



Semi-automated approach for generation of biological networks on drug-induced cholestasis, steatosis, hepatitis, and cirrhosis

Hyun Kil Shin^{1,2}  · Oana Florean³ · Barry Hardy³ · Tatyana Doktorova³ · Myung-Gyun Kang¹ 

Received: 9 December 2021 / Revised: 2 February 2022 / Accepted: 11 February 2022 / Published online: 3 March 2022
© The Author(s) 2022

Abstract

Drug-induced liver injury (DILI) is one of the leading reasons for discontinuation of a new drug development project. Diverse machine learning or deep learning models have been developed to predict DILI. However, these models have not provided an adequate understanding of the mechanisms leading to DILI. The development of safer drugs requires novel computational approaches that enable the prompt understanding of the mechanism of DILI. In this study, the mechanisms leading to the development of cholestasis, steatosis, hepatitis, and cirrhosis were explored using a semi-automated approach for data gathering and associations. Diverse data from ToxCast, Comparative Toxicogenomic Database (CTD), Reactome, and Open TG-GATEs on reference molecules leading to the development of the respective diseases were extracted. The data were used to create biological networks of the four diseases. As expected, the four networks had several common pathways, and a joint DILI network was assembled. Such biological networks could be used in drug discovery to identify possible molecules of concern as they provide a better understanding of the disease-specific key events. The events can be target-tested to provide indications for potential DILI effects.

Keywords DILI · Drug-induced liver disease · Hepatotoxicity · Computational toxicology

Introduction

To protect human health and the environment from undesired adverse effects, safety information on the toxic potential of new molecules has to be generated and revised to regulate their use [1]. Safety assessments usually involve a substantial amount of animal experimentation. This is financially burdensome and time-consuming [2], is not always ethically performed, and sometimes, the relevance for humans is questionable. To address these shortcomings, regulatory agencies worldwide are encouraging the implementation of the 3Rs

(reduction, refinement, and replacement) of animal experimentation. Among the numerous alternative approaches being developed and tested, those using toxicogenomics or computational toxicology in general are considered especially promising because they will likely facilitate faster hypothesis generation covering data-rich historical sources as input information. This will provide a more detailed and precise understanding of the mechanisms of toxicity.

Drug-induced liver injury (DILI) is one of the common reasons for terminating drug development projects. Approaches for the early detection of such alerts are needed. Numerous machine learning (ML) and deep learning (DL) models have been developed for the detection of DILI and resulting withdrawal of drug candidates in the early phases of development [3]. Although ML or DL models can accurately predict DILI, their interpretability is still an issue for researchers because the process of decision-making by the models is not explainable in most cases [4]. Therefore, a different approach is required to determine the mechanisms of the development of liver disease due to drug-mediated damage. The adverse outcome pathway (AOP) concept partially addresses the issue. The AOP is a formalized, transparent, and quality-controlled way to describe mechanistic

✉ Tatyana Doktorova
tatyana@edelweissconnect.com

✉ Myung-Gyun Kang
myung-gyun.kang@kitox.re.kr

¹ Toxicoinformatics Group, Department of Predictive Toxicology, Korea Institute of Toxicology, Daejeon 34114, Republic of Korea

² Human and Environmental Toxicology, University of Science and Technology, Daejeon 34113, Republic of Korea

³ Edelweiss Connect GmbH, Hochbergerstrasse 60C, 4057 Basel, Switzerland

information to endpoints for regulatory purposes [5, 6]. However, a potential problem of the AOP is that it requires a lot of time for experts to manually search the literature and identify key events and the links between them. Furthermore, the suggested constructs are linear and are sometimes too simplistic to cover the entire spectrum of events leading to an adverse outcome. Thus, another approach is needed to broaden the understanding of how to deploy knowledge in the development of safer drugs.

Biological networks are one of the most common ways to represent a sequence of molecular and cellular events leading to toxic effects following exogenous exposure [7]. Such networks provide a systemic approach for key events identification and targeted testing by means of alternative approaches [8]. For some complex diseases such as DILI, biological networks are an appropriate way of capturing events and relationships as they cover the entire spectrum of relationships, which can also be non-linear. In evidence collection for biological network development, use of historical data and novel computational approaches for data analysis and assembly have proven very useful for hypothesis generation [9].

In this study, we used a semi-automated approach for diverse data gathering, harmonization, and integration to identify important key events and their relationships in the development of four DILI diseases: cholestasis, steatosis, hepatitis, and cirrhosis. Diverse data from different levels of biological organization extracted from ToxCast, Comparative Toxicogenomics Database (CTD), Reactome, and Open TG-GATEs were subjected to association analysis. Top-ranked genes and pathways associated with the respective diseases were identified and interlinked. As expected, the four studied diseases had a significant number of overlapping biological processes. Finally, a joint DILI network was assembled. Oxidative stress was an obvious major factor contributing to the development of DILI.

Materials and methods

Chemical-gene-pathway-disease mapping

Following a previously described approach [9], frequent item set mining methods were applied to high-throughput screening, gene expression, in vivo studies, and disease data present in ToxCast and the CTD. ToxCast provides high-throughput toxicity screening in vitro assay data, and active/inactivity calls on one target gene were collected. CTD provided chemical-gene and chemical-disease interactions. Chemical-gene interactions were defined as activity or expression variation due to chemical treatment, and chemical-disease interactions were inferred associations between chemicals and diseases through genes commonly

associated between the chemical and the disease. Reactome database clustered pathways according to function similarity, and pathway hierarchy data was used in the mapping.

Selection of reference chemicals for the diseases

In this study, cholestasis, steatosis, hepatitis, and cirrhosis were selected as the most common diseases among DILI cases in post-market phase based on PharmaPendium® database (<https://www.pharmapendium.com/>, accessed Dec. 2018) which curated the post-market reports based on disease terms. After the removal of not-specific adverse endpoints such as liver injury and hepatotoxicity, post market reports for these four DILI cases were more than 10,000 in numbers, respectively. Some drugs were withdrawn from the market due to development of severe DILI even though there were no sign of DILI in preclinical and clinical test, therefore, the diseases were selected based on post-market reports only. A set of reference compounds specific for each disease of interest were selected. Positive reference chemicals induce the disease of interest. Negative chemicals are those for which there is no evidence in the literature of an association to the disease of interest. This is a challenging task as most of the known chemicals coincidentally cause several of the studied diseases, rather than being disease-specific. As this was a very important part of the analysis and forms the basis of the developed biological networks, the following strategy was applied. First, search of publicly available literature sources was performed to identify chemicals that were specific to one of the four diseases. Once the disease-specific chemicals had been identified, those also present in Open TG-GATEs [10] were selected for further analysis. The list of OpenTG-GATEs compounds was compared with the list of LiverTox compounds. The LiverTox DB provides up-to-date comprehensive and unbiased information about DILI (<https://www.ncbi.nlm.nih.gov/books/NBK547852/>). The idea was to obtain a list of compounds that lead to DILI and have available gene expression data. Furthermore, the probable mechanisms of action as listed according to LiverTox were extracted whenever possible or available. The chemicals were split into groups according to the cause of the diseases. Finally, compounds in each group were compared to post-market reports on the four diseases.

Refinement of chemical-gene-pathway-disease associations

Transcriptomics data from Open TG-GATEs of the selected set of chemicals were applied to customize chemical-gene-pathway-disease associations to the reference chemicals. The fold changes between treatment and control were used from transcriptomics data based on p values. Genes were labeled as toxic group-specific if they showed statistically significant

deregulation ($-1 < \text{Log}_2 \text{ fold change} > 1$ and p value 0.05 by the Welch t test) over more than three chemicals. This approach was applied identically to both positive and negative groups at each time point and dose combination. The genes were further selected only if they were deregulated mostly by the positive reference chemicals. Finally, to ensure that all the identified specific genes were differentially expressed compared to that in response to the negative group of chemicals, the individual gene profiles were manually examined using a bar chart time-series to acknowledge changes in expression over time and by heatmap validation to acknowledge grouping and differential expression in comparison to the negative group of chemicals. This additional curation was done to capture group-specific genes involved in the development of a disease and to remove the individual chemical-specific behaviors. For example, if there is a gene deregulated in response to three chemicals only, it would fulfill the above-described cutoff criteria and would automatically be selected and initially appear in the gene list. This would be a captured chemical-specific behavior, rather than disease-specific. Therefore, an additional step of individual gene examination was performed. Only the group-specific genes were used in the further analysis.

The goal was to extract associations between genes and diseases that co-occur across datasets. The four diseases were used as input queries, and the output data of the high-probability of a disease of interest to be connected to a

set of genes was extracted. Figure 1 describes the workflow of this study.

Biological networks assembly

Lastly, biological networks were defined as a set of biological events (e.g., at a gene, protein, and pathway level) and relationships assembled. To accomplish this, the first step was to search for direct links between the selected genes. This did not yield very promising results. Thus, the next step was to examine the possibility of interlinking through connector genes. This was done using the GeneMANIA (<https://genemania.org/>) web-based application for gene function prediction [11]. Once the biological networks were compiled, the processes in which these direct and indirect interacting genes were involved were identified. This was done mostly through a literature search and manual curation.

