Transcriptomics-based prediction of human hepatotoxic blood concentrations of chemicals.

DISSERTATION

ZUR ERLANGUNG DES AKADEMISCHEN GRADES DES DOKTORS
DER NATURWISSENSCHAFTEN (DR. RER. NAT.)

DER CHEMISCHEN FALKULTÄT
DER TECHNISCHEN UNIVERSITÄT DORTMUND

VORGELEGT VON

Regina Stöber, M.Sc.

DORTMUND 2016

1. GUTACHTER: PROF. DR. JAN G. HENGSTLER
2. GUTACHTER: PROF. DR. FRANK WEHNER

Für meine Familie

Eidesstattliche Versicherung (Affidavit)

Name, Vorname
(Surname, first name)

Matrikel-Nr.
(Enrolment number)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.
Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Table of contents

Summary

Drug-induced liver injury represents one of the most critical issues during drug development and leads to failure of many drug candidates in preclinical or clinical studies. Currently, the common model for safety evaluation and human health risk assessment is repeated dose toxicity (RDT) testing in rodents. However, RDT studies require numerous animals and the capacity for this conventional testing is limited. There is an urgent need for the development of novel test systems, where complex *in vivo* processes and different mechanisms of toxicity can be addressed. In recent years, numerous research groups have focused on the identification and development of biomarkers of hepatotoxicity. In this context, genomic approaches are used to identify patterns in mRNA expression changes, referred to as toxicogenomics. Emerging databases provide a vast amount of transcriptomics data from compound-exposed hepatocytes, as well as rodent livers. This large amount of publically available genome wide expression data provides valuable information for the identification and development of novel biomarkers of hepatotoxicity. However, a comprehensive analysis summarizing the key features of chemically-influenced gene expression has not yet been performed.

The first part of this thesis focusses on the definition of key principles of global expression alterations in compound-exposed hepatocytes. Therefore, genome wide expression data from the Open Toxicogenomics Project-Genomics Assisted Toxicity Evaluation System (TG-GATES) database were used. This database comprises gene array data from primary human hepatocytes that were incubated with 150 compounds for several time points and concentrations. Before analyzing the structure of the database, a number of curation steps were performed to improve the data set. Genes were only considered to be up or down regulated when the mean alteration was at least 3 fold compared to the untreated control condition. Furthermore, the concentration progression of each compound was analyzed and compounds that followed an implausible concentration progression were excluded from the data set. With the final optimized dataset, a toxicotranscriptomics directory was developed, which indicates whether a particular gene is altered upon chemical exposure. If there are gene expression changes, the type and number of compounds inducing this change could also be identified. The directory further provides information on whether a gene is also altered in human liver diseases, such as hepatocellular carcinoma (HCC), non-alcoholic steatohepatitis (NASH) or cirrhosis, thus implying *in vivo* relevance. Genes that are influenced by the hepatocyte isolation and cultivation procedures are highlighted and defined as unstable baseline genes. Finally biomarker candidates were chosen that are altered by a large set of chemicals that simultaneously overlap with those deregulated in liver diseases, but not by the hepatocyte isolation and cultivation procedures. Based on these criteria, the toxicogenomics directory was used to identify a set of seven potential biomarker candidates: The cytochrome P450 isoenzymes 1B1 (CYP1B1) and 3A7 (CYP3A7), the cytoskeletal protein tubulin 2B (TUBB2B), sulfotransferase 1C2 (SULT1C2), the stress response gene FBXO32, regu-

lator of cell cycle (RGCC), and glucose-6 phosphate dehydrogenase (G6PD). These genes cover a broad range of toxicological motifs, such as the metabolism of xenobiotics, energy and lipid metabolism, cytoskeleton, cell cycle and protein degradation.
The second part of this thesis focusses on the applicability of the selected genes to predict human hepatotoxicity. In a pilot study, a biomarker and cytotoxicity-based *in vitro* system was established, which predicts human blood concentrations that are associated with an increased risk of hepatotoxicity. A set of 12 hepatotoxic compounds, as well as 9 non-hepatotoxic compounds were identified. The former are associated with an increased risk of hepatotoxicity when administered at therapeutic doses; whereas, the latter are considered harmless in the therapeutic concentration range. For each compound, a literature search was performed to identify the resulting blood concentrations from therapeutic doses. HepG2 cells, as well as primary human hepatocytes were treated with each compound in a concentration range that included the peak plasma concentration identified for the therapeutic dose, in addition to doses that resulted in a slightly cytotoxic concentration. Two readouts – biomarker expression and cytotoxicity tests – were used to identify critical concentrations *in vitro*. The lowest observed effect concentrations (LOECs) *in vitro* were finally compared to peak plasma concentrations of therapeutic doses *in vivo*.

In HepG2 cells, the biomarker-based *in vitro* system was able to adequately discriminate between the two sets of compounds. The prediction sensitivity improved in primary human hepatocytes, because the model was able to identify hepatotoxic effects at even lower concentrations. The results revealed that for a large amount of compounds, the *in vitro* model precisely predicted human blood concentrations that are associated with an increased risk of hepatotoxicity. However, the model is not yet applicable to all compounds, because for many of them it still underestimates the risk of hepatotoxicity. Future studies should identify further biomarkers that are able to capture more compounds and allow a more precise prediction.

Based on the so far available biomarkers, the presented model allows for an approximation whether a therapeutic dose would be associated with a high or a low risk of hepatotoxicity *in vivo*. Although it is still in its developmental stage, the model shows promise as it identifies a number of idiosyncratic hepatotoxic compounds, which are distinguishable from non-hepatotoxic compounds. The clustering within the set of hepatotoxic or non-hepatotoxic compounds allows the estimation of the hepatotoxic potential of an unknown compound. In conclusion, the novel prediction system represents a promising tool to assess a putative risk of hepatotoxicity for unknown compounds and provides valuable knowledge that contributes to, for example the ranking and prioritization of compounds in early drug development.

Zusammenfassung

Eine der größten Herausforderungen bei der Entwicklung neuer Medikamente sind Chemikalien-induzierte Leberschäden. Oftmals wird neben dem gewünschten therapeutischen Effekt auch Lebertoxizität beobachtet, wodurch vielversprechende Kandidaten während der vorklinischen und klinischen Phase scheitern oder auch häufig nach der Zulassung noch vom Markt genommen werden. In der gängigen Praxis werden zur Sicherheitseinstufung und Risikobewertung von Medikamenten vor allem Tiermodelle genutzt, bei welchen die Tiere nach wiederholter Applikation auf Anzeichen von Toxizität untersucht werden. Dieses Verfahren bedarf jedoch einer sehr großen Anzahl an Tieren, ist sehr kostenintensiv und übersteigt die Prüfkapazität für neue Substanzen um ein Vielfaches. Demnach stellt die Entwicklung neuartiger Testsysteme eine dringende Notwendigkeit dar.

In den letzten zwei Jahrzehnten konzentriert sich ein Großteil der Forschungsvorhaben auf die Identifizierung und die Entwicklung von Biomarkern, welche einen hepatotoxischen Effekt frühzeitig signalisieren. In diesem Zusammenhang stellt die Entwicklung von *–omics* Technologien, insbesondere Toxicogenomics, einen prominenten Ansatz dar. Genomweite Analysen werden herangezogen, um Muster in Chemikalien-induzierten Genexpressionsveränderungen zu detektieren. Transkriptomdaten von Substanz-exponierten Zellen und Nagetier-Lebern sind in Datenbanken im Internet öffentlich zugänglich und bieten einen großen Informationspool für die Entwicklung genomischer Biomarker. Um diese umfangreichen Datenmengen jedoch optimal für die Entwicklung neuartiger Biomarker nutzen zu können, ist das Verständnis von Schlüsseleigenschaften Chemikalien-induzierter Genexpressionsveränderungen von großem Vorteil. Dennoch gibt es bislang keine umfassenden Studien, die sich mit typischen Merkmalen Chemikalien-induzierter Transkriptionsveränderungen beschäftigen.

Um ein generelles Verständnis globaler Expressionsveränderungen in Substanz-exponierten Hepatozyten zu erlangen, beschäftigt sich der erste Teil dieser Doktorarbeit mit der Definition von Schlüsselprinzipien, welche Chemikalien-induzierten Transkriptionsmustern unterliegen. Dazu wurden globale Expressionsstudien der öffentlichen Datenbank *Toxicogenomics Project-Genomics Assisted Toxicity Evaluation System* (TG-GATES) herangezogen, in welcher *Gene Array* Daten primärer humaner Hepatozyten von 150 getesteten Substanzen zusammengefasst sind. Um besagte Schlüsselmerkmale optimal herausarbeiten zu können, wurden zunächst einige Optimierungsschritte am Datensatz vorgenommen. Es wurden z.B. nur Gene als hoch- oder herunter reguliert betrachtet, wenn eine Deregulation im Vergleich zur unbehandelten Kontrolle um mindestens den Faktor 3 vorlag. Weiterhin wurden die Konzentrationsverläufe aller Substanzen analysiert. Substanzen, die beispielsweise Gene bei einer niedrigen, nicht jedoch einer höheren Konzentration deregulieren, weisen einen unlogischen Konzentrationsverlauf auf und wurden von der weiteren Analyse ausgeschlossen. Mit dem optimierten Datensatz wurde anschließend ein Toxicotranskriptom-Verzeichnis entwickelt. Dieses gibt Auskunft darüber, ob ein Gen durch Chemikalien beeinflusst wird und wenn ja, durch wie viele und welche Art von Substanzen. Weiterhin werden Gene gekennzeichnet,

deren Expression auch in Leberkrankheiten wie Zirrhose, hepatozelluläres Karzinom oder bei einer nicht-alkoholischen Fettleber verändert ist. Ein derartiger Überlapp impliziert eine mögliche Relevanz des Gens *in vivo* und minimiert die Wahrscheinlichkeit, dass sich bei der Chemikalien-induzierten Expressionsveränderung um einen *in vitro* Artefakt handelt. Gene, welche durch die Isolierungs- und Kultivierungsbedingungen beeinflusst werden, sind ebenfalls hervorgehoben.

Gene, welche von möglichst vielen unterschiedlichen Substanzen dereguliert werden, ebenfalls in Leberkrankheiten verändert sind, jedoch nicht durch die Isolierungs- und Kultivierungsbedingungen beeinflusst werden, repräsentieren potentielle Biomarker-Kandidaten. Basierend auf diesen Kriterien wurde das Toxicotranskriptom-Verzeichnis genutzt, um sieben mögliche Kandidaten zu identifizieren: Die Cytochrom P450 Isoenzyme CYP1B1 und CYP3A7, das zytoskeletale Protein Tubulin 2 B (TUBB2B), die Sulfotransferase SULT1C2, das Stress-induzierte Gen FBXO32, das Zellzyklus-regulierende Protein RGCC und das Gen der Glucose-6-Phosphat Dehydrogenase (G6PD). Diese Gene decken eine Vielzahl möglicher Toxizitätsmechanismen ab, nämlich den Metabolismus von Xenobiotika, den Energie- und Lipidstoffwechsel, das Zytoskelett, den Zellzyklus und den Abbau von Proteinen.

Der zweite Teil dieser Arbeit konzentriert sich auf eine mögliche Anwendbarkeit der ausgewählten Gene, um humane Hepatotoxizität vorher zu sagen. In einer Teststudie wurde ein Biomarker- und Zytotoxizität-basiertes *in vitro* System entwickelt, was die Vorhersage humaner Blutkonzentrationen ermöglicht, welche mit einem erhöhten Risiko für Lebertoxizität assoziiert sind. Dazu wurden 12 hepatotoxische sowie 9 nicht-hepatotoxische Substanzen ausgewählt. Hepatotoxische Substanzen weisen bei therapeutisch wirksamer Dosierung ein erhöhtes Risiko für Lebertoxizität auf, während bei nicht-hepatotoxischen Substanzen in dieser Konzentrationsspanne keine Gefahr für einen Leberschaden besteht. Für alle Substanzen wurden im Rahmen einer Literaturrecherche die Plasmakonzentrationen einer therapeutischen Dosis identifiziert. HepG2 Zellen sowie primäre humane Hepatozyten wurden mit den jeweiligen Substanzen inkubiert, wobei sowohl therapeutisch wirksame, bis hin zu leicht zytotoxischen Konzentrationen getestet wurden. Um die jeweils niedrigste Konzentration zu ermitteln, bei welcher *in vitro* ein hepatotoxischer Effekt auftritt, wurden sowohl Zytotoxizitätsexperimente durchgeführt, als auch die Expression der ausgewählten Biomarker Gene analysiert. Anschließend wurden diese kritischen Konzentrationen *in vitro* mit der Plasmakonzentration einer therapeutischen Dosis *in vivo* verglichen.

Sowohl für HepG2 Zellen, als auch in primären humanen Hepatozyten, konnte eine Separierung hepatotoxischer und nicht-hepatotoxischer Medikamente beobachtet werden. Mit den primären Zellen wurde zudem eine wesentlich sensitivere Vorhersagbarkeit für einen möglichen Leberschaden erzielt, da hepatotoxische Effekte *in vitro* bereits bei niedrigeren Konzentrationen auftraten. Erste Ergebnisse zeigen, dass das beschriebene *in vitro* Modell bereits für eine große Anzahl an Substanzen humane Blutkonzentrationen relativ genau vorhersagen kann, bei denen ein erhöhtes Risiko für einen Leberschaden besteht. Dennoch ist das Modell noch nicht vollständig ausgereift und für alle Substanzen anwendbar, da es das Risiko einer hepatotoxischen Wirkung für manche Medikamente noch unterschätzt. Zukünftige Experimente werden sich mit der Identifizierung weiterer Biomarker beschäftigen, die

einen weiteren Bereich an Substanzen abdecken und eine genauere Vorhersagbarkeit ermöglichen.

Basierend auf den bisher vorliegenden Biomarkern ist das entwickelte *in vitro* Modell in der Lage, einzuschätzen, ob eine therapeutisch wirksame Dosis eines Medikaments mit einem hohen oder einem niedrigen Risiko für einen Leberschaden einhergeht. Die Gruppierung innerhalb der Klasse hepatotoxischer oder nicht-hepatotoxischer Substanzen kann dafür genutzt werden, das Risiko für einen hepatotoxischen Effekt einer noch unbekannten Substanz abzuschätzen. Demnach stellt das im Rahmen dieser Arbeit entwickelte Modell einen erfolgreichen Ansatz dar, um bei der Entwicklung neuer Medikamente vielversprechende Kandidaten zu sondieren und somit das Risiko für einen möglichen Leberschaden zu minimieren.

Abbreviations

AA	Allyl alcohol
ABC	ATP-binding cassette
ADME	Absorption, distribution, metabolism, excretion
AFB1	Aflatoxin B1
Akt	Protein kinase B
ALDH	Alcohol dehydrogenase
ALT	Alanine aminotransferase
ALP	Alkaline phosphatase
ANGPTL4	Angiopoietin-like 4
APAP	Acetaminophen
ASP	Aspirin
AST	Aspartate aminotransferase
ATP	Adenosine triphosphate
ATF3	Activating transcription factor 3
AXL	AXL receptor tyrosine kinase
BEA	Bromoethylamine
BCL2A1	BCL2-related protein A1
BPR	Buspirone
Bsep	Bile salt export pump
Ca	Calcium
CBX4	E3 SUMO-protein ligase CBX4
CBZ	Carbamazepine
CCL2	Chemokine (C-C motif) ligand 2
CCl_4	Carbon tetrachloride
CCNE2	Cyclin E2
CDK	Cycline dependent kinase
CDKN2C	Cyclin-Dependent Kinase Inhibitor 2C
cDNA	Coding deoxyribonucleic acid
CHL	Chlorpheniramine
CHX	Cycloheximide
CLON	Clonidine
CLRN	Clarin
CoA	Coenzyme A
CO_2	Carbon dioxide
CPS	Carbamoyl phosphatate synthase
Ct	Cycle threshold
CUX2	Cut-Like Homeobox 2
CYP	Cytochrome P450 enzymes

DEPC	Diethylpyrocarbonate
DFN	Diclofenac
DHFR	Dihydrofolate reductase
DILI	Drug-induced liver injury
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
Dntp	Di-deoxyribonucleic acid
EDTA	Ethylene diamine tetra acidic acid
EFNA1	Ephrin-A1
EGTA	Ethylene glycol tetraacetic acid
EtOH	Ethanol
FAM	Famotidine
FBXO32	F-Box Protein 32
FC	Fold change
FCS	Fetal calf serum
FDR	False discovery rate
FRET	Fluorescence resonance energy transfer
g	Standard gravity
g	Gram
G6PD	Glucose-6-hosphate dehydrogenas
GAPDH	Glycerinaldehyd-3-phosphat-Dehydrogenase
GATA	Erythroid transcription factor also known as GATA-binding factor 1
GDF15	Growth differentiation factor 15
GO	Gene Ontology
h	Hour
H_2O	Water
HCC	Hepatocellular carcinoma
HNF4	Hepatocyte nuclear factor 4
HNMT	Histamine N-methyl transferase
HOGA1	4-hydroxy-2-oxoglutarate aldolase
HSPA6	Heat shock protein 6
HYZ	Hydroxyzine
ID1	Inhibitor Of DNA Binding 1
INAH	Isoniazid
INSIG	Insulin-induced gene
KC	Ketoconazole
KCl	Potassium chloride
KCNJ8	Potassium channel, inwardly rectifying subfamily J, member 8
kg	Kilogram
KH_2PO_4	Potassium dihydrogen phosphate
L	Liter

LAB	Labetalol
LEV	Levofloxacine
LOEC	Lowest observed effect concentration
LPS	Lipopolysaccharide
M	Molar
MEF2	Myocyte enhancer factor 2
MEL	Melatonin
mg	Milligram
min	Minute
mL	Milliliter
mM	Millimolar
mRNA	Messenger RNA
NaCl	Sodium chloride
NADPH	Nicotinamide adenine dinucleotide phosphate
NaH_2PO_4	Sodium dihydrogen phosphate
NaOH	sodium hydroxide
NASH	Non-alcoholic steatohepatitis
NAT	N-acetyltransferase
NF-kB	Nuclear factor kappa B
NFT	Nitrofurantoin
NIM	Nimesulide
nM	Nanomolar
n-Myc	n-Myc proto-oncogenic transcription factor
NREP	Neuronal regeneration related protein
NSAID	Non-steroidal anti-inflammatory drug
PBLD	phenazine biosynthesis-like protein domain containing protein
PBS	Phosphate buffered saline
PCA	Principal component analysis
PCK	Phosphoenolpyruvate carboxykinase
PCR	Polymerase chain reaction
PDK	Pyruvate dehydrogenase kinase
PhB	Phenylbutazone
PHH	Primary human hepatocytes
PHO	Phorone
PI3K	Phosphoinositide 3 kinase
pM	Picomolar
PMZ	Promethazine
PPL	Propranolol
PPM1L	protein phosphatase, Mg2+/Mn2+ dependent 1L
PS	Probe set
qRT-PCR	Quantitative real time PCR
RAR	Retinoic acid receptor

RDT	Repeated dose toxicity
Rep	Replicate
RGCC	Regulator of cell cycle
RIF	Rifampicin
RMA	Robust Multi-Array Average
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
ROS	Reactive oxygen species
RXR	Retinoic X receptor
sec	Second
SD	Standard deviation
SLC	Solute carrier
SULT	Sulfotransferase
SV	Selection value
TCGA	The Cancer Genome Atlas
THRSP	Thyroid hormone responsive
TFBS	Transcription factor binding sides
TG-GATES	Toxicogenomics Project-Genomics Assisted Toxicity Evaluation System
TMD	Trimethadione
TNFSF	Tumor necrosis factor superfamily
TOP2A	Topoisomerase 2
TRM22	Putative transposase of insertion sequence ISRm22 protein
TUBB	Tubulin
TXNIP	Thioredoxin interacting protein
TZM	Triazolam
U	Unit
UGT2B15	UDP glucuronosyltransferase 2 family, polypeptide B15
VPA	Valproic acid
WDR72	WD Repeat Domain 72
ZCCHC6	Zinc finger, CCHC domain containing 6
µg	Microgram
µM	Micromolar

1 Introduction

1.1 Critical aspects of drug-induced liver injury and models for hepatotoxicity prediction

The liver represents the central organ of metabolism and detoxification in the body (Bandara and Kennedy 2002). Its primary functions comprise intermediary and energetic metabolism, as well as biotransformation of various substances, which makes the liver the major target of drug toxicity (Gomez-Lechon et al. 2010). Consequently, drug-induced liver injury (DILI) is one of the most critical issues during drug development and leads to failure of many drug candidates during preclinical or clinical studies (Jaeschke et al. 2002). In addition, hepatotoxicity is the main reason for drug withdrawal from the market. It is a reported side effect of more than 900 drugs and is responsible for 5 % of all hospital admissions and for 50 % of all acute liver failures, (Ostapowicz et al. 2002; Wilke et al. 2007; Pandit et al. 2012). Unfortunately, up to 10 % of DILI patients will develop jaundice and eventually die (Navarro and Senior 2006). For this reason, a major goal of the pharmaceutical industry is to market safer drugs with less adverse effects, predictable pharmacokinetic properties and quantifiable drug-drug interactions. In order to achieve this, the evaluation of potential hepatotoxic effects represents a critical step in drug development (Gomez-Lechon et al. 2010).

During the last decades, several animal models have been used to study cytological, physiological, metabolic and morphological endpoints to illustrate clinical and pathophysiological injury (Suter et al. 2004). Among the most frequently used tools in preclinical evaluation are two year repeated dose toxicity rodent studies, as well as conventional toxicity tests, which focus on transaminase levels and histopathological findings (Cheng et al. 2011). Currently, animal *in vivo* studies represent the best model to mimic the physiological microenvironment in humans, but do not allow high-throughput screenings with a large number of compounds. For practical, as well as ethical reasons, only a small number of preselected compounds can be examined *in vivo* (Cheng et al. 2011). In addition, screening large sets of chemicals is limited due to high costs, and the large number of animals and extensive time needed to conduct such experiments.

