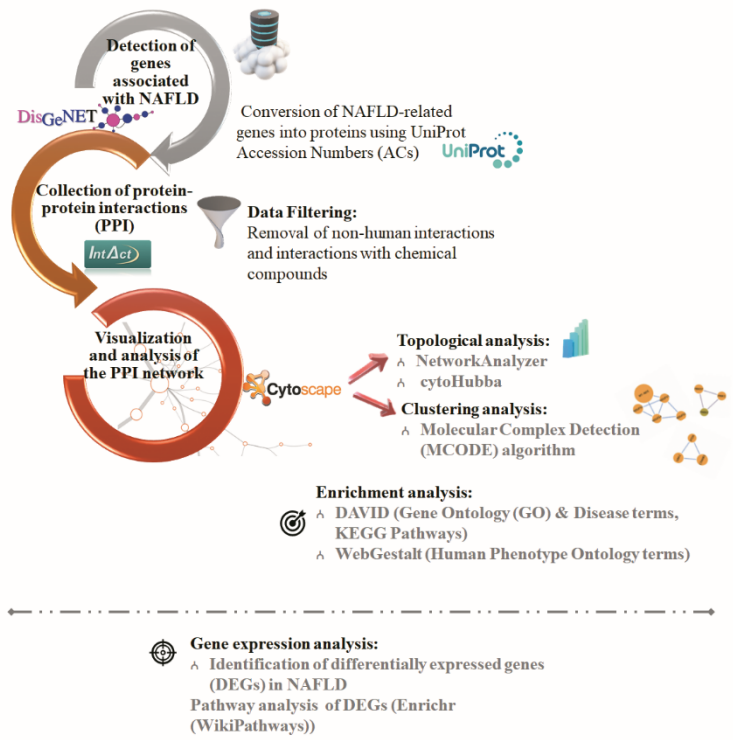

Fig. 1: The schematic diagram of the research methodology.

2.1 Detection of genes associated with NAFLD

NAFLD and its subtype NASH have been queried using "Non-alcoholic Fatty Liver Disease" and "NASH - Nonalcoholic steatohepatitis" terms in a DisGeNET search panel which is a discovery platform containing one of the largest collections of genes and variants associated with human diseases [12]. All the NAFLD-related genes are either genetic associations or under/over expressed in the gene transcription levels or are present at low/high protein levels in patient's plasma/serum. Eventually, the disease-related genes were manually confirmed for their association with NAFLD.

2.2 Collection of protein-protein interactions (PPI)

The NAFLD-related genes were then converted to proteins using UniProt Accession Numbers (ACs) via UniProt database [13]. A query was then conducted in IntAct [14], a molecular interaction database with highly curated data, using the ACs of the proteins, to retrieve all experimentally confirmed interactions of these proteins and their first neighbors. Interaction data were obtained in a MI-TAB 2.7 format file [15] in which any non-human interactions and interactions with chemical compounds were removed.

2.3 Visualization and analysis of the PPI network

Cytoscape (version) 3.7.2 software, a popular open source bioinformatics platform for the data integration and network analysis [16], was used to visualize and analyze the PPI network. In this network, every node corresponds to a protein and the edges represent interactions, where the latter were treated as undirected for this analysis. Additionally, browser-based web application was generated to visualize interactive networks via the CyNetShare tool (http://idekerlab.github.io/cy-net-share/). Links are provided in the legends of the respective figures.

Afterwards a **topological analysis** was conducted using the NetworkAnalyzer [17], a handy Cytoscape plugin, to estimate simple and complex topology parameters. The three important metrics – degree, betweenness and closeness centrality – were utilized to evaluate the importance of nodes in a network [10, 18]. *Hub* proteins were identified by their very high degree of connectivity. Proteins with high betweenness centrality, namely *bottlenecks*, are key connectors in the PPI network, controlling the flow of information within a network [19]. For the identification of proteins - from which the flow of information passes faster to other network's proteins - are those with high closeness centrality, hereby referred to as *PHC (proteins with high closeness centrality)* [10]. The top scoring proteins corresponding to about the 5% of the network's proteins were then selected for each of the three aforementioned network centralities. A Venn diagram was subsequently applied to identify **candidate**

NAFLD-related proteins that were on the three high scoring protein lists but did not belong to the list of the NAFLD-related proteins.

Given the heterogeneous nature behind biological networks, it is advisable to use more than one approach to capture essential proteins. Therefore, a newly proposed method Maximal Clique Centrality (MCC) was estimated using the *cytoHybba* software [20], that has been proven for its great performance in predicting important proteins from the PPI network. The 10 top ranked proteins based on MCC algorithm were also identified as **candidate NAFLD-related proteins.**

Subsequently, Molecular Complex Detection (MCODE) algorithm was utilized to perform a **clustering analysis [21]**. The selection parameters were set as follows: MCODE scores>5, degree cut-off=2, node-score cut-off=0.2 and k-core=2.

Afterwards, an **enrichment analysis** was performed with the use of two bioinformatics tools, DAVID [22] and WebGestalt [23]. DAVID was used for functional enrichment analysis, disease association as well as pathway analysis and WebGestalt was utilized for human phenotype ontology (HPO) analysis. **Functional enrichment analysis** was applied to detect statistically significant overrepresented Gene Ontology (GO) [24] terms in the network. **Disease association analysis** was used to uncover the association of network's proteins with disease terms from Gene Association Database (GAD) [25]. **Pathway analysis** was applied to detect the KEGG pathways from KEGG PATHWAY Database [26] and **HPO analysis** [27] used to detect the phenotype of network proteins'. P-value<0.05 was defined as statistical significance.

2.4 Gene expression data and pathway analyses of candidate NAFLD-related proteins

To detect differentially expressed genes (DEGs) in NAFLD compared to normal condition, the human gene expression data set GSE151158 [28] was downloaded from the Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/) [29], including 21 control liver samples, 40 NAFLD samples – 23 of which have NAFLD Activity Score (NAS) ≤ 3 and 17 have NAS

≥ 5. The analysis was performed through GEO2R [30] tool which applies *limma* (Linear Models for Microarray Analysis) [31] and *GEOquery* [32] R packages from the Bioconductor project. The data were log-transformed, and P-values were adjusted based on the Benjamini & Hochberg (False discovery rate, FDR) method for multiple testing. The significantly DEGs were defined with an adjusted P-value<0.05 and were then subjected to discover whether it contains any of the candidate NAFLD-related proteins resulting from the topological network analysis. The list of significantly DEGs were further analyzed against the WikiPathways [33] database by using the Enrichr [34] tool. P-value cutoff of 0.05 was selected to identify significantly enriched terms as well.

3. Results

3.1 Construction and analysis of NAFLD Interactome

The data set of NAFLD-related proteins is comprised of 254 proteins (**Supplementary Table 1**). They were then inserted into IntAct to collect their PPI, 226 of which have stored PPI data (**Supplementary Table 2**). Subsequently, the collected PPI data (**Supplementary Table 3**) were imported into Cytoscape 3.7.2 to construct a PPI network, refer to as '**NAFLD Interactome'**, comprising of 2624 proteins (nodes) and 20259 interactions (edges) (**Fig. 2**).

After conducting a **topological analysis** with the utilization of NetworkAnalyzer in NAFLD Interactome, important information regarding the network's topology and the biological value of its proteins was revealed. The network's density (show how sparse/dense is a network) is estimated as 0.006, a value lower than 0.1, which denotes that the NAFLD Interactome is a sparsely connected network, as other biological networks [35]. The clustering coefficient, the propensity of the network to grouped into clusters, is measured as 0.110 and the characteristic path length (CPL) [36] is 3.285.

The node degree distribution *P(k)* [37], follows the power-law $\mathrm{P(k)} = Ak^{-\gamma}$, where A is constant and γ is the degree exponent. In our case, the distribution is of the following form:

$$\mathrm{P(k)} = 2485.86k^{-1.597} \quad (1)$$

PPI networks are scale-free and its main feature is that they follow the power law node degree distribution [38]. Since this network also follows the power law distribution; it is characterized by a small number of highly connected proteins, while the majority of the other proteins have few interactions with others [37].

To quantify the importance of network's proteins, metrics for the degree, betweenness and closeness centrality were applied for all NAFLD interactome's proteins. Specifically, the proteins were ranked based on the three afore mentioned centrality measures and then the top 5% of the network's proteins with the highest values were chosen. Considering the overlapping

proteins among the protein lists of each network centrality, a total of 208 proteins were finally selected (**Supplementary Table 4**). Particularly, in the NAFLD Interactome, 25 proteins are *hubs* (**Fig. 2**, triangles), 22 proteins are *bottlenecks* (**Fig. 2**, rectangles), 17 proteins are *hubs* and *bottlenecks* (**Fig. 2**, diamonds), 40 proteins are *PHCs* (**Fig. 2**, V-shaped nodes), 11 proteins are *hubs* and *PHCs* (**Fig. 2**, hexagons), 14 proteins are *bottlenecks* and *PHCs* (**Fig. 2**, octagons), and 79 proteins are *hubs*, *bottlenecks* and *PHCs* (**Fig. 2**, parallelograms). It is noteworthy that 30 NAFLD-related proteins play an essential role in the NAFLD Interactome.