Results

Genes associated with the four liver diseases

The reference chemicals (Fig. 2) were used in chemical-gene-pathway-disease mapping. Querying the assembled chemical-gene-pathway-disease output for the selected disease of interest resulted in 200 to 300 characteristic genes

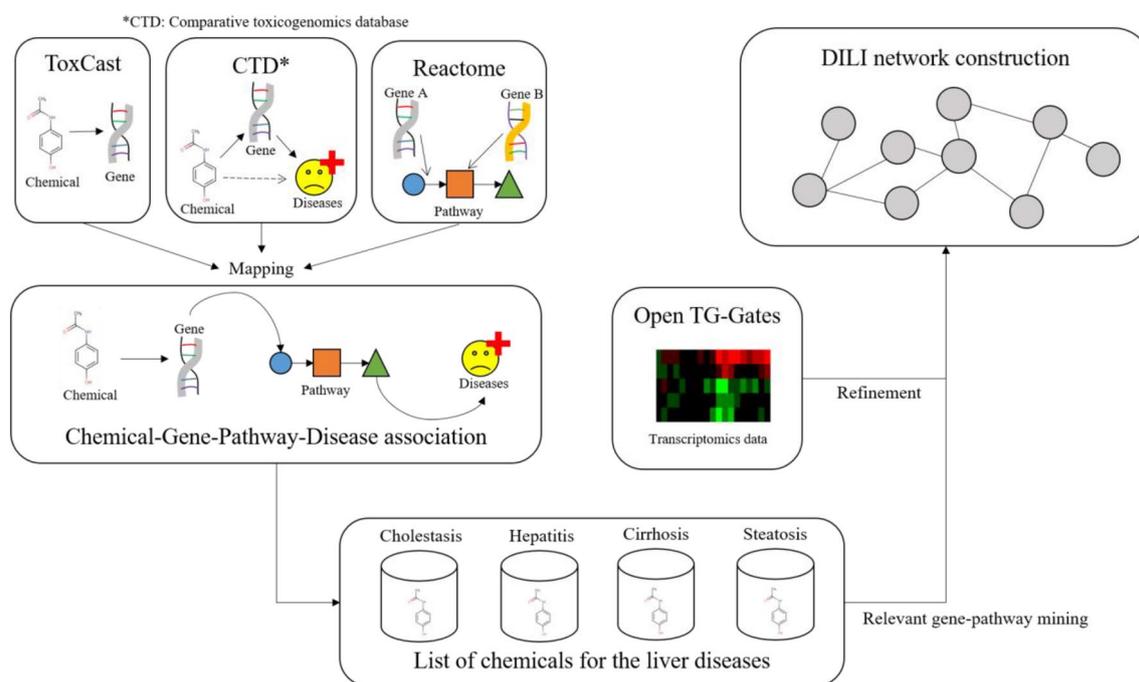


Fig. 1 Workflow of this study. Associations between chemicals, genes, pathways, and diseases were established based on data from ToxCast, CTD, and Reactome. This association was further refined

with TG-Gates data. Query molecules for each disease were used to extract relevant associations

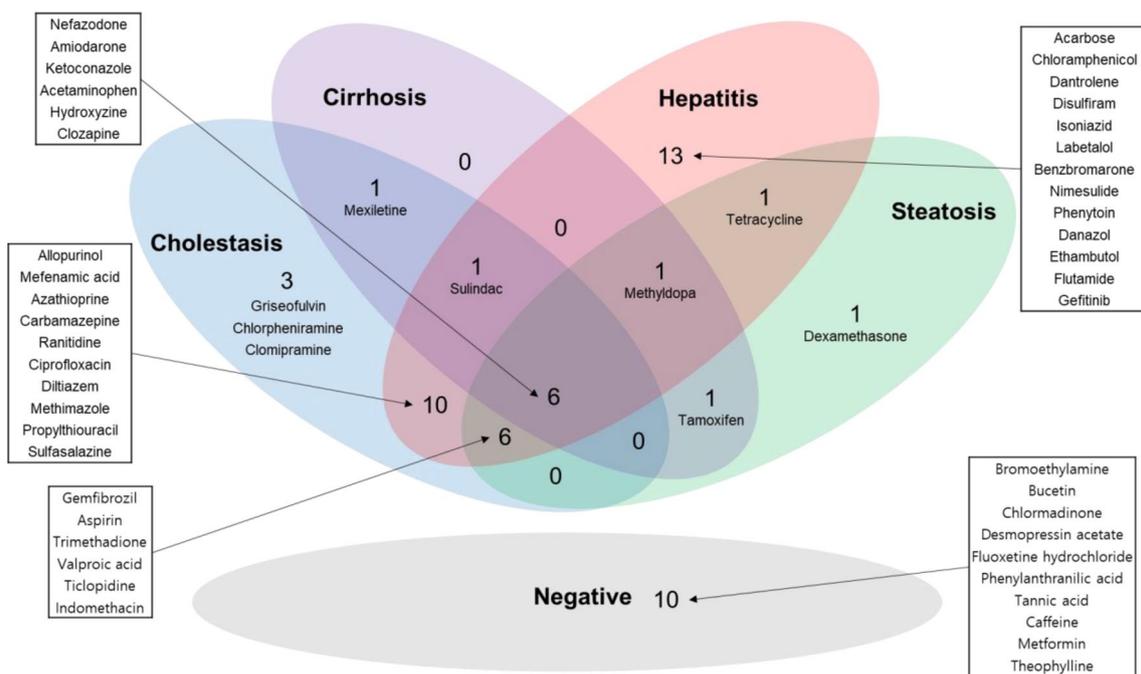
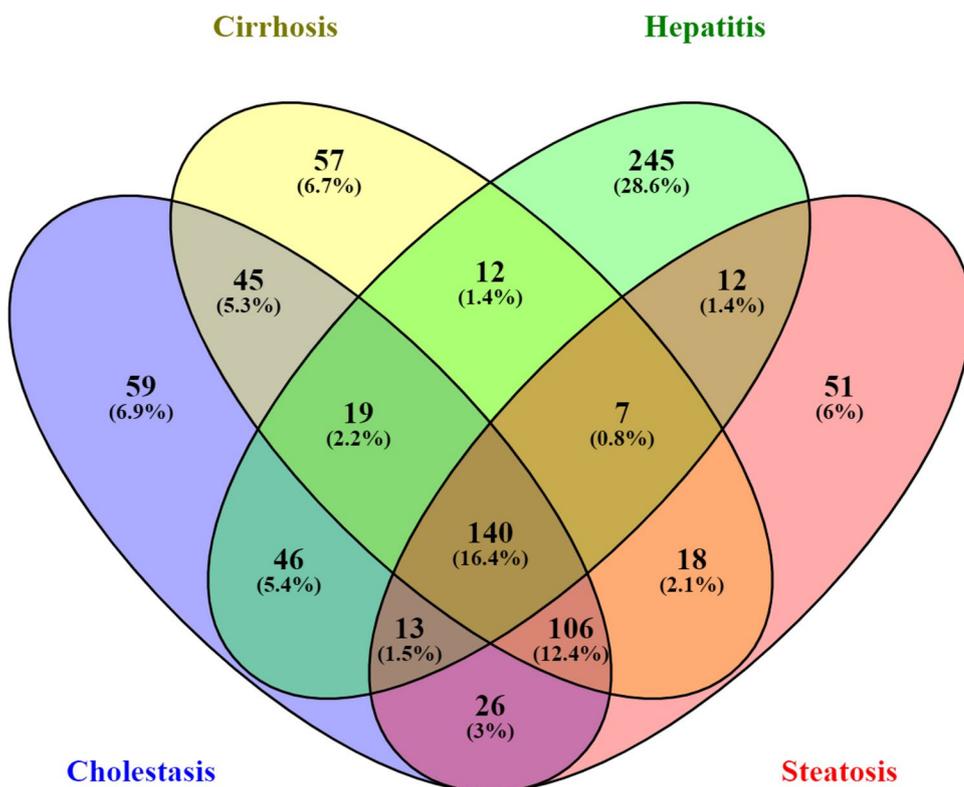


Fig. 2 Venn diagram for chemicals selected for each disease. Most of the compounds were labeled over multiple DILI subtypes