However, due to interspecies differences in hepatocellular function, pharmacokinetics, as well as administration, distribution, metabolism and excretion (ADME) information for a particular test compound or drug gained from animal models cannot simply be transferred to the human system. It is estimated that preclinical evaluation of drug candidates using conventional clinical pathology and animal testing fails to detect up to 40% of potentially hepatotoxic drugs in humans (Aubrecht et al. 2013). Since human *in vivo* studies cannot be performed for ethical reasons, human hepatocyte *in vitro* systems are frequently applied to mimic the human *in vivo* situation. These *in vitro* models generally comprise immortalized human hepatic cell lines, such as HepG2 or HepaRG cells, primary hepatocytes, liver slices, stem cell derived hepatocytes and 3D systems, such as liver spheroids. Furthermore, co-

cultures with non-parenchymal liver cells are used to enable cross-talk between hepatocytes and further liver cells, thus stabilizing hepatic functionality and thereby minimizing discrepancies between *in vitro* and *in vivo* models (Jiang et al. 2015; Nibourg et al. 2012). These systems offer the possibility to investigate specific parameters in a controlled environment (Tuschl et al. 2008). On the other hand, *in vitro* test systems do not fully reflect systemic influences and hepatocellular toxic effects, such as transaminase induction and toxicity related to *in vivo* metabolites or mitochondrial dysfunction (Cheng et al. 2011).

To reduce the number of animals and to minimize the risk of hepatotoxicity in humans, the early detection of drug-induced hepatotoxicity is essential before compounds are tested in animals or clinical trials (O'Brien et al. 2006). In the current clinical practice, liver injury is detected by measuring circulating molecules, indicating alterations in liver function and homeostasis, or changes in tissue or cell integrity (Aubrecht et al. 2013). These biomarkers encompass for example, total bilirubin, total bile acids, alanine- (ALT) and aspartate aminotransferase (AST) levels, alkaline phosphatase (ALP), lactate dehydrogenase and γ-glutamyl transpeptidase concentrations (Aubrecht et al. 2013; Giannini et al. 2005; Navarro and Senior 2006). However, these clinical biomarkers detect liver injury only after a significant injury has already occurred, but not before liver function is compromised. Total bilirubin levels for example, increase only after the liver has lost approximately half of its excretory capacity (Navarro and Senior 2006). In general, the listed biomarkers are often sensitive, but not necessarily specific for the target organ. Some markers are more sensitive than others or are elevated by non-hepatic injury. For example, ALT is not necessarily specific for liver injury and ALT levels do not always correlate to the extent of liver injury (Sun et al. 2014; Yang et al. 2012).

Since the currently-available toxicity test systems are obviously not sufficient to predict human hepatotoxicity, and because current serum markers indicate hepatotoxicity only at a progressed state of liver injury, there is an urgent need for novel tools to predict human hepatotoxicity. Ideally, new test systems should be robust, cheaper, faster and more convenient for screening than the so far available test systems and cover even complex *in vivo* processes, such as ADME and different mechanisms of toxicity. The overall aim of this work is therefore to identify novel biomarkers which are organ specific and can identify the hepatotoxic potential of compounds prior to the development of clinical signs. Optimally, these biomarkers are applicable in *in vitro* systems to predict the risk of hepatotoxicity of a particular compound *in vivo*.

1.2 Toxicogenomics for the identification of novel biomarkers of toxicity

Technological advances in the field of omics technologies have shown promise in the area of biomarker development. Genomics, proteomics and metabolomics play an important role in uncovering novel biochemical pathways and are used in preclinical animal studies, as well as clinical investigations to evaluate markers of hepatotoxicity in tissues and in easily-obtained body fluids, such as urine or serum (Yang et al. 2012).

While proteomics are used as a tool to identify cytokines and cellular stress markers of hepatotoxicity, metabolomics analyze for example, bile acid metabolism or hepatic glutathione depletion (Yang et al. 2012). Another important aspect is the use of genomics, particularly toxicogenomics to get insight into the molecular mechanisms of drug toxicity.

Toxicogenomics combines conventional toxicology with genomics and bioinformatics to study adverse effects of chemicals. Genome wide expression data are analyzed for gene expression changes that influence, predict or help to define drug toxicity (Suter et al. 2004). Although the relationship between changes of gene expression and adverse effects *in vivo* are not yet fully understood, the evaluation and characterization of differentially expressed genes in chemically-exposed cells can be used to predict toxicologic outcomes and to identify mechanisms of action. Several studies demonstrate that compounds, which cause the same toxic end points, also generate a unique gene expression pattern (Gomez-Lechon et al. 2010). For example, a previous study successfully showed the usefulness of clustering hepatotoxins by gene expression profiling (Ellinger-Ziegelbauer et al. 2008). In this study gene expression profiles of drug-exposed rat livers were analyzed and the authors were able to distinguish between genotoxic and non-genotoxic carcinogens. Another study has shown that gene expression profiles can discriminate between hepatotoxic and non-hepatotoxic compounds in rats (Zidek et al. 2007). This study identified a set of marker genes, which reflected typical hepatotoxic responses and allowed the prediction of compound classes. Furthermore, differently acting hepatotoxins can be distinguished according to their gene expression profile. This was for example shown by a research group who identified highly discriminating genes which differentiated between enzyme inducing compounds and peroxisome proliferators in exposed rat livers (Hamadeh et al. 2002a; Hamadeh et al. 2002b). However, although the results of the aforementioned studies are promising, it is unclear how relevant biomarkers identified in *in vivo* animal models are representative of the situation in the human liver.

Besides the rodent *in vivo* studies, other groups focused on toxicogenomics-based hepatotoxicity prediction in human hepatic cell lines. Cha et al. identified 77 specific genes, which may be indicative of early, as well as the later onset of non-steroidal anti-inflammatory drug (NSAID) - induced hepatotoxicity in HepG2 cells (Cha et al. 2010). A set of hepatotoxic and non-hepatotoxic compounds were used to validate the model and 100 % prediction accuracy was achieved. However, gene expression of HepG2 cells does not represent the real situation of gene expression in the human liver *in vivo* and the reliability of the prediction system has not yet been confirmed in primary hepatocytes or *in vivo* studies (Godoy et al. 2013).

Currently, the best available modeling systems to identify novel biomarkers for the prediction of hepatotoxicity include a combination of *in vivo* animal data and data from exposed human cell lines and cultivated primary cells *in vitro*. It is a long term goal to identify biomarkers in *in vitro* systems, which are capable to predict mechanisms of toxicity *in vivo*. However, this requires comprehensive knowledge of ideally all mechanisms leading to adverse effects, as well as an *in vitro* system that reflects critical mechanisms of *in vivo* toxicity. Since the link between gene expression alteration and adverse effects *in vivo* is not completely understood, it is of great importance to understand which of the responses observed

in the *in vitro* systems are relevant for the situation *in vivo*. It was shown that clusters of genes are up or down regulated simply by the hepatocyte isolation and cultivation procedure (Zellmer et al. 2010). Consequently, this response represents a pure *in vitro* artifact. Likewise, it is reported that cultivated primary hepatocytes become resistant to apoptosis in culture, which might result in a repression of certain *in vivo* relevant responses (Godoy et al. 2009; Godoy et al. 2010a; Godoy et al. 2010b). In contrast, a systematic comparison of gene expression profiles from exposed rat livers *in vivo* and cultivated rat hepatocytes *in vitro* revealed a good correlation for some cellular stress, as well as DNA damage and metabolism associated genes (Heise et al. 2012).

To bridge the gap between biomarkers of toxicity identified from *in vitro* systems and their potential function *in vivo*, one research group focused on a set of genes that are associated with elevated serum ALT levels after exposure to six heterogeneous compounds (Cheng et al. 2011). Thirty-two genes were used as a multi gene expression signature to predict hepatotoxicity in rats *in vivo*, and in HepG2 cells, as well as primary human hepatocytes *in vitro*. Different degrees of toxicity in response to drug concentrations were evaluated, allowing the estimation of the general hepatotoxic potential of a compound and its toxic concentration. However, pharmacokinetic differences between the *in vivo* and *in vitro* systems might lead to discrepancies in the drug-induced gene expression alterations (Schug et al. 2013). In general it is recommended that *in vitro* concentrations are selected, which reflect critical concentrations *in vivo*. For instance, physiologically-based pharmacokinetic (PBPK) models are used to predict doses that result in critical concentrations in the target cells *in vivo*. One study was able to show that the gene expression pattern induced by a histamine 3 receptor inverse agonist was comparable between exposed rat livers *in vivo* and corresponding concentrations in primary human hepatocytes *in vitro*, representing the maximal blood concentration (Roth et al. 2011). The group focused on genes that were critical for the hepatotoxicity induced by the compound, and was able to extrapolate the toxic effects to an unknown compound of the same compound class. This example demonstrates the identification of specific biomarkers for a selected compound and the applicability of the toxicogenomics tool to predict hepatotoxicity for uncharacterized compounds. Nevertheless, the identification of predictive biomarkers of toxicity remains challenging, since different compounds may induce different forms of liver toxicity (such as metabolic perturbations, cell death or mitochondrial dysfunction), which could result in unique gene expression profiles (Cheng et al. 2011).

Ideally, novel biomarkers of hepatotoxicity will cover a broad range of toxic mechanisms to capture as many compounds as possible. To identify potentially hepatotoxic drugs in preclinical studies, these biomarkers should give alerts independent of the chemical structure or the toxic mode of action. To evaluate which biomarker candidate genes might be of interest, it is critical to obtain a comprehensive overview of chemically-induced gene expression alterations.

1.3 Publically available transcriptomics databases – challenges and limitations

In recent years, several public databases, such as DrugMatrix, diXa and Toxicogenomics Project-Genomics Assisted Toxicity Evaluation System (TG-GATES), have emerged, providing gene array data of chemically-exposed hepatocytes and other cells from different organs (Jiang et al. 2015). All three databases encompass *in vitro* and *in vivo* transcriptomics data of compound-exposed rat organs or cultivated primary cells with multiple doses and time points (Chen et al. 2012; Hendrickx et al. 2015). Hundreds of compounds acting via various mechanisms were tested, including therapeutic, industrial, and environmental chemicals at both non-toxic and toxic doses. In addition to the transcriptomics data, some of the databases also provide additional information for each compound, also collecting including toxicity data and relevant sources from literature, together with available proteomics, metabolomics and epigenetics data (Hendrickx et al. 2015).

The scope of this thesis utilizes the transcriptomics data set of the publically available database, Open-TG-GATE. The database consists of transcriptomics data from 158 chemicals tested in cultivated primary human and primary rat hepatocytes, as well as *in vivo* data of exposed rat livers. Hepatotoxic and non-hepatotoxic drugs and some experimental hepatotoxic compounds were tested at three different time points, in three different concentrations, with the highest dose approaching cytotoxicity.

Although the vast amount of transcriptomics data may provide useful insights into various toxic mechanisms, the handling of this huge amount of data is not trivial. On the one hand, working with large data sets, especially when generated by several research consortia with independent contributors, is challenging because experimental errors and artifacts cannot be excluded (Grinberg et al. 2014). Having to combine several analytical batches, which will undoubtedly contain experimental errors in a subset of samples, is often unavoidable, and may lead to misinterpretation of the data. Exclusion of implausible data may improve the reliability of the Open TG GATEs transcriptomics data and form a basis for the identification of novel biomarkers of toxicity. On the other hand, the extraction of specific biomarkers of toxicity from such a large amount of data requires a general understanding, not only of possible mechanisms of action, but also of the typical changes in the cells as they undergo chemically-induced stress. Understanding the patterns of up or down regulated genes of chemically-exposed cells *in vitro* could provide valuable information for the extraction of potential biomarker genes and for the identification of toxic mechanisms. However, despite the frequent use of the previously mentioned *in vitro* test systems, a comprehensive analysis of genes altered by chemicals *in vitro* has not been performed. Therefore, in order to obtain a better understanding of global gene expression profiles after chemical exposure, this thesis summarizes key features of chemically-influenced genes and provides a guideline for the identification of novel biomarkers of hepatotoxicity.

1.4 Aim of this work

The aim of this thesis was to establish a guideline to describe how transcriptomics data of large data sets can be used to extract specific biomarkers of human hepatotoxicity.
For this purpose, genome wide expression data obtained from chemical-exposed primary human hepatocytes from the Open TG-GATES database is considered. The first part of this thesis focusses on the *in silico* characterization and curation of the Open TG-GATES database. To improve the reliability of the data, batch effects are identified and controlled, the data reproducibility across replicates is assessed and compounds following an implausible concentration are excluded from further analysis. With the curated data set, comprehensive bio-statistical analysis is performed and a novel toxicogenomics directory is established. The goal of establishing such a directory is to improve the understanding of how genes are typically altered by chemicals *in vitro*, which may contribute towards the identification of potential biomarkers of toxicity.
Since the heterogeneity of compounds involves various mechanisms of toxicity, it was assumed that the database comprised a comprehensive overview of all genes that could be deregulated in primary human hepatocytes after compound exposure. To enable the extraction of potential biomarker candidate genes, the structure of chemical-induced gene expression is analyzed and the altered genes are categorized using the following strategy:

- Identification of genes which are altered by many compounds. A change in the expression of these frequently altered genes represents a stereotypical response to cellular stress.
- Identification of genes which are also associated with human liver diseases.
- Exclusion of unstable baseline genes, which are altered because of the hepatocyte isolation and cultivation conditions.
- Identification of biological motifs to cover the most relevant toxic mechanisms.

Based on these key principles, the second part of the thesis focusses on the identification of novel biomarkers of human hepatotoxicity. Two different *in vitro* systems, namely HepG2 cells and cultivated primary human hepatocytes will be used to analyze the expression of the selected marker genes and to evaluate their applicability to predict human hepatotoxic blood concentrations that are associated with an increased risk of hepatotoxicity *in vivo*.
A set of hepatotoxic, as well as non-hepatotoxic chemicals is used to validate the expression of the selected biomarkers at concentrations, which reflect critical, as well as therapeutic doses *in vivo*. In the event that the set of biomarkers is able to differentiate between hepatotoxic and non-hepatotoxic compounds at therapeutic doses, the novel prediction system will provide a promising tool to identify hazardous compounds during early screening processes in drug development. Furthermore, predicting the blood concentrations that are associated with an increased risk of hepatotoxicity *in vivo* will provide a valuable method to evaluate the safety of novel drugs at therapeutic doses in humans.

2 Material and methods

2.1 Material

2.1.1 Technical equipment

Table 2.1: Technical equipment in the laboratory

Equipment	Company
Balance	EW, Kern
Bunsen Burner	IBS Fireboy Plus, Integra Bioscences
Bright Field Microscope	Primo Vert, Zeiss, Software ZEN from Zeiss
Casy®	Innovatis
Centrifuge	Megafuge 1.0R, Thermo Scientific
Centrifuge	Centrifuge 5415 R, Eppendorf
Centrifuge with cooling function	5424R, Eppendorf
Centrifuge with cooling function	Biofuge Fresco, Heraeus
Incubators	CO_2 Incubator C150 R Hinge 230, Binder
Laminar Flow Hood	Electronics FAZ 2, Waldner
Magnetic stirrer	IKAMAG RCT, IKA
Microcentrifuge	Mini Spin Plus, Eppendorf
Microscope CCD-Camera	AxioCam ICm 1
Minicentrifuge	FVL-2400N Combi-Spin, Biosan
pH meter	CG 842, Schott
Pipetteboy	Integra
Pipettes	Research and Reference, Eppendorf
Infinite M200 Pro Plate reader	Tecan
Precision balance	EW 150-3M, Kern
Real Time PCR System	7500 Real-Time PCR System, Applied Biosystems
Real Time PCR System	7900 HT, Applied Biosystems
Sonicator	Bandelin, SONOPLUS
Spectrometer	NanoDrop 2000, Thermo Scientific
Thermocycler	TGRADIENT, Biometra
Vacuum pump	Diaphragm Vacuum Pump, Vacuumbrand
Vortex	Vortex-Genie 2, Bender&Hobein
Water purification system	Maxima Ultra-Pure Water, ELGA
Waterbath	GFL 1083, Gesellschaft für Labortechnik

2.1.2 Chemicals and kits

Table 2.2: Compounds and kits

Compound	Company	Catalog number
2-Propanol	Carl Roth, Karlsruhe, Germany	7590.1
Acetaminophen	Sigma-Aldrich Corp., St. Louis, MO, USA	A7085
Acetic acid	Carl Roth, Karlsruhe, Germany	3738.5
Aspirin	Sigma-Aldrich Corp., St. Louis, MO, USA	A5376
Buspirone	Sigma-Aldrich Corp., St. Louis, MO, USA	B7148
Carbamazepine	Sigma-Aldrich Corp., St. Louis, MO, USA	C4024
Cell Titer Blue Assay	Promega	G8081
Chloroform	Carl Roth, Karlsruhe, Germany	7331.2
Chlorpheniramine	Sigma-Aldrich Corp., St. Louis, MO, USA	C3025
Clonidine	Sigma-Aldrich Corp., St. Louis, MO, USA	C7897
DEPC sterile water	Invitrogen	
Diclofenac	Sigma-Aldrich Corp., St. Louis, MO, USA	D6899
Disodium hydrogen phosphate	Carl Roth, Karlsruhe, Germany	T876.2
Ethanol	VWR Chemicals, Germany	20821.33
Famotidine	Sigma-Aldrich Corp., St. Louis, MO, USA	F6889
High Capaccity cDNA Reverse Transcription Kit	Applied Biosystems	4368813
Hydroxyzine	Sigma-Aldrich Corp., St. Louis, MO, USA	H8885
Isoniazid	Sigma-Aldrich Corp., St. Louis, MO, USA	I3377
Ketoconazole	Sigma-Aldrich Corp., St. Louis, MO, USA	K1003
Labetalol	Sigma-Aldrich Corp., St. Louis, MO, USA	L1011
Levofloxacin	Sigma-Aldrich Corp., St. Louis, MO, USA	40922
Melatonin	Sigma-Aldrich Corp., St. Louis, MO, USA	M5250
Nimesulide	Sigma-Aldrich Corp., St. Louis, MO, USA	N1016
Nitrofurantoin	Sigma-Aldrich Corp., St. Louis, MO, USA	N7878
Phenylbutazone	Sigma-Aldrich Corp., St. Louis, MO, USA	P8386
Potassium chloride	Fluka Chemie AG, Switzerland	60129
Potassium dihydrogen phosphate	Merk, Darmstadt, Germany	1.04873.1000
Promethazine	Sigma-Aldrich Corp., St. Louis, MO, USA	P4651
Propranolol	Sigma-Aldrich Corp., St. Louis, MO, USA	P0884
Qiazol® Lysis Reagent	Qiagen Sciences, Maryland, USA	79306
Rifampicin	Sigma-Aldrich Corp., St. Louis, MO, USA	R3501
Sodium chloride	Carl Roth, Karlsruhe, Germany	3957.2
Sodium hydroxid	Merk, Darmstadt, Germany	1.06482
Valproic acid	Sigma-Aldrich Corp., St. Louis, MO, USA	PHR1061

2.1.3 Consumables

Table 2.3: Consumables

Compound	Company	Catalog number
Biosphere Filtered Tip, 1000uL	Sarstedt, Numbrecht, Germany	70.762.211
Biosphere Filtered Tip, 100uL	Sarstedt, Numbrecht, Germany	70.760.212
Biosphere Filtered Tip, 200uL	Sarstedt, Numbrecht, Germany	70.760.211
Biosphere Filtered Tip, 20uL	Sarstedt, Numbrecht, Germany	70.1116.210
Cell culture microtiter plate, 96 well	Greiner bio-one	655986
Cell Scraper, 25cm	Sarstedt, Numbrecht, Germany	83.183
Falcon tube, 15mL	Sarstedt, Numbrecht, Germany	62.554.512
Falcon tube, 50mL	Sarstedt, Numbrecht, Germany	62.547.254
Parafilm Wrap	Cole-Parmer, Kehl/Rhein, Germany	PM-992
Pipette Tips, 1000uL	Sarstedt, Numbrecht, Germany	70.762
Pipette Tips, 200uL	Sarstedt, Numbrecht, Germany	70.760.002
Pipette Tips, 20uL	Sarstedt, Numbrecht, Germany	70.1116
RNase-free Microfuge Tubes 1.5 mL	Ambion, Thermo Fischer Scientific, USA	AM12400
RNaseZap® RNase Decontamination Solution	Ambion, Thermo Fischer Scientific, USA	AM9780/AM9782
SafeSeal 0.5mL microtube	Sarstedt, Numbrecht, Germany	72.699
SafeSeal 1.5mL micotube	Sarstedt, Numbrecht, Germany	72.706
SafeSeal 2.0mL microtube	Sarstedt, Numbrecht, Germany	72.695.500
Serological Pipette, 10mL	Sarstedt, Numbrecht, Germany	86.1254.001
Serological Pipette, 25mL	Sarstedt, Numbrecht, Germany	86.1685.001
Serological Pipette, 5mL	Sarstedt, Numbrecht, Germany	86.1253.001
Tissue Culture Plate Flat-Bottom 12-Well Plate	VWR Chemicals, Germany	734-2324
Tissue Culture Plate Flat-Bottom 24-Well Plate	Sarstedt, Numbrecht, Germany	83.1836
Tissue Culture Plate Flat-Bottom 6-Well Plate	Sarstedt, Numbrecht, Germany	83.1839
Vacuum Filtration Unit, 0.22um, 250mL	Sarstedt, Numbrecht, Germany	83.1822.001

2.1.4 Cell culture material and buffers

Table 2.4: Cell culture supplies

Compound	Company	Catalog number
Casyton solution	Roche Diagnostics GmbH, Manheim	5651808001
Collagen lyophilize (rat-tail), 10mg	Roche Diagnostics GmbH, Manheim	11171179001
Dexamethason	Sigma-Aldrich Corp., St. Louis, MO, USA	D4902-25MG
Dimethyl sulfoxid (DMSO)	Sigma-Aldrich Corp., St. Louis, MO, USA	472301
DMEM low glucose 1.0 g/L 10x	BioConcept, Allschwil, Switzerland	1-25K03-I
Dulbecco's modified eagle's medium (DMEM)	PAN Biotech GmbH, Aidenbach, Germany	P04-04500
Gentamicin	PAN Biotech GmbH, Aidenbach, Germany	P06-13001
Insulin supplement (ITS)	Sigma-Aldrich Corp., St. Louis, MO, USA	3146
Penicillin/Streptomycin	PAN Biotech GmbH, Aidenbach, Germany	P06-07100
Sera Plus (Special Processed FBS)	PAN Biotech GmbH, Aidenbach, Germany	3702-P103009
Stable L-Glutamin	PAN Biotech GmbH, Aidenbach, Germany	P04-82100
Trypan blue solution	Sigma-Aldrich Corp., St. Louis, MO, USA	T8154
Trypsin/EDTA	Sigma-Aldrich Corp., St. Louis, MO, USA	P10-023100
William's E medium	PAN Biotech GmbH, Aidenbach, Germany	P04_29510

2.1.4.1 Phosphate buffered saline (PBS) buffer for cell culture

For 5 L 10x PBS:

KCl	10 g
KH_2PO4	10 g
NaCl	400 g
Na_2HPO4	46 g

All reagents were dissolved in double distilled water and the pH was adjusted to pH 7.4. Afterwards the buffer was sterile filtered. For application in the cell culture, 10x PBS was diluted to 1x PBS with double distilled water and autoclaved before usage.