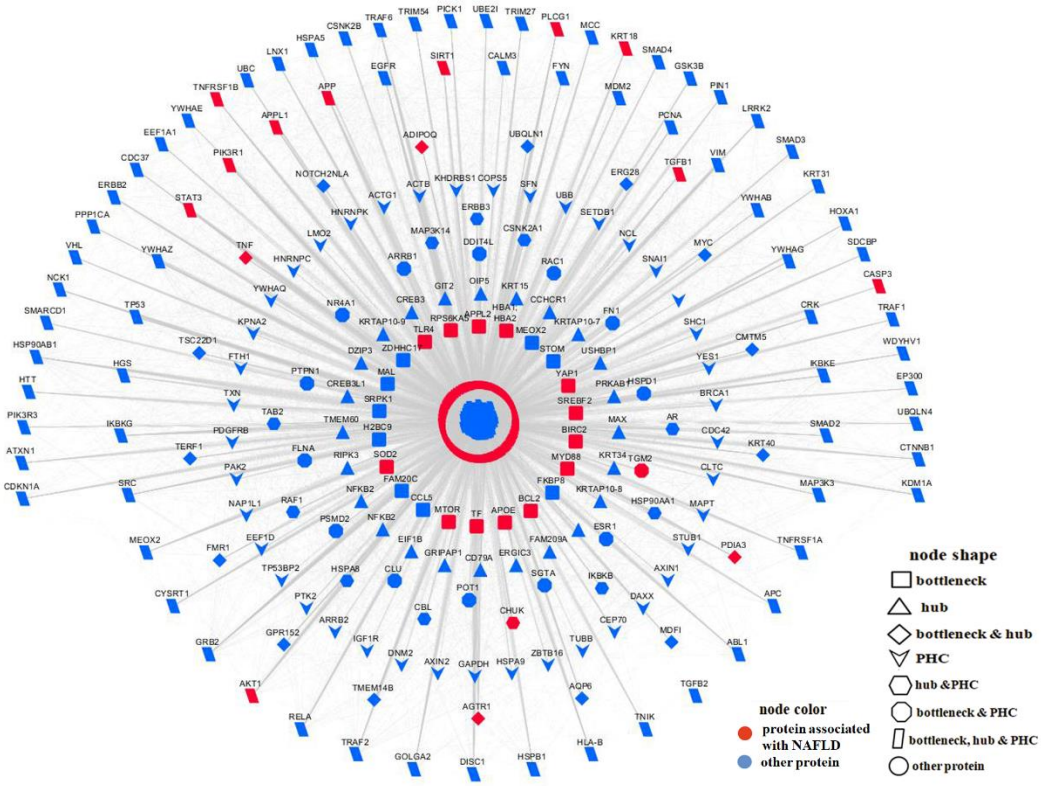

Fig. 2: The NAFLD Interactome. A web visualization of this network is available at /NAFLDInteractome.

The **enrichment analysis** in NAFLD Interactome (2624 proteins) was performed to uncover the role of the network's proteins (more details are given in **Supplementary Tables 5-8**). Among of the most statistically significant over-represented **GO terms** are the following: negative (GO:0043066) (P-value: 5.22E-37) and positive regulation of apoptotic process (GO:0043065) (P-value: 3.78E-34), positive regulation of transcription from RNA polymerase

II promoter (GO:0045944) (P-value: 4.19E-34) and inflammatory response (GO:0006954) (P-value: 9.55E-28).

The **KEGG pathways terms** in which most proteins were found to be involved are pathways in cancer (hsa05200) (P-value: 4.11E-41), PI3K-Akt signaling pathway (hsa04151) (P-value: 1.15E-25), proteoglycans in cancer (hsa05205) (P-value: 1.11E-29), MAPK signaling pathway (hsa04010) (P-value: 5.67E-18) and focal adhesion (hsa04510) (P-value: 3.89E-24). The **disease association analysis** shows that type 2 diabetes (P-value: 1.91E-52), chronic kidney failure (P-value: 3.90E-38), Alzheimer's disease (P-value: 9.41E-23), lung (P-value: 5.84E-53), bladder (P-value: 1.12E-48) and breast (P-value: 9.83E-54) cancer, as well as multiple sclerosis (P-value: 8.62E-27) and schizophrenia (P-value: 4.56E-17) are among of the numerous identified disease terms. Moreover, several phenotypic abnormalities were identified from **HPO analysis** including abnormality of the digestive system (HP: 0025031) (P-value: 5.23E-08), metabolism/homeostasis (HP: 0001939) (P-value: 2.26E-07), cardiovascular system (HP: 0001626) (P-value: 2.53E-04), skin morphology (HP: 0011121) (P-value: 1.96E-07) and immune system (HP: 0002715) (P-value: 6.89E-10).

Two different approaches were applied to identify candidate NAFLD-related proteins, as previously described in the Methods section. In the first approach, in order to find which proteins are present in the list of 79 high scoring proteins (hubs, bottlenecks and PHCs) and already associated with NAFLD, the list of high scoring proteins was combined with the list of 226 NAFLD-related proteins using Venn diagram. Thusly, 68 proteins were recognized as belonging only to the list of high scoring proteins, called candidate NAFLD-related proteins (**Table 1a)**. In the second approach, the 10 top-ranked proteins were found applying MCC algorithm, which are given in **Table 1b**. While CLOCK belongs to the list of 226 NAFLD-related proteins, the remaining 9 proteins were identified as candidate NAFLD-related proteins.

Table 1a: Identification of candidate NAFLD-related proteins. The column "Centrality measures" shows the proteins' ranking in Degree-D, Betweenness-B and Closeness-C network centrality measures. The rank of each protein is given inside the parenthesis of the corresponding centrality measure in the top 140 rankings (approximately the top 5% of the network's proteins).

UniProt AC	Gene	Protein name	Centrality measures (Ranking)
P62993	GRB2	Growth factor receptor-bound protein 2	D (1), B(2), C(3)
P00533	EGFR	Epidermal growth factor receptor	D (3), B(4), C(1)
P63104	YWHAZ	14-3-3 protein zeta/delta	D (5), B(7), C(2)
Q9Y4K3	TRAF6	TNF receptor-associated factor 6	D (8), B(12), C(12)
Q9NRI5	DISC1	Disrupted in schizophrenia 1 protein	D (9), B(8), C(13)
P08238	HSP90AB1	Heat shock protein HSP 90-beta	D (10), B(9), C(5)
Q04206	RELA	Transcription factor p65	D (12), B(18), C(15)
Q9Y6K9	IKBKG	NF-kappa-B essential modulator	D (11), B(15), C(9)
P04637	TP53	Cellular tumor antigen p53	D (13), B(11), C(7)
P16333	NCK1	Cytoplasmic protein NCK1	D (14), B(69), C(134)
P06241	FYN	Tyrosine-protein kinase Fyn	D (15), B(38), C(35)
P12931	SRC	Proto-oncogene tyrosine-protein kinase Src	D (16), B(25), C(14)
P46108	CRK	Adapter molecule crk	D (18), B(40), C(42)
Q14164	IKBKE	Inhibitor of nuclear factor kappa-B kinase subunit epsilon	D (17), B(17), C(17)
Q12933	TRAF2	TNF receptor-associated factor 2	D (20), B(21), C(8)
P04626	ERBB2	Receptor tyrosine-protein kinase erbB-2	D (21), B(20), C(10)
Q08379	GOLGA2	Golgin subfamily A member 2	D (23), B(42), C(33)
A8MQ03	CYSRT1	Cysteine-rich tail protein 1	D (24), B(65), C(102)
Q8TBB1	LNX1	E3 ubiquitin-protein ligase LNX	D (25), B(29), C(18)
O60341	KDM1A	Lysine-specific histone demethylase 1A	D (28), B(32), C(41)
P00519	ABL1	Tyrosine-protein kinase ABL1	D (26), B(58), C(28)
Q6FHY5	MEOX2	MEOX2 protein	D (29), B(13), C(45)
Q99759	MAP3K3	Mitogen-activated protein kinase 3	D (27), B(67), C(38)
P01889	HLA-B	HLA class I histocompatibility antigen, B alpha chain	D (30), B(43), C(78)
Q96HA8	WDYHV1	Protein N-terminal glutamine amidohydrolase	D (31), B(23), C(46)
Q5S007	LRRK2	Leucine-rich repeat serine/threonine-protein kinase 2	D (32), B(22), C(36)
P12004	PCNA	Proliferating cell nuclear antigen	D (34), B(19), C(32)
P35222	CTNNB1	Catenin beta-1	D (35), B(36), C(31)
P61981	YWHAG	14-3-3 protein gamma	D (36), B(46), C(24)

P38936	CDKN1A	Cyclin-dependent kinase inhibitor 1	D (38), B(28), C(25)
Q16543	CDC37	Hsp90 co-chaperone Cdc37	D (39), B(33), C(11)
P08670	VIM	Vimentin	D (43), B(31), C(23)
P19438	TNFRSF1A	Tumor necrosis factor receptor superfamily member 1A	D (42), B(75), C(93)
P23508	MCC	Colorectal mutant cancer protein	D (44), B(48), C(39)
P0CG48	UBC	Polyubiquitin-C	D (47), B(53), C(20)
P49639	HOXA1	Homeobox protein Hox-A1	D (49), B(97), C(75)
Q15323	KRT31	Keratin, type I cuticular Ha1	D (48), B(68), C(90)
Q00987	MDM2	E3 ubiquitin-protein ligase Mdm2	D (51), B(72), C(27)
Q13526	PIN1	Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1	D (50), B(39), C(19)
Q13077	TRAF1	TNF receptor-associated factor 1	D (53), B(95), C(53)
P04792	HSPB1	Heat shock protein beta-1	D (55), B(49), C(29)
P14373	TRIM27	Zinc finger protein RFP	D (58), B(74), C(48)
Q9BYV2	TRIM54	Tripartite motif-containing protein 54	D (57), B(59), C(83)
O00560	SDCBP	Syntenin-1	D (60), B(47), C(71)
P42858	HTT	Huntingtin	D (59), B(73), C(58)
P84022	SMAD3	Mothers against decapentaplegic homolog 3	D (61), B(34), C(30)
P63279	UBE2I	SUMO-conjugating enzyme UBC9	D (62), B(63), C(51)
P54253	ATXN1	Ataxin-1	D (64), B(45), C(54)
P31946	YWHAB	14-3-3 protein beta/alpha	D (67), B(111), C(49)
Q15796	SMAD2	Mothers against decapentaplegic homolog 2	D (66), B(70), C(70)
P40337	VHL	von Hippel-Lindau disease tumor suppressor	D (69), B(114), C(141)
P49841	GSK3B	Glycogen synthase kinase-3 beta	D (70), B(57), C(37)
Q9NRD5	PICK1	PRKCA-binding protein	D (77), B(82), C(88)
P0DP25	CALM3	Calmodulin-3	D (84), B(66), C(40)
P25054	APC	Adenomatous polyposis coli protein	D (82), B(90), C(85)
Q09472	EP300	Histone acetyltransferase p300	D (83), B(54), C(26)
Q9UKE5	TNIK	TRAF2 and NCK-interacting protein kinase	D (81), B(115), C(73)
P67870	CSNK2B	Casein kinase II subunit beta	D (92), B(77), C(80)
O14964	HGS	Hepatocyte growth factor-regulated tyrosine kinase substrate	D (97), B(125), C(92)
P62136	PPP1CA	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit	D (94), B(56), C(100)
Q13485	SMAD4	Mothers against decapentaplegic homolog 4	D (103), B(106), C(126)
Q92569	PIK3R3	Phosphatidylinositol 3-kinase regulatory subunit gamma	D (100), B(121), C(115)
P11021	HSPA5	Endoplasmic reticulum chaperone BiP	D (106), B(79), C(34)
P68104	EEF1A1	Elongation factor 1-alpha 1	D (111), B(86), C(124)
P62258	YWHAE	14-3-3 protein epsilon	D (123), B(132), C(98)