Fig. 3 Venn diagram of the genes associated with the diseases



per disease. There was a substantial gene overlap of approximately 140 genes between all four diseases (Fig. 3). The largest difference at the gene level was evident for hepatitis, with 28.6% of the genes specific to hepatitis. The follow-up pathway analysis also indicated a substantial overlap in the number of affected pathways between all four diseases. Thus, a second tier was added. By incorporating data on disease-specific reference chemicals, identification of genes highly specific to the disease of interest was possible. After the application of the second tier, 32 genes highly associated with the development of cholestasis were initially identified (Table 1). The expression of each gene was individually examined and compared to the negative control group to identify whether these genes could serve as biomarkers to characterize and identify cholestasis-specific mechanisms. This further refinement showed that most of the identified genes were also deregulated in a similar fashion in the negative control group, with the exception of five genes (BHLHE40: basic helix-loop-helix family member e40, CSRP1: cysteine and glycine rich protein 1, NQO1: NAD(P) H quinone dehydrogenase 1, SLC16A10: solute carrier family 16 member 10, and AGT: angiotensinogen) that were consistently deregulated in the positive group as opposed to in the negative group (Fig. 4). In the case of steatosis, 20 genes were initially identified as highly associated with its development (Table 2). Further examination revealed that only four genes (CCL2: C-C motif chemokine ligand 2, ICAM1: intracellular adhesion molecule 1, ME1: malic enzyme 1, and SGK1: serum/glucocorticoid regulated kinase 1) were consistently upregulated in the positive group, and they exhibited a different trend in the negative group (Fig. 5). Twenty-five genes were initially identified as being highly

associated with the development of hepatitis (Table 3). Further examination revealed that only nine genes (SLC6A6: solute carrier family 6 member 6, CSRP1, RAB30: RAS oncogene family member, APOM: apolipoprotein M, PPP2R1B: protein phosphatase 2 scaffold subunit abeta, HMOX1: heme oxygenase 1, TSR1: ribosome maturation factor, EBNA1BP2: EBNA1 binding protein 2, and WDR77: WD repeat domain 77) were consistently deregulated in the positive reference group, whereas they were unaffected or less affected by the negative reference group of chemicals (Fig. 6). In the case of cirrhosis, 37 genes were initially identified as highly associated with the development of cirrhosis (Table 4). Further examination showed that only eight genes (CSF1R: colony stimulating factor 1 receptor, EGR1: early growth response 1, GPNMB: glycoprotein Nmb, LPL: lipoprotein lipase, HSD11B2: hydroxysteroid 11-beta dehydrogenase 2, MGAT2: alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase, NR1D1: nuclear receptor subfamily 1 group D member 1, and S100A9: S100 calcium binding protein A9) were consistently deregulated (Fig. 7).

Biological network compilation

The identified disease-specific genes were assembled in a network by using the GeneMANIA tool along with manual curation to identify directly and indirectly connected genes. Using this approach, four networks comprising one network each for cholestasis, steatosis, cirrhosis, and hepatitis were generated. In the case of cholestasis (Fig. 8A), the main processes triggering the development of the disease were NRF2 (nuclear factor erythroid 2-related factor 2) activation, oxidative stress, increase in expression of detoxification

Table 1 Genes identified as of highest probability to be associated with the development of cholestasis

Time	29Days	15 days	8 days	4 days	1 day	9 h	6 h	3 h
Genes	ABCC2	ABCC2	CD63	A1BG	ABHD3	A1BG	CARS	A1BG
	ADSSL1	ACLY	CYP26B1	ABHD3	CYP26B1	ABHD3	EGR1	ABHD3
	BHLHE40	AGT	CYP7A1	ACLY	EGR1	BHLHE40	HSPB8	IFIT1
	CD63	CSRP1		BHLHE40	MID1IP1	CEBPB	NARS	NQO1
	CFD	CYP27A1		C1R	UHRF1	CREM	S100A9	USP18
	CREM	IFIT1		MID1IP1		CYP7A1		
	CSRP1	IL13RA1		TM7SF2		MID1IP1		
	CYP7A1	S100A9				NQO1		
	IP6K2					SREBF1		
	MID1IP1							
	NQO1							
	PPP2R1B							
	S100A9							
	SLC16A10							
	SREBF1							
	USP18							

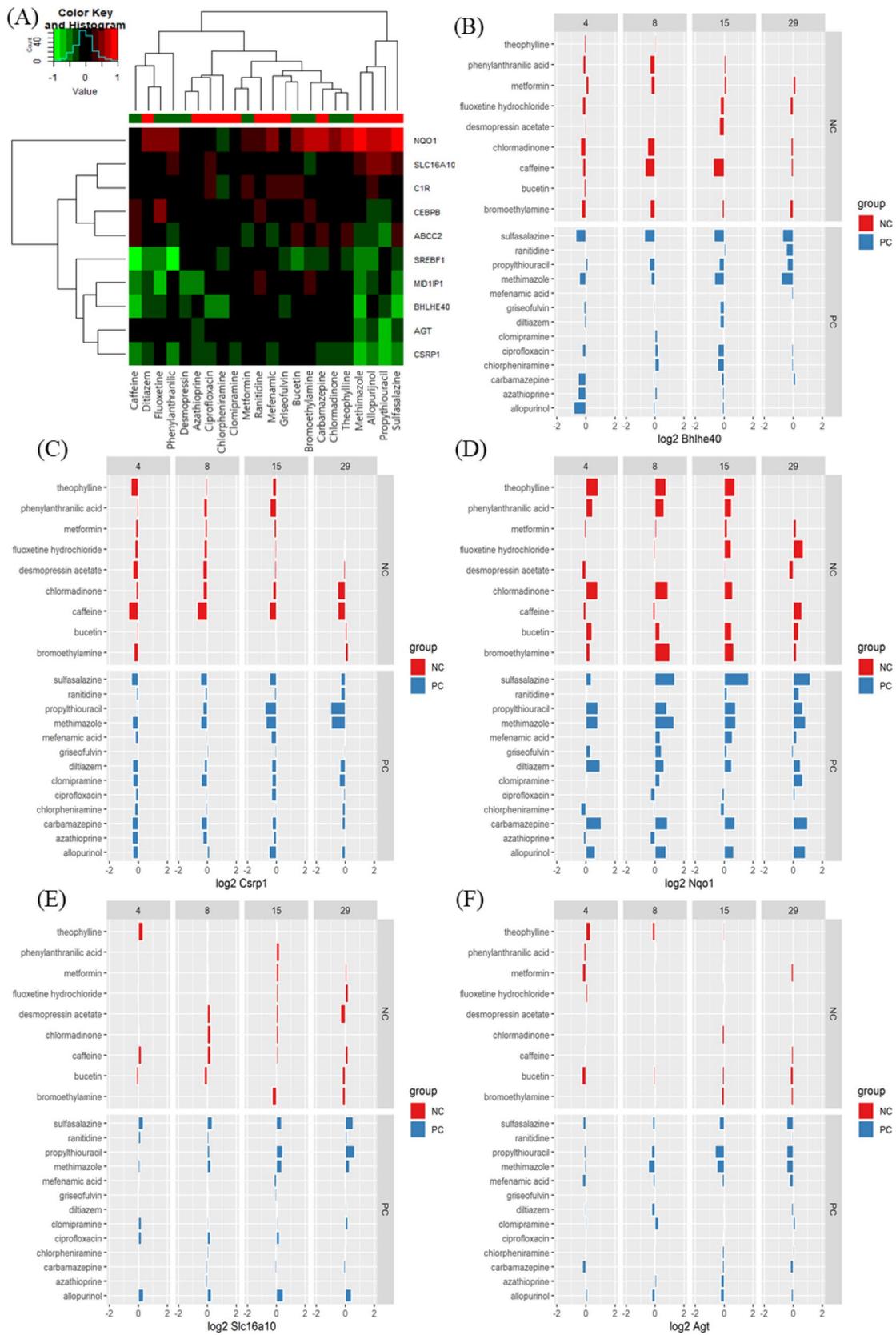


Fig. 4 Heat map (A) and bar chart (B–F) analyses of genes associated with drug-induced cholestasis by comparison between positive compounds (PCs) and negative compounds (NCs). Five genes were consistently deregulated among PCs and displayed a different trend in NCs: BHLHE40 (downregulated, B), CSRP1 (downregulated, C), NQO1 (upregulated, D), SLC15A10 (upregulated, E), and AGT (downregulated, F)

enzymes, and cytoprotection. The decrease in the level of AGT, which is involved in the regulation of glucose and homeostasis, indirectly leads to an increase in oxidative stress [12]. This increase stimulates the natural defenses of the organism, with increased activities of detoxification enzymes. This is also supported by NRF2 activation, represented by an increase of NQO1 level [13, 14] and increased transport of amino acids by SLC16A10 upregulation [15]. Although the goal of the responses is cytoprotection, a general stress response to the downregulation of BHLHE40 also occurs [16]. This is further reflected in the decrease of responses leading to hepatic differentiation by a decrease in CSRP1 level [17]. The findings indicate that oxidative stress plays an important role in the development of cholestasis. In the steatosis network (Fig. 8B), the main identified triggers were an increase in the stress response, as indicated by an increase in SGK1 level [18, 19]. This can be associated with increased *de novo* fatty acid biosynthesis, as verified by the increase in ME1 level [20–22]. This is linked to an overall increase in apoptosis (upregulation of ICAM1) [23] which leads to the recruitment of immune cells and inflammation via the upregulation of CCL2 [24]. For the hepatitis network (Fig. 8C), important identified pathways were, again, mainly involved in responses to increased oxidative stress,