2.1.4.2 HepG2 cell line

HepG2 liver cells were purchased from ATCC LGC Standards, product number HB-8065. The cell line was generated from a 15 year old Caucasian male with hepatocellular carcinoma.

2.2 Methods

2.2.1 Cell culture of HepG2 cells

2.2.1.1 Cultivation of HepG2 cells

HepG2 cells were cultivated in Dulbecco's modified eagle's medium (DMEM) containing 4.5 % glucose, 1 % penicillin/streptomycin mixture and 10 % heat inactivated FCS. The FCS heat inactivation was performed at 56 °C for 30 minutes in a water bath. The cells were seeded in conventional T75 or T175 flask and kept at 37°C with constant humidity and 5 % CO_2 content.

2.2.1.2 Thawing and freezing HepG2 cells

For thawing, the frozen cell suspension was thawed in a water bath (37 °C) and immediately transferred into a Falcon tube. Afterwards the suspension was diluted in 7-8 mL medium and centrifuged for 5 minutes at 600 rpm at room temperature to remove the freezing medium. The cell pellet was re-suspended in 1 mL medium and given into a T75 cell culture flask with 9 mL medium.

For storage, cells were preserved in freezing media containing the regular media plus 10 % DMSO. Cells were usually frozen when reaching 80-90 % confluency. The cells were trypsinized by adding 1 mL trypsin per T75 flask or 2 mL per T175 flask and subsequently re-suspended in 5 mL (for T75 flask) or 10 mL media (for T175 flasks). The cell suspension was then transferred into a Falcon tube and centrifuged at 800 xg for 5 minutes to form a clear pellet. The supernatant was aspirated and the cells were re-suspended in freezing media (3 mL per T75 flask, 6 mL per T175 flask). 1 mL cryo vial aliquots were prepared and kept on ice for 20 minutes before storage at -80°C. For long time incubation the cells were stored in liquid nitrogen.

2.2.1.3 Passaging HepG2 cells

Upon 80-90 % confluency, the HepG2 cells were either sub-cultured or seeded in multi well plates for further experiments. For splitting, the cells were washed once with 10 mL sterile PBS, the PBS was aspirated and 4 mL Trypsin/EDTA were added. The cells were incubated for 7 minutes in Trypsin/EDTA at 37°C in the incubator to detach from the plastic surface. The enzymatic reaction was stopped by adding 20 mL of warm cultivation medium. The trypsin was removed by centrifugation at 600 rpm for 5 minutes at room temperature. The obtained cell pellet was re-suspended initially in 1 mL cultivation medium and gently pipetted up and down with a 1 mL tip of an Eppendorf pipette. Subsequently the dense cell suspension was diluted with further 9 mL medium and distributed 1:3 or 1:10 into new T175 cell culture flasks. In a total volume of 25 mL, cells were kept at 37°C until 80-95 % confluency was reached again.

2.2.1.4 Seeding and treatment of HepG2 cells

For counting cells with the CASY Cell Counter, 100 µL of the cell suspension diluted in 10 mL Casyton.

For gene expression analysis (two days in culture, 24h compound exposure) 500,000 cells per well were seeded in conventional 6-well plates in 2 mL medium per well. For cytotoxicity tests (three days in culture, 48 h compound exposure) the cells were cultivated in 24-well plates, 62,500 cells seeded per well in 500 µL medium. Compound exposure was started the next morning after plating the cells.

For gene expression analysis as well as cytotoxicity experiments each compound was tested in 5 different concentrations plus vehicle control. The chemical amount for the highest concentration was weighed and dilution series were prepared for the lower concentrated solutions. In case of water soluble compounds, substances were dissolved in medium and sterile filtered before adding to the cells. If the compound amount for the highest concentration was below 1 mg, 100x higher concentrated stock solutions in sterile water were prepared. DMSO soluble compounds were dissolved in a higher concentrated DMSO stock solution and dilution series were prepared (see Table 3.17). Cells were exposed for 24 hours at 37°C at constant humidity and 5 % CO_2.

2.2.2 Cell culture of primary human hepatocytes

2.2.2.1 Medium for cultivated primary human hepatocytes

Primary hepatocytes were cultivated in William's E medium (PAN Biotech, P04_29510) with 100 U/mL penicillin, 0,1 mg/mL streptomycin, 10 µg/mL gentamicin, 2 mM stable glutamin, 100 nM dexa-methasone and 2 nM insulin-transferrin- selenite (ITS) supplement. When plating cells, 10 % fetal calf serum was added for the first 3-4 h of cultivation.

2.2.2.2 Isolation of primary human hepatocytes

Primary human hepatocytes were isolated from liver sections of patients undergoing surgical liver resection. Prior to the resection, informed content was obtained from each patient. The isolation procedure was performed in three cooperating clinics, the Charité Berlin, and the university hospitals Munich and Regensburg.
Resected tissue samples were immediately transported into a sterile vessel containing PBS or culture medium, in order to prevent warm ischemia. Upon arrival in the laboratory, the tissue was placed on a sterile Petri dish and prepared for perfusion: Remaining blood was removed by using an aseptic gauze and buttoned cannula were placed into several vessels of

the resection side and fixed with tissue glue. Depending on the size of the tissue and by using the biological blood vessel architecture, 3-8 cannula are sufficient to perfuse the whole piece of liver (Godoy et al. 2013). Liver perfusion for hepatocyte isolation was implemented by a two-step isolation procedure which was developed by Seglen et al. and processed as recently described by Shinde et al. (Seglen 1976; Shinde et al. 2015). During the first perfusion step, the piece of liver is rinsed for approximately 10 minutes with a pre-warmed EGTA containing buffer (Godoy et al. 2013). EGTA is added to prevent coagulation, to remove the residual blood and to deplete calcium from the vessels. Calcium is important for cellular adhesion, therefore, the washing out of these ions depletes adhesion factors, results in loosening of the tissue and promotes the perfusion process (Moscona et al. 1956 and Gingell et al. 1970). The flow rate of the perfusion solution through the tissue is about 15mL/min. An optimal rinsing of the tissue with the buffer is accompanied by a tissue color change from brownish red to beige. In the second perfusion step, the piece of liver is perfused with a pre-warmed collagenase containing buffer (Godoy et al. 2013). For optimal enzymatic activity, calcium has to be added to the perfusion solution. In this perfusion step the extracellular matrix of the liver tissue is gradually digested within 5-15 minutes. Cannula have to be pulled out quickly before the piece of digested liver is placed in a Petri dish with stop solution for enzyme inactivation. By cutting the perfused liver piece into two halves and gently shaking the tissue, hepatocytes are released into the stop solution. The cell suspension was passed through a funnel of gauze in order to remove tissue debris. A centrifugation step was included to separate non parenchymal liver cells from the more heavy hepatocytes. The centrifugation was carried out at 50-100xg at 4°C for 5 minutes. The cell pellet was re-suspended in PBS or hepatocyte culture medium and placed on ice. The transport of the cells from the surgical departments to our laboratory was accomplished overnight in cold stored suspensions on ice. Upon arrival, cells were re-suspended in fresh cultivation medium and the viability was determined using trypan blue exclusion method.

2.2.2.3 Determination of cell viability and cell yield with trypan blue vital stain

Trypan blue is a commonly used dye to selectively stain dead cells or tissue. Vital cells with intact cell membranes cannot incorporate the dye, but dead cells with perforated, destroyed cell membranes easily take up the stain. In order to determine cell yield and viability, an aliquot of cell suspension was diluted 1:10 in hepatocyte culture medium and mixed 1:2 with a 0,4 % sterile filtered trypan blue solution. The obtained mixture (1:20 dilution) was filled into the chamber of a hemocytometer and the cells in the outer four counting grid squares were counted. Vital, unstained cells as well as dead, blue colored cells were counted and the cell yield as well as the cell viability was calculated as follows:

(i) Total amount of cells per m = (counted cells/number of counted square grids) x 10^4 x dilution factor

(ii) Cell viability (%) = number of vital cells x 100 / total number of cells

2.2.2.4 Cultivation of primary hepatocytes in collagen sandwich system

Primary hepatocytes were seeded in conventional 6-well plates between two soft layers of collagen gel. For the gel preparation, a bottle of 10 mg lyophilized rat tail collagen was dissolved overnight in 9 mL 0.2 % acetic acid at 4°C. Before plating cells dissolved collagen was placed on ice, mixed 1:10 with 10x DMEM and drop wise neutralized with 1M NaOH until the color turned from yellow to pink. The obtained gel has now a concentration of 1 mg/mL.

Using conventional 6-well plates, each well was coated with 350 µL of collagen gel for the first layer (Godoy et al. 2013) and left for polymerization for 30-45 minutes at 37°C. After successful gelation, cells were plated in FCS containing medium into the wells of each plate and kept at 37°C in the incubator for at least 3 hours. During this time the cells attach to the collagen matrix. For homogenous distribution of the cells, the plate was carefully shaken every 5-10 minutes during the first half hour of incubation. After the incubation period, the cells were carefully washed 3 x with warm sterile PBS before the second layer of collagen was added. The gel of the second layer polymerized at 37 °C for 30-45 min in the incubator. Afterwards, 2mL warm FCS free medium was added per well and the cells were kept at 37°C in the incubator overnight.

2.2.2.5 Treatment of cultivated primary human hepatocytes

Gene expression experiments were carried out using 1,500,000 cells per well plated in conventional 6-well plates between the two soft collagen gel layers and 2 mL medium per well. Analogue to HepG2 cells, compound exposure was started the morning after the day of plating. For gene expression analysis, the cells were exposed for 24 hours before RNA was collected. Cytotoxicity experiments followed an incubation period of 48 hours.

2.2.3 RNA sample collection and isolation procedure

For RNA sample collection, the plates were transferred on ice and the medium supernatant was immediately aspirated. QIAzol lysis reagent was applied according to the manufacturer's protocol (1mL QIAzol/well in a 6-well plate format) and the cells were lysed by mechanical scraping with a cell scraper. After transferring the liquid into a sterile 2 mL Eppendorf tube, the lysates were sonicated on ice for 30 seconds (5 sec pulse, 2 sec break). Samples from freshly isolated hepatocytes were processes similarly: 1-1.5 Mio cells were transferred in a reaction tube on ice for some minutes. During this time, the cells accumulate at the bottom of the reaction tube and the supernatant can be removed carefully. 1mL QIAzol was added and the lysate was sonicated as described.

RNA was isolated by using QIAzol lysis reagent for phenol-chloroform extraction. QIAzol comprises a guanidinium-thiocyanat-phenol mixture, which lyses the cells and supports degradation of proteins. Addition of chloroform results into a phase separation. Under acidic conditions the denaturated proteins and DNA partition in the organic phase and interphase, while the RNA remains soluble and accumulates in the aqueous phase. Subsequently, the aqueous phase was separated and RNA was precipitated by addition of isopropanol. Several washing and centrifugation steps with ethanol were performed to increase the RNA purity. The amount of QIAzol as well as the following sample procession was modified depending on the plate format.

Samples were further processed as followed: 200 µL chloroform was added to each reaction tube and the samples were strongly shaken for approximately 15 seconds. After 2-3 minutes incubation at room temperature, a first phase separation was observed. After a centrifugation step at 4°C for 15 minutes at 12,000xg the upper aqueous phase was transferred into a new reaction tube and the RNA was precipitated with adding 500 µL of isopropanol. The samples were incubated at room temperature for 10 minutes, centrifuged at 12,000xg for 15 minutes at 4°C and the supernatant was removed from the RNA pellet. 1 mL of 100% ethanol was added to wash the RNA pellet, followed by a centrifugation step at 7,500xg and 4°C for 5 minutes. After two more washing and centrifugation steps with 80 % and 75 % ethanol (each 1 mL), the supernatant was again removed carefully and the RNA pellet was air dried for some minutes. Depending on the RNA pellet size, 7.5-15 µL RNAse free DEPC water was used to re-suspend the RNA. Isolated RNA was stored at -80°C until further usage. The RNA quantity of each sample was determined photometrically with the NanoDrop 2000.

2.2.4 cDNA synthesis

For quantification of gene expression, the isolated RNA had to be reversely transcribed into cDNA. This reaction step is catalyzed by the enzyme reverse transcriptase, which is capable to create single stranded DNA from a RNA template. For this purpose the *High Capacity cDNA Reverse Transcription Kit* from Applied Biosystems was used. 500 ng – 2 µg RNA were reversely transcribed according to the manufacturer's protocol. The volumes for the reaction mixture are listed in Table 2.5.

Table 2.5: Reaction mixture for cDNA synthesis

Compound	Volume per reaction
10x RT buffer	2 µL
Random primers	2 µL
dNTPs	0.8 µL
Reverse Transcriptase	1 µL
DEPC H_2O	4.2 µL
Mastermix volume 1	**10 µL**
RNA	500ng - 2µg
DEPC H_2O	up to 10µL
Total volume 2	**10 µL**
Final volume in total	**20 µL**

Thermal cycling conditions were chosen as follows:

Table 2.6: Conditions for the thermal cycling program

Step	Temperature	Time
Incubation	25°C	10 min
Reverse Transcription	37°C	120 min
Inactivation	85°C	5 sec
	4°C	hold

The concentration of the cDNA was adjusted with DEPC- treated water to a final concentration of 10 ng/µL. The samples were stored at -20°C until further usage.

2.2.5 Quantitative Real Time PCR (qRT-PCR)

Quantitative real time PCR is applied to detect and quantify expression of target genes, which are altered in comparison to a stable expressed housekeeping gene. The procedure follows the conventional PCR method: Based on thermal cycling, primers bind to defined areas of a single DNA template strand and function as starting point for the enzyme Taq polymerase, which elongates the strand with deoxy nucleotide triphosphates (dNTPs) until the complementary strand is completed. Within 40 cycles of the reaction program, the template DNA is exponentially amplified. In contrast to the conventional PCR, qRT-PCR additionally enables the quantification of the synthesized DNA by combining each amplification cycle with a fluorescence signal. The PCR product concentration correlates with the fluorescence intensity (Wong and Medrano 2005). Formats for the fluorescence detection are for example

hydrolysis probes, such as TaqMan probes or molecular beacons, hybridization probes or double stranded DNA binding fluorescing dyes such as SYBR Green (Holzapfel and Wickert, 2007). In this thesis, the template strand was hybridized to TaqMan probes, which are primers carrying a fluorescence resonance energy transfer (FRET) pair. This FRET pair consists of a fluorescence donor, emitting light of a particular wave length, and a quencher, whose absorption spectrum overlaps with the emission spectrum of the donor and thereby quenches its fluorescence. In every amplification step, primers bind to the template strand and the polymerase elongates the complementary strand in 5'→ 3' direction. During this extension step, the enzyme separates the TaqMan probe from the template strand with its exonuclease activity and splits it into single nucleotides (Holland et al. 1991). By splitting the probe, donor and quencher fluorescence molecules of the FRET pair are separated and a fluorescent signal occurs. The fluorescence signal increases proportional to the amount of newly synthesized DNA product. During 40 cycles, the fluorescence is detected as a function of time and so called Cycle threshold (Ct) values are registered. The Ct-value is defined as the number of cycles that is required to cross a certain fluorescence threshold. Consequently, with an initial high amount of template DNA, a significant change in the fluorescence signal is observable at an early time point.

The measurements were performed using TaqMan –PCR technology with a 7500 Real-Time PCR System. Buffers and reagents for the PCR reaction were used as a TaqMan Universal PCR Master Mix from Applied Biosystems. Table 2.7 shows the volumes for the reaction mixture for one sample:

Table 2.7: qRT-PCR reaction mixture per sample

Compound	Volume
Universal PCR Master Mix	10 µL
DEPC treated H_2O	6.5 µL
Taqman probe	1 µL
25 ng cDNA	2.5 µL
Final volume	**20 µL**

Selected TaqMan probes for all genes are listed in Table 2.8.

Table 2.8: TaqMan probes from Applied Biosystems for gene expression quantification.

Probe Number	Gene
Hs00164383_m1	Cyp1B1
Hs00426361_m1	Cyp3A7
Hs00602560_m1	SULT1C2
Hs00166169_m1	G6PD
Hs00603550_g1	TUBB2B
Hs00204129_m1	RGCC
Hs01041408_m1	FBXO32
Hs01037712_m1	PDK4
Hs01650979_m1	INSIG1
Hs01572978_g1	PCK1
Hs00930058_m 1	THRSP

Samples were measured in technical duplicates. Negative controls with water instead of template DNA were included. The conditions for the PCR reactions are shown in Table 2.9.

Table 2.9: Thermal conditions for qRT-PCR measurements

Stage	Temperature	Time	Repetitions
1	50 °C	2 min	1
2	95 °C	10 min	1
3	94 °C 60 °C 72 °C	15 sec 30 sec 35 sec	40-45
4	95 °C 60 °C 95 °C 60 °C	15 sec 20 sec 15 sec 15 sec	1

PCR products were analyzed with the 7500 Real-Time PCR System software. To determine the expression of a target gene, the expression of an endogenous reference, so called housekeeping gene needs to be quantified. An ideal reference gene features a consistent gene expression level, independent of the cell cycle state, of the cell type or any treatments and conditions during the experimental procedure (Holzapfel and Wickert, 2007). Common reference genes are for example ubiquitine, glyceraldehyde 3-phosphate dehydrogenase (GAPDH), b-actin or 18 S RNA (Goidin et al. 2001). To quantify the gene expression levels of a target gene, several quantification methods are available: The standard curve method, the relative quantification method and the comparative threshold method (= $\Delta\Delta$Ct method). In this thesis, the $\Delta\Delta$Ct method was applied for data analysis and GAPDH expression was used as endogenous control. The $\Delta\Delta$Ct method is a relative method which shows a correlation between the gene expression in compound exposed cells compared to untreated controls, but not giving absolute values. In a first step the expression level of a target gene is normal-

ized to the expression level of a stable expressed housekeeping gene. Next, the normalized data of untreated cells is subtracted from data of exposed cells. Relative expression is determined according to the following calculation steps:

1. ΔCt1 = Ct target gene – Ct house keeper
2. ΔCt2 = Ct target gene control samples – Ct house keeper control samples
3. $\Delta\Delta$Ct = ΔCt1 – ΔCt2
4. Calculation of $2^{-\Delta\Delta Ct}$

Finally, the $2^{-\Delta\Delta Ct}$ is depicted in a histogram, showing the relative genes expression values for every experimental condition.

2.2.6 Cytotoxicity tests with the CellTiter-Blue® Cell Viability Assay

For cytotoxicity testing in HepG2 cells as well as primary human hepatocytes, the CellTiter-Blue® Cell Viability Assay from Promega was applied. The assay is based on the metabolic capacity of cells and analyzes the reduction rate of the dark blue indicator dye resazurin to the pink and highly fluorescent dye resofurin. While resazurin is only slightly fluorescent, resofurin is highly fluorescent and can be quantified at its emission maximum at a wavelength of 584 nm (Figure 2.1). Vital cells exhibit a strong metabolization capacity for resazurin, and the fluorescence intensity in the medium supernatant is increasing. In contrast, non-viable cells show a decreased resazurin metabolization capacity, resulting in lower fluorescence intensity or, in case of dead cells the dark blue resazurin is not even metabolized.

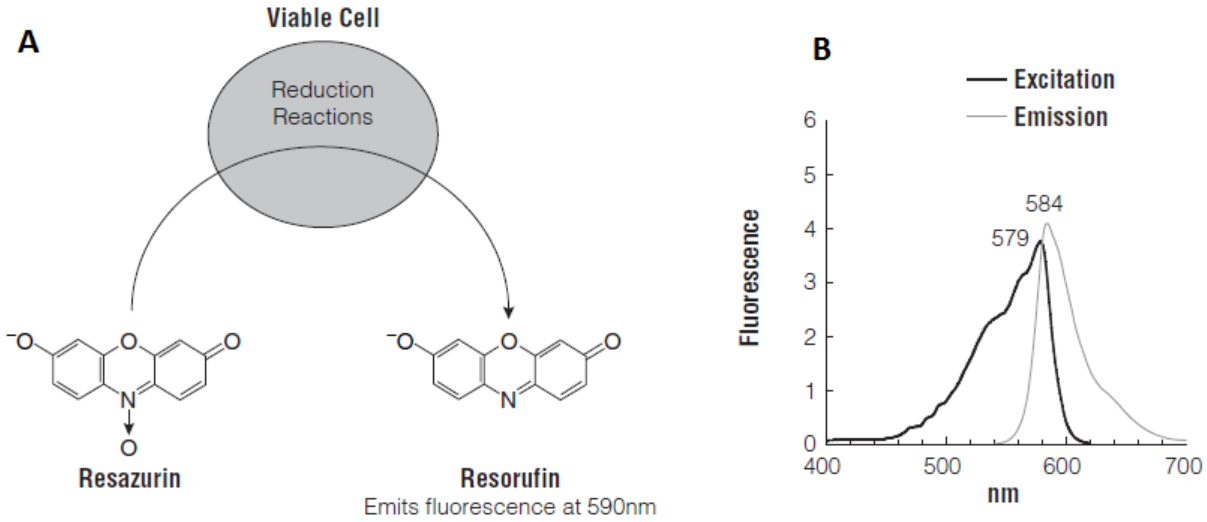


Figure 2.1: A The CellTiter-Blue® Cell reaction is based on the metabolization of resazurin to the highly fluorescent dye resofurin. The fluorescence intensity correlates with the vitality of the cell system metabolizing the dye. **B** Resofurin is excitated at wavelength of 579 nm and exhibtis an emission maximum at 584 nm. Reference: promega.com.