Q96GM5	SMARCD1	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 1	D (122), B(119), C(62)
Q9NRR5	UBQLN4	Ubiquilin-4	D (125), B(64), C(122)
intact:EBI-4399559	-	-	D (45), B(27), C(22)

Table 1b: Identification of candidate NAFLD-related proteins. The 10 top-ranked proteins based on MCC method in NAFLD Interactome. CLOCK protein, highlighted in bold, is already in the list of NAFLD-related proteins.

UniProt AC	Gene	Protein name
O15516	**CLOCK**	**Circadian locomoter output cycles protein kaput**
Q9UKL0	RCOR1	REST corepressor 1
Q9NNX1	TUFT1	Tuftelin
Q96BD5	PHF21A	PHD finger protein 21A
O43482	OIP5	Opa-interacting protein 5
Q86Y13	DZIP3	E3 ubiquitin-protein ligase DZIP3
Q9NP66	HMG20A	High mobility group protein 20A
O95619	YEATS4	YEATS domain-containing protein 4
Q96JG6	VPS50	Syndetin
Q567U6	CCDC93	Coiled-coil domain-containing protein 93

The results of the **enrichment analysis** of candidate NAFLD-related proteins are shown in **Supplementary Table 9**.

3.2 Clustering and enrichment analysis

Clustering analysis. The base of this study is the NAFLD Interactome, a large interconnected network with interactive embedded subnetworks. Hence, with a valuable applying of clustering analysis via MCODE algorithm, the detection of 6 clusters with MCODE score>5 was achieved (**Fig. 3**). The first cluster (MCODE score=29.655) consists of 30 proteins, including 1 NAFLD-related protein: CLOCK (**Fig. 3, 1st Cluster-red node**). It is of utmost importance for our analysis to note that 9 of which are candidate NAFLD-related proteins: **RCOR1, TUFT1, PHF21A, OIP5, DZIP3, HMG20A, YEATS4, VPS50 and CCDC93** (**Fig. 3, 1st Cluster-magenta nodes**). Also, the second cluster (MCODE score=15.412) integrates 18 proteins 2 of which are candidate NAFLD-related proteins: **HOXA1 and CYSRT1** (**Fig. 3, 2nd Cluster-magenta nodes**).

Subsequently, the third (MCODE score=12.500) and fourth (MCODE score=10.273) cluster comprise of 13 and 23 proteins, respectively, containing 1 NAFLD-related protein: PDIA3 (**Fig. 3, 3rd Cluster-red node**) and 2 NAFLD-related proteins: CXCL10 and PF4 (**Fig. 3, 4th Cluster-red nodes**), correspondingly. The fifth cluster (MCODE score=6.200) integrates 11 proteins, 2 of which are NAFLD-related proteins: CHUK and PLCG1 (**Fig. 3, 5th Cluster-red nodes**) and 3 are candidate NAFLD-related proteins: **RELA, IKBKG and EGFR** (**Fig. 3, 5th Cluster-magenta nodes**). Finally, the sixth cluster (MCODE score=5.125) encompasses 17 proteins, involving 2 NAFLD-related proteins: LUM and TGFB1 (**Fig. 3, 6th Cluster-red nodes**) and 3 candidate NAFLD-related proteins: **MEOX2, LNX1 and PIN1** (**Fig. 3, 6th Cluster-magenta nodes**).

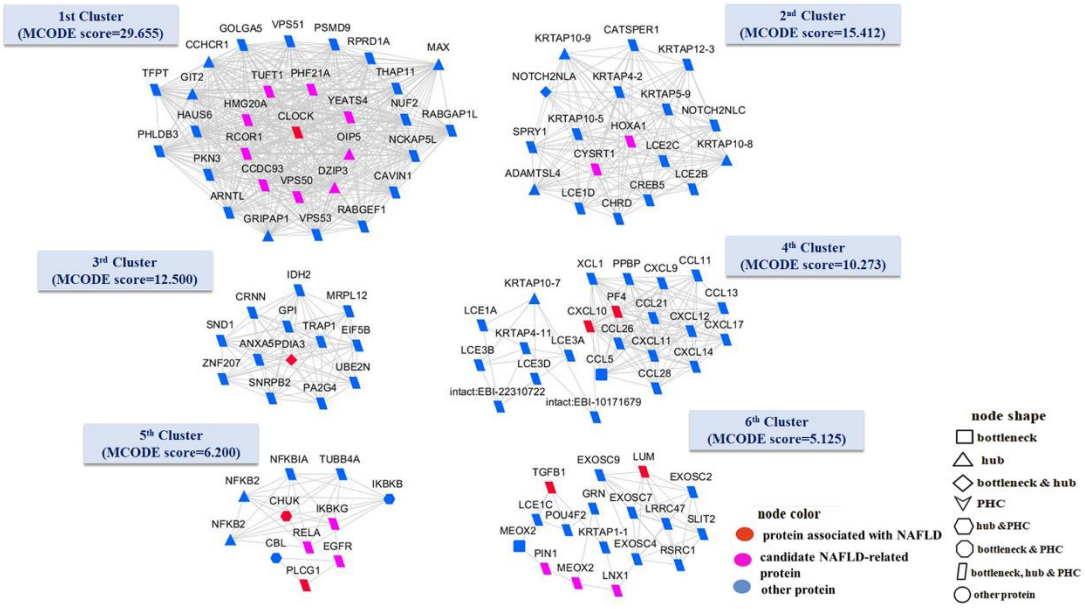

Fig. 3: Clustering analysis of the NAFLD Interactome. A web visualization of this network is available at /ClusteringAnalysisNAFLDInteractome.

Functional enrichment analysis. GO terms were detected for each cluster. Specifically, BP terms could be extracted for the 1st, 2nd, 4th, 5th and 6th clusters (**Supplementary Table 10**), while the MF and CC terms are identified for all clusters (**Supplementary Table 11-12**).

Pathway analysis. The pathway analysis brings to light information regarding the common pathways in which each cluster's proteins partake. Results were detected for all clusters except for the 2nd cluster. Circadian rhythm (hsa04710) (P-value: 0.0223) was found present in the 1st

cluster. Chemokine signaling pathway (hsa04062) (P-value: 2.72E-21) and cytokine-cytokine receptor interaction (hsa04060) (P-value: 9.73E-20) dominated in the 4th cluster. Moreover, the majority of 5th cluster's proteins were found to be involved in epithelial cell signaling in Helicobacter pylori infection (hsa05120) (P-value: 5.58E-11) and NF-kappa B signaling pathway (hsa04064) (P-value: 2.80E-10). Finally, only RNA degradation (hsa03018) (P-value: 4.57E-05) was detected in 6th cluster. No results were returned for the 2nd cluster. More details of pathway analysis are given in **Supplementary Table 13**.

Disease association analysis. Statistically significant disease terms were retrieved for each cluster, although no results were detected for the 2nd cluster (**Supplementary Table 14**). Interestingly, depression (P-value: 0.0193) and sleep disorders (P-value: 0.0368) are associated with the 1st cluster's proteins. Acquired immunodeficiency syndrome (P-value: 0.0147) is the only statistically significant term of the 3rd cluster and respiratory syncytial virus bronchiolitis (P-value: 3.77E-11) is highly related to the 4th cluster's proteins. Also, rheumatoid arthritis (P-value: 1.86E-09) and benzene haematotoxicity (P-value: 3.82E-07) are among the highly statistical terms associated with proteins of the 5th cluster. Lastly, vesico-ureteral reflux (P-value: 0.0055) was found to be the most statistically significant term of the 6th cluster's proteins.

HPO analysis. Phenotypic abnormality terms are detected for all clusters apart from 4th cluster. Please refer to **Supplementary Table 15** for more details.

3.3 Gene expression data and pathway analyses of candidate NAFLD-related proteins

Identification of DEGs. A gene expression analysis was performed to detect DEGs that were differentially expressed between 23 NAFLD-NAS ≤ 3 samples and 21 controls (NAFLD-NAS ≤ 3 vs. Controls), between 17 NAFLD-NAS ≥ 5 samples and 21 controls (NAFLD-NAS ≥ 5 vs. Controls), and between 40 NAFLD samples and 21 controls (NAFLD-all vs. Controls). A total of 55 DEGs, 249 DEGs and 223 DEGs were identified between NAFLD-NAS ≤ 3 vs.