which initiates cytoprotective action of upregulated HMOX1 [25] leading to mitochondrial dysfunction associated with downregulation of the SLC6A6 taurine transporter [26]. Mitochondrial dysfunction triggers autophagy events by increasing RAB30 level [27] and its linked genes. Increased PPP2R1B level is associated with cell cycle arrest [28] and decreased TSR1 level [29] is associated with DNA damage. Cell cycle arrest causes hepatic differentiation, which maps to the decreased expression of CSRP1 and its linked genes. Upregulated APOM [30] and WDR77 [31] and their associated genes are related to hepatitis virus replication and re-entry. Similarly, in the case of the cirrhosis network (Fig. 8D), the results revealed a central role for oxidative stress. Increased oxidative stress leads to apoptosis associated with upregulated EGR1 [32, 33] and activation of the NF-κB (nuclear factor-kappa B) pathway through the downregulation of HSD11B2 [34, 35]. The genes surrounding EGR1 and HSD11B2 are also important for the mapped events. Activation of apoptosis and the NF-κB pathway leads to production, differentiation, and function of macrophages connected to decreased CSF1R level [36] and its interacting genes. This event promotes the activation of Kupffer cells (evident as increased GPNMB level) [37] and macrophage response (evident as increased NR1D1 level) [38]. The macrophage response causes inflammation (upregulated S100A9) [39], which promotes uncontrolled wound healing and hepatic stellate cell (HSC) activation (evident upregulation of lipoprotein lipase) [40]. HSC activation produces hepatic fibrosis, and glycoproteins appear as indicators of the disease (increased MGAT2 level) [41]. After analysis of each network, the four networks were combined to create

Table 2 Genes identified as of highest probability to be associated with the development of steatosis

Time	29 days	15 days	8 days	4 days	1 day	9 h	6 h	3 h
Genes	CROT ME1	CROT ME1 S100A9	A1BG CROT EGR1 ME1 SLC22A8	ME1 SLC22A8	ACLY CCL2 CROT EGR1 ICAM1 KLKB1 LAMC2 ME1 MYC S100A9 SOD2 SREBF1 TARS USP18 XBP1	ACLY EGR1 ME1 PHGDH SREBF1	ACLY	INSIG1 SGK1

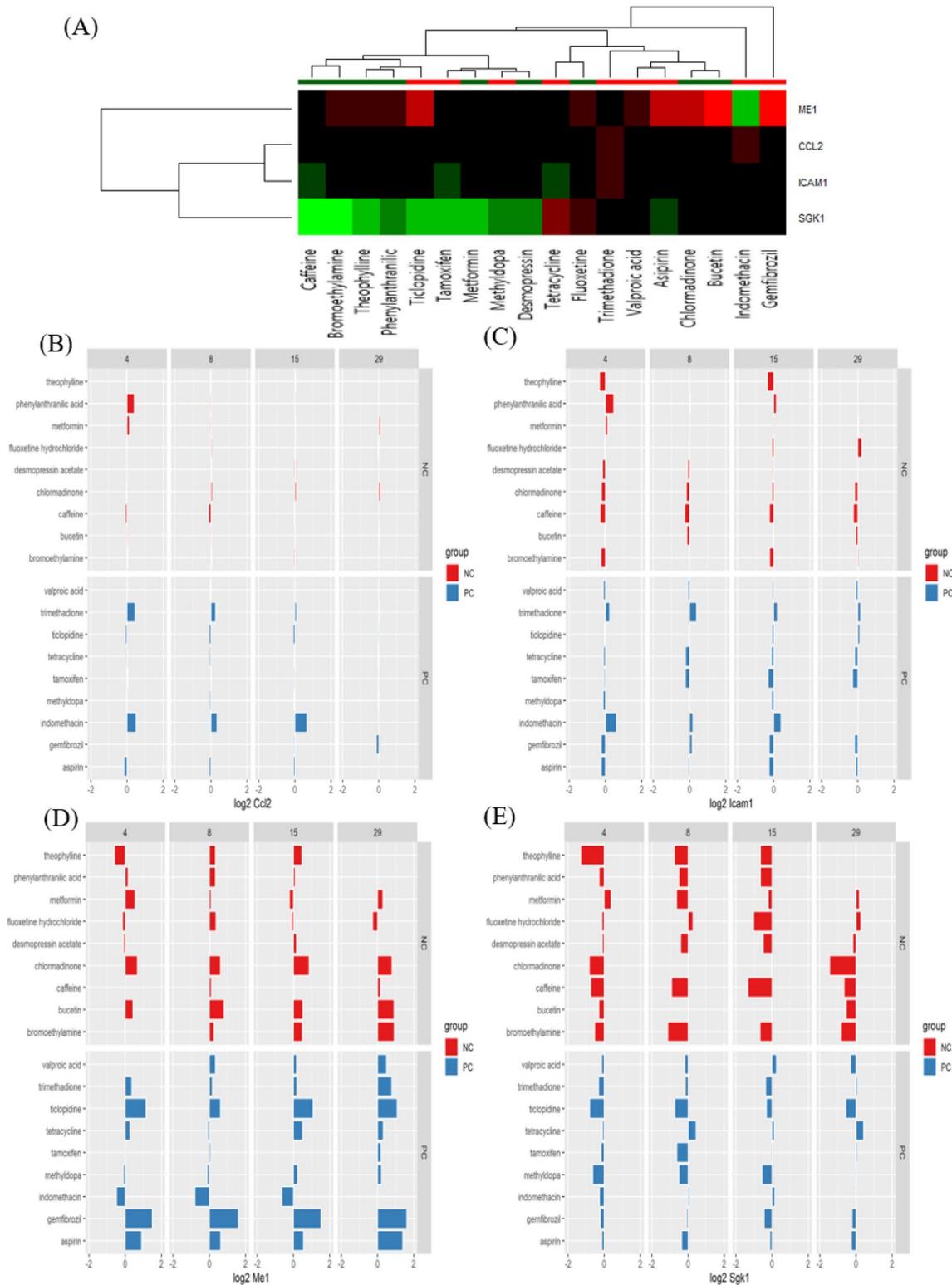


Fig. 5 Heat map (A) and bar chart (B–E) analyses of genes associated with drug-induced steatosis by comparison between positive compounds (PCs) and negative compounds (NCs). Four genes were

consistently upregulated among PCs and displayed a different trend in NCs: CCL2 (B), ICAM1 (C), ME1 (D), and SGK1 (E)

a general DILI network to identify overlapping or common events (Fig. 9). Oxidative stress was evident in all four diseases. Cholestasis and hepatitis shared cytoprotection and hepatic differentiation. Cholestasis and steatosis featured a general stress response as the most common process. Steatosis and cirrhosis seem to be involved in apoptosis.

Discussion

Drug-induced hepatic injury is the most common reason for the termination of drug development projects or withdrawal of market-approved drugs. Characteristic DILI-associated diseases include cholestasis, steatosis, hepatitis, and cirrhosis. Cholestasis is a condition in which the flow of bile from the liver is slowed or blocked. Steatosis features abnormal retention of fat within liver cells. Impaired hepatic lipid storage usually leads to liver metabolic dysfunction and inflammation. Cirrhosis is an irreversible scarring of the liver that leads to impaired liver function. It is considered a terminal stage of chronic liver impairment. Hepatitis is characterized by general inflammation of the liver, triggered by either chemical exposure or viral infection. Although all these diseases have specific characteristics, they are interconnected in their manifestation. Discrimination of these diseases can be difficult as they can progress from one of the diseases to another. For example, steatosis, cholestasis, and hepatitis can progress to cirrhosis, which is considered an end-stage liver disease. Better understanding of the mechanisms behind the progression of these diseases is important. Many predictive tools have tried to provide clarity using numerous cell-based systems, animal models and *in silico* algorithms. However, progress has been unsatisfactory.

We used semi-automated approaches for data gathering and integration to try to capture the major events responsible for the progression and development of the four selected diseases. The added value of the selected approach includes its fast and robust performance. Diverse data from high-throughput screening, gene expression, *in vivo* observations,

and various diseases can be harmonized, analyzed, and assembled to generate the complex DILI biological network. The collective findings indicate that oxidative stress is a central mechanism in the pathogenesis of DILI. Results of the applied semi-automated approach for biological network assembly revealed an increase in the release of reactive oxygen species (ROS) in all four diseases. ROS are associated with an increase in cytoprotective mechanisms and decrease in hepatic differentiation. Furthermore, when such damage is persistent, the natural defenses of the organism are over-ruled and cell damage occurs. This leads to the generation of oxidative stress, mitochondrial dysfunction, and endoplasmic reticulum stress, which may result in scarring and necrosis. At the same time, increased cellular damage stimulates the immune response, whereby inflammasomes trigger an adaptive immune response. Normally, the liver has natural defense mechanisms. However, in DILI, these natural defense mechanisms are compromised, resulting in deteriorated oxidative stress, mitochondrial damage, cell death, and inflammation. These series of events have recently been described in the literature as biomarkers for early identification of DILI in clinical practice [42].