For cytotoxicity testing the cells were isolated and plated as described. HepG2 cells were seeded in 24-well plates, 62,500 cells per well whereas primary human hepatocytes were plated in a collagen sandwich matrix, 1,000,000 cells per well in a 6-well format. The day after plating, the cells were exposed to appropriate concentration series of selected compounds. After 48 h of compound exposure, the medium was aspirated and fresh medium with 20 µL CellTiter-Blue® reagent per 100 µL medium were added to the cells. HepG2 cells were exposed to the resazurin reagent for 2 hours whereas primary human hepatocytes were exposed for four hours. After wards, the supernatant was transferred to 96-well plates and the fluorescence intensity was read out with the Tecan Infinite M200 Pro plate reader using the i-control software (version 1.7.1.12). Untreated cells and cells incubated with vehicle controls only underwent the same procedure as the compound exposed cells and were used as a reference for 100 % viability. Cell viability was calculated after background subtraction and expressed as percentage of control. The incubations were performed for three independent experiments (3 biological replicates) for HepG2 cells and preliminary results with one biological replicate (cells from one donor) were obtained in primary human hepatocytes. From each biological replicate, three technical replicates were applied for the fluorescence read out.

2.2.7 Statistical analysis

The statistical part of this thesis was performed in close cooperation with the statisticians Jörg Rahnenführer, Marianna Grinberg and Eugen Rempel from the Technical University of Dortmund. The following context corresponds in large parts to the publication Grinberg et al. (2014).

2.2.7.1 Download and processing of the OPEN TG GATEs data

The Open TG-GATEs (Toxicogenomics Project—Genomics-Assisted Toxicity Evaluation System) database (NIBIO 2013) compiles publically available Affymetrix HG U133 Plus 2.0 gene expression microarray data (54,675 probe sets, corresponding to 19,945 uniquely annotated Gene Symbol IDs) from monolayer cultured primary human hepatocytes. 158 compounds were tested in total and for each compound corresponding untreated controls were generated. A subset of 52 compounds was tested using three concentrations (low, middle and high) at three incubation periods (2 h, 8 h, and 24 h). For the additional 106 compounds the concentration and time sets are incomplete. The compounds were tested either for only one or two incubation periods, or with only two concentrations (Table 3.1 and Supplemental Table 1).

Based on LDH release, the highest tested concentration (or the only tested concentration if only one was analyzed) was chosen as a slightly cytotoxic concentration yielding an 80–90 % relative survival rate. In case of non-cytotoxic compounds, a concentration of 10 mM or the highest soluble concentration was defined as the highest concentration. Solvent controls were routinely used at 0.1 % DMSO concentration, which was increased to maximum 0.5 % DMSO for compounds with low solubility. In total, six batches of human hepatocytes were used. Together with the gene expression raw data in Open TG-GATEs, information on the gender of the donor is given. Hepatocytes from 'male' and 'female' were specified by the columns 'sex_type' in the 'Attribute.tsv' file.

For 155 of the 158 compounds two replicates were available. Single experiments without replication were not considered in this study. For all compounds and conditions the raw microarray data (CEL files) were downloaded from the Open TG-GATEs website (http://toxico.nibio.go.jp/) . Robust Multi-Array Average (RMA) algorithm was used to normalize the entire set of expression arrays. This algorithm uses background correction, log2 transformation, quantile normalization, and a linear model fit to the normalized data to obtain a value for each probe set (PS) on each array (Harbron et al. 2007; Krug et al. 2013). For each compound, concentration and incubation period, the fold change of gene expression between compound exposed samples and corresponding untreated controls was calculated based on the average of replicate values. Data preprocessing and all subsequent analyses were performed using the statistical programming language R, version 3.0.1 (R Development Core Team 2013).

2.2.7.2 Visualization of high dimensional gene expression data

Unsupervised hierarchical clustering was applied to visualize matrices of gene expression values. The expression alterations in these heat maps range from low expression (blue color) to high expression (red color) (Figure 3.11). To visualize expression data in two dimensions, principal component analysis (PCA) was used. The first two principal components represent the two orthogonal directions of the data with the highest variance. Both, heat maps and PCA were generated on the basis of the 100 top-ranking genes with highest fold change (absolute values) across all compounds. These genes were selected separately for all nine combinations of concentration and exposure periods.

2.2.7.3 Gene set enrichment methods

The topGO package (Alexa and Rahnenführer 2010) was applied for gene ontology enrichment analysis. This package uses the Fisher's exact test and considers only results from the biological process ontology. The cutoff for the enrichment *p* value was set to 0.001.

Transcription factor binding site (TFBS) enrichment was analyzed using the PRIMA algorithm (http://acgt.cs.tau.ac.il/prima/) (Elkon et al. 2003) provided in the Expander software suite (version 6.04; 43 http://acgt.cs.tau.ac.il/expander/) (Ulitsky et al. 2010). The cutoff for the enrichment *p* value was set to 0.01.

2.2.7.4 Definition of indices for concentration progression

Two indices were introduced to analyze the progression of gene expression alterations with increasing concentrations – the 'progression profile index' and the 'progression profile error indicator'. For each compound and for each adjacent concentration, both indices were calculated.

The 'progression profile index' was defined as the proportion of genes which is at least 2 fold up or down regulated (compared to control) at a higher compared to a respective lower concentration. If only a few additional genes were deregulated at the next higher concentration, the value is close to zero. A value close to one indicates that the number of deregulated genes increases concentration dependently.

In contrast, the second index, the 'progression profile error indicator', determines the proportion of genes deregulated exclusively at a lower compared to a respective higher concentration. A value above 0.5 indicates an implausible concentration progression of a compound. However, if only a few genes in total were deregulated exclusively at the lower compared to the higher concentration, these genes were considered as outliers. Compounds following a very implausible concentration progression were excluded from the study. In case only a few genes followed an implausible concentration progression, meaning the 'progression profile error indicator' value is above 0.5 and at most 20 genes in total are altered at the respective lower concentration, the 'modified progression profile error indicator' was introduced. The 'modified progression profile error indicator' is an adjustment of the 'progression profile error indicator' and sets the index to zero to decrease the influence of a few outlying genes. For all three incubation periods of 2, 8 and 24 h the three progression indices were calculated separately.

2.2.7.5 Principles for differentiation between stereotypic and compound specific gene expression responses

To distinguish between stereotypic and compound specific gene expression responses of chemical exposed hepatocytes, the selection value concept is introduced. If a gene is deregulated by many compounds, this expression response is considered as a stereotypic effect. In case a gene is altered by only a few compounds, the response is defined as rather compound specific. The selection value is defined as the number of compounds which induces the expression of a given probe set at least three fold. It relates to a specific concentration and incubation time. For the selected condition, compounds are ranked in order of their fold

change for each probe set. For up regulated probe sets, the compounds are ranked from a high to a low fold change whereas compounds are ranked in reverse order to obtain the set of down regulated probe sets. Accordingly, the selection value x for a single probe set corresponds to compound on rank x, indicating that the probe set is altered at least 3 fold by at least x compounds. Considering selection value 20 (SV 20), a list of genes is obtained which are at least 3 fold up or down regulated by at least 20 compounds. In contrast, a selection value of 5 (SV 5) gives a list of genes which are at least 3 fold deregulated by at least 5 compounds. The threshold of 3-fold induction is chosen arbitrarily to keep the number of also positive genes relatively low.

2.2.7.6 Liver disease dataset analysis

To establish a link between genes that are deregulated by compounds *in vitro* and a possible relevance *in vivo*, human gene array data from patients with liver diseases was investigated. Microarray datasets comprising global gene expression alterations in liver diseases were obtained from public data repositories ArrayExpress (E-MEXP-3291) and Gene Expression Omnibus (GSE25097). E-MEXP-3291 (Lake et al. 2011) was analyzed on Affymetrix GeneChip Human 1.0 ST arrays and used to compare 16 samples of non-alcoholic steatohepatitis (NASH) liver tissue to data from 19 samples of healthy liver tissue. GSE25097 (Tung et al. 2011) was analyzed on Human RSTA Affymetrix 1.0 Custom CDF microarrays and was used to compare cirrhotic liver (40 samples) to non-tumor liver tissue (243 samples). Normalized RNA sequencing (RNA-Seq) data was obtained from The Cancer Genome Atlas (TCGA) (http://cancergenome.nih.gov/) and analyzed on the Illumina HiSeq platform to study gene expression alterations in hepatocellular carcinoma (HCC) (163 samples) as compared to matched non-tumor liver tissue (49 samples). Processing and quantile normalization of the microarray data was performed with the Piano R package (Varemo et al. 2013). This package was also applied to analyze differential gene expression; *p* values were corrected for multiple testing by the method of Benjamini and Hochberg (Benjamini and Hochberg 1995). The RNA-Seq data was analyzed for differential expression using the R package DESeq (Anders and Huber 2010). Comparing healthy/non-tumor tissue to diseased tissue, genes were considered to be differentially expressed when having a fold change of minimum 1.3 and a false discovery rate (FDR) adjusted *p* value of ≤0.05.

For a direct comparison of differentially expressed genes in liver diseases and genes altered by chemicals in human hepatocytes *in vitro*, probe sets included on the Affymetrix arrays were converted into uniquely annotated Ensembl Gene IDs. 18,809 genes were considered for the Open TG-GATEs dataset (originally 54,675 probe sets), 19,477 genes for E-MEXP-3291 (originally 32,321 probe sets), and 25,426 genes for GSE25097 (originally 37,582 probe sets). Only genes being present in both, the Open TG-GATEs dataset and E-MEXP-3291 (17,663 genes) or GSE25097 (16,514 genes), respectively, were included for the final comparison. A direct comparison of the TCGA dataset (20,471 genes, as recognized by a unique Entrez Gene ID) to the probe sets of the Open TG-GATEs data was enabled by converting the ap-

propriate Affymetrix probe sets into Entrez IDs. Therefore, manufacturer mapping after duplicate removal was used, resulting in 19,944 uniquely annotated genes and 17,895 genes included in the final comparison.

3 Results

Excerpts of this thesis have been published in Grinberg et al. (2014). The content of pages 25 to 68 corresponds largely to the mentioned publication.

3.1 Establisment of a toxicogenomics directory for compound-exposed primary human hepatocytes based on the Open TG-GATEs transcriptomics data

3.1.1 *In silico* characterization and curation of the Open TG GATES data

The Open TG-GATES toxicogenomics database is comprised of gene expression profiles derived from cultivated and compound-exposed primary human hepatocytes. 158 chemicals and drugs were tested with corresponding untreated controls in different concentrations and for three incubation periods (2h, 8h and 24h) and two replicates per experiment were generated. For 52 of the 158 compounds gene array data is available. These compounds were tested in a high, middle and a low dosage. The concentrations were selected based on cytotoxicity: the high dose represents a 80-90% relative survival ratio or was defined by the maximal solubility of the compound, the ratio of the concentrations for the low, middle and high dose was 1:5:25 (Igarashi et al. 2015). For the remaining 106 compounds, the data set is incomplete in terms of time points and concentration sets. For one compound, pherone (PHO), gene expression data for the 24h time point with the highest concentration is missing. Three further compounds (bromoethylamine (BEA), lipopolysaccharide (LPS) and trimethadione (TMD) were tested with only 1 replicate for the highest dose. The 24h time point, highest concentration was therefore excluded for the four mentioned compounds. Seven of the tested compounds were cytokines (interferon alpha, interleukin 1 beta, interleukin 6, transforming growth factor beta 1, hepatocyte growth factor, tumor necrosis factor) and LPS, which were all excluded from further analysis. In total, 151 chemicals were included in the analysis. An overview of the compounds and tested concentrations of the available data set is given in Table 3.1. A detailed overview of the data set is shown in Supplemental Table 1.

Table 3.1: Matrix of the tested compounds. The tables provide the numbers of compounds tested under the indicated conditions for each combination of concentration and exposure period

A. Before excluding cytokines and LPS from the analysis.

	2hr	8hr	24hr	Overlap time points
Low	53	82	81	52
Middle	53	153	157	52
High	53	153	153	52
Overlap concentrations	53	82	78	52

B. After excluding cytokines and LPS from the analysis.

	2hr	8hr	24hr	Overlap time points
Low	48	75	75	48
Middle	48	146	151	48
High	48	146	148	48
Overlap concentrations	48	75	72	48

C. Overview of the experimental design including the number of all available time and concentration sets, as well as replicates.

Number of chemicals	Number of concentrations	Number of time points	Number of replicates
71	2	2	2
52	3	3	2
26	3	2	2
5	2	1	2
3	3	2	High, 24h only 1 sample*
1	3	2	High, 24h not available**

3.1.2 Identification and control of batch effects

To illustrate the alterations of gene expression in primary human hepatocytes under the exposure of chemicals, principal component analysis (PCA) was performed. The 100 top ranking genes with the highest fold changes (absolute values) across all compounds were included in the analysis. Supplemental Figure 1, Supplemental Figure 2 and Supplemental Figure 3 show nine combinations of concentration and incubation periods. The strongest effect on gene expression was observed for the 24h time point at the highest concentration (Figure 3.1 A).

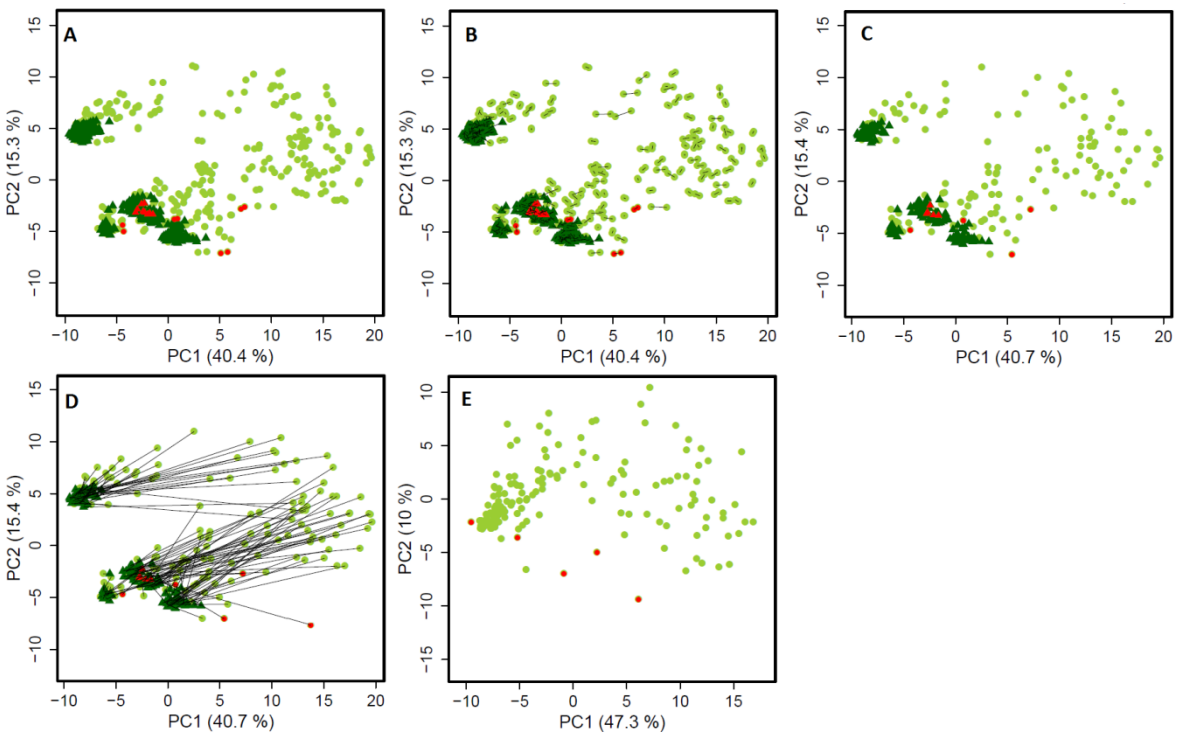


Figure 3.1: Principal component analysis of gene expression data from primary human hepatocytes after 24h incubation with 148 chemicals (green) and 7 cytokines (red) at the highest concentration. **A** Overview of all samples and replicates. Light green samples illustrate the exposed samples, dark green are the corresponding controls. Cytokines are illustrated in red. **B** Connecting lines show the degree of variability between 2 replicates. **C** Data points represent the mean values of the replicates. **D** Connecting lines illustrate the distance between the exposed samples and the corresponding controls. **E** Distribution of the exposed samples after subtraction of the corresponding controls.

The controls are located within two main clusters whereas the majority of compound-exposed samples move into the direction of the first principal component. Connecting lines between replicates (Figure 3.1 B) demonstrate a low degree of technical variability, since most of the paired replicates are located close to each other. For this reason, the following PCAs include only the mean values of two matching replicates (Figure 3.1 C). Connecting lines between compound-exposed samples and corresponding controls illustrate that treatment-control pairs are located in the same main cluster (Figure 3.1 D), suggesting that the reason for the formation of the two main clusters can be explained by experimental variability, a so called batch effect. The subtraction of the controls from the corresponding compound-exposed samples reverses the cluster formation, which supports the assumption that there is no effect related to a biological or scientific variable (Figure 3.1 E).

3.1.3 Evaluation of data reproducibility across replicates

In order to assess the reproducibility between the two replicates of a sample, the Euclidean distances between all pairs of replicates (replicate sample pairs, as well as control-treatment sample pairs) in the PCA plot were determined. For the samples that were tested at the highest concentration for 24 h, the median distance between control-treatment replicates was 4.9-fold higher than the median distances between the two identically treated replicate pairs. The frequency distribution of Euclidean distances of the replicate sample pairs is illustrated in Figure 3.2A.

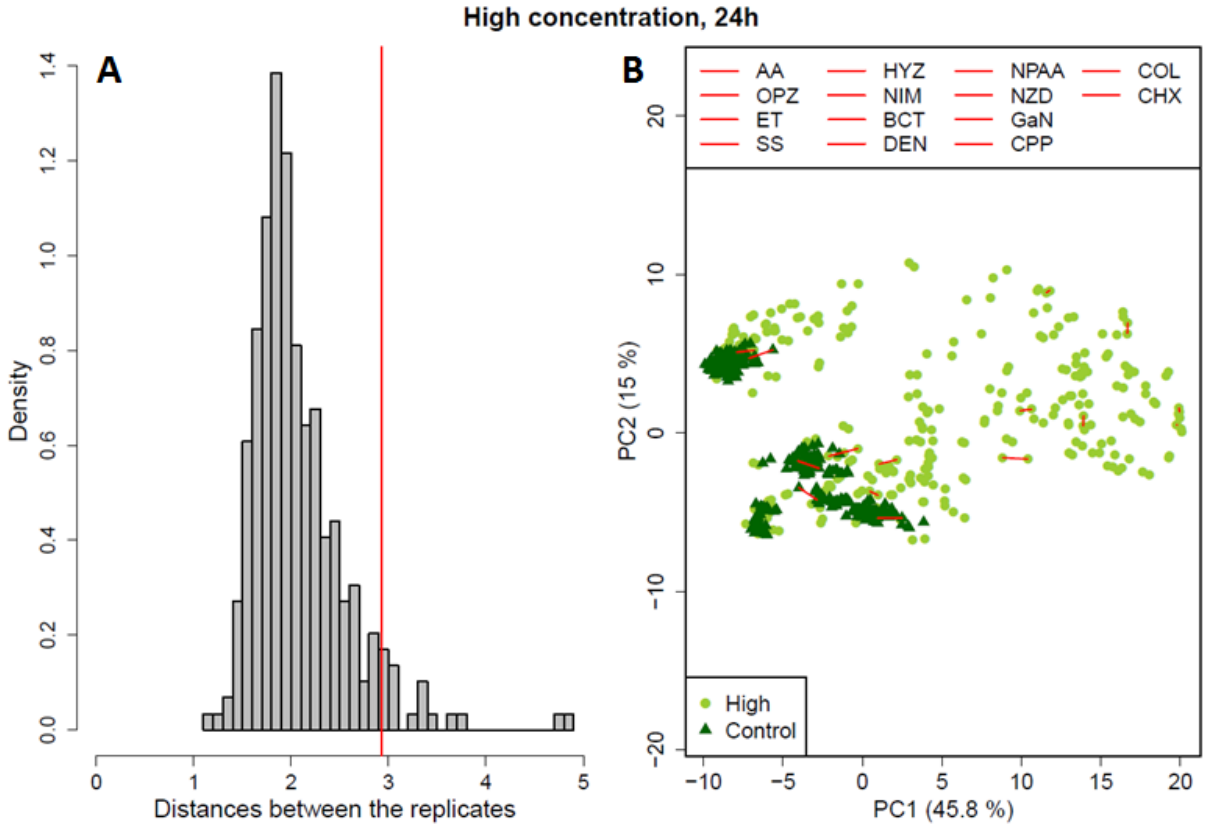


Figure 3.2: Reproducibility between replicates. **A** Frequency distribution of the Euclidean distance between all pairs of sample replicates. The red line shows the 5 % largest observed distances between the replicates. **B** PCA analysis of the 24h highest concentration subset. The connecting lines indicate that 14 out of 148 (9.5 %) tested compounds belong to the five percent of the replicate sample pairs with the highest Euclidean distance in the PCA plot.

The red line in the histogram in Figure 3.2 A separates the five percent largest observed distances from the main distribution, representing 14 out of 148 tested compounds (9.5 %) tested in the 24 h, high concentration subset. Related replicate pairs are shown by connected lines in the PCA plot (Figure 3.2 B). Compared to the much larger compound effects (**Figure 3.1**), the variability among replicates with the highest Euclidean distance is relatively small. Therefore, the reproducibility among replicates is most widely very high and the degree of variability in gene expression of identically treated samples is in an acceptable range.

3.1.4 Number of deregulated genes per compound

Comparison of the number of deregulated genes among different compounds revealed that the majority of gene expression effects can be attributed to a relatively small subset of compounds. There is a time and concentration dependent increase in the number of significantly up regulated genes per compound for the fold changes 1.5, 2 and 3 fold (Figure 3.3).