Controls, NAFLD-NAS ≥ 5 vs. Controls and NAFLD-all vs. Controls, respectively. In accordance with our results, **TRAF1, HLA-B, IKBKE and SRC** are the genes that previously were identified as candidate NAFLD-related proteins and were also found as differentially expressed between NAFLD-NAS ≤ 3, NAFLD-NAS ≥ 5, NAFLD-all and Controls. Likewise, **TRAF2, CDKN1A and TP53** were found common between NAFLD-NAS ≥ 5, NAFLD-all and Controls. Please refer to the **Supplementary Table 17** for further details.

Pathway analysis of DEGs. In NAFLD-NAS ≤ 3 vs. Controls, NAFLD-NAS ≥ 5 vs. Controls and NAFLD-all vs. Controls contrast groups, DEGs were significantly enriched in 93, 186 and 185 pathways, respectively (**Supplementary Tables 18 A-C**). The top 10 enriched pathways of DEGs that were most statistically significant between NAFLD-NAS ≤ 3, NAFLD-NAS ≥ 5 and Controls are shown in **Table 2**. Interestingly, **IKBKE** is involved in several pathways such as regulation of toll-like receptor signaling pathway and RIG-I-like Receptor Signaling; **SRC** is implicated in Fibrin Complement Receptor 3 Signaling Pathway and Viral Acute Myocarditis; **HLA-B** is enriched in Allograft Rejection and Type II interferon signaling; **TRAF1, TRAF2 and TP53** are associated with apoptosis; **CDKN1A, SRC and TP53** are implicated in Senescence and Autophagy in Cancer.

Table 2: The top 10 most significantly enriched pathways of DEGs between NAFLD-NAS ≤ 3, NAFLD-NAS ≥ 5 and Controls. The genes that previously identified as candidate NAFLD-related proteins are highlighted in bold.

Term	P-value	Count	Genes
NAFLD-NAS ≤ 3			
Regulation of toll-like receptor signaling pathway (WP1449)	4.25E-12	10	CXCL10, CXCL9, CASP8, SYK, IRF7, SPP1, LY96, CD14, TNF, **IKBKE**
Fibrin Complement Receptor 3 Signaling Pathway (WP4136)	7.73E-12	7	CXCL10, SYK, **SRC,** ITGB2, LY96, CD14, TNF
Toll-like Receptor Signaling Pathway (WP75)	9.32E-12	9	CXCL10, CXCL9, CASP8, IRF7, SPP1, LY96, CD14, TNF, **IKBKE**
Apoptosis (WP254)	7.14E-11	8	CASP8, CASP3, CASP1, IRF7, BAX, FAS, **TRAF1,** TNF
Viral Acute Myocarditis (WP4298)	7.14E-11	8	CASP8, **SRC,** CASP3, ITGB2, CASP1, BAX, NOD2, TNF
Allograft Rejection (WP2328)	1.59E-07	6	CXCL9, CASP8, CASP3, **HLA-B,** FAS, TNF
Nanomaterial induced apoptosis (WP2507)	2.40E-07	4	CASP8, CASP3, FAS, BAX
RIG-I-like Receptor Signaling (WP3865)	6.36E-07	5	CXCL10, CASP8, IRF7, TNF, **IKBKE**
Type II interferon signaling (IFNG) (WP619)	3.16E-06	4	CXCL10, CXCL9, **HLA-B,** PSMB9
Amyotrophic lateral sclerosis (ALS) (WP2447)	3.52E-06	4	CASP3, CASP1, BAX, TNF
NAFLD-NAS ≥ 5			
Allograft Rejection (WP2328)	1.05E-38	32	CD86, CXCL9, ABCB1, CD80, PRF1, CXCL13, HLA-DMB, **HLA-B**…
Regulation of toll-like receptor signaling pathway (WP1449)	4.69E-26	28	CD86, CXCL9, CD80, LY96, TNFAIP3, TNF, CASP8, CCL5, CCL4, **IKBKE**…
Viral Acute Myocarditis (WP4298)	6.47E-25	23	TGFB1, **SRC**, STAT1, CD80, ITGB2, CXCR4, NOD2…
Toll-like Receptor Signaling Pathway (WP75)	4.24E-24	24	CD86, CXCL9, STAT1, CD80, LY96, TNF, **IKBKE**, TLR3…
Ebola Virus Pathway on Host (WP4217)	4.95E-19	22	**HLA-B**, ICAM3, HLA-C, HLA-A, NFKB2, HLA-DMA, HLA-DMB, IRF7, HLA-DPB1, **IKBKE**…
Chemokine signaling pathway (WP3929)	1.02E-16	22	CCR1, CX3CR1, CXCL9, CCL22, CCL20, STAT1…
Human Complement System (WP2806)	3.49E-15	17	SELPLG, C1R, ITGB2, PLAUR, C8A, C2, C5…
Apoptosis (WP254)	5.80E-15	16	**TRAF2**, **TRAF1**, TNF, CASP8, CASP10, **TP53**…
Senescence and Autophagy in Cancer (WP615)	1.39E-14	17	**CDKN1A**, TGFB1, **SRC**, ATG10, IFI16, IL1B, **TP53**…
T-Cell antigen Receptor (TCR) Signaling Pathway (WP69)	1.82E-14	16	MAP4K1, CD83, TGFB1, PRKCD, NFATC1…

Discussion

PPI networks are widely accepted for their valuable contribution to the identification of candidate disease-related proteins in several diseases such as hepatocellular carcinoma, blood-cell targeting autoimmune diseases, breast cancer, etc [10, 39, 40]. In the present study, a topological analysis of the NAFLD Interactome was conducted by applying two different approaches (as presented throughout the Methods section), thusly a **total of 77 candidate NAFLD-related proteins were identified**. Surprisingly, about 50% of these proteins are previously verified in human and animal studies, as well as in other bioinformatics studies regarding their implication in NAFLD and in liver-related manifestations. The validation of our results through literature, which are described bellow, shows that the approach followed in this study is effective in identifying candidate NAFLD-related proteins. Therefore, the remaining unconfirmed proteins should be further investigated for their possible association with NAFLD.

The findings of our literature survey confirmed the implication of the following: **HSP90AB1** has been suggested as a possible biomarker in overweight and obese children with NAFLD [41]; **HLA-B** [42], **CTNNB1** [43] and **HSPA5** [44] are found to be abnormally expressed in NAFLD patients; **CDKN1A** polymorphism is associated with the development of human NAFLD [45]; **TRAF1** has been also detected in NAFLD patients [46]; **HSPB1** phosphorylation site has been differed between NAFLD cohorts [47]; **SMAD4** was overexpresed in NASH patients [48]; **SMAD2/3** phosphorylation and nuclear translocation documented in the liver of NASH patients[49]; **RELA** is well-known to cause inflammatory responses in NAFLD [50]; **PIK3R3** has been proposed as an effective candidate target for the development of NAFLD [51]; **GSK3B** inhibition has been proposed as a possible therapeutic target to manipulate the NAFLD [52].

Remarkably, our findings are in aggreement with previous animal studies as mentioned below: **EGFR** inhibition has been proved to attenuate NAFLD in obese mice model, playing an

essential role in NAFLD as a possible therapeutic target [53]; **TP53** inhibition in a NAFLD mice model resulting in decreased steatosis and liver injury [54]; **PIN1** was essentially involved in NASH development in a rodent model [55]; **SMAD3** overexpression was identified in the liver of monkeys with simple steatosis (SS) and fibrosing NASH [56]; **KDM1A** elevated expression was found in NASH-related hepatocarcinogenesis in a mice model [57]; **EEF1A1** inhibition has been shown to reduce lipotoxicity in obese mice with NAFLD [58]; **TNFRSF1A** has been identified as a potentially effective target factor to prevent the attenuation of SS progression to a more complex phenotype with many NASH features in a mice model [59]; **IKBKE** has been found to specifically expressed in hepatic stellate cells (HSCs) in which inhibition by amlexanox in a NAFLD mice model resulted in improved insulin signal pathway in hepatocytes [60]; **FYN** is implicated in fatty acid oxidation and hepatic steatosis development under chronic ethanol intake in mice model [61]; the increased expression of **VIM** has been found during hepatic steatosis development to NASH in mice, suggesting it as a valuable prognostic factor of liver disease severity [62]; **VIM** and **MAP3K3** were identified upregulated by decreased liver miR-122, possibly contributing in NASH-induced hepatic fibrosis in mice [63]; **ABL1** is implicated in axis which regulates a murine hepatic steatosis, serving as candidate anti-steatosis target [64]; **EP300** inhibition could be effective in hepatic steatosis in mice [65].

In light of the literature review, our results seem to be promising regarding their possible implication in NAFLD development and progression. Recently, **YWHAZ** has been defined as a new regulator of several genes which are dysregulated in NAFLD development [66]. Remarkably, the genetic dysfunction of **MDM2** in adipocytes activates apoptotic and senescent **TP53**-mediated programs causing lipodystrophy and its related several metabolic diseases such as NAFLD [67]. Also, **VHL** disruption resulted in significant lipid accumulation, hepatic inflammation and fibrosis in the liver [68]. Lately, **SRC** has been found upregulated during the hepatic HSCs activation and liver fibrosis [69]. Also, **IKBKG** (or NEMO) deletion in liver parenchymal cells results in steatohepatitis and hepatocellular carcinoma [70]. Furthermore,

GRB2 suppression has been shown to improve hepatic steatosis, glucose metabolism, apoptosis and oxidative stress [71]. Moreover, the decreased expression of **SMARCD1** activates lipid accumulation and cellular senescence, denoting its preventative role regarding lifestyle-related diseases [72]. The phospho-**UBE2I** has been suggested to potentially enhance NF-kB signaling, revealing a possible new mechanism that deregulates inflammatory signaling of the liver [73]. The **GOLGA2** inhibition is found to induce fibrosis with autophagy in the liver and lung of mice [74]. **ERBB2** (also known as HER2) is closely linked to many enzymes, e.g. fatty acid synthase, which play essential regulatory roles in lipid metabolism or lipogenic pathways [75] and its hepatic expression has been identified in liver diseases [76, 77]. Remarkably, the hepatic gene expression of **SDCBP** has been found differentially expressed in steatotic liver [78]. Also, **CDC37** was defined with a modulatory role of INK4A activity in rat hepatic carcinogenesis and human hepatic cancer [79].