Another important identified mechanism is apoptosis. It seems to be a common event between steatosis and cirrhosis. However, in the case of cirrhosis, the inflammatory response is markedly more pronounced due to the simultaneous action of macrophage response and activation. Cell death is an important mechanism contributing to the development of DILI [43]. Although important, apoptosis is not a central mechanism for DILI development, and therefore, inhibiting apoptosis is insufficient to prevent liver injury. Furthermore, it is not an individual event, but rather an interconnected series of chain events involving increased oxidative stress, refractory increase of cytoprotective mechanisms, and failure of cytoprotection due to the DILI-compromised mechanisms, and eventually mitochondrial dysfunction and apoptosis.

As this approach identified the significant biological events related with the diseases, it can be used to suggest

Table 3 Genes identified as of highest probability to be associated with the development of hepatitis

Time	29 days	15 days	8 days	4 days	1 day	9 h	6 h
Genes	HBB	HBB	SLPI	SLPI	EGR1	EGR1	EGR1
	CROT	SLPI	CROT	CROT	HMOX1	HMOX1	MID1IP1
	FOXA2	EGR1	SLC6A6	CSRP1	MID1IP1	XBP1	RAB30
	SLC6A6	CROT	RAB30	ABHD3	TSR1	MID1IP1	CEBPB
	CSRP1	SLC6A6	TM7SF2	IP6K2	AURKA	RAB30	CROT
	REEP5	CSRP1		SREBF1	RAB30	CYP26B1	
	RAB30	PPP2R1B			CYP26B1	SREBF1	
	APOM	RAB30			DNAJB9		
	SREBF1	CEBPB			EBNA1BP2		
					WDR77		

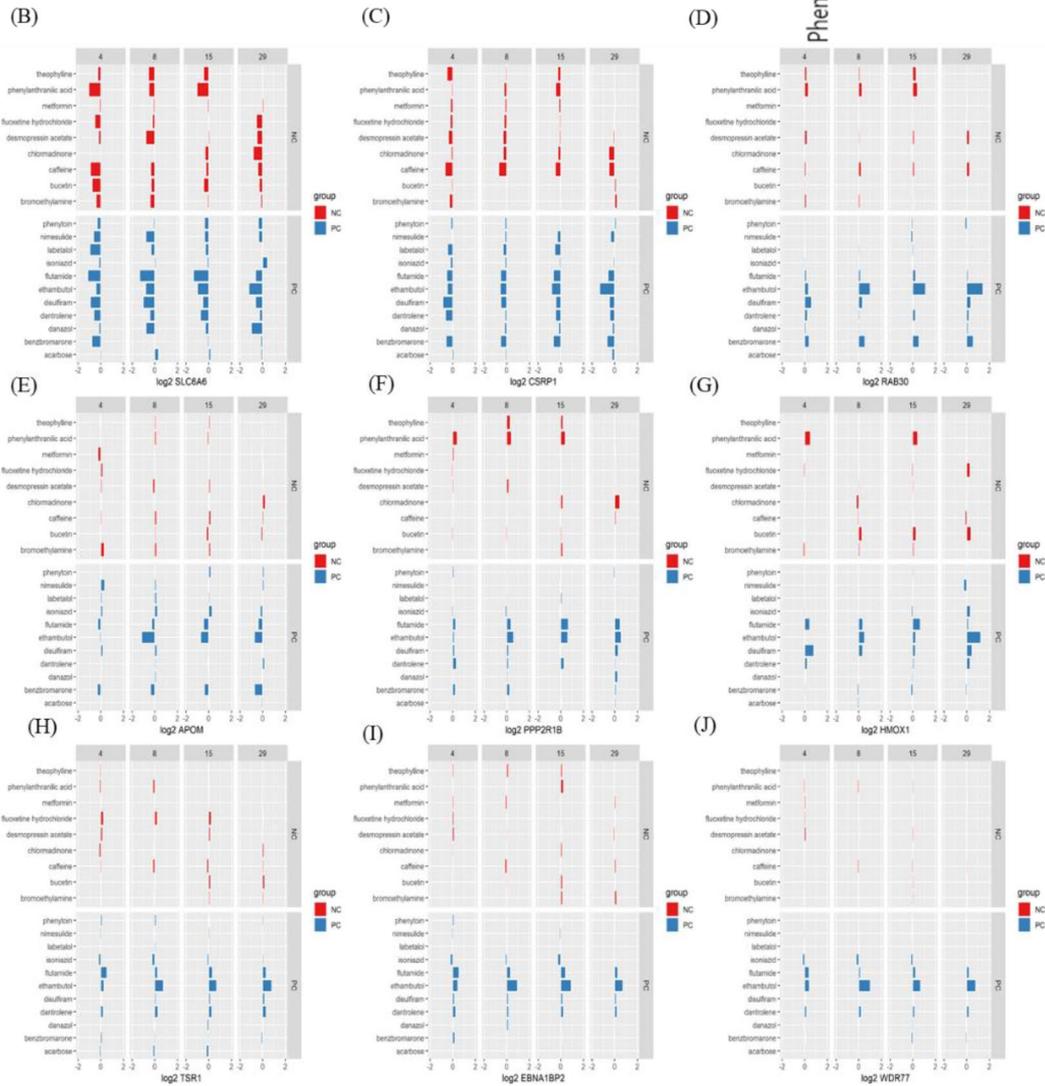
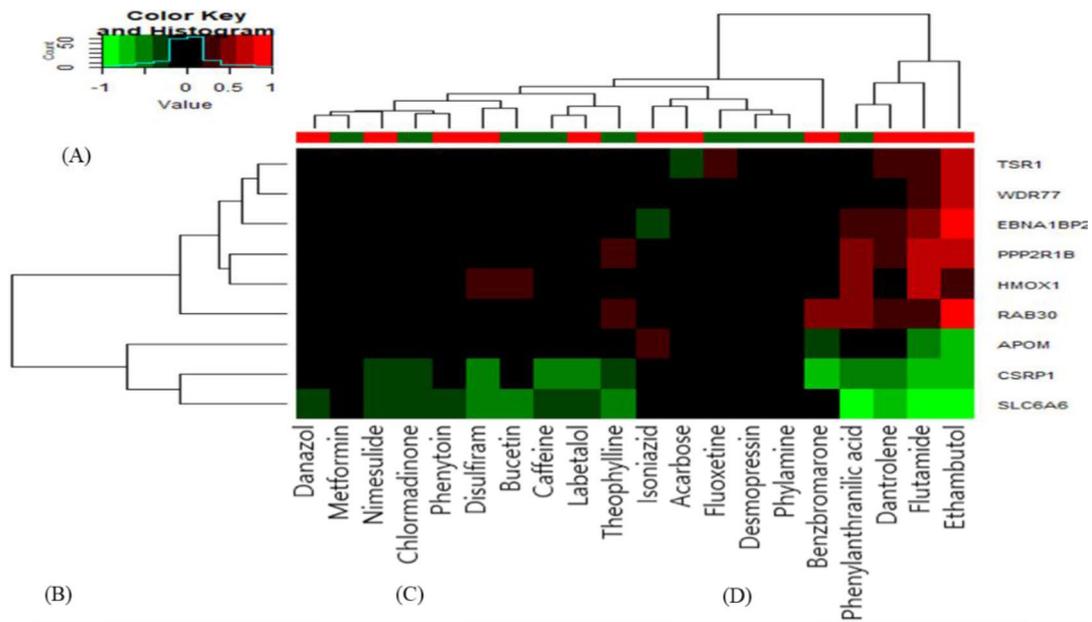


Fig. 6 Heat map (A) and bar chart (B–J) analyses of genes associated with drug-induced hepatitis in comparison between positive compounds (PCs) and negative compounds (NCs). Nine genes were consistently deregulated among PCs and displayed a different trend in NCs: SLC6A6 (B), CSRP1 (C), RAB30 (D), APOM (E), PPP2R1B (F), HMOX1 (G), TSR1 (H), EBNA1BP2 (I), and WDR77 (J)

possible molecular initiating events (MIEs), key events (KEs), and key event relationships (KERs) in AOP development. Currently AOP development requires massive manual curation to identify MIEs, KEs, and KERs. Furthermore, AOPs that have been suggested seem to be over-simplified to cover the whole process of adverse outcome development. Therefore, this study can help enrich the understanding of the adverse outcome development. Another possible application of this approach is to suggest biological descriptors and interpret the biological significance of the descriptors. Since hepatotoxicity is generally caused by reactive metabolites rather than drug molecule itself, sometimes molecular descriptors calculated from the parent compounds show a tendency of low correlation with the target endpoint; therefore, most of *in silico* models used complicated ML or DL algorithms to improve prediction accuracy with the price of renouncing the interpretability of the models. To overcome the limitation of molecular structure-based prediction models, toxicogenomics profiles were used in prediction model. Yen Low et al. used toxicogenomics profiles in hepatotoxicity prediction model development and achieved higher

prediction accuracy compared to the model developed with molecular descriptors alone [44]. As toxicogenomics profiles cover wide range of genes, descriptor selection is needed in order to make model building process efficient, and model interpretation also requires analysis of selected descriptors. The approach suggested in this work can be used both in selection and interpretation of the model. As Xi Chen et al. recently developed deep generative adversarial network (GAN) to generate toxicogenomics profile [45], GAN-driven toxicogenomics profile data can be increased, and thus, the semi-automated approach in this study is expected to be applied in wider range of chemicals.