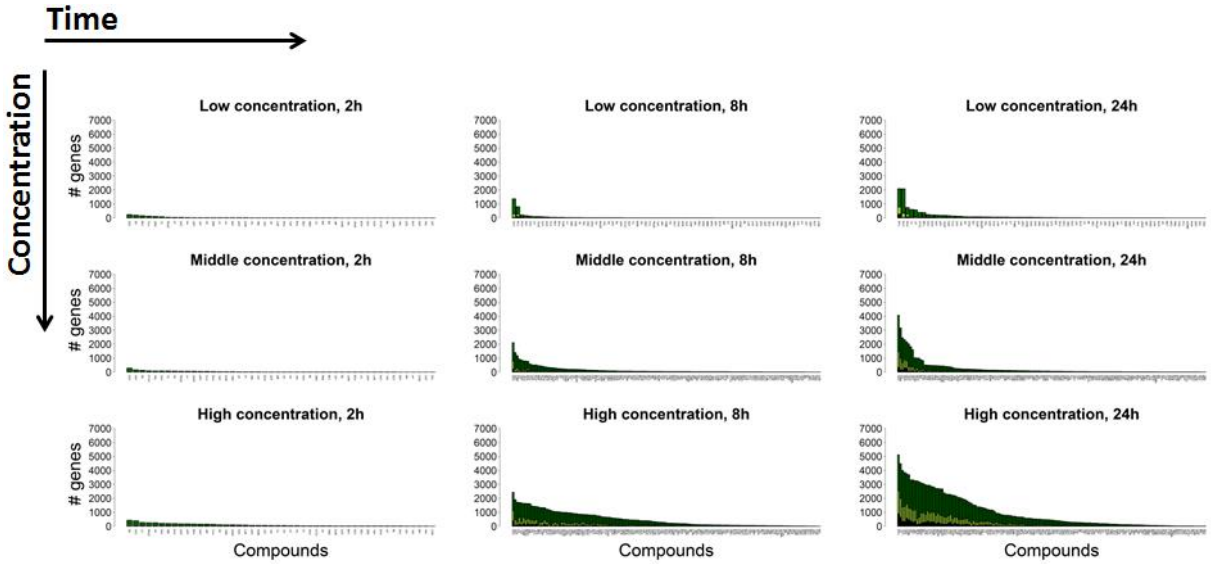


Figure 3.3: Number of significantly up regulated genes per compound. For all concentration and time series, all compounds are listed on the x-axis. The y-axis illustrates the number of up regulated genes with at least 1.5, 2 or 3 fold up regulation. Dark green: more than 1.5 fold up regulated; light green: more than 2 fold up regulated; black: more than 3 fold up regulated.

Substantial differences between the compounds cycloheximide (CHX) and triazolam (TZM) were observed for the latest time point (24h) at the highest concentration. While CHX up regulated expression of 8,558 genes (5,124 genes with at least 1.5 fold induction, 2,547 genes with at least 2 fold induction and 887 genes which are at least 3 fold up regulated), TZM deregulated under the same conditions only 38 genes in total – 6 genes were at least 1.5 fold up regulated, 31 down regulated and only one gene was at least 2 fold down regulated.

The situation for the down regulated genes was similar (Figure 3.4). There was a time and concentration dependent increase in the number of significantly down regulated genes. The strongest effects on gene expression can be attributed to a relatively small subset of compounds, whereas the majority of compounds show less strong effects.

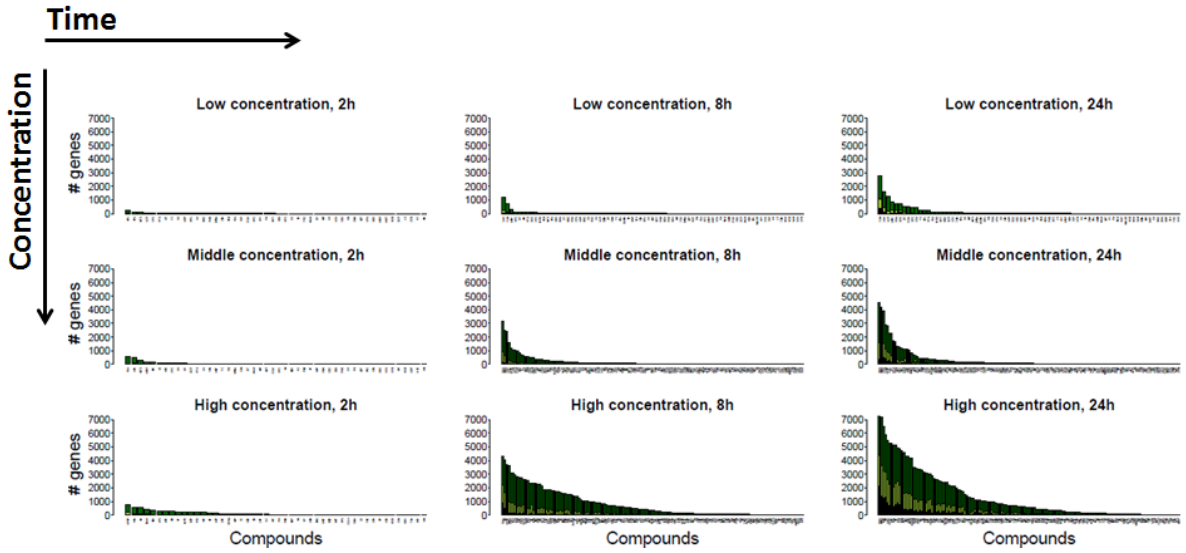


Figure 3.4: Number of significantly down regulated genes per compound. For all concentration and time series, all compounds are listed on the x-axis. The y-axis illustrates the number of down regulated genes with at least 1.5, 2 or 3 fold down regulation. Dark green: more than 1.5 fold down regulated; light green: more than 2 fold down regulated; black: more than 3 fold down regulated.

Among the 151 tested compounds, 48 substances showed a very weak effect on gene expression and up or down regulated not more than 20 genes in total for each time point at any concentration (Supplemental Table 2). Surprisingly, among these compounds was also carbon tetrachloride (CCl_4), a compound that is well documented to be hepatotoxic (Bauer et al. 2009; Hoehme et al. 2010; Weber et al. 2003). Eleven of the 48 compounds exhibited even less effect on gene expression and deregulated at most 20 genes in total, independent of the direction (induction or down regulation) and time period. These compounds are amiodarone, bromobenzene, cimetidine, clofibrate, coumarin, gemfibrozil, glibenclamide, haloperidol, hexachlorobenzene, phenytoin and sulfasalazine.

In contrast, the strongest effects with the highest fold changes among all genes and across all compounds were attributed to a small set of compounds. Figure 3.5 and Figure 3.6 show to which degree the compounds contribute to the 100 most up or down regulated genes for all incubation periods. The black bars give the number of genes which belong to the TOP 100 fraction, whereas the white bars indicate how many of the TOP 100 genes are at least two fold up or down regulated. For all conditions, the number of compounds with the strongest effects on gene expression was comparably low. The amount of at least twofold induced or down regulated genes increased concentration dependently. The strongest induction was observed for the 24h high-concentration time point, where all TOP 100 genes were at least twofold deregulated. However, only 32 compounds were responsible for the 100 most up regulated genes at this time point and only 23 compounds were responsible for the 100 strongest down regulations.

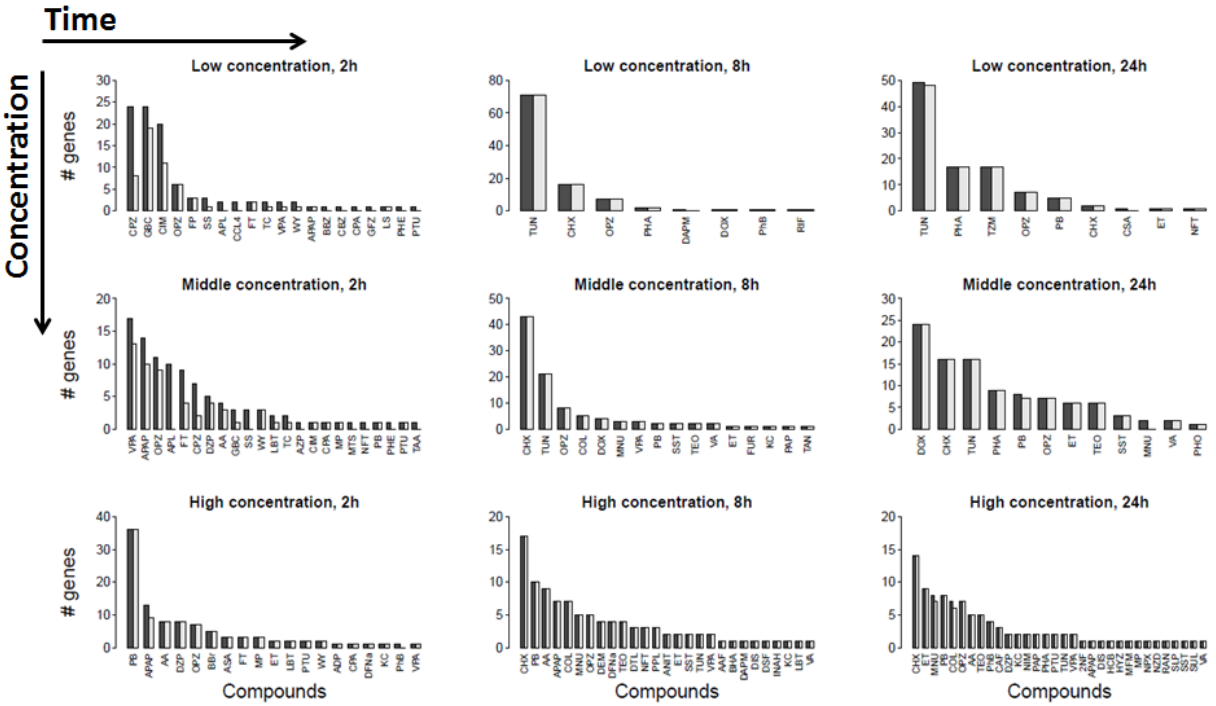


Figure 3.5: Analysis of the strongest up regulated genes with the highest fold change across all compounds. The x-axis lists the compounds which are responsible for the 100 induced genes with the highest fold change. The y-axis gives the number of significantly up regulated genes for the listed compounds. The black bars illustrate the contribution of genes for the appropriate compound that is among the 100 genes with the strongest fold change. How many of these genes are up regulated with a fold change of at least 2 is demonstrated by the white bars.

This result indicates that either a large fraction of the 151 compounds cause only weak gene expression alterations or that strong effects on gene expression requires a higher concentration than the slightly cytotoxic one that was tested. Previous studies have shown that the identification of close to cytotoxic concentrations is challenging, especially for compounds with steep dose response curves (Krug et al. 2013; Waldmann et al. 2014). In addition, the method by which toxicity is determined may play a role.

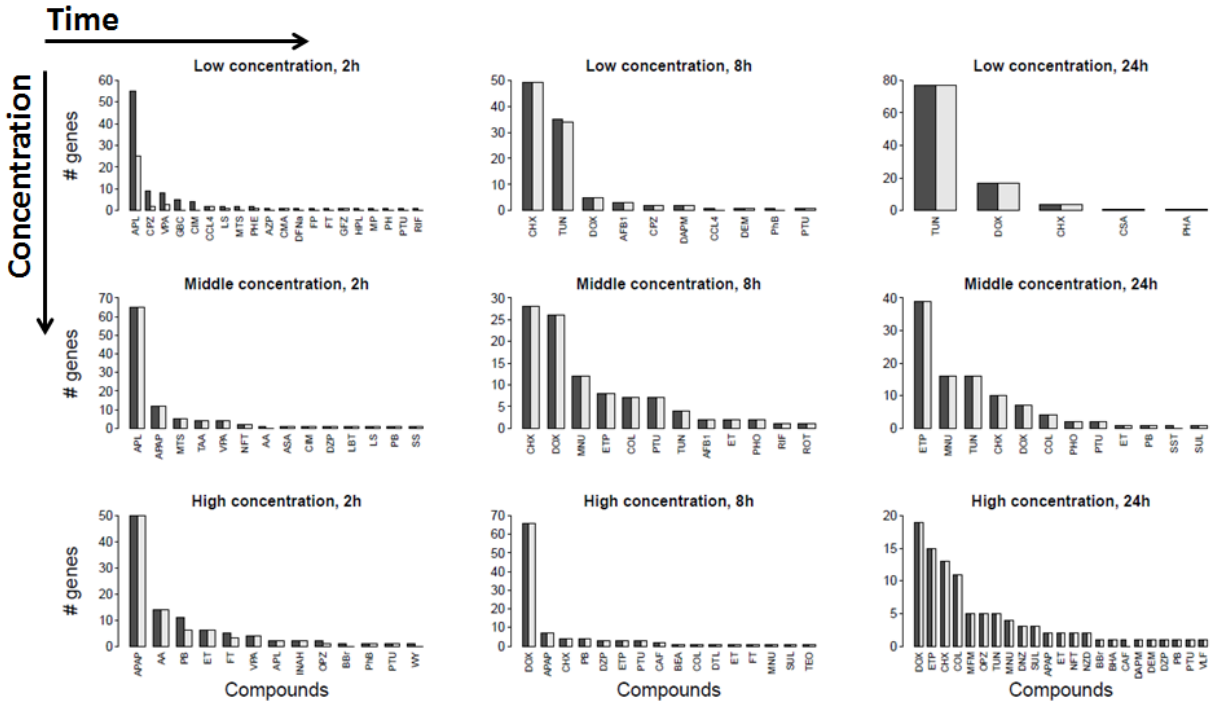


Figure 3.6: Analysis of the strongest down regulated genes with the highest fold change across all compounds. The x-axis lists the compounds which are responsible for the 100 strongest down regulations with the highest fold change. The y-axis gives the number of significantly down regulated genes for the listed compounds. The black bars illustrate the contribution of genes for the appropriate compound that is among the 100 genes with the strongest fold change. How many of these genes are up regulated with a fold change of at least 2 is demonstrated by the white bars.

3.1.5 Exclusion of compounds following an implausible concentration progression

A logical assumption is that genes, which are deregulated at a low concentration of a particular compound, are also deregulated at a respective higher concentration. If such a dose response relationship is not given, further analysis is required to interpret the data. Two types of analysis were performed to elucidate the concentration progression across the data base. To describe, at which concentration (low, middle or high) the deregulation of a gene occured, the ‘progression profile index’ was created. Second, compounds following an implausible concentration progression (deregulating a large fraction of genes at a lower, but not at a higher dose) were identified by introducing the ’progression profile error indicator’.

The analysis of concentration progression is exemplary shown for the 4 compounds valproic acid (VPA), propranolol (PPL), triazolam (TZM) and allyl alcohol (AA) (Figure 3.7). All 4 compounds exhibited different concentration progression profiles. Whereas for VPA a large fraction of genes was deregulated at the middle concentration and even more genes at the highest concentration, AA and PPL deregulated the largest fraction of genes only at the highest, but not at any lower concentration. TZM exhibited an unusual concentration progression and showed an effect on gene expression only at the lowest concentration, but no additional genes were altered with increasing concentrations. The graphs shown comprise exclusively genes which are at least twofold significantly up or down regulated, independent

of the direction of deregulation. Corresponding Venn diagrams show the amount and overlap of deregulated genes across the concentrations.

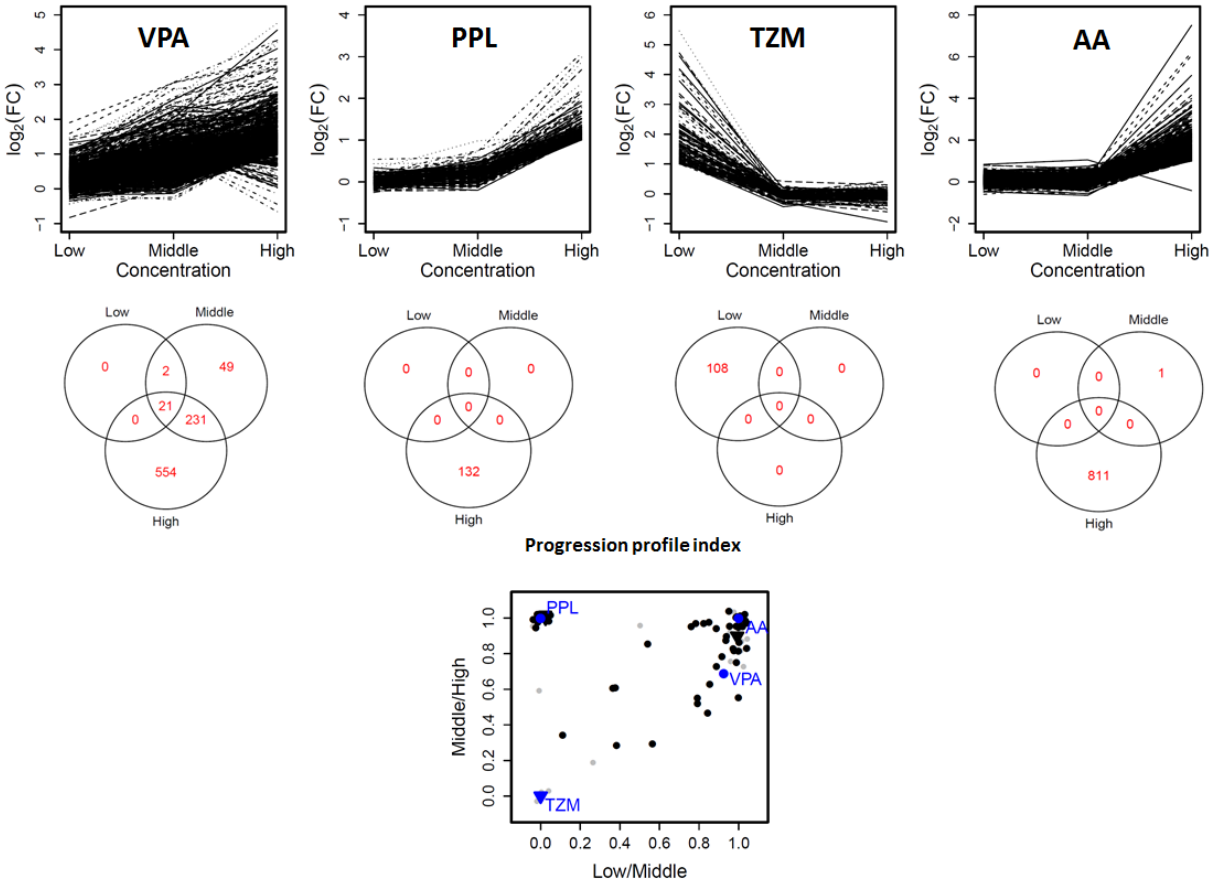


Figure 3.7: Analysis of concentration progression with the 'principal progression profile index' and 'error indicator', shown for the compounds valproic acid (VPA), propranolol (PPL), triazolam (TZM) and allyl alcohol (AA) after 24h of exposure. The first row shows the expression course of all at least 2 fold significantly deregulated genes by the considered compounds across the 3 concentrations. The corresponding Venn diagrams are shown in the middle, illustrating the overlap of at least 2 fold up or down regulated genes at the different concentrations. The lowest panel shows the distribution of the compounds in the progression profile index. In blue, the 4 mentioned compounds are marked; triangles represent the later excluded compounds.

The 'progression profile index' illustrates the fraction of genes, which were at least twofold up or down regulated at a higher concentration compared to a lower concentration. Every symbol in the diagram represents one compound. For each substance the proportions middle vs. high (y-axis) and low vs. middle (x-axis) were calculated (Figure 3.7). A value close to one indicates that a large fraction of genes is altered at a higher, compared to a respective lower concentration. If only a few additional genes were deregulated with an increased concentration, the value is closer to zero. VPA is positioned in the upper right of the panel. With each concentration step, additional genes became deregulated. However, TZM deregulated genes only at the lowest concentration and therefore clusters in the lower right corner. This compound follows an implausible concentration progression, since no additional genes were deregulated with an increasing concentration. PPL and AA exhibited a pattern of concentration dependent deregulation where most genes were altered at the highest concentration.

Nevertheless, these compounds cluster in different regions in the 'progression profile index' plot, due to one outlying gene altered by the middle concentration of allyl alcohol.

An overview of the 'progression profile indices' for all 151 compounds across all time points is shown in Figure 3.8. In a first step, the genes up or down regulated (at least twofold deregulated compared to the controls) at a higher concentration were determined. Second, the proportion of genes that were not altered at the lower concentration was determined. This was achieved by comparing the middle versus low (x-axis) and the high versus middle (y-axis) concentration.

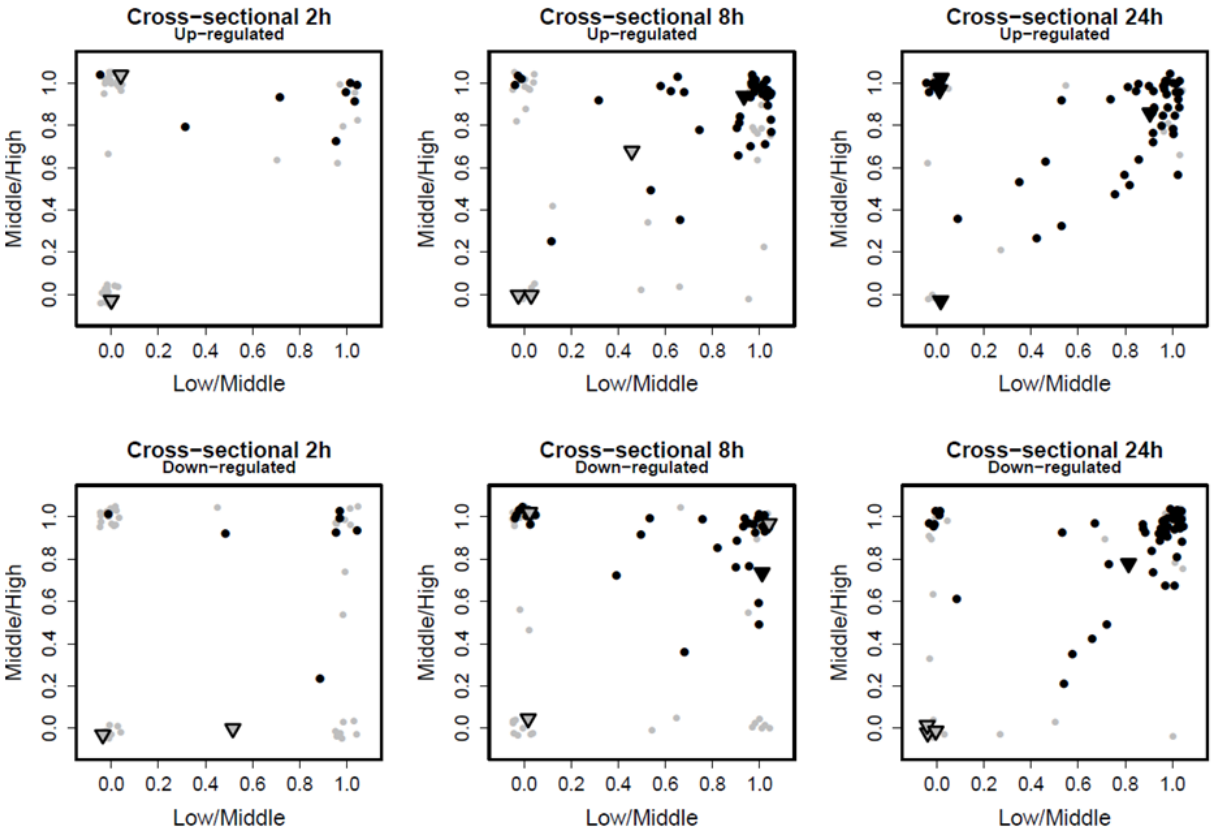


Figure 3.8: Progression profile index for all compounds which have been tested at the three concentrations across all time points. Each point represents one compound. Triangles show the latter excluded compounds. Gray symbols: compounds which deregulate at most 20 genes in total. Black symbols: compounds which deregulate more than 20 genes in total.