Interestingly, several studies applying bioinformatics analyses are in consistensy with our findings, revealing the possible implication of **UBQLN4** [80], **UBC** [81] and **PCNA** [82] in NAFLD development as potential biomarkers. Likewise, a bioinformatics analysis in a PPI network of steatosis highlights **CRK** and **MDM2** among of the top 10 important genes [83].

It is a well-known fact that disease-related proteins are clustered together and are also centrally located within a network [84]. As demonstrated from our results, the identified candidate NAFLD-related proteins: **RCOR1, TUFT1, PHF21A, OIP5, DZIP3, HMG20A, YEATS4, VPS50, CCDC93 (Fig. 3, 1st Cluster-magenta nodes), RELA, IKBKG, EGFR (Fig. 3, 5th Cluster-magenta nodes), MEOX2, LNX1 and PIN1 (Fig. 3, 6th Cluster-magenta nodes),** are found in the same clusters with already known NAFLD-related proteins, enhancing their potential implication in NAFLD. Notably, **RELA, IKBKG, EGFR** and **PIN1**, as already mentioned, are literally confirmed for their possible association with NAFLD.

Worthwhille to mention that the 7 candidate NAFLD-related proteins: **TRAF1, TRAF2, HLA-B, IKBKE, SRC, CDKN1A** and **TP53** are validated through the gene expression analysis. At first glance, this will probably not seem very prominent but it does show that the network

approach followed in this study is complementary to gene expression analysis by identifying more candidates associated with NAFLD that would otherwise not be detected. After performing pathway analysis of DEGs, **IKBKE** was found to be involved in toll-like receptor signaling pathway that play an important role in the NAFLD development [85]. Moreover, **TRAF1, TRAF2** and **TP53** are implicated in apoptosis which seems to be important in NAFLD and NASH progression [86]. Reportedly, **CDKN1A, SRC** and **TP53** are participated in senescence and autophagy in cancer. Interestingly, considerable associations have been established between regulation of autophagy and obesity-related liver complications, NAFLD [87]. It is important to mention that human clinical studies revealed the association of senescence with NAFLD [88]. Thereby, the aforementioned genes might play pivotal roles in the development and progression of NAFLD via regulating the pathways involved in this disease.

The **enrichment analysis** of the NAFLD Interactome was performed to examine the functional and biological interactions among the proteins, as well as to uncover their associations with diseases and several phenotypic abnormalities in human. Pathway analysis revealed that proteins are significantly enriched among others in pathways in cancer, PI3K-Akt signaling pathway, proteoglycans in cancer, MAPK signaling pathway and focal adhesion. It has been demonstated that PI3K-Akt and MAPK signaling pathways have been shown to be involved in NAFLD [89, 90]. Moreover, focal adhestion kinase regulates the activation of HSCs and liver fibrosis [91]. Interestingly, in the wound healing response, focal adhesion and proteoglycans in cancer pathways are implicated. As stated by other research works, these wound healing and cell migration pathways have been shown to be dysregulated in NASH leading to fibrosis [92]. Disease association analysis showed that proteins are associated with a number of diseases such as type 2 diabetes [93], chronic kidney failure [94], Alzheimer's disease [95], multiple sclerosis, schizophrenia [96], lung, bladder [97] and breast cancer [98], most of which are associated with NAFLD. Also, the phenotypic abnormalities of proteins such as those of

digestive system, metabolism/homeostasis, cardiovascular system, skin morphology and immune system are linked with NAFLD [99-102].

In conclusion, applying a systemic approach to this study, we were able to identify **77 candidate NAFLD-related proteins,** out of which 41 *(HSP90AB1, HLA-B, CTNNB1, HSPA5, CDKN1A, SMAD4, SMAD2, SMAD3, TRAF1, HSPB1, RELA, PIK3R3, GSK3B, VHL, SRC, EGFR, TP53, PIN1, KDM1A, EEF1A1, UBQLN4, UBC, PCNA, CRK, MDM2, VIM, MAP3K3, TNFRSF1A, YWHAZ, IKBKG, FYN, ABL1, GRB2, SMARCD1, UBE2I, GOLGA2, IKBKE, EP300, ERBB2, SDCBP,CDC37)* are confirmed through literature searches. The novelty of our findings lies in the remaining 36 proteins *(TRAF6, DISC1, NCK1, TRAF2, CYSRT1, LNX1, MEOX2, WDYHV1, LRRK2, YWHAG, MCC, HOXA1, KRT31, TRIM27, TRIM54, HTT, ATXN1, YWHAB, PICK1, CALM3, APC, TNIK, CSNK2B, HGS, PPP1CA, YWHAE, RCOR1, TUFT1, PHF21A, OIP5, DZIP3, HMG20A, YEATS4, VPS50, CCDC93, intact:EBI-4399559)* that could may be involved in NAFLD. It should be pointed out that the implementation of clustering analysis revealed the importance of 15 candidate NAFLD-related proteins in NAFLD *(RCOR1, TUFT1, PHF21A, OIP5, DZIP3, HMG20A, YEATS4, VPS50, CCDC93, RELA, IKBKG, EGFR, MEOX2, LNX1 and PIN1)* in light of the fact that are clustered together with known NAFLD-related proteins. Also, 9 of which (*RCOR1, TUFT1, PHF21A, OIP5, DZIP3, HMG20A, YEATS4, VPS50 and CCDC93)* had not been published before in other research works. Noteworthy, we subsequently achieved via gene expression analysis the verification of 7 candidate NAFLD-related proteins: *TRAF1, TRAF2, HLA-B, IKBKE, SRC, CDKN1A and TP53*, while TRAF2 is one of the proteins that has not been found previously in the literature. Several of the results obtained in the present study are also reported by many other studies, as outlined in the Discussion section of this manuscript. We hope that our research will serve as a base for further experimental works.

Acknowledgement

The authors thank the Harokopio University of Athens for use of premises and equipment.

Funding

The research work was financially supported by the Hellenic Foundation for Research and Innovation (HFRI) under the HFRI PhD Fellowship grant (Fellowship Number: 1529).

CRediT author statement

Athina I. Amanatidou: Conceptualization, Methodology, Software, Validation, Formal analysis, Data Curation, Investigation, Visualization, Writing - Original Draft. **George. V. Dedoussis:** Supervision, Writing - Review & Editing.

Conflict of Interest:

None declared.

Abbreviations

NAFLD: non-alcoholic fatty liver disease, NASH: nonalcoholic steatohepatitis, PMIDs: PubMed IDs, PPI: Protein-protein interaction, MCODE: Molecular Complex Detection, PHC: Proteins with high closeness centrality, HPO: Human Phenotype Ontology, DEGs: Differentially expressed genes, CPL: Characteristic path length, NAS: NAFLD Activity Score, HSP90AB1: Heat shock protein HSP 90-beta, HLA-B: histocompatibility antigen, B alpha chain, SRC: Proto-oncogene tyrosine-protein kinase Src, TRAF1: TNF receptor-associated factor 1, TRAF2: TNF receptor-associated factor 2, CTNNB1: Catenin beta-1, HSPA5: Endoplasmic reticulum chaperone BiP, CDKN1A: Cyclin-dependent kinase inhibitor 1, SMAD4: Mothers against decapentaplegic homolog 4, SMAD2: Mothers against decapentaplegic homolog 2, HSPB1: Heat shock protein beta-1, RELA: Transcription factor p65, PIK3R3: Phosphatidylinositol 3-kinase regulatory subunit gamma, GSK3B: Glycogen

synthase kinase-3 beta, VHL: von Hippel-Lindau disease tumor suppressor, EGFR: Epidermal growth factor receptor, TP53: Cellular tumor antigen p53, PIN1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1, SMAD3: Mothers against decapentaplegic homolog 3, KDM1A: Lysine-specific histone demethylase 1A, EEF1A1: Elongation factor 1-alpha 1, UBQLN4: Ubiquilin-4, UBC: Polyubiquitin-C, PCNA: Proliferating cell nuclear antigen, CRK: Adapter molecule crk, MDM2: E3 ubiquitin-protein ligase Mdm2, TP53: Cellular tumor antigen p53, VIM: Vimentin, MAP3K3: Mitogen-activated protein kinase 3, TNFRSF1A: Tumor necrosis factor receptor superfamily member 1A, YWHAZ: 14-3-3 protein zeta/delta, IKBKG: NF-kappa-B essential modulator, FYN: Tyrosine-protein kinase Fyn, ABL1: Tyrosine-protein kinase ABL1, GRB2: Growth factor receptor-bound protein 2, SMARCD1: SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 1, UBE2I: SUMO-conjugating enzyme UBC9, GOLGA2: Golgin subfamily A member 2, IKBKE: Inhibitor of nuclear factor kappa-B kinase subunit epsilon, EP300: Histone acetyltransferase p300, ERBB2: Receptor tyrosine-protein kinase erbB-2, SDCBP: Syntenin-1, CDC37: Hsp90 co-chaperone Cdc37, HSCs: hepatic stellate cells