Conclusions

Semi-automated approaches for data gathering, analysis, and biological network assembly can be valuable as they provide a robust and unbiased way to handle large amounts of diverse data. In this study, the approaches were applied to identify genes deregulated by the reference chemicals that cause the four liver diseases (cholestasis, steatosis, hepatitis, and cirrhosis). Biological networks of each disease was compiled based on the selected genes, and these were combined to deduce significant events in the progression of the four chemical-induced liver diseases. The assembled network can be used to identify important key events and their relationships in the development of four DILI diseases:

Table 4 Genes identified as of highest probability to be associated with the development of cirrhosis

Time	29 days	15 days	8 days	4 days	1 day	9 h	6 h	3 h
Genes	CSRP1	ABHD3	CROT	C1R	A1BG	ABHD3	CSF1R	CYP26B1
	DUSP6	CSRP1	HSD11B2	DUSP6	CYP26B1	BHLHE40	CYP26B1	NR1D1
	EGR1	CYP27A1	LPL	HBB	CYP27A1	CEBPB	GPNMB	NR1D2
	HBB	EBP	PTPRF		EGR1	CYP26B1		SLC16A10
	LPL	GPNMB			GPNMB	DUSP6		
	PHGDH	HBB			KCNJ8			
	PPP2R1B	LPL			KDR			
	S100A9	NCALD			LAMC2			
		PPP2R1B			LPL			
		S100A9			MGAT2			
					NARS			
					NR3C1			
					NT5E			
					PHLDA1			
					POLR3G			
					SLC1A4			
					TARS			
					WARS			

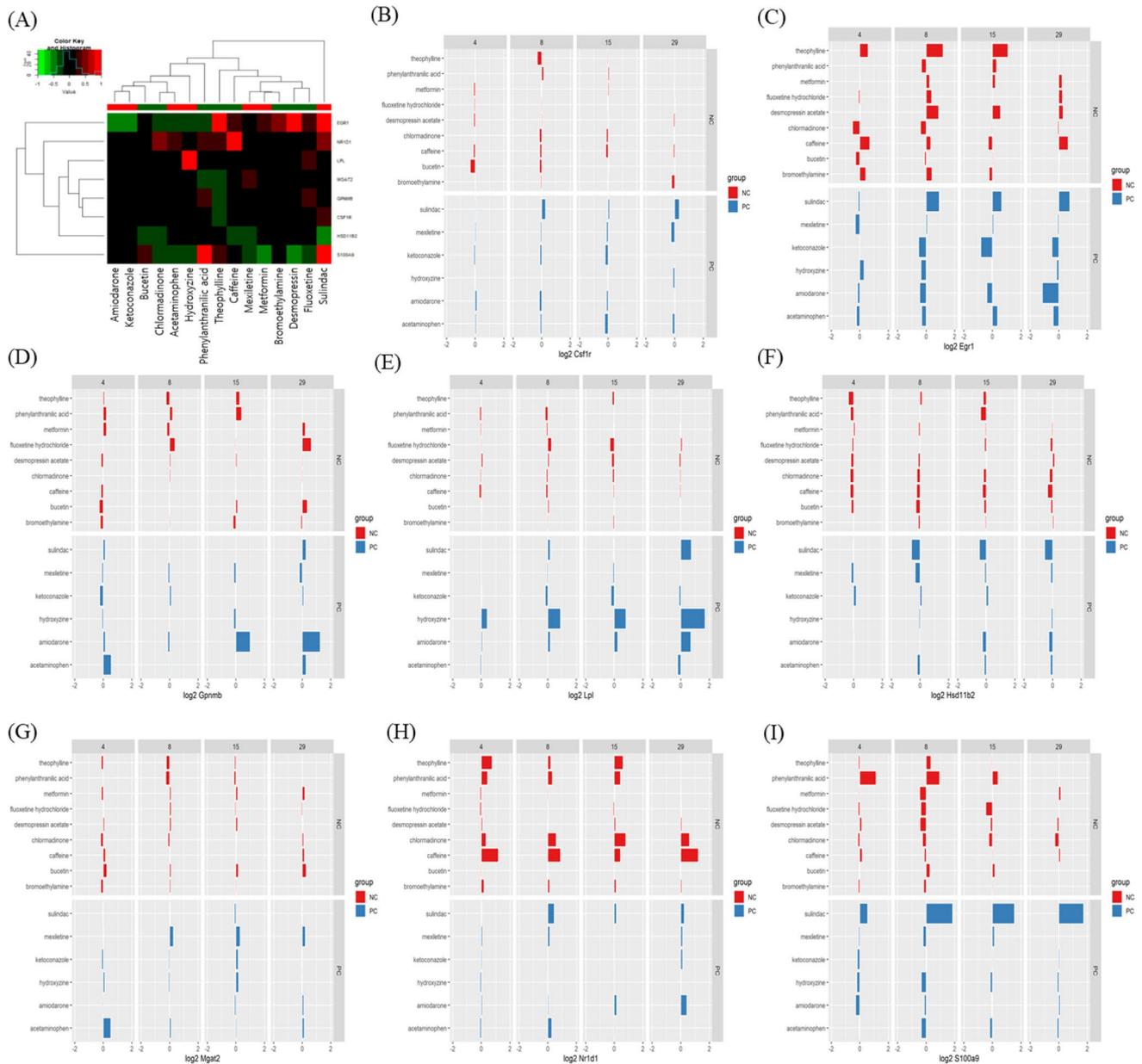


Fig. 7 Heat map (A) and bar chart (B–I) analyses of genes associated with drug-induced cirrhosis in comparison between positive compounds (PCs) and negative compounds (NCs). Nine genes were

consistently deregulated among PCs and displayed a different trend in NCs: CSF1R (B), EGR1 (C), GPNMB (D), LPL (E), HSD11B2 (F), MGAT2 (G), NR1D1 (H), and S100A9 (I)

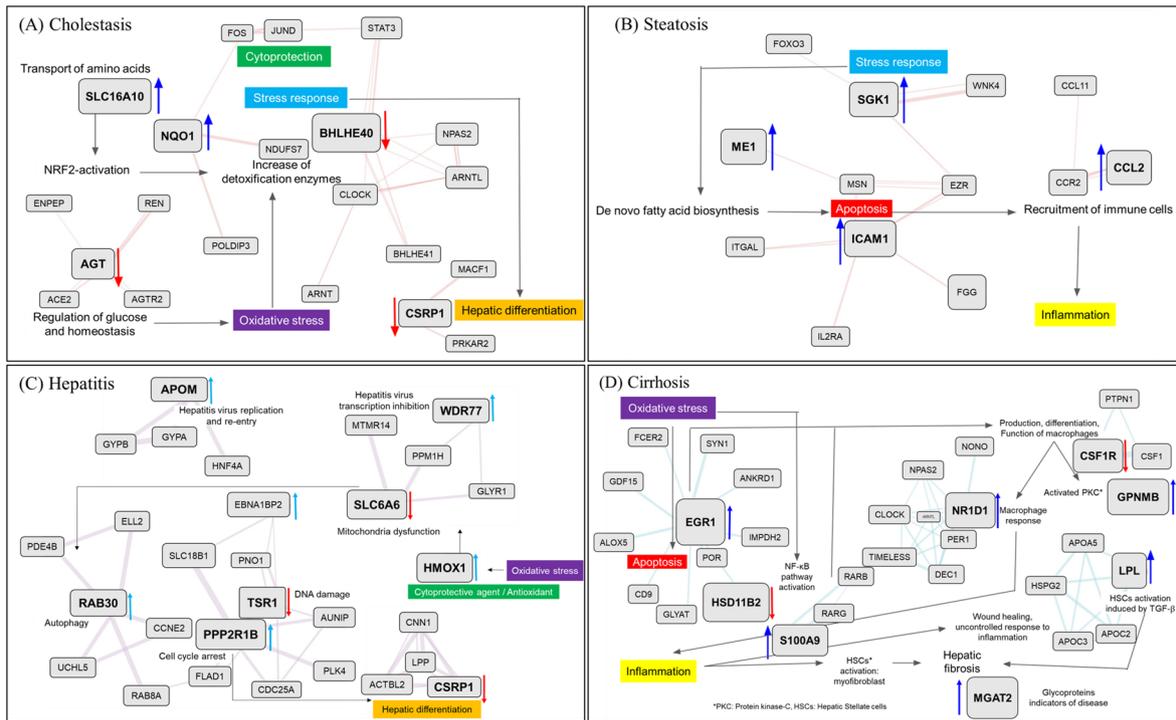


Fig. 8 Each biological network for cholestasis (A), steatosis (B), hepatitis (C), and cirrhosis (D)