The number of compounds causing up and down regulation of target genes, increased over time. Gradually, from 2h to 24h of exposure, most compounds cluster in the upper right corner, indicating that additional genes become up or down regulated with increasing concentration steps. Compounds that up or down regulate more than 20 genes are shown in black; the gray symbols represent the fraction of relatively weak compounds that deregulate at most 20 genes in total (Supplemental Table 2). The second biggest cluster is located in the upper left corner and represents the compounds where the largest fraction of genes was deregulated by increasing the concentration from middle to high, whereas no genes were altered at a lower concentration. Compounds that deregulate genes solely at the lowest, but

only a few additional genes at a higher concentration, cluster in the lower left corner. All of these compounds show only weak effects in gene expression and deregulate at most 20 genes in total. In general, the clustering of weak compounds mainly occurs in the left part of the 'progression profile index' plot.

To elucidate whether a compound mainly deregulates genes at a lower, but not at a respective higher concentration, the 'progression profile error indicator' was introduced (Figure 3.9). In this case the x-axis reflects the ratio middle/low, whereas the y-axis represents the ratio high/middle. Ideally both values are below 0.5. Compounds cluster in the left part of the plot if they mainly up and/or down regulate genes at the middle dose compared to the lowest dose. When even more genes are deregulated by increasing the middle to the highest dose, the compounds cluster in the lower left corner. Values above 0.5 indicate an implausible concentration progression; in these cases the compounds deregulate genes mainly at a lower, but not at a respective higher concentration. If only a few genes are altered at lower, but not with increasing concentrations, they can be interpreted as outliers.

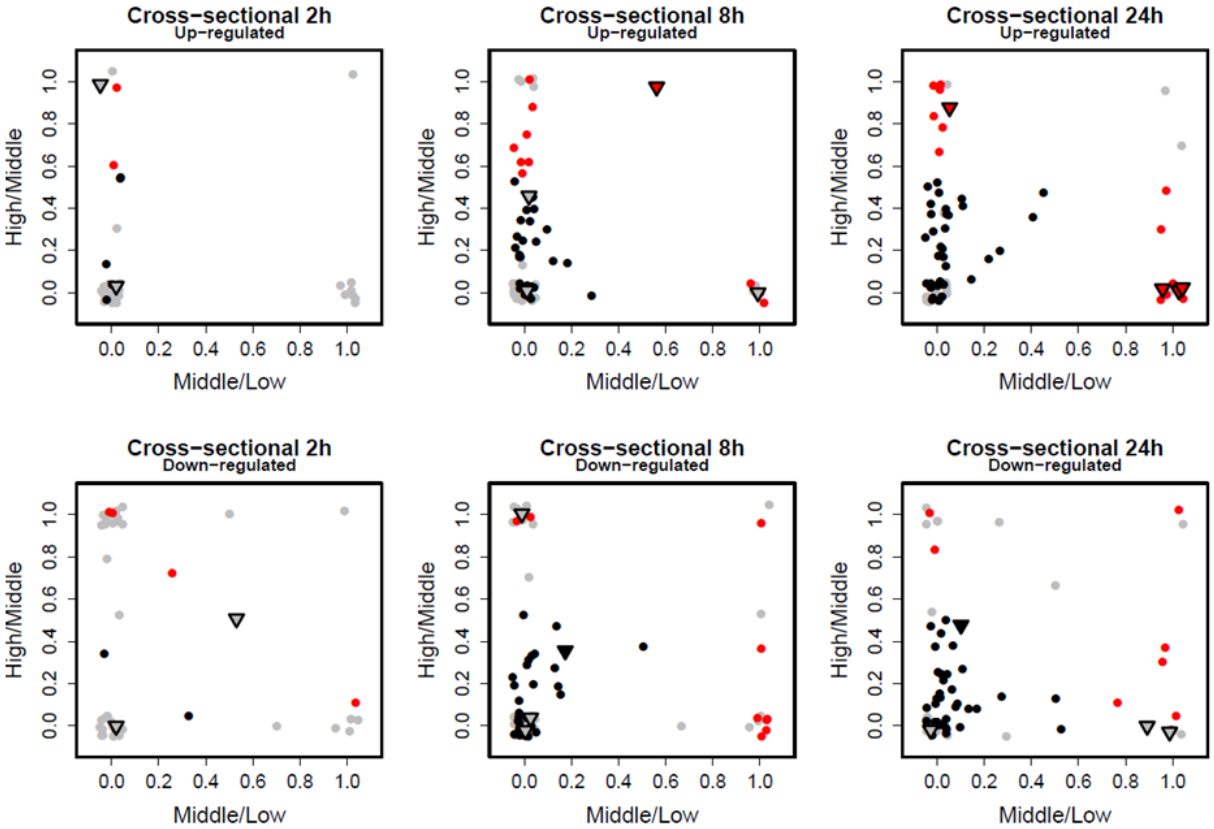


Figure 3.9: Progression profile error indicator for up and down regulated genes at different time points. A high value means that a high fraction of genes is deregulated exclusively at a lower compared to a respective higher concentration. Each point represents one compound. Black symbols indicate that a compound deregulates more than 20 genes in total and that both values are ≥ 0.5. Gray symbols represent compounds that deregulate at most 20 genes in total. Red marked compounds deregulate more than 20 genes in total but exhibit at least one error indicator value above 0.5. Triangles show mark compounds that are excluded from further analysis.

Again the time dependent increase in the amount of compounds that cause gene expression becomes obvious. For the up and down regulated genes, more and more compounds cluster within time in the left part of the 'progression profile error indicator' plot, indicating that

genes are predominantly deregulated at the middle and the highest concentration. Although the majority of substances clusters in the lower left part, exhibiting a small high/middle error indicator value, a large subset of compounds obviously induces gene expression at the medium concentration, meaning that only a few more genes are induced at the highest concentration (red symbols in Figure 3.9). Compounds which cluster in the lower right part of the error indicator plot follow a non-monotonous concentration progression. The largest amount of genes that are deregulated by these compounds is found to be altered at low concentration already, stays deregulated at the middle and the highest concentration, but does not increase dose dependently.

Due to the calculation procedure of the 'progression profile indices', the influence of outliers is immense. Single genes that are deregulated at a lower, but not at a respective higher concentration, may increase the error indicator value to a large extent. This can be demonstrated by comparing the progression profile indices of the previously mentioned compounds valproic acid, triazolam, propranolol and allyl alcohol (Figure 3.10).

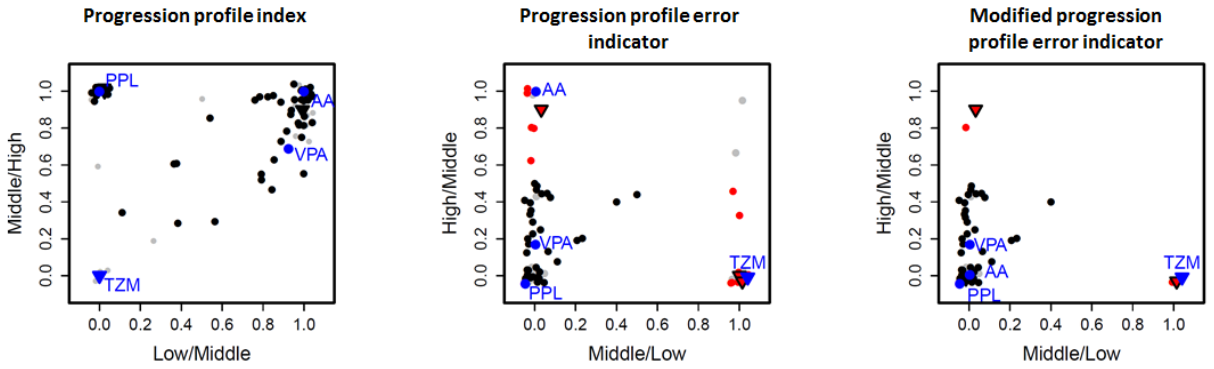


Figure 3.10: Progression profile indices for all 151 compounds after 24h of exposure. In blue, the 4 mentioned compounds are marked, red shows the 11 compounds which deregulate at most 20 genes in total. Triangles represent the later excluded compounds.

As an example, the gene expression profiles and Venn diagrams of propranolol and allyl alcohol (Figure 3.7) show that both compounds dose dependently increase the number of deregulated genes after 24 h of exposure. The largest fraction of genes is not altered at the lowest and the middle concentration and most events happen only at the highest, slightly cytotoxic concentration. However, propranolol and allyl alcohol cluster in different regions of the 'progression profile index' and the 'progression profile error indicator' plot. According to the Venn diagrams, a single gene is deregulated by allyl alcohol at the middle, but not at the highest concentration – which leads to a high high/middle ratio in the 'progression profile error indicator'. To decrease the influence of such outliers, the modified 'progression profile error indicator' was introduced. It is an adjustment of the 'progression profile error indicator' which sets the index of a compound to zero if the error indicator value of two compared concentrations is ≥0.5 and if the number of genes, which are deregulated at the lower concentration, is at most 20. The modified 'progression profile error indicator' evaluates all compounds and compares the low/middle ratio and the middle/high ratio of deregulated genes and it considers the amount of at least twofold up and down regulated genes.

Meaning, if the error indicator value for low vs. middle is ≥ 0.5 and if the compound does not alter more than 20 genes in total at the lower concentration, the index is set to zero. Analogously, if the error indicator middle vs. high is ≥0.5 and if the compound does not alter more than 20 genes at the middle concentration, the index is set to zero. In case of allyl alcohol, normalizing the 'progression profile error indicator' value artificially to zero results into the exclusion of the single outlying gene from the analysis and allyl alcohol re-clusters close to propranolol in the modified 'progression profile error indicator´ plot (Figure 3.10). However, triazolam exhibits an ambiguous concentration progression. One hundred and eight genes were up regulated at the lowest concentration, but no genes were altered at the middle and the high concentration (Fig. AA7). This results in a high value for the middle to low ratio and triazolam clusters in the lower right corner in the 'progression profile error indicator' plot, as well as in the modified 'progression profile error indicator' plot. Since this compound follows an implausible concentration progression, it was excluded for further analysis. In contrast, valproic acid and propranolol directly exhibit a small error indicator value, indicating that the deregulation of genes follows a monotonous concentration progression. Their position does not change in the modified 'progression profile error indicator' plot, as well.

On the basis of the modified 'progression profile error indicator', a 'progression error profile' was developed for each compound. The following labels are introduced to annotate the concentration progression for each time point. NA: the compound was not tested for the respective time point. OO: the number of differentially expressed genes is zero for all concentrations. o: indicates that the number of differentially expressed genes is ≤ 20 for the tested time point. +: the number of differentially expressed genes is ≥ 20 and that both 'progression profile error indicator' values are above 0.5. -: the number of differentially expressed genes is ≥ 20 but at least 1 error indicator value is above 0.5. The concentration progression for each individual compound is given as follows: 2h up │8h up │24h up │ 2h down │ 8h down │24h down (Table 3.2 and Supplemental Table 3). Compounds with the same concentration progression profile can be assigned to a group. In total, 63 different profiles were observed. The largest group of compounds with the same profile comprised 35 compounds, representing the concentration progression pattern NA │ + │ + │NA │ + │ +, which exhibits a plausible, monotonous concentration progression for the 8 h and 24 h incubation period. For the 2 h time point, no data is available. Five compounds follow an implausible concentration progression and were excluded from further analysis: doxorubicin (NA │ - │ - │ NA │ + │ +), triazolam (NA │ o │ - │ NA│ OO │ OO), tetracycline (o │o │ - │ OO │ OO), ticlopedine (NA │ o │o │ NA │ o │ -) and carbon tetra chloride (CCl_4). CCl_4 features the concentration progression profile o │ o │ + │o │ o │ o that is also found by the compounds aspirin, indomethacin and methyltestosterone. According to the strong hepatotoxic potential of CCl_4, this expression response is unexpected, since the compound represents a strong hepatotoxin with a well characterized mechanism of toxicity (Bauer et al. 2009; Hoehme et al. 2010; Weber et al. 2003). However, CCl_4 is a highly lipophilic compound and dissolving it in aqueous solutions is technically challenging. An inhomogeneous distribution of the compound in the medium supernatant of the cells may explain the experimental results. Follow-up studies will be re-

quired to reanalyze the compounds which have been excluded due to implausible concentration progression.

Table 3.2: Progression error profiles for all compounds. Based on the modified error indicator values the compounds were assigned to the labels “NA”, “OO”, “o”, “+” and “-“.”NA”: the compound was not tested for the respective time point. “OO”: the number of differentially expressed genes is zero for all concentrations. o: indicates that the number of differentially expressed is ≤ 20 for the tested time point. “+”: the number of differentially expressed genes is ≥ 20 and that both ‘progression profile error indicator’ values are above 0.5. “-“: the number of differentially expressed genes is ≥ 20 but at least 1 error indicator value is above 0.5. For each time point (up- and down) one label was annotated so that the profile for one compound is composed is designed as follows: "2h Up| 8h Up|24h Up|2h Down|8h Down|24h Down". Compounds marked in red follow an implausible concentration progression and were excluded from further analysis.

Abbr.	Compound name	Progression error profile
2NF	2-nitrofluorene	NA|o|o|NA|o|+
AA	allyl alcohol	+|+|+|+|+|+
AAA	acetamide	NA|OO|OO|NA|OO|OO
AAF	acetamidofluorene	NA|+|+|NA|+|o
ACA	acarbose	NA|o|o|NA|OO|o
ACZ	acetazolamide	NA|OO|OO|NA|OO|OO
ADM	alpidem	NA|NA|o|NA|NA|+
ADP	adapin	o|+|+|o|+|+
AFB1	aflatoxin B1	NA|+|+|NA|+|+
AJM	ajmaline	NA|+|+|NA|o|+
AM	amiodarone	OO|OO|o|o|o|o
AMB	amphotericin B	NA|+|+|NA|+|+
AMT	amitriptyline	NA|o|o|NA|o|o
ANIT	naphthyl isothiocyanate	OO|+|+|OO|+|+
APAP	acetaminophen	o|+|+|+|+|+
APL	allopurinol	OO|o|+|+|o|+
ASA	aspirin	o|o|+|o|o|o
AZP	azathioprine	OO|+|+|o|+|+
BBr	benzbromarone	+|+|+|o|+|+
BBZ	bromobenzene	OO|o|OO|o|o|o
BCT	bucetin	NA|o|o|NA|o|o
BDZ	bendazac	NA|o|+|NA|o|+
BEA	bromoethylamine	NA|+|+|NA|+|+
BHA	butylated hydroxyanisole	NA|+|+|NA|+|+
BPR	buspirone	NA|NA|+|NA|NA|+
BSO	buthionine sulfoximine	NA|OO|o|NA|o|o
BZD	benziodarone	NA|o|+|NA|o|+

Abbr.	Compound name	Progression error profile
CAF	caffeine	NA|+|+|NA|+|+
CAP	captopril	NA|+|+|NA|+|+
CBZ	carbamazepine	o|o|+|o|OO|+
CCL4	carbon tetrachloride	o|o|+|o|o|o
CFB	clofibrate	o|o|o|o|o|o
CHL	chlorpheniramine	NA|OO|o|NA|OO|o
CHX	cycloheximide	NA|+|+|NA|+|+
CIM	cimetidine	o|o|o|OO|OO|o
CLM	chlormadinone	NA|+|+|NA|+|+
CMA	coumarin	OO|o|o|o|o|o
CMN	chlormezanone	NA|OO|o|NA|o|o
CMP	chloramphenicol	NA|o|o|NA|o|+
COL	colchicine	NA|+|+|NA|+|+
CPA	cyclophosphamide	o|o|+|o|o|+
CPM	clomipramine	NA|OO|OO|NA|o|o
CPP	chlorpropamide	NA|o|o|NA|o|o
CPX	ciprofloxacin	NA|o|OO|NA|OO|o
CPZ	chlorpromazine	+|o|+|o|o|+
CSA	cyclosporine A	NA|+|-|NA|o|+
CZP	clozapine	NA|NA|+|NA|NA|+
DAPM	methylene dianiline	NA|+|+|NA|+|+
DEM	diethyl maleate	NA|+|+|NA|+|+
DEN	nitrosodiethylamine	NA|OO|o|NA|o|+
DEX	dexamethasone	NA|o|+|NA|o|+
DFN	diclofenac	o|+|+|o|+|+
DIL	diltiazem	NA|o|+|NA|+|+
DIS	disopyramide	NA|+|+|NA|+|+

Abbr.	Compound name	Progression error profile
DNP	2,4-dinitrophenol	NA|+|+|NA|+|+
DNZ	danazol	NA|+|+|NA|+|+
DOX	doxorubicin	NA|-|-|NA|+|+
DSF	disulfiram	NA|+|o|NA|+|o
DTL	dantrolene	NA|+|o|NA|+|OO
DZP	diazepam	o|+|+|+|+|+
EBU	ethambutol	NA|+|+|NA|+|+
EE	ethinylestradiol	NA|o|o|NA|o|o
EME	erythromycin ethylsuccinate	NA|o|o|NA|OO|o
ENA	enalapril	NA|+|+|NA|+|+
ET	ethionine	o|+|+|o|+|+
ETH	ethionamide	NA|o|+|NA|o|+
ETN	ethanol	NA|o|o|NA|o|o
ETP	etoposide	NA|+|+|NA|+|+
FAM	famotidine	NA|OO|o|NA|OO|o
FFB	fenofibrate	NA|o|o|NA|OO|o
FLX	fluoxetine hydrochloride	NA|+|+|NA|+|+
FP	fluphenazine	o|o|+|o|o|+
FT	flutamide	+|+|+|o|+|+
FUR	furosemide	NA|+|+|NA|+|+
GaN	galactosamine	NA|o|+|NA|+|+
GBC	glibenclamide	o|o|o|OO|o|o
GF	griseofulvin	o|o|o|o|o|+
GFZ	gemfibrozil	o|o|o|o|OO|o
HCB	hexachlorobenzene	o|o|o|o|o|o
HPL	haloperidol	OO|o|o|o|o|o
HYZ	hydroxyzine	NA|o|+|NA|+|+
IBU	ibuprofen	NA|o|+|NA|+|o
IM	indomethacin	o|o|+|o|o|o
IMI	imipramine	NA|OO|o|NA|OO|OO
INAH	isoniazid	OO|+|+|+|+|+
IPA	iproniazid	NA|+|+|NA|o|+
KC	ketoconazole	o|+|+|o|+|+
LBT	labetalol	o|+|+|o|+|+
LNX	lornoxicam	NA|o|OO|NA|OO|OO
LS	lomustine	o|+|o|o|+|+
MCT	monocrotaline	NA|o|o|NA|o|o
MDP	methyldopa	NA|o|o|NA|o|OO
MEF	mefenamic acid	NA|o|o|NA|+|o
MEX	mexiletine	NA|+|+|NA|o|+
MFM	metformin	NA|+|+|NA|+|+
MLX	meloxicam	NA|OO|o|NA|o|OO
MNU	N-methyl-N-nitrosourea	NA|+|+|NA|+|+
MP	methapyrilene	+|+|+|o|+|+
MTS	methyltestosterone	o|o|+|o|o|o

Abbr.	Compound name	Progression error profile
MTZ	methimazole	NA|+|+|NA|+|+
MXS	moxisylyte	NA|+|+|NA|+|+
NFT	nitrofurantoin	o|+|+|o|+|+
NFZ	nitrofurazone	NA|o|+|NA|o|+
NIC	nicotinic acid	NA|OO|o|NA|+|o
NIF	nifedipine	NA|+|+|NA|+|+
NIM	nimesulide	NA|+|+|NA|+|+
NMOR	N-nitrosomorpholine	NA|o|o|NA|OO|o
NPAA	phenylanthranilic acid	NA|+|+|NA|+|+
NPX	naproxen	NA|o|+|NA|+|+
NZD	nefazodone	NA|NA|+|NA|NA|+
OPZ	omeprazole	+|-|+|o|+|+
PAP	papaverine	NA|+|+|NA|+|+
PB	phenobarbital	+|+|-|+|+|+
PCT	phenacetin	NA|o|o|NA|o|o
PEN	penicillamine	NA|o|+|NA|+|+
PH	perhexiline	o|+|+|o|+|+
PHA	phalloidin	NA|+|+|NA|+|+
PhB	phenylbutazone	o|+|+|o|o|+
PHE	phenytoin	OO|o|o|o|OO|o
PHO	phorone	NA|-|o|NA|+|+
PML	pemoline	NA|o|OO|NA|OO|OO
PMZ	promethazine	NA|+|+|NA|+|+
PPL	propranolol	NA|+|+|NA|+|+
PTU	propylthiouracil	o|+|+|o|+|+
QND	quinidine	NA|o|+|NA|o|+
RAN	ranitidine	NA|+|+|NA|+|+
RGZ	rosiglitazone maleate	NA|+|+|NA|+|+
RIF	rifampicin	OO|o|+|o|o|o
ROT	rotenone	NA|o|OO|NA|o|o
SLP	sulpiride	NA|+|+|NA|+|+
SS	sulfasalazine	o|o|o|o|o|o
SST	simvastatin	NA|o|+|NA|o|+
SUL	sulindac	NA|+|-|NA|+|+
TAA	thioacetamide	OO|o|+|o|o|o
TAC	tacrine	NA|o|+|NA|+|+
TAN	tannic acid	NA|+|+|NA|+|+
TBF	terbinafine	NA|o|o|NA|OO|o
TC	tetracycline	o|o|-|OO|o|o
TCP	ticlopidine	NA|o|o|NA|o|-
TEO	theophylline	NA|-|+|NA|+|+
TIO	tiopronin	NA|o|OO|NA|OO|o
TLB	tolbutamide	NA|o|+|NA|o|+
TMD	trimethadione	NA|o|OO|NA|OO|OO
TMX	tamoxifen	NA|OO|o|NA|OO|OO

Abbr.	Compound name	Progression error profile
TRZ	thioridazine	o|o|o|OO|o|+
TUN	tunicamycin	NA|+|+|NA|+|+
TZM	triazolam	NA|o|-|NA|OO|OO
VA	vitamin A	NA|o|+|NA|-|o

Abbr.	Compound name	Progression error profile
VLF	venlafaxine	NA|NA|+|NA|NA|+
VPA	valproic acid	+|+|+|+|+|+
WY	WY-14643	o|o|+|OO|o|o

Compounds which deregulate more than 20 genes and have an error indicator value > 0.5 are listed in Table 3.3. Small error indicator values between 0.1 and 0.4 across all time points are observed for all 32 (23) compounds that contribute to the 100 most up (down) regulated genes (Figure 3.5 and Figure 3.6). In contrast, compounds with weak effects on gene expression that deregulate less than 20 genes in total, exhibit the highest error indicator values of all compounds (Supplemental Table 3). Since these compounds deregulate the majority of genes at the lowest concentration only, they cluster in the right part of the 'progression profile error indicator' plot and indicate a high low/middle ratio, as well as middle/high ratio. Neither the concentration progression from the middle to low, nor from the high to middle concentration follows a monotonous course. In conclusion, implausible concentration progression profiles were mainly observed for compounds with weak gene expression responses. For this reason the rest of the study focusses on the compounds with strong effects on gene expression, whereas compounds with an unusual concentration progression have been removed.