References

[1] M. Blencowe, T. Karunanayake, J. Wier, N. Hsu, X. Yang, Network Modeling Approaches and Applications to Unravelling Non-Alcoholic Fatty Liver Disease, Genes, 10 (2019).
[2] S.K. Asrani, H. Devarbhavi, J. Eaton, P.S. Kamath, Burden of liver diseases in the world, Journal of hepatology, 70 (2019) 151-171.
[3] G.B. Goh, M.R. Pagadala, J. Dasarathy, A. Unalp-Arida, R. Sargent, C. Hawkins, A. Sourianarayanane, A. Khiyami, L. Yerian, R.K. Pai, S. Dasarathy, A.J. McCullough, Clinical spectrum of non-alcoholic fatty liver disease in diabetic and non-diabetic patients, BBA clinical, 3 (2015) 141-145.
[4] C.D. Byrne, G. Targher, NAFLD: a multisystem disease, Journal of hepatology, 62 (2015) S47-64.
[5] Z.M. Younossi, D. Blissett, R. Blissett, L. Henry, M. Stepanova, Y. Younossi, A. Racila, S. Hunt, R. Beckerman, The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe, Hepatology, 64 (2016) 1577-1586.
[6] S. Sookoian, C.J. Pirola, Personalizing care for nonalcoholic fatty liver disease patients: what are the research priorities?, Personalized medicine, 11 (2014) 735-743.
[7] S.M. Abd El-Kader, E.M. El-Den Ashmawy, Non-alcoholic fatty liver disease: The diagnosis and management, World journal of hepatology, 7 (2015) 846-858.
[8] M. Benedict, X. Zhang, Non-alcoholic fatty liver disease: An expanded review, World journal of hepatology, 9 (2017) 715-732.
[9] G.V. Dedoussis, A.I. Amanatidou, From Transcriptomic to Metabolomic in the Development of Biomarkers in NAFLD/NASH, in: M. Romero-Gomez (Ed.) NAFLD and NASH: Biomarkers in Detection, Diagnosis and Monitoring, Springer International Publishing, Cham, 2020, pp. 181-190.
[10] A.I. Amanatidou, K.C. Nastou, O.E. Tsitsilonis, V.A. Iconomidou, Visualization and analysis of the interaction network of proteins associated with blood-cell targeting autoimmune diseases, Biochimica et biophysica acta. Molecular basis of disease, 1866 (2020) 165714.
[11] G. Fiscon, F. Conte, L. Farina, P. Paci, Network-Based Approaches to Explore Complex Biological Systems towards Network Medicine, Genes, 9 (2018).
[12] J. Pinero, N. Queralt-Rosinach, A. Bravo, J. Deu-Pons, A. Bauer-Mehren, M. Baron, F. Sanz, L.I. Furlong, DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes, Database : the journal of biological databases and curation, 2015 (2015) bav028.
[13] C. UniProt, The universal protein resource (UniProt), Nucleic acids research, 36 (2008) D190-195.
[14] S. Orchard, M. Ammari, B. Aranda, L. Breuza, L. Briganti, F. Broackes-Carter, N.H. Campbell, G. Chavali, C. Chen, N. del-Toro, M. Duesbury, M. Dumousseau, E. Galeota, U. Hinz, M. Iannuccelli, S. Jagannathan, R. Jimenez, J. Khadake, A. Lagreid, L. Licata, R.C. Lovering, B. Meldal, A.N. Melidoni, M. Milagros, D. Peluso, L. Perfetto, P. Porras, A. Raghunath, S. Ricard-Blum, B. Roechert, A. Stutz, M. Tognolli, K. van Roey, G. Cesareni, H. Hermjakob, The MIntAct project--IntAct as a common curation platform for 11 molecular interaction databases, Nucleic acids research, 42 (2014) D358-363.
[15] M. Sivade Dumousseau, D. Alonso-Lopez, M. Ammari, G. Bradley, N.H. Campbell, A. Ceol, G. Cesareni, C. Combe, J. De Las Rivas, N. Del-Toro, J. Heimbach, H. Hermjakob, I. Jurisica, M. Koch, L. Licata, R.C. Lovering, D.J. Lynn, B.H.M. Meldal, G. Micklem, S. Panni, P. Porras, S. Ricard-Blum, B. Roechert, L. Salwinski, A. Shrivastava, J. Sullivan, N. Thierry-Mieg, Y. Yehudi, K. Van Roey, S. Orchard, Encompassing new use cases - level 3.0 of the HUPO-PSI format for molecular interactions, BMC bioinformatics, 19 (2018) 134.

[16] S. Lotia, J. Montojo, Y. Dong, G.D. Bader, A.R. Pico, Cytoscape app store, Bioinformatics, 29 (2013) 1350-1351.
[17] Y. Assenov, F. Ramirez, S.E. Schelhorn, T. Lengauer, M. Albrecht, Computing topological parameters of biological networks, Bioinformatics, 24 (2008) 282-284.
[18] C. Chen, H. Shen, L.G. Zhang, J. Liu, X.G. Cao, A.L. Yao, S.S. Kang, W.X. Gao, H. Han, F.H. Cao, Z.G. Li, Construction and analysis of protein-protein interaction networks based on proteomics data of prostate cancer, International journal of molecular medicine, 37 (2016) 1576-1586.
[19] H. Yu, P.M. Kim, E. Sprecher, V. Trifonov, M. Gerstein, The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics, PLoS computational biology, 3 (2007) e59.
[20] C.H. Chin, S.H. Chen, H.H. Wu, C.W. Ho, M.T. Ko, C.Y. Lin, cytoHubba: identifying hub objects and sub-networks from complex interactome, BMC systems biology, 8 Suppl 4 (2014) S11.
[21] G.D. Bader, C.W. Hogue, An automated method for finding molecular complexes in large protein interaction networks, BMC bioinformatics, 4 (2003) 2.
[22] G. Dennis, Jr., B.T. Sherman, D.A. Hosack, J. Yang, W. Gao, H.C. Lane, R.A. Lempicki, DAVID: Database for Annotation, Visualization, and Integrated Discovery, Genome biology, 4 (2003) P3.
[23] B. Zhang, S. Kirov, J. Snoddy, WebGestalt: an integrated system for exploring gene sets in various biological contexts, Nucleic acids research, 33 (2005) W741-748.
[24] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, G. Sherlock, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium, Nature genetics, 25 (2000) 25-29.
[25] K.G. Becker, K.C. Barnes, T.J. Bright, S.A. Wang, The genetic association database, Nature genetics, 36 (2004) 431-432.
[26] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, Nucleic acids research, 28 (2000) 27-30.
[27] S. Kohler, N.A. Vasilevsky, M. Engelstad, E. Foster, J. McMurry, S. Ayme, G. Baynam, S.M. Bello, C.F. Boerkoel, K.M. Boycott, M. Brudno, O.J. Buske, P.F. Chinnery, V. Cipriani, L.E. Connell, H.J. Dawkins, L.E. DeMare, A.D. Devereau, B.B. de Vries, H.V. Firth, K. Freson, D. Greene, A. Hamosh, I. Helbig, C. Hum, J.A. Jahn, R. James, R. Krause, F.L. SJ, H. Lochmuller, G.J. Lyon, S. Ogishima, A. Olry, W.H. Ouwehand, N. Pontikos, A. Rath, F. Schaefer, R.H. Scott, M. Segal, P.I. Sergouniotis, R. Sever, C.L. Smith, V. Straub, R. Thompson, C. Turner, E. Turro, M.W. Veltman, T. Vulliamy, J. Yu, J. von Ziegenweidt, A. Zankl, S. Zuchner, T. Zemojtel, J.O. Jacobsen, T. Groza, D. Smedley, C.J. Mungall, M. Haendel, P.N. Robinson, The Human Phenotype Ontology in 2017, Nucleic acids research, 45 (2017) D865-D876.
[28] M. Kriss, L. Golden-Mason, J. Kaplan, F. Mirshahi, V.W. Setiawan, A.J. Sanyal, H.R. Rosen, Increased hepatic and circulating chemokine and osteopontin expression occurs early in human NAFLD development, PloS one, 15 (2020) e0236353.
[29] R. Edgar, M. Domrachev, A.E. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository, Nucleic acids research, 30 (2002) 207-210.
[30] T. Barrett, S.E. Wilhite, P. Ledoux, C. Evangelista, I.F. Kim, M. Tomashevsky, K.A. Marshall, K.H. Phillippy, P.M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C.L. Robertson, N. Serova, S. Davis, A. Soboleva, NCBI GEO: archive for functional genomics data sets--update, Nucleic acids research, 41 (2013) D991-995.
[31] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic acids research, 43 (2015) e47.