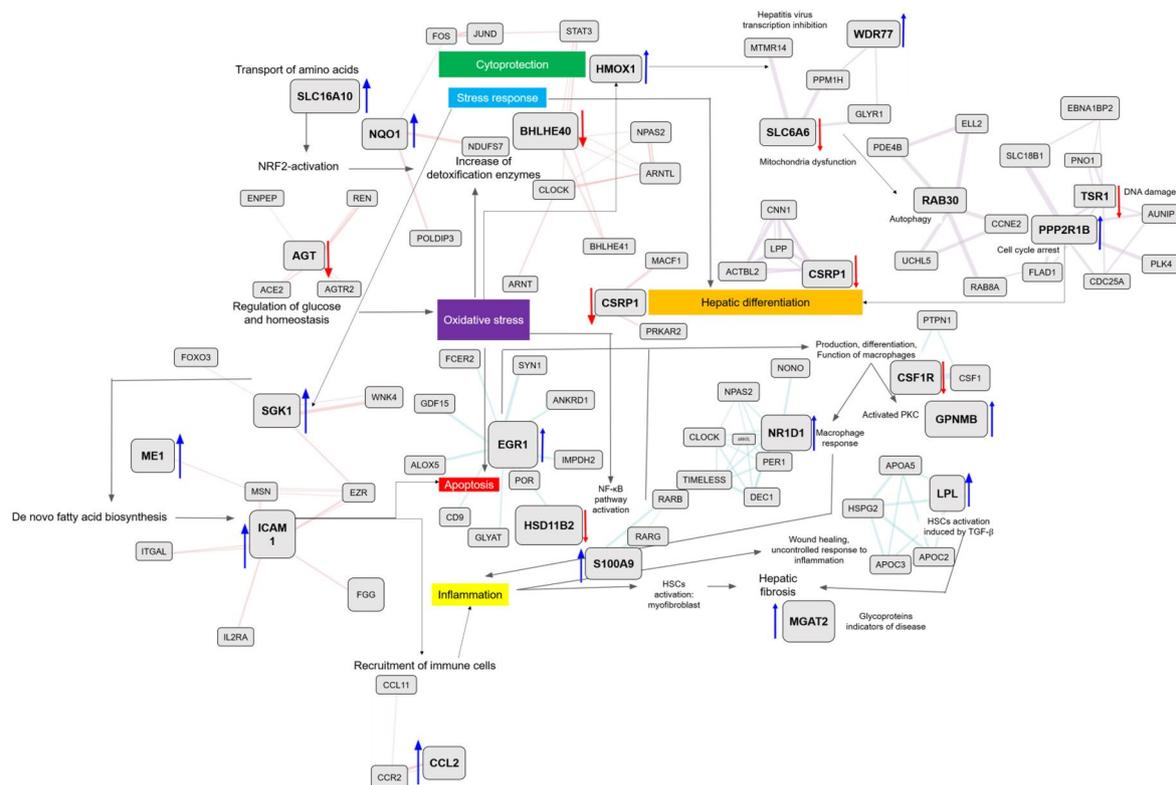


Fig. 9 Integrated biological networks on four liver diseases

cholestasis, steatosis, hepatitis, and cirrhosis. Suggested key events are basis for initial testing of chemicals responsible for the development of any of the diseases, and they can be further refined once the patterns of deregulation are consistently recorded.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s43188-022-00124-6>.

Author contributions Conceptualization, writing-original draft, and validation: HK and TD; Investigation and formal analysis: OF; Data curation: HK and OF; Methodology, software, and visualization: TD; Writing-editing and review: HK, TD, and MG; Project administration: BH; supervision and funding acquisition: MG.

Funding This study was supported by the National Research Foundation of Korea (NRF) funded by the Korean government (NRF-2019R1F1A1061955), and by the Korea Institute of Toxicology (KIT), Republic of Korea (1711133838).

Declarations

Conflict of interest The authors have no conflicts of interest to disclose.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

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

References

- Silbergeld EK, Mandrioli D, Cranor CF (2015) Regulating chemicals: law, science, and the unbearable burdens of regulation. *Annu Rev Public Health* 36:175–191. <https://doi.org/10.1146/annurev-publhealth-031914-122654>
- Van Norman GA (2019) Limitations of animal studies for predicting toxicity in clinical trials: is it time to rethink our current approach? *JACC Basic Transl Sci* 4:845–854. <https://doi.org/10.1016/j.jacbts.2019.10.008>
- Vall A, Sabnis Y, Shi J, Class R, Hochreiter S, Klambauer G (2021) The promise of AI for DILI prediction. *Front Artif Intell* 4:638410. <https://doi.org/10.3389/frai.2021.638410>
- Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 1:206–215. <https://doi.org/10.1038/s42256-019-0048-x>
- Vinken M (2013) The adverse outcome pathway concept: a pragmatic tool in toxicology. *Toxicology* 312:158–165. <https://doi.org/10.1016/j.tox.2013.08.011>
- Villeneuve DL, Crump D, Garcia-Reyero N, Hecker M, Hutchinson TH, LaLone CA, Landesmann B, Lettieri T, Munn S, Nepelska M, Ottinger MA, Vergauwen L, Whelan M (2014) Adverse outcome pathway (AOP) development I: strategies and principles. *Toxicol Sci* 142:312–320. <https://doi.org/10.1093/toxsci/kfu199>
- Hardt C, Bauer C, Schuchhardt J, Herwig R (2018) Computational network analysis for drug toxicity prediction. In: von Stechow L, Santos Delgado A (eds) *Computational cell biology: methods and protocols*. Springer New York, New York, pp 335–355. https://doi.org/10.1007/978-1-4939-8618-7_16
- Kim KT (2021) Application of the adverse outcome pathway framework to predict the toxicity of chemicals in the semiconductor manufacturing industry. *Mol Cell Toxicol* 17:325–345. <https://doi.org/10.1007/s13273-021-00139-4>
- Doktorova TY, Oki NO, Mohorič T, Exner TE, Hardy B (2020) A semi-automated workflow for adverse outcome pathway hypothesis generation: the use case of non-genotoxic induced hepatocellular carcinoma. *Regul Toxicol Pharmacol* 114:104652. <https://doi.org/10.1016/j.yrtph.2020.104652>
- Igarashi Y, Nakatsu N, Yamashita T, Ono A, Ohno Y, Urushidani T, Yamada H (2015) Open TG-GATEs: a large-scale toxicogenomics database. *Nucleic Acids Res* 43:D921–D927. <https://doi.org/10.1093/nar/gku955>
- Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q (2008) GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol* 9:S4. <https://doi.org/10.1186/gb-2008-9-s1-s4>
- Ramalingam L, Menikdiwela K, LeMieux M, Dufour JM, Kaur G, Kalupahana N, Moustaid-Moussa N (2017) The renin-angiotensin system, oxidative stress and mitochondrial function in obesity and insulin resistance. *Biochim Biophys Acta Mol Basis Dis* 1863:1106–1114. <https://doi.org/10.1016/j.bbdis.2016.07.019>
- Aleksunes LM, Manautou JE (2007) Emerging role of Nrf2 in protecting against hepatic and gastrointestinal disease. *Toxicol Pathol* 35:459–473. <https://doi.org/10.1080/01926230701311344>
- Wang G, Xiu P, Li F, Xin C, Li K (2014) Vitamin A supplementation alleviates extrahepatic cholestasis liver injury through Nrf2 activation. *Oxid Med Cell Longev* 2014:273692. <https://doi.org/10.1155/2014/273692>
- Kim DK, Kanai Y, Matsuo H, Kim JY, Chairoungdua A, Kobayashi Y, Enomoto A, Cha SH, Goya T, Endou H (2002) The human T-type amino acid transporter-1: characterization, gene organization, and chromosomal location. *Genomics* 79:95–103. <https://doi.org/10.1006/geno.2001.6678>
- Chung SY, Kao CH, Villarroja F, Chang HY, Chang HC, Hsiao SP, Liou G-G, Chen SL (2015) Bhlhe40 represses PGC-1 α activity on metabolic gene promoters in myogenic cells. *Mol Cell Biol* 35:2518–2529. <https://doi.org/10.1128/MCB.00387-15>
- Erdel M, Weiskirchen R (1998) Assignment of CSRPI encoding the LIM domain protein CRP1, to human chromosome 1q32 by fluorescence in situ hybridization. *Cytogenet Cell Genet* 83:10–11. <https://doi.org/10.1159/000015152>
- Liu L, Wang Q, Wang Q, Zhao X, Zhao P, Geng T, Gong D (2018) Role of miR29c in goose fatty liver is mediated by its target genes that are involved in energy homeostasis and cell growth. *BMC Vet Res* 14:325. <https://doi.org/10.1186/s12917-018-1653-3>
- Sierra-Ramos C, Velazquez-Garcia S, Vastola-Mascolo A, Hernández G, Faresse N, de la Alvarez D (2020) SGK1 activation exacerbates diet-induced obesity, metabolic syndrome and