Table 3.3: Compounds that deregulated (2-fold up or down compared to control) more than 20 genes in total at any concentration and yield at least one error indicator value which is greater than 0.5.

Up regulation	Down regulation
2 h	
allyl alcohol valproic acid	allyl alcoholc acetaminophen allopurinol phenobarbital
8 h	
azathioprine benzbromarone bromoethylamine doxorubicin methapyrilene pmeprazole phenobarbital perhexilene phenylbutazone phorone rosiglitazone maleate	allyl alcohol diclofenac flutamide labetalol lomustine methapyrilene omeprazole phenobarbital phorone

Up regulation	Down regulation
24 h	
allyl alcohol adapin benbromarone carbon tetrachloride chlorpromazine cyclosporine A diethyl maleate 2,4-dinitrophenol doxorubicin fluotexine hydrochloride fluphenazine galactosamine methapyrilene phenobarbital phenylbutazone tetracycline triazolam	chlorpromazine cyclosporine A dexamethasone diclofenac 2,4-dinitrophenol griseofulvin isoniazid methapyrilene perhexilene

3.1.6 Reproduction of the gene expression effects observed by TG GATES *in vitro*

Microarray technology provides a powerful tool for gene expression analysis, but the specificity, sensitivity and reproducibility is often controversial (Draghici et al. 2006; Kothapalli et al. 2002). In order to evaluate the reproducibility of the TG GATES data on chemically-induced gene expression effects, quantitative real time PCR (qPCR) of compound-exposed primary human hepatocytes was performed.

Table 3.4: Comparison of TG-GATES gene array data with qPCR data from compound treated primary human hepatocytes. Quantitative gene expression was performed for 2-5 replicates (cells from different donors). Only the highest concentration at time point 24 h was validated.

Compound	Gene	TG-GATES data: fold changes				Quantitative real time PCR data: fold changes						
		Probe set 1	Probe set 2	Probe set 3	Probe set 4	Rep. 1	Rep. 2	Rep. 3	Rep. 4	Rep. 5	Mean	SD
Valproic acid 5mM	G6PD	23.71				37.16	8.61	39.61	21.59	43.30	**30.05**	14.57
	PDK4	4.23	3.74	1.01		2.93	3.25	1.57	2.16	5.85	**3.15**	1.64
	PCK1	2.27				5.45	13.10	4.99	3.30	7.43	**6.86**	3.79
	INSIG	2.08	2.44	2.16	2.27	1.10	1.39	0.39	0.68	0.54	**0.82**	0.41
Ketoconazole 15µM	G6PD	3.40				2.50	1.00		1.66	1.30	**1.61**	0.65
	PDK4	0.99	0.87	0.94		0.47	2.12		0.77	0.33	**0.92**	0.82
	PCK1	0.06				0.28	2.67		0.80	0.15	**0.97**	1.16
	INSIG	0.89	10.79	5.43	10.73	1.79	1.89		1.82	1.66	**1.79**	0.09
Acetamino-phen 5mM	G6PD	7.30					3.00	4.67	4.02	7.44	**4.78**	1.90
	PDK4	3.37	3.34	1.08			0.53	0.36	1.77	0.82	**0.87**	0.63
	PCK1	0.13					0.68	0.56	0.81	0.35	**0.60**	0.19
	INSIG	0.93	2.10	1.14	1.36		0.28	0.50	1.02	0.34	**0.54**	0.34
Galacto-samine 10mM	G6PD	1.02						16.71		0.95	**8.83**	11.14
	PDK4	0.84	0.43	1.09				1.77		0.58	**1.18**	0.85
	PCK1	1.03						2.56		1.11	**1.84**	1.03
	INSIG	0.61	1.52	1.10	1.18			0.33		0.13	**0.23**	0.14
Isoniaziode 10mM	G6PD	2.42				1.08	0.87				**0.98**	0.15
	PDK4	0.71	0.38	0.95		0.08	0.34				**0.21**	0.18
	PCK1	0.13				1.33	1.00				**1.17**	0.23
	INSIG	0.97	2.61	3.24	4.63	0.37	0.35				**0.36**	0.02

Replicates listed in one column are from cells of the same donor. Probe set numbers of respective genes are: G6PD: 202275_at, PDK4 no. 1, 2 and 3 in stated order 205960_at; 225207_at; 1562321_at, PCK1: 208383_s_at and INSIG1 no. 1, 2, 3 and 4 in stated order 209566_at; 201625_s_at; 201626_at; 201627_at

Hepatocytes were obtained from five different donors and isolated and cultivated as described. Five compounds out of the Open TG GATES compound set were selected for treatment: valproic acid, ketoconazole, isoniazid, galactosamine, and acetaminophen. The treatment was carried out under the same conditions as in TG GATES but only the latest time point of 24 h incubation was analyzed for the highest concentration. To assess qualitative agreement between the Open TG-GATES data and the more sensitive qPCR, four genes involved in energy and lipid metabolism were selected to validate the compound-induced effect *in vitro*: glucose-6-phosphate-dehydrogenase (G6PD), phosphoenolpyruvate carboxykinase 1 (PCK1), pyruvate dehydrogenase kinase 4 (PDK4) and insulin-induced gene 1 (INSIG1)(Table 3.4). Up to four probe sets were analyzed in the gene array to measure alterations of a target gene. A gene was considered to be deregulated in the gene array, when the majority of probe sets indicated a distinct direction of gene expression alteration. In the qPCR analysis a gene was considered to be deregulated with a change of minimum twofold, which implies a $2^{-\Delta\Delta Ct}$ value of at least two for up regulation and a $2^{-\Delta\Delta Ct}$ value of maximum 0.5 for down regulation. Between two and five replicates, meaning samples from independent experiments with cells from different donors were measured for each gene. Although fold change quantities between qPCRs of different replicates alter strongly, a distinct direction of gene expression alteration is observed in most cases.
The strong induction of G6PD by valproic acid, acetaminophen and ketoconazole (23.71 fold, 7.3 fold and 3.4 fold) that was observed in the gene array was qualitatively confirmed, in contrast to the induction of G6PD by isoniazid. Galactosamine caused large differences in gene expression among the different donors. Up regulation of PDK4 (4.2 fold) in compound-exposed hepatocytes was reproduced for valproic acid, but not for acetaminophen (3.2 fold induction in the gene array).

Although the set of selected genes is very small, the majority of the results obtained for gene expression induction in compound-exposed primary hepatocytes qualitatively agree with the qPCR of the tested donors and the TG GATEs data. Especially for genes that are found to be up-regulated in the gene array, many expression changes were confirmed by qPCR. Nevertheless, independent confirmation is necessary to obtain higher reliability.

Table 3.5: Comparison of TG-GATES gene array data with qPCR data from treated primary human hepatocytes. Gene expression levels for THRSP were measured for 2-4 replicates (cells from different donors). Cells were treated for 24 h before sample collection.

Compound	Conc.	TG-GATES data: fold changes			quantitative real time PCR data: fold changes					
		1553583_a_at	229476_s_at	229477_at	Rep. 1	Rep. 2	Rep. 3	Rep. 4	Mean	SD
Valproic acid	0.2mM	1.172	1.482	1.142	1.833	1.502	1.398	1.805	**1.634**	0.22
	1mM	3.002	6.802	5.610	1.705	2.074	3.107	1.613	**2.125**	0.68
	5mM	13.595	27.209	26.246	1.506	2.035	1.694	0.492	**1.432**	0.66
Ketocon-azole	0,6µM	1.358	0.755	0.834	1.064	0.960	0.264		**0.763**	0.43
	3µM	1.204	1.080	1.214	0.832	1.134	0.554		**0.840**	0.29
	15µM	17.077	41.614	41.413	1.429	0.794	0.447		**0.890**	0.50
Isoniazide	0.4mM	1.120	1.111	1.107	1.111	1.538			**1.324**	0.30
	2mM	1.311	1.619	1.724	2.682	1.854			**2.268**	0.58
	10mM	4.497	10.600	8.718	0.698	0.417			**0.557**	0.20

Particularly, expression changes of the gene thyroid hormone responsive spot 14 (THRSP) illustrate the importance of independent confirmation. THRSP is a gene involved in hepatic lipogenesis and biosynthesis of triglycerides; hence it was suggested that it plays a role in hepatic steatosis (Wu et al. 2013; Zhu et al. 2001). In the Open TG-GATES gene array data of cultivated primary human hepatocytes, THRSP was strongly up regulated by a large set of chemicals after 24 h of exposure at a slightly cytotoxic concentration. For instance, THRSP was strongly induced by valproic acid, ketoconazole and isoniazid. However, none of the inductions were confirmed by qRT-PCR (Table 3.5). Although valproic acid slightly elevated THRSP levels in primary human hepatocytes, the quantities are not comparable to the gene array results. The majority of reproduced gene expression alterations observed by TG-GATES were confirmed by qRT-PCR, however, this example demonstrates that a high risk of false positive results cannot be excluded. This underlines the importance of independent confirmation of gene expression alterations before focusing on single compound effects that were identified by microarray technology.

3.1.7 Characterization of unstable baseline genes

Isolation and cultivation of primary hepatocytes is known to cause strong effects in gene expression. More than 3000 genes are up or down regulated as a consequence of isolation stress and because of the culture conditions (Godoy et al. 2013). The majority of gene expression changes occur during the first 24 h of cultivation (Zellmer et al. 2010). Upregulated genes are predominantly associated with inflammation whereas down regulated genes mainly affect xenobiotic and endogenous metabolism.

To identify stress and cultivation-induced genes, gene expression profiles of freshly isolated hepatocytes were compared to those of cultivated hepatocytes after 1, 2, 3, 5, 7, 10 and 14 days in collagen sandwich culture. Gene array analysis identified 1086 genes (1509 probe

sets) that were up regulated and 988 genes (1754 probe sets) which were found to be down regulated in culture. These 'unstable baseline genes' may be responsible for false positive effects among compound-induced gene expression changes. Moreover, unstable baseline genes may cover compound-induced effects, if a chemical induces the same set of genes. In this case, a biological relevance of the respective gene may not become obvious, since the compound-induced effect is difficult to distinguish from the stress-induced alteration. In any case, genes belonging to the set of unstable baseline genes should be considered with caution and are therefore for highlighted in the toxicotranscriptomics directory.

3.1.8 Detection of biological motifs

After the aforementioned curation steps, gene array data of 143 compounds remained for further analysis. For the 24 h incubation time point at the highest concentration, unsupervised cluster analysis was performed to characterize the 100 strongest deregulated genes across all remaining compounds (Table 3.6 and Figure 3.11). Cluster analysis revealed a pattern where compounds with strong effects on gene expression cluster in the lower part of the heat map whereas rather weak compounds cluster in the upper part. Three of the clusters could be manually associated with biological motifs: a large set of proliferation associated genes in the left corner of the heat map (blue color) is strongly down regulated (compounds AFB_1 – CHX; genes CXCL6 to CDK1). This list of genes includes well characterized genes coding for proteins involved in cell cycle progression, such as cyclin dependent kinases (CDK) and cyclins (CCN proteins), proteins involved in DNA uncoiling and replication, such as topoisomerase 2 (TOP2A), as well as several genes associated with the mitotic spindle formation. The second cluster is located in the middle of the heat map and illustrates compound-induced phase I metabolism of xenobiotics. It comprises the cytochrome P450 isoenzymes (CYP 3A4, 3A7, 1B1, 1A1 and 1A2), which are predominantly induced by a large set of chemicals. Finally, a third biological motif was identified, representing genes that are associated with different forms of cellular stress. This set of genes includes, for example, thioredoxin interacting protein (TXNIP), a gene coding for an oxidative stress mediator protein or the ER stress inducible activating transcription factor 3 (ATF3). Furthermore, heat shock 70-KD protein 6 (HSPA6) is upregulated by various compounds, as well as regulator of cell cycle (RGCC), which is induced by p53 in response to DNA damage. Pyruvate dehydrogenase kinase 4 (PDK4) is also detected among the up regulated genes – it is a gene involved in energy and lipid metabolism and reported to be induced upon starvation or hypoxia (Lee et al. 2012).

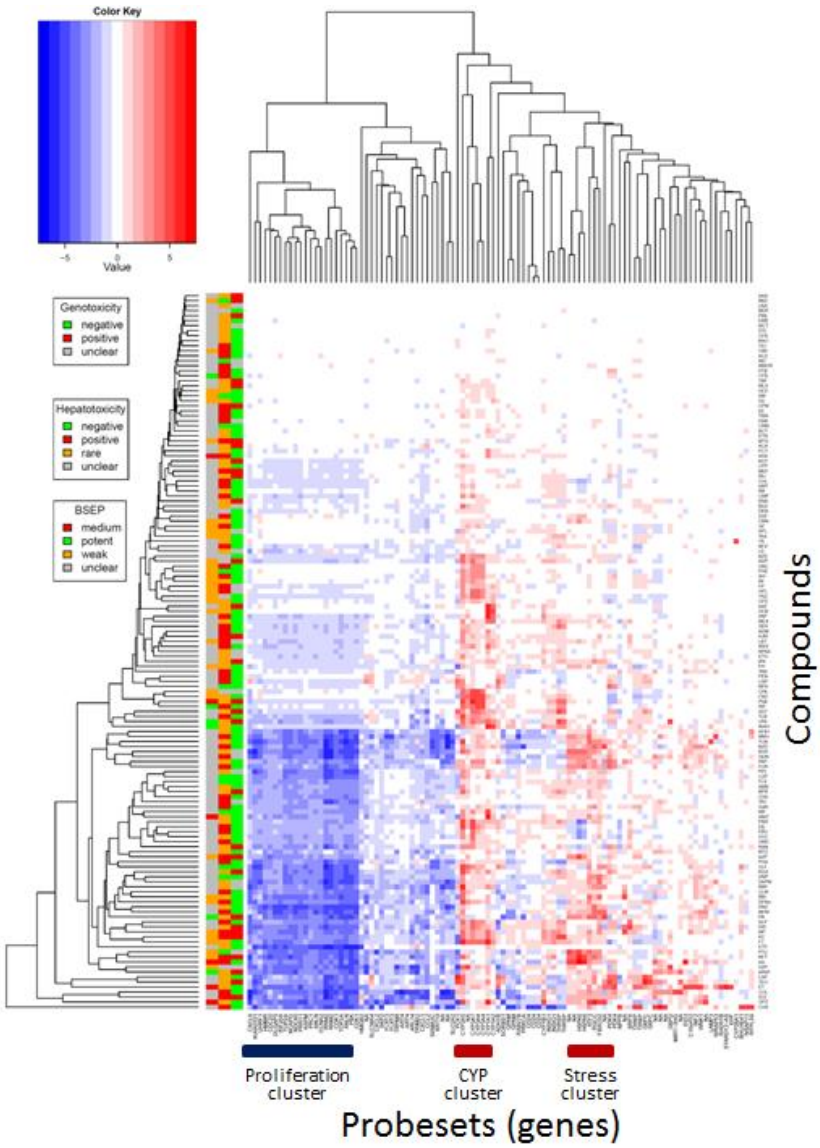


Figure 3.11: Unsupervised Clustering of the 100 most deregulated genes across all compounds tested at the highest concentration for 24h of incubation. Compounds are listed in lines whereas columns represent the 100 strongest deregulated genes. Up- regulated genes are marked in red, down regulated genes are shown in blue. The left column of the heat map shows a further classification of the compounds in terms of their potential in genotoxicity, hepatotoxicity and Bsep inhibiting capacity.

The compounds were further classified according to their potential in genotoxicity and hepatotoxicity, as well as inhibition of the bile salt export pump (Bsep) (Figure 3.11). Bsep is responsible for the active transport of bile acids from the cytosol across the hepatocyte canalicular membrane into the bile. Therefore, inhibited Bsep activity may indicate a cholestatic effect of a compound. Reduced Bsep activity has been reported for various cholestasis-inducing drugs, such as troglitazone and glibenclamide (Kis et al. 2012).

Table 3.6: The 100 strongest deregulated genes at the highest tested concentration for the incubation period of 24h across all compounds. For each probe set the compounds were ranked in order of their fold change and the top 100 probe sets with the highest absolute value of fold change were included.

Probe sets	Gene Symbol
209613_s_at	ADH1B
222608_s_at	ANLN
1552619_a_at	ANLN
220468_at	ARL14
242496_at	ART4
207220_at	ART4
219918_s_at	ASPM
202672_s_at	ATF3
221530_s_at	BHLHE41
209183_s_at	C10orf10
216598_s_at	CCL2
203418_at	CCNA2
266_s_at	CD24
216379_x_at	CD24
209771_x_at	CD24
202870_s_at	CDC20
203213_at	CDK1
210559_s_at	CDK1
204470_at	CXCL1
209774_x_at	CXCL2
206336_at	CXCL6
205749_at	CYP1A1
207608_x_at	CYP1A2
202437_s_at	CYP1B1
206424_at	CYP26A1
205999_x_at	CYP3A4
208367_x_at	CYP3A4
243015_at	CYP3A5
205939_at	CYP3A7
203764_at	DLGAP5
225645_at	EHF
231292_at	EID3
225803_at	FBXO32
222853_at	FLRT3

Probe sets	Gene Symbol
219250_s_at	FLRT3
206952_at	G6PC
1555612_s_at	G6PC
202275_at	G6PD
225420_at	GPAM
225424_at	GPAM
208808_s_at	HMGB2
207165_at	HMMR
213418_at	HSPA6
117_at	HSPA6
224469_s_at	INF2
201626_at	INSIG1
201627_s_at	INSIG1
202503_s_at	KIAA0101
218755_at	KIF20A
201650_at	KRT19
205569_at	LAMP3
223913_s_at	MIR7-3HG
232325_at	NA
235456_at	NA
244567_at	NA
AFFX-M27830_5_at	NA
215078_at	NA
202581_at	NA
210387_at	NA
235681_at	NA
230554_at	NA
237031_at	NA
200800_s_at	NA
243631_at	NA
235102_x_at	NA
218663_at	NCAPG
204162_at	NDC80
206801_at	NPPB

Probe sets	Gene Symbol
1555366_at	NSAP11
219148_at	PBK
208383_s_at	PCK1
225207_at	PDK4
205960_at	PDK4
218009_s_at	PRC1
213093_at	PRKCA
228273_at	PRR11
228708_at	RAB27B
216880_at	RAD51B
214409_at	RFPL3S
218723_s_at	RGCC
201890_at	RRM2
209773_s_at	RRM2
203789_s_at	SEMA3C
203625_x_at	SKP2
206535_at	SLC2A2
220786_s_at	SLC38A4
218087_s_at	SORBS1
204955_at	SRPX
233194_at	STARD13-AS
229476_s_at	THRSP
229477_at	THRSP
201291_s_at	TOP2A
201292_at	TOP2A
213293_s_at	TRIM22
243483_at	TRPM8
201008_s_at	TXNIP
201010_s_at	TXNIP
225655_at	UHRF1
1555068_at	WNK1
224185_at	WRAP53

3.1.9 Stereotypic versus compound specific gene expression responses

Whereas some clusters of genes are deregulated by a large set of chemicals (see Figure 3.5), other genes are affected by individual compounds only. One intention of the toxicogenomics directory is to systematically distinguish between consensus or stereotypical gene expression effects and more compound specific effects. Therefore, the *selection value* concept was introduced, which was based on the sample subset of 24 h exposure at the highest concentration. Each probe set was considered individually and a list was generated that ranks the compounds in the order to their fold change. The list of up regulated probe sets ranks the compounds from the highest to the lowest fold change whereas the list of down regulated probe sets considers the compound ranking from the lowest to the highest fold change.

With this ranking concept, selection value (x) was defined as the list of probe sets that are at least 3 fold up or down regulated by at least x compounds. Accordingly, selection value 1 (SV 1) delivers the list of genes that are at least threefold deregulated by at least one compound; each gene on the list is represented by the particular compound with the strongest impact on this gene. SV 5, for example, delivers the list of genes which are at least threefold deregulated by at least five compounds. Figure 3.12 illustrates the number of deregulated genes for each selection value in a time and concentration dependent manner. As previously described, most genes are deregulated in a time and concentration dependent manner (Figure 3.3 and Figure 3.4). For this reason, the 24 h time point with the highest concentration for cell exposure is most suitable for considering high selection values representing a stereotypic gene expression response. Applying this concept to the Open TG GATES data, 4,135 probe sets were found to be up regulated (4,479 down regulated) for SV 1; 1,101 probe sets were up regulated (1,713 down regulated) for SV 3; 531 (857) for SV 5 and 31 probe sets (179) were induced for SV 20. Figure 3.13 gives an overview in how far genes overlap within the different SV lists.