[32] S. Davis, P.S. Meltzer, GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor, Bioinformatics, 23 (2007) 1846-1847.
[33] D.N. Slenter, M. Kutmon, K. Hanspers, A. Riutta, J. Windsor, N. Nunes, J. Melius, E. Cirillo, S.L. Coort, D. Digles, F. Ehrhart, P. Giesbertz, M. Kalafati, M. Martens, R. Miller, K. Nishida, L. Rieswijk, A. Waagmeester, L.M.T. Eijssen, C.T. Evelo, A.R. Pico, E.L. Willighagen, WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research, Nucleic acids research, 46 (2018) D661-D667.
[34] M.V. Kuleshov, M.R. Jones, A.D. Rouillard, N.F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S.L. Jenkins, K.M. Jagodnik, A. Lachmann, M.G. McDermott, C.D. Monteiro, G.W. Gundersen, A. Ma'ayan, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, Nucleic acids research, 44 (2016) W90-97.
[35] R.D. Leclerc, Survival of the sparsest: robust gene networks are parsimonious, Molecular systems biology, 4 (2008) 213.
[36] G. Mao, N. Zhang, Analysis of Average Shortest-Path Length of Scale-Free Network, Journal of Applied Mathematics, 2013 (2013) 865643.
[37] A.L. Barabasi, Z.N. Oltvai, Network biology: understanding the cell's functional organization, Nature reviews. Genetics, 5 (2004) 101-113.
[38] R. Albert, Scale-free networks in cell biology, Journal of cell science, 118 (2005) 4947-4957.
[39] W. Chen, J. Jiang, P.P. Wang, L. Gong, J. Chen, W. Du, K. Bi, H. Diao, Identifying Hepatocellular Carcinoma Driver Genes by Integrative Pathway Crosstalk and Protein Interaction Network, DNA and cell biology, 38 (2019) 1112-1124.
[40] Y. Wang, Y. Zhang, Q. Huang, C. Li, Integrated bioinformatics analysis reveals key candidate genes and pathways in breast cancer, Molecular medicine reports, 17 (2018) 8091-8100.
[41] A. Balanescu, I. Stan, I. Codreanu, V. Comanici, E. Balanescu, P. Balanescu, Circulating Hsp90 Isoform Levels in Overweight and Obese Children and the Relation to Nonalcoholic Fatty Liver Disease: Results from a Cross-Sectional Study, Disease markers, 2019 (2019) 9560247.
[42] M. Celikbilek, H. Selcuk, U. Yilmaz, A new risk factor for the development of non-alcoholic fatty liver disease: HLA complex genes, The Turkish journal of gastroenterology : the official journal of Turkish Society of Gastroenterology, 22 (2011) 395-399.
[43] K. Enooku, M. Kondo, N. Fujiwara, T. Sasako, J. Shibahara, A. Kado, K. Okushin, H. Fujinaga, T. Tsutsumi, R. Nakagomi, T. Minami, M. Sato, H. Nakagawa, Y. Kondo, Y. Asaoka, R. Tateishi, K. Ueki, H. Ikeda, H. Yoshida, K. Moriya, H. Yotsuyanagi, T. Kadowaki, M. Fukayama, K. Koike, Hepatic IRS1 and ss-catenin expression is associated with histological progression and overt diabetes emergence in NAFLD patients, Journal of gastroenterology, 53 (2018) 1261-1275.
[44] E. Rodriguez-Suarez, A.M. Duce, J. Caballeria, F. Martinez Arrieta, E. Fernandez, C. Gomara, N. Alkorta, U. Ariz, M.L. Martinez-Chantar, S.C. Lu, F. Elortza, J.M. Mato, Non-alcoholic fatty liver disease proteomics, Proteomics. Clinical applications, 4 (2010) 362-371.
[45] A. Aravinthan, G. Mells, M. Allison, J. Leathart, A. Kotronen, H. Yki-Jarvinen, A.K. Daly, C.P. Day, Q.M. Anstee, G. Alexander, Gene polymorphisms of cellular senescence marker p21 and disease progression in non-alcohol-related fatty liver disease, Cell cycle, 13 (2014) 1489-1494.
[46] M. Xiang, P.X. Wang, A.B. Wang, X.J. Zhang, Y. Zhang, P. Zhang, F.H. Mei, M.H. Chen, H. Li, Targeting hepatic TRAF1-ASK1 signaling to improve inflammation, insulin resistance, and hepatic steatosis, Journal of hepatology, 64 (2016) 1365-1377.
[47] J. Wattacheril, K.L. Rose, S. Hill, C. Lanciault, C.R. Murray, K. Washington, B. Williams, W. English, M. Spann, R. Clements, N. Abumrad, C.R. Flynn, Non-alcoholic fatty liver disease

phosphoproteomics: A functional piece of the precision puzzle, Hepatology research : the official journal of the Japan Society of Hepatology, 47 (2017) 1469-1483.
[48] G. Qin, G.Z. Wang, D.D. Guo, R.X. Bai, M. Wang, S.Y. Du, Deletion of Smad4 reduces hepatic inflammation and fibrogenesis during nonalcoholic steatohepatitis progression, Journal of digestive diseases, 19 (2018) 301-313.
[49] L. Yang, Y.S. Roh, J. Song, B. Zhang, C. Liu, R. Loomba, E. Seki, Transforming growth factor beta signaling in hepatocytes participates in steatohepatitis through regulation of cell death and lipid metabolism in mice, Hepatology, 59 (2014) 483-495.
[50] J.A. Willy, S.K. Young, J.L. Stevens, H.C. Masuoka, R.C. Wek, CHOP links endoplasmic reticulum stress to NF-kappaB activation in the pathogenesis of nonalcoholic steatohepatitis, Molecular biology of the cell, 26 (2015) 2190-2204.
[51] X. Yang, Y. Fu, F. Hu, X. Luo, J. Hu, G. Wang, PIK3R3 regulates PPARalpha expression to stimulate fatty acid beta-oxidation and decrease hepatosteatosis, Experimental & molecular medicine, 50 (2018) e431.
[52] J. Cao, X.X. Feng, L. Yao, B. Ning, Z.X. Yang, D.L. Fang, W. Shen, Saturated free fatty acid sodium palmitate-induced lipoapoptosis by targeting glycogen synthase kinase-3beta activation in human liver cells, Digestive diseases and sciences, 59 (2014) 346-357.
[53] S. Choung, J.M. Kim, K.H. Joung, E.S. Lee, H.J. Kim, B.J. Ku, Epidermal growth factor receptor inhibition attenuates non-alcoholic fatty liver disease in diet-induced obese mice, PloS one, 14 (2019) e0210828.
[54] Z. Derdak, K.A. Villegas, R. Harb, A.M. Wu, A. Sousa, J.R. Wands, Inhibition of p53 attenuates steatosis and liver injury in a mouse model of non-alcoholic fatty liver disease, Journal of hepatology, 58 (2013) 785-791.
[55] Y. Nakatsu, Y. Otani, H. Sakoda, J. Zhang, Y. Guo, H. Okubo, A. Kushiyama, M. Fujishiro, T. Kikuch, T. Fukushima, H. Ohno, Y. Tsuchiya, H. Kamata, A. Nagamachi, T. Inaba, F. Nishimura, H. Katagiri, S. Takahashi, H. Kurihara, T. Uchida, T. Asano, Role of Pin1 protein in the pathogenesis of nonalcoholic steatohepatitis in a rodent model, The Journal of biological chemistry, 287 (2012) 44526-44535.
[56] P. Chen, Q. Luo, C. Huang, Q. Gao, L. Li, J. Chen, B. Chen, W. Liu, W. Zeng, Z. Chen, Pathogenesis of non-alcoholic fatty liver disease mediated by YAP, Hepatology international, 12 (2018) 26-36.
[57] K. Dreval, V. Tryndyak, A. de Conti, F.A. Beland, I.P. Pogribny, Gene Expression and DNA Methylation Alterations During Non-alcoholic Steatohepatitis-Associated Liver Carcinogenesis, Frontiers in genetics, 10 (2019) 486.
[58] A.M. Hetherington, C.G. Sawyez, B.G. Sutherland, D.L. Robson, R. Arya, K. Kelly, R.L. Jacobs, N.M. Borradaile, Treatment with didemnin B, an elongation factor 1A inhibitor, improves hepatic lipotoxicity in obese mice, Physiological reports, 4 (2016).
[59] M. Aparicio-Vergara, P.P. Hommelberg, M. Schreurs, N. Gruben, R. Stienstra, R. Shiri-Sverdlov, N.J. Kloosterhuis, A. de Bruin, B. van de Sluis, D.P. Koonen, M.H. Hofker, Tumor necrosis factor receptor 1 gain-of-function mutation aggravates nonalcoholic fatty liver disease but does not cause insulin resistance in a murine model, Hepatology, 57 (2013) 566-576.
[60] Q. He, X. Xia, K. Yao, J. Zeng, W. Wang, Q. Wu, R. Tang, X. Zou, Amlexanox reversed non-alcoholic fatty liver disease through IKKepsilon inhibition of hepatic stellate cell, Life sciences, 239 (2019) 117010.
[61] S. Fukunishi, Y. Tsuda, A. Takeshita, H. Fukui, K. Miyaji, A. Fukuda, K. Higuchi, p59fyn is associated with the development of hepatic steatosis due to chronic ethanol consumption, Journal of clinical biochemistry and nutrition, 49 (2011) 20-24.
[62] S.J. Lee, J.D. Yoo, S.Y. Choi, O.S. Kwon, The expression and secretion of vimentin in the progression of non-alcoholic steatohepatitis, BMB reports, 47 (2014) 457-462.