- hypertension. *Endocrinology* 244:149–162. <https://doi.org/10.1530/JOE-19-0275>
20. Liang YJ, Jiang JG (2015) Characterization of malic enzyme and the regulation of its activity and metabolic engineering on lipid production. *RSC Adv* 5:45558–45570. <https://doi.org/10.1039/C5RA04635A>
 21. Lahey R, Carley AN, Wang X, Glass CE, Accola KD, Silvestry S, O'Donnell JM, Lewandowski ED (2018) Enhanced redox state and efficiency of glucose oxidation with miR based suppression of maladaptive NADPH-dependent malic enzyme 1 expression in hypertrophied hearts. *Circ Res* 122:836–845. <https://doi.org/10.1161/CIRCRESAHA.118.312660>
 22. Zhu BH, Zhang RH, Lv NN, Yang GP, Wang YS, Pan KH (2018) The role of malic enzyme on promoting total lipid and fatty acid production in *Phaeodactylum tricornutum*. *Front Plant Sci* 9:826. <https://doi.org/10.3389/fpls.2018.00826>
 23. Sookoian S, Castaño GO, Burgueño AL, Rosselli MS, Gianotti TF, Mallardi P, Martino JS, Pirola CJ (2010) Circulating levels and hepatic expression of molecular mediators of atherosclerosis in nonalcoholic fatty liver disease. *Atherosclerosis* 209:585–591. <https://doi.org/10.1016/j.atherosclerosis.2009.10.011>
 24. Haukeland JW, Damås JK, Konopski Z, Løberg EM, Haaland T, Goverud I, Torjesen PA, Birkeland K, Bjørø K, Aukrust P (2006) Systemic inflammation in nonalcoholic fatty liver disease is characterized by elevated levels of CCL2. *Hepatology* 44:1167–1174. <https://doi.org/10.1016/j.jhep.2006.02.011>
 25. Bansal S, Biswas G, Avadhani NG (2013) Mitochondria-targeted heme oxygenase-1 induces oxidative stress and mitochondrial dysfunction in macrophages, kidney fibroblasts and in chronic alcohol hepatotoxicity. *Redox Biol* 2:273–283. <https://doi.org/10.1152/ajprenal.00160.2013>
 26. Warskulat U, Borsch E, Reinehr R, Heller-Stilb B, Mönnighoff I, Buchczyk D, Donner M, Flögel U, Kappert G, Soboll S, Beer S, Pfeffer K, Marschall H-U, Gabrielsen M, Amiry-Moghaddam M, Ottersen OP, Dienes HP, Häussinger D (2006) Chronic liver disease is triggered by taurine transporter knockout in the mouse. *FASEB J* 20:574–576. <https://doi.org/10.1096/fj.05-5016jfe>
 27. Banworth MJ, Li G (2018) Consequences of Rab GTPase dysfunction in genetic or acquired human diseases. *Small GTPases* 9:158–181. <https://doi.org/10.1080/21541248.2017.1397833>
 28. Chou HC, Chen CH, Lee HS, Lee CZ, Huang GT, Yang PM, Lee PH, Sheu JC (2007) Alterations of tumour suppressor gene PPP2R1B in hepatocellular carcinoma. *Cancer Lett* 253:138–143. <https://doi.org/10.1016/j.canlet.2007.01.016>
 29. Upadhyay A, Dixit U, Manvar D, Chaturvedi N, Pandey VN (2013) Affinity capture and identification of host cell factors associated with hepatitis C virus (+) strand subgenomic RNA. *Mol Cell Proteomics* 12:1539–1552. <https://doi.org/10.1074/mcp.M112.017020>
 30. Gu JG, Zhu CI, Cheng DZ, Xie Y, Liu F, Zhou X (2011) Enhanced levels of apolipoprotein M during HBV infection feedback suppresses HBV replication. *Lipids Health Dis* 10:154. <https://doi.org/10.1186/1476-511X-10-154>
 31. Lubyova B, Hodek J, Zabransky A, Prouzova H, Hubalek M, Hirsch I, Weber J (2017) PRMT5: a novel regulator of hepatitis B virus replication and an arginine methylase of HBV core. *PLoS ONE* 12:e0186982. <https://doi.org/10.1371/journal.pone.0186982>
 32. Havis E, Duprez D (2020) EGR1 transcription factor is a multifaceted regulator of matrix production in tendons and other connective tissues. *Int J Mol Sci* 21:1664. <https://doi.org/10.3390/ijms21051664>
 33. Li TZ, Kim SM, Hur W, Choi JE, Kim J-H, Hong SW, Lee EB, Lee JH, Yoon SK (2017) Elk-3 contributes to the progression of liver fibrosis by regulating the epithelial–mesenchymal transition. *Gut Liver* 11:102–111. <https://doi.org/10.5009/gnl15566>
 34. Kostadinova RM, Nawrocki AR, Frey FJ, Frey BM (2005) Tumor necrosis factor alpha and phorbol 12-myristate-13-acetate down-regulate human 11 β -hydroxysteroid dehydrogenase type 2 through p50/p50 NF- κ B homodimers and Egr-1. *FASEB J* 19:1–30. <https://doi.org/10.1096/fj.04-2820fje>
 35. Escher G, Nawrocki A, Staub T, Vishwanath BS, Frey BM, Reichen J, Frey FJ (1998) Down-regulation of hepatic and renal 11 β -hydroxysteroid dehydrogenase in rats with liver cirrhosis. *Gastroenterology* 114:175–184. [https://doi.org/10.1016/S0016-5085\(98\)70645-6](https://doi.org/10.1016/S0016-5085(98)70645-6)
 36. Rojo R, Raper A, Ozdemir DD, Lefevre L, Grabert K, Wollscheid-Lengeling E, Bradford B, Caruso M, Gazova I, Sánchez A, Lisowski ZM, Alves J, Molina-Gonzalez I, Davtyan H, Lodge RJ, Glover JD, Wallace R, Munro DAD, David E, Amit I, Miron VE, Priller J, Jenkins SJ, Hardingham GE, Blurton-Jones M, Mabbott NA, Summers KM, Hohenstein P, Hume DA, Pridans C (2019) Deletion of a Csf1r enhancer selectively impacts CSF1R expression and development of tissue macrophage populations. *Nat Commun* 10:3215. <https://doi.org/10.1038/s41467-019-11053-8>
 37. Katayama A, Nakatsuka A, Eguchi J, Murakami K, Teshigawara S, Kanzaki M, Nunoue T, Hida K, Wada N, Yasunaka T, Ikeda F, Takaki A, Yamamoto K, Kiyonari H, Makino H, Wada J (2015) Beneficial impact of Gpnb and its significance as a biomarker in nonalcoholic steatohepatitis. *Sci Rep* 5:16920. <https://doi.org/10.1038/srep16920>
 38. Oishi Y, Hayashi S, Isagawa T, Oshima M, Iwama A, Shimba S, Okamura H, Manabe I (2017) Bmal1 regulates inflammatory responses in macrophages by modulating enhancer RNA transcription. *Sci Rep* 7:7086. <https://doi.org/10.1038/s41598-017-07100-3>
 39. Wu R, Zhang Y, Xiang Y, Tang Y, Cui F, Cao J, Zhou L, You Y, Duan L (2018) Association between serum S100A9 levels and liver necroinflammation in chronic hepatitis B. *Transl Med* 16:83. <https://doi.org/10.1186/s12967-018-1462-2>
 40. Teratani T, Tomita K, Furuhashi H, Sugihara N, Higashiyama M, Nishikawa M, Irie R, Takajo T, Wada A, Horiuchi K, Inaba K, Hanawa Y, Shibuya N, Okada Y, Kurihara C, Nishii S, Mizoguchi A, Hozumi H, Watanabe C, Komoto S, Nagao S, Yamamoto J, Miura S, Hokari R, Kanai T (2019) Lipoprotein lipase up-regulation in hepatic stellate cells exacerbates liver fibrosis in nonalcoholic steatohepatitis in mice. *Hepatology* 3:1098–1112. <https://doi.org/10.1002/hep4.1383>
 41. Yang M, Nickels JT (2015) MOGAT2: a new therapeutic target for metabolic syndrome. *Diseases* 3:176–192. <https://doi.org/10.3390/diseases3030176>
 42. Villanueva-Paz M, Morán L, López-Alcántara N, Freixo C, Andrade RJ, Lucena MI, Cubero FJ (2021) Oxidative stress in drug-induced liver injury (DILI): from mechanisms to biomarkers for use in clinical practice. *Antioxidants* 10:390. <https://doi.org/10.3390/antiox10030390>
 43. Iorga A, Dara L (2019) Chapter Two—cell death in drug-induced liver injury. In: Ramachandran A, Jaeschke H (eds) *Advances in Pharmacology*, vol 85. Academic Press, New York, pp 31–74. <https://doi.org/10.1016/bs.apha.2019.01.006>
 44. Low Y, Uehara T, Minowa Y, Yamada H, Ohno Y, Urushidani T, Sedykh A, Muratov E, Kuz'min V, Fourches D, Zhu H, Rusyn I, Tropsha A (2011) Predicting drug-induced hepatotoxicity using QSAR and toxicogenomics approaches. *Chem Res Toxicol* 24:1251–1262. <https://doi.org/10.1021/tx200148a>
 45. Chen X, Roberts R, Tong W, Liu Z (2021) Tox-GAN: an artificial intelligence approach alternative to animal studies—a case study with toxicogenomics. *Toxicol Sci*. <https://doi.org/10.1093/toxsci/kfab157>