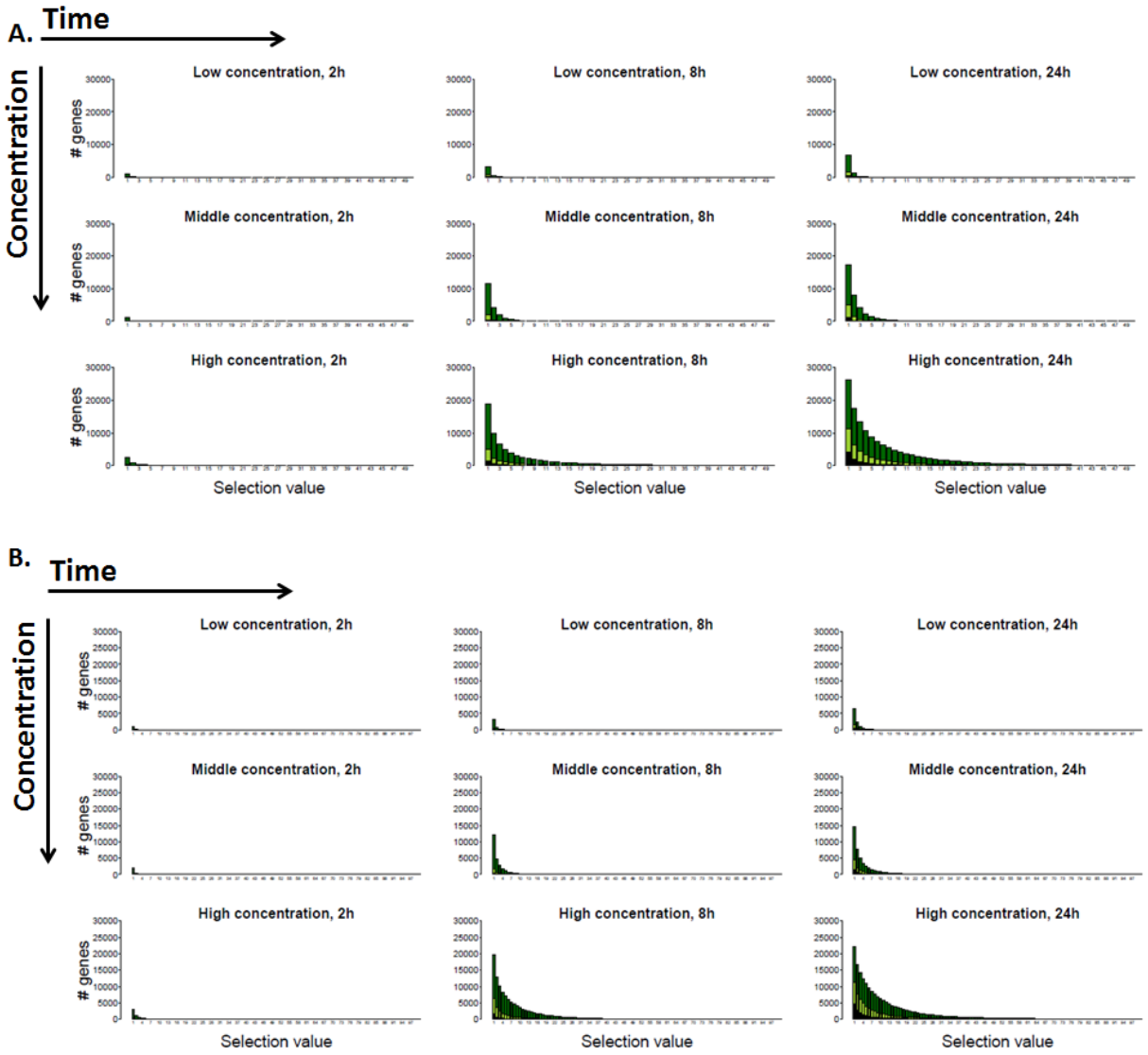


Figure 3.12: A: Selection values for the up regulated genes and **B**: for the down regulated genes. The number of deregulated probe sets per selection value increases time and concentration dependently. A selection value of for example 5 means that at least 5 compounds up or down regulate the indicated gene.

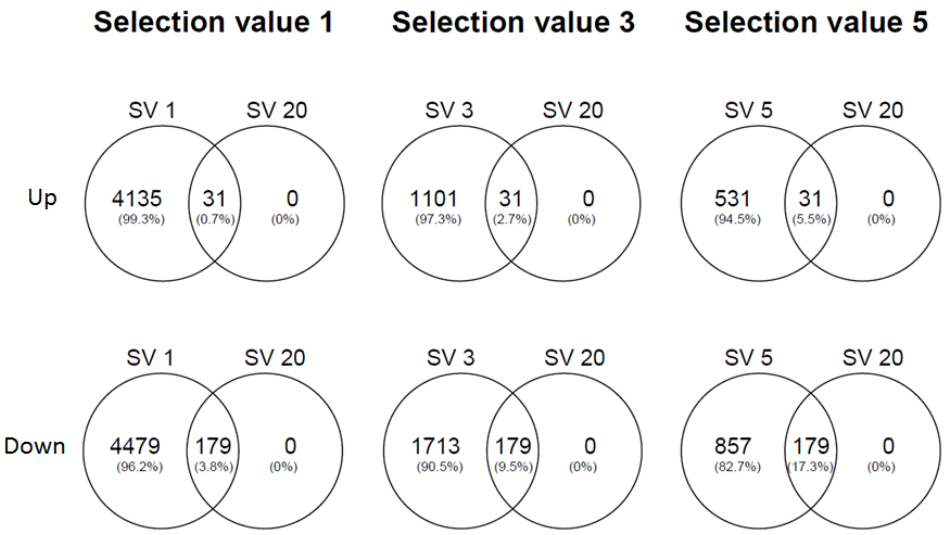


Figure 3.13: Overview of the selection value 1, 3, 5 and 20 genes. Each selection value (x) delivers the list of genes that are at least threefold up or down regulated by at least x compounds.

Accordingly, selection value 20 (SV 20) comprises the list of genes, which are at least threefold deregulated by at least twenty compounds. This list delivers the genes that are altered as a stereotypical response to chemical exposure. Keeping in mind that only 32 of the 143 compounds show strong effects on gene expression, 20 compounds represent a large fraction for defining a stereotypic response. 31 probe sets were identified as up regulated, whereas 179 probe sets were at least threefold down regulated by at least 20 compounds (Figure 3.13). The SV 20 genes were studied individually and assigned to biological categories (Table 3.7).

Table 3.7: Consensus genes deregulated in human hepatocytes by chemical exposure. The listed genes are at least 3-fold up (**A**) or down regulated (**B**) by at least 20 of the 148 studied chemicals (selection value 20). [1]Gene deregulated in liver disease (NASH, cirrhosis and/or HCC). [2]Unstable baseline gene. [3]Not annotated, functionally unclear probe set.

A. Upregulated consensus genes

Symbol	Gene	Probe set	Function of the Gene Product	
Category: Metabolism, Xenobiotics				
CYP1A1	cytochrome P450, sub-fam. 1A, polypeptide 1	205749_at	metabolic enzyme in the ER	phase I enzymes
CYP2C9	cytochrome P450, sub-fam. 2C, polypeptide 9	217558_at		
CYP3A4	cytochrome P450, sub-fam. 3A, polypeptide 4	205999_x_at 208367_x_at		
CYP3A5[2]	cytochrome P450, sub-fam. 3A, polypeptide 5	214235_at 243015_at		
CYP3A7[1]	cytochrome P450, sub-fam. 3A, polypeptide 7	205939_at 211843_x_at		
SULT1C2[1]	sulfotransferase 1C2	205342_s_at	cytosolic enzyme; catalyzes sulfonation	phase II enzymes
SULT2A1	sulfotransferase 2A1	206292_s_at 206293_at		
Category: Differentiation and Development				
FGF21	fibroblast growth factor 21	221433_at	secreted growth factor; mitosis and survival	growth factor
GDF15[1]	growth/differentiation factor 15	221577_x_at	secreted growth factor; inflammation and apoptosis	
IFRD1	interferon-related developmental regulator 1	202147_s_at	nuclear protein; regulation of gene expression in proliferative and differentiative pathways	interferon-related signaling
EFNA1[1]	ephrin-A1	202023_at	receptor tyrosine kinase; migration and adhesion	other
Category: Protein Modification and Degradation				
CBX4[1]	E3 SUMO-protein ligase CBX4	227558_at	protein ligase; SUMO1 conjugation and proteasomal degradation	
FBXO32[1]	F-box protein 32	225803_at	cytosolic protein; ubiquitination and proteasomal degradation	
KLHL24	kelch-like protein 24	221985_at 221986_s_at	cytosolic protein; role in protein degradation	
Category: Stress Response				
ATF3	activating transcription factor 3	202672_s_at 1554980_a_at	transcription factor; stress response, further involved in cell cycle regulation, DNA repair, apoptosis	cell cycle arrest
RGCC[1]	regulator of cell cycle	218723_s_at	cytosolic protein; induced by p53 modulates the activity of cell cycle specific kinases in response to DNA damage	
CREBRF	CREB3 regulatory factor	225956_at	nuclear protein; regulates transcription, negative regulator of the ER stress response	stress response/ ER stress
PPM1E	phosphoprotease 1E	205938_at	serine/threonine protein phosphatase; negative regulator of cell stress response pathways	
Category: Energy and Lipid Metabolism				
PDK4[1]	pyruvate dehydrogenase kinase	225207_at	mitochondrial membrane enzyme; increased PDK4 leads to enhanced gluconeogenesis	glucose metabolism/ homeostasis
PPM1L[2]	protein phosphatase 1L	228108_at	membrane bound enzyme; regulation of blood-glucose	
Category: Other				
SLC7A11[1]	solute carrier fam. 7 member 11	217678_at	membrane anchored protein; cysteine and glutamate transport	
ZCCHC6[1]	terminal uridylyltransferase 7	242776_at	enzyme involved in RNA processing	

[3]236542_at[1], 237031_at

B. Downregulated consensus genes. The 100 probe sets (74 genes) with the highest fold change are given below.

Symbol	Gene	Probe set	Function of the Gene Product
Category: Cell Cycle Progression and Regulation			
ASPM	abnormal spindle protein homolog	219918_s_at	cytosolic protein; role in mitotic spindle regulation and coordination of mitotic processes
AURKA	aurora kinase A	208079_s_at	cytosolic kinase; regulation of cell cycle progression
AURKB	aurora kinase B	209464_at	cytosolic kinase; regulation of cell cycle progression
BIRC5	baculoviral IAP repeat-containing protein 5	202095_s_at	cytosolic protein; chromosome alignment and segregation during mitosis, further role as apoptotic factor
CENPK	centromere protein K	222848_at	nuclear protein; assembly of kinetochore proteins, mitotic progression and chromosome segregation
FAM83D	protein FAM83D	225687_at	cytosolic protein; chromosome alignment on the spindle
MLF1IP	centromere protein U	218883_s_at	component of a nucleosome-associated complex; assembly of kinetochore proteins
NCAPG	condensin-2 complex subunit G2	218662_s_at 218663_at	nuclear protein; regulation of mitotic chromosome architecture
OIP5	Opa-interacting prot. 5	213599_at	nuclear protein; chromosome segregation
TTK	dual specificity protein kinase TTK	204822_at	protein kinase; centrosome duplication and mitosis progression, associated with cell proliferation
CCNA2	cyclin A2	203418_at 213226_at	nuclear protein; cell cycle control at the G1/S and G2/M transitions
CCNB1	cyclin B1	214710_s_at 228729_at	cytosolic and nuclear protein; cell cycle control at the G2/M transition
CCNB2	cyclin B2	202705_at	
CCNE2[2]	cyclin E2	205034_at 211814_s_at	nuclear protein; controls the cell cycle at the late G1 and early S phase
CDC20	cell division cycle 20 homolog	202870_s_at	cytosolic protein;regulation of anaphase initiation and mitotic exit
CDC6	cell division control protein 6 homolog	203967_at 203968_s_at	cytosolic and nuclear protein; control and initiation of DNA replication
CDCA3	cell division cycle-associated protein 3	223307_at	cytosolic protein; required for entry into mitosis
CDK1	cyclin-dependent kinase 1	203213_at 203214_x_at 210559_s_at	kinase, cell cycle control by modulation of the centrosome cycle and mitotis initiation
CDKN3	cyclin-dependent kinase inhibitor 3	209714_s_at 1555758_a_a	cytosolic protein; cell cycle regulation
DLGAP5	discs, large homolog-associated protein 5	203764_at	cytosolic and nuclear protein; cell cycle regulator
DTL	denticleless protein homolog	218585_s_at 222680_s_at	cytosolic and nuclear protein; cell cycle control, DNA repair
HMMR	hyaluronan-mediated motility receptor	207165_at 209709_s_at	cell surface receptor; required for entry and regulation of mitosis
MELK	maternal embryonic leucine zipper kinase	204825_at	serine/threonine-protein kinase; modulator of intracellular signaling, further role in apoptosis
TRIP13	thyroid receptor-interacting protein 13	204033_at	transcription factor interacting protein; checkpoint arrest, chromosome re-combination and structure development during meiosis, role in DNA double strand break repair
BUB1	serine/threonine-protein kinase BUB1	209642_at	cytoplasmic and nuclear kinase; mitotic checkpoint, required for normal mitosis progression
BUB1B	serine/threonine-protein kinase BUB1b	203755_at	
CASC5	cancer susceptibility candidate 5	228323_at	nuclear protein; spindle-assembly checkpoint signaling and chromosome alignment

INSC	protein inscuteable homolog	237056_at	cytosolic protein; spindle orientation
MAD2L1	mitotic arrest-deficient 2L1	203362_s_at 1554768_a_at	cytosolic and nuclear protein; spindle-assembly checkpoint
NDC80	kinetochore protein NDC80 homolog	204162_at	nuclear protein; chromosome segregation and spindle checkpoint activity
NEK2	never in mitosis A-related kinase 2	204641_at	mitotic kinase; controls centrosome separation and bipolar spindle formation in mitotic cells
SPC25	kinetochore protein Spc25	209891_at	nuclear protein; chromosome segregation, spindle checkpoint activity
TPX2	targeting protein for Xklp2	210052_s_at	spindle associated protein; spindle assembly factor, colocalises with apoptotic microtubules
ZWINT	ZW 10 interactor	204026_s_at	nuclear protein of the MIS12 complex; kinetochore formation and spindle checkpoint activity
GINS1	GINS complex subunit 1	206102_at	nuclear protein; initiation of DNA replication and progression of DNA replication forks
GINS2	GINS complex subunit 2	221521_s_at	
KIAA0101	PCNA-associated factor	202503_s_at	cytosolic and nuclear protein; regulates DNA repair during DNA replication
LMNB1	lamin-B1	203276_at	membrane protein in the nuclear laminar; DNA replication, stress response and development
MCM10	minichromosome maintenance 10	220651_s_at	nuclear protein; functions as replication initiation factor
TOP2A	topoisomerase II alpha	201291_s_at 201292_at	nucleoplasm enzyme; transiently breaks and reunites double strand DNA during replication
KIF23	kinesin-like protein KIF23	204709_s_at	protein required for the myosin contractile ring formation during cytokinesis
KIF4A	chromosome-associated kinesin KIF4A	218355_at	microtubulus motor protein; spindle formation
NUSAP1	nucleolar and spindle-associated protein 1	218039_at 219978_s_at	cytosolic and nuclear protein; stabilization of microtubules
ANLN	anillin	222608_s_at 1552619_a_at	nuclear, actin binding protein; role in cytokinesis, deregulated in many cancers
CEP55	centrosomal protein 55 kDa	218542_at	cytosolic and nuclear protein; mitotic exit and cytokinesis
PRC1	protein regulator of cytokinesis 1	218009_s_at	cytosolic and nuclear protein; regulator of cytokinesis, cross-links antiparallel microtubules
PBK	PDZ binding kinase	219148_at	mitotic kinase; phosphorylates MAP kinase p38, only active during mitosis
PRR11	proline rich 11	228273_at	cytosolic protein; expression increases from G1 to G2/M phase
E2F8	transcription factor E2F8	219990_at	transcription factor; regulates transcription of other transcription factors with role in cell cycle
MYBL1	myeloblastosis-related protein A	213906_at	transcriptional activator; master regulator of meiotic genes, role in proliferation and differentiation
HELLS	lymphoid-specific helicase	223556_at 227350_at	helicase with role in normal development and survival, chromatin remodeling and DNA methylation

Category: DNA Synthesis, Recombination and Repair

DHFR[1,2]	dihydrofolate reductase	202533_s_at	enzyme in folate metabolism; synthesis of purines	DNA synthesis
RRM2	ribonucleotide reductase M2 polypeptide	201890_at 209773_s_at	cytosolic enzyme; biosynthesis of deoxyribonucleotides, inhibits Wnt signaling	
TK1	thymidine kinase	202338_at 1554408_a_at	cytosolic kinase for DNA synthesis; phosphorylation of thymidine to deoxythymidine monophosphate	
TYMS	thymidylate synthase	202589_at 1554696_s_at	enzyme which contributes to the de novo mitochondrial thymidylate biosynthesis pathway	
FANCI	fanconi-associated nuclease 1	213007_at	nuclease; required for maintenance of chromosomal stability, key role in DNA repair	DNA repair/ recombination
RAD51AP1	RAD51-associated protein 1	204146_at	nuclear, DNA binding protein; DNA damage response	

Category: Immune Response and Inflammation

CCL2[1]	C-C motif chemokine 2	216598_s_at	secreted protein; chemotactic factor attracting monocytes and basophils
CXCL6[1]	chemokine (C-X-C motif) ligand 6	206336_at	secreted protein with chemotactic activity for neutrophils; inflammation and development
FSTL1[1]	follistatin-related protein 1	208782_at	secreted glycoprotein; inflammatory protein, enhancing synthesis of pro-inflammatory cytokines and chemokines by immune cells
HAVCR1	hepatitis A virus cellular receptor 1	207052_at	membrane protein receptor; role in T-helper cell development, cell-surface receptor for the virus
Symbol	Gene	Probe set	Function of the Gene Product

Category: Cytoskeleton and Trafficking

ARL14	ADP-ribosylation factor-like 14	220468_at	cytoplasmic vesicle GTPase; controls transport of vesicles along microtubules	transport/ trafficking
KIF20A	kinesin-like protein KIF20A	218755_at	cytosolic protein; controls transport along microtubules	
KRT7[1]	keratin type II cytoskeletal 7	209016_s_at	cytoplasmic intermediate filament protein; responsible for structural cell integrity, stimulates DNA synthesis	cytoskeleton
PALMD	palmdelphin	218736_s_at 222725_s_at	cytosolic protein; role in the cell shape control and cell dynamics	

Category: Metabolism

ALDH8A1[1]	aldehyde dehydrogenase fam. 8 member A1	220148_at	cytosolic enzyme; converts 9-cis-retinal to 9-cis-retinoic acid
CPS1[1,2]	carbamoyl-phosphate synthase	204920_at 217564_s_at	mitochondrial protein; role in the urea cycle, removes excess ammonia from the cell

Category: Other

ABCA8[1,2]	ATP-binding cassette sub-family A, member 8	204719_at	membrane located, ATP-dependent lipophilic drug transporter
BCL2A1[1,2]	Bcl-2-related protein A1	205681_at	cytoplasmic protein involved in apoptosis regulation
DEPDC1	DEP domain-containing protein 1A	222958_s_at 232278_s_at	nuclear protein; involved in transcriptional regulation as a transcriptional corepressor
SHCBP1	SHC SH2 domain-binding protein 1	219493_at	cytosolic protein; signaling pathways in proliferation, cell growth and differentiation
UBE2C	ubiquitin-conjugating enzyme E2 C	202954_at	enzyme involved in protein ubiquitinylation and degradation
UHRF1	ubiquitin-like PHD and RING finger domain-containing protein 1	225655_at	nuclear epigenetic regulator enzyme; bridges DNA methylation and chromatin modification
WDR72[1,2]	WD repeat-containing protein 72	227174_at	mutations in this gene have been associated with amelogenesis imperfecta hypomaturation type 2A3

[3]204962_s_at, 225834_at, 229490_s_at, 230554_at, 232325_at[1], 244567_at

Most of the SV 20 genes that are up regulated as a stereotypical expression response were associated with biological functions, such as phase I and II metabolism of xenobiotics, development and differentiation, protein modification and degradation, stress response and energy and lipid metabolism (Table 3.7 A). Down regulated consensus genes are mainly involved in cell cycle progression, but a smaller fraction of genes can be attributed to categories, such as DNA repair and synthesis, immune response, cytoskeleton and intracellular trafficking and metabolism (Table 3.7 B).

In order to consider rather compound specific gene expression alterations, detailed analysis was also performed for the SV 3 genes. This list comprises the genes which are at least threefold up or down regulated by at least three chemicals. More individual gene expression responses may be obtained with the SV 1 list, because the SV 1 list gives each top gene with the highest fold change for each compound. However, there are only 2 replicates per samples available and the amount of false positive results within the multiple testing may be relatively high. Considering the genes that are deregulated by at least three compounds, offers a compromise between individuality and reliability in order to identify rather compound specific effects. This list of genes also comprises the SV 20 genes, but covers an even more diverse pattern of biological functions (Table 3.8). Up regulated SV 3 genes were assigned to categories, such as energy and lipid metabolism, inflammatory response and the immune system, development and differentiation, protein modification and degradation, regulation of transcription, metabolism, stress response and apoptosis, transport, as well as cytoskeletal factors (Table 3.8 A).

Table 3.8: SV 3 genes deregulated in human hepatocytes by chemical exposure. The listed genes are **A** up or **B** down regulated by at least 3 of the 143 studied chemicals (selection value 3). The genes PCK1, ADH1B and CPS1 are among the TOP150 up- and downregulated genes, the deregulations are caused by different chemicals. [1]Gene deregulated in liver disease (NASH, cirrhosis and/or HCC). [2]Unstable baseline gene. [3]Not annotated probe set.

A. Up regulated SV 3 genes. The 118 probe sets (85 genes) with the highest fold changes are given below. 32 probe sets (17 genes) are overlapping with upregulated genes in SV20 and are not listed here.

Symbol	Gene	Probe set	Function of the Gene Product	
Category: Energy and Lipid Metabolism				
G6PC	glucose-6-phosphatase, catalytic subunit	202275_at 206952_at	enzyme in regulation of blood glucose levels; role in pentose phosphate way, fatty acid and nucleic acid synthesis	glucose metabolism/ homeostasis
SDS	serine dehydratase	205695_at	enzyme with role in gluconeogenesis; catalyzes serin into ammonia and pyruvate	
PCK1	phosphoenolpyruvate carboxykinase 1 (soluble)	208383_s_at	gluconeogenic enzyme; production of glucose from lactate and other precursors derived from the citric acid cycle	
PDK4	pyruvate dehydrogenase kinase, isozyme 4	205960_at	mitochondrial membrane enzyme; increased PDK4 leads to enhanced gluconeogenesis	
PPARGC1A	peroxisome proliferator-activated receptor gamma, coactivator 1 alpha	219195_at	transcriptional coactivator for steroid and nuclear receptors and genes involved in glucose and