[63] T. Csak, S. Bala, D. Lippai, A. Satishchandran, D. Catalano, K. Kodys, G. Szabo, microRNA-122 regulates hypoxia-inducible factor-1 and vimentin in hepatocytes and correlates with fibrosis in diet-induced steatohepatitis, Liver international : official journal of the International Association for the Study of the Liver, 35 (2015) 532-541.
[64] D.H. Kim, J. Kim, J.S. Kwon, J. Sandhu, P. Tontonoz, S.K. Lee, S. Lee, J.W. Lee, Critical Roles of the Histone Methyltransferase MLL4/KMT2D in Murine Hepatic Steatosis Directed by ABL1 and PPARgamma2, Cell reports, 17 (2016) 1671-1682.
[65] J. Bricambert, J. Miranda, F. Benhamed, J. Girard, C. Postic, R. Dentin, Salt-inducible kinase 2 links transcriptional coactivator p300 phosphorylation to the prevention of ChREBP-dependent hepatic steatosis in mice, The Journal of clinical investigation, 120 (2010) 4316-4331.
[66] C. Desterke, F. Chiappini, Lipid Related Genes Altered in NASH Connect Inflammation in Liver Pathogenesis Progression to HCC: A Canonical Pathway, International journal of molecular sciences, 20 (2019).
[67] Z. Liu, L. Jin, J.K. Yang, B. Wang, K.K.L. Wu, P. Hallenborg, A. Xu, K.K.Y. Cheng, The Dysfunctional MDM2-p53 Axis in Adipocytes Contributes to Aging-Related Metabolic Complications by Induction of Lipodystrophy, Diabetes, 67 (2018) 2397-2409.
[68] E. Paschetta, P. Belci, A. Alisi, D. Liccardo, R. Cutrera, G. Musso, V. Nobili, OSAS-related inflammatory mechanisms of liver injury in nonalcoholic fatty liver disease, Mediators of inflammation, 2015 (2015) 815721.
[69] H.Y. Seo, S.H. Lee, J.H. Lee, Y.N. Kang, J.S. Hwang, K.G. Park, M.K. Kim, B.K. Jang, Src Inhibition Attenuates Liver Fibrosis by Preventing Hepatic Stellate Cell Activation and Decreasing Connetive Tissue Growth Factor, Cells, 9 (2020).
[70] T. Luedde, N. Beraza, V. Kotsikoris, G. van Loo, A. Nenci, R. De Vos, T. Roskams, C. Trautwein, M. Pasparakis, Deletion of NEMO/IKKgamma in liver parenchymal cells causes steatohepatitis and hepatocellular carcinoma, Cancer cell, 11 (2007) 119-132.
[71] X. Shan, Y. Miao, R. Fan, C. Song, G. Wu, Z. Wan, J. Zhu, G. Sun, W. Zha, X. Mu, G. Zhou, Y. Chen, Suppression of Grb2 expression improved hepatic steatosis, oxidative stress, and apoptosis induced by palmitic acid in vitro partly through insulin signaling alteration, In vitro cellular & developmental biology. Animal, 49 (2013) 576-582.
[72] C. Inoue, C. Zhao, Y. Tsuduki, M. Udono, L. Wang, M. Nomura, Y. Katakura, SMARCD1 regulates senescence-associated lipid accumulation in hepatocytes, NPJ aging and mechanisms of disease, 3 (2017) 11.
[73] M.L. Tomasi, K. Ramani, M. Ryoo, Ubiquitin-Conjugating Enzyme 9 Phosphorylation as a Novel Mechanism for Potentiation of the Inflammatory Response, The American journal of pathology, 186 (2016) 2326-2336.
[74] S. Park, S. Kim, M.J. Kim, Y. Hong, A.Y. Lee, H. Lee, Q. Tran, M. Kim, H. Cho, J. Park, K.P. Kim, J. Park, M.H. Cho, GOLGA2 loss causes fibrosis with autophagy in the mouse lung and liver, Biochemical and biophysical research communications, 495 (2018) 594-600.
[75] A. Ray, Tumor-linked HER2 expression: association with obesity and lipid-related microenvironment, Hormone molecular biology and clinical investigation, 32 (2017).
[76] P. Doring, G.M. Pilo, D.F. Calvisi, F. Dombrowski, [Nuclear Her2 expression in hepatocytes in liver disease], Der Pathologe, 38 (2017) 211-217.
[77] J.H. Shi, W.Z. Guo, Y. Jin, H.P. Zhang, C. Pang, J. Li, P.D. Line, S.J. Zhang, Recognition of HER2 expression in hepatocellular carcinoma and its significance in postoperative tumor recurrence, Cancer medicine, 8 (2019) 1269-1278.
[78] N. Guillen, M.A. Navarro, C. Arnal, E. Noone, J.M. Arbones-Mainar, S. Acin, J.C. Surra, P. Muniesa, H.M. Roche, J. Osada, Microarray analysis of hepatic gene expression identifies new genes involved in steatotic liver, Physiological genomics, 37 (2009) 187-198.
[79] R.M. Pascale, M.M. Simile, D.F. Calvisi, M. Frau, M.R. Muroni, M.A. Seddaiu, L. Daino, M.D. Muntoni, M.R. De Miglio, S.S. Thorgeirsson, F. Feo, Role of HSP90, CDC37, and CRM1 as

modulators of P16(INK4A) activity in rat liver carcinogenesis and human liver cancer, Hepatology, 42 (2005) 1310-1319.
[80] S. Qi, C. Wang, C. Li, P. Wang, M. Liu, Candidate genes investigation for severe nonalcoholic fatty liver disease based on bioinformatics analysis, Medicine, 96 (2017) e7743.
[81] L. Li, H. Liu, X. Hu, Y. Huang, Y. Wang, Y. He, Q. Lei, Identification of key genes in nonalcoholic fatty liver disease progression based on bioinformatics analysis, Molecular medicine reports, 17 (2018) 7708-7720.
[82] J. Liu, B. Lin, Z. Chen, M. Deng, Y. Wang, J. Wang, L. Chen, Z. Zhang, X. Xiao, C. Chen, Y. Song, Identification of key pathways and genes in nonalcoholic fatty liver disease using bioinformatics analysis, Archives of medical science : AMS, 16 (2020) 374-385.
[83] R. Wang, X. Wang, L. Zhuang, Gene expression profiling reveals key genes and pathways related to the development of non-alcoholic fatty liver disease, Annals of hepatology, 15 (2016) 190-199.
[84] T. Ideker, R. Sharan, Protein networks in disease, Genome research, 18 (2008) 644-652.
[85] K. Miura, H. Ohnishi, Role of gut microbiota and Toll-like receptors in nonalcoholic fatty liver disease, World journal of gastroenterology, 20 (2014) 7381-7391.
[86] T. Kanda, S. Matsuoka, M. Yamazaki, T. Shibata, K. Nirei, H. Takahashi, T. Kaneko, M. Fujisawa, T. Higuchi, H. Nakamura, N. Matsumoto, H. Yamagami, M. Ogawa, H. Imazu, K. Kuroda, M. Moriyama, Apoptosis and non-alcoholic fatty liver diseases, World journal of gastroenterology, 24 (2018) 2661-2672.
[87] V.J. Lavallard, P. Gual, Autophagy and non-alcoholic fatty liver disease, BioMed research international, 2014 (2014) 120179.
[88] A.M. Papatheodoridi, L. Chrysavgis, M. Koutsilieris, A. Chatzigeorgiou, The Role of Senescence in the Development of Nonalcoholic Fatty Liver Disease and Progression to Nonalcoholic Steatohepatitis, Hepatology, 71 (2020) 363-374.
[89] S. Matsuda, M. Kobayashi, Y. Kitagishi, Roles for PI3K/AKT/PTEN Pathway in Cell Signaling of Nonalcoholic Fatty Liver Disease, ISRN endocrinology, 2013 (2013) 472432.
[90] A. Lawan, A.M. Bennett, Mitogen-Activated Protein Kinase Regulation in Hepatic Metabolism, Trends in endocrinology and metabolism: TEM, 28 (2017) 868-878.
[91] X.-K. Zhao, L. Yu, M.-L. Cheng, P. Che, Y.-Y. Lu, Q. Zhang, M. Mu, H. Li, L.-L. Zhu, J.-J. Zhu, M. Hu, P. Li, Y.-D. Liang, X.-H. Luo, Y.-J. Cheng, Z.-X. Xu, Q. Ding, Focal Adhesion Kinase Regulates Hepatic Stellate Cell Activation and Liver Fibrosis, Scientific Reports, 7 (2017) 4032.
[92] A. Chatterjee, A. Basu, K. Das, P. Singh, D. Mondal, B. Bhattacharya, S. Roychoudhury, P.P. Majumder, A. Chowdhury, P. Basu, Hepatic transcriptome signature correlated with HOMA-IR explains early nonalcoholic fatty liver disease pathogenesis, Annals of hepatology, 19 (2020) 472-481.
[93] S. Tomah, N. Alkhouri, O. Hamdy, Nonalcoholic fatty liver disease and type 2 diabetes: where do Diabetologists stand?, Clinical diabetes and endocrinology, 6 (2020) 9.
[94] M. Marcuccilli, M. Chonchol, NAFLD and Chronic Kidney Disease, International journal of molecular sciences, 17 (2016) 562.
[95] L.D. Estrada, P. Ahumada, D. Cabrera, J.P. Arab, Liver Dysfunction as a Novel Player in Alzheimer's Progression: Looking Outside the Brain, Frontiers in aging neuroscience, 11 (2019) 174.
[96] Ó. Soto-Angona, G. Anmella, M.J. Valdés-Florido, N. De Uribe-Viloria, A.F. Carvalho, B.W.J.H. Penninx, M. Berk, Non-alcoholic fatty liver disease (NAFLD) as a neglected metabolic companion of psychiatric disorders: common pathways and future approaches, BMC Medicine, 18 (2020) 261.
[97] C.L. Chiang, H.H. Huang, T.Y. Huang, Y.L. Shih, T.Y. Hsieh, H.H. Lin, Nonalcoholic Fatty Liver Disease Associated With Bladder Cancer, The American journal of the medical sciences, 360 (2020) 161-165.

[98] Y.S. Lee, H.S. Lee, S.W. Chang, C.U. Lee, J.S. Kim, Y.K. Jung, J.H. Kim, Y.S. Seo, H.J. Yim, C.H. Lee, S.U. Woo, J.H. Seo, J.E. Yeon, S.H. Um, K.S. Byun, Underlying nonalcoholic fatty liver disease is a significant factor for breast cancer recurrence after curative surgery, Medicine, 98 (2019) e17277.
[99] M. Kosmalski, L. Mokros, P. Kuna, A. Witusik, T. Pietras, Changes in the immune system - the key to diagnostics and therapy of patients with non-alcoholic fatty liver disease, Central-European journal of immunology, 43 (2018) 231-239.
[100] A. Scalera, M.N. Di Minno, G. Tarantino, What does irritable bowel syndrome share with non-alcoholic fatty liver disease?, World journal of gastroenterology, 19 (2013) 5402-5420.
[101] R. Prussick, L. Prussick, D. Nussbaum, Nonalcoholic Fatty liver disease and psoriasis: what a dermatologist needs to know, The Journal of clinical and aesthetic dermatology, 8 (2015) 43-45.
[102] C. Tana, S. Ballestri, F. Ricci, A. Di Vincenzo, A. Ticinesi, S. Gallina, M.A. Giamberardino, F. Cipollone, R. Sutton, R. Vettor, A. Fedorowski, T. Meschi, Cardiovascular Risk in Non-Alcoholic Fatty Liver Disease: Mechanisms and Therapeutic Implications, International journal of environmental research and public health, 16 (2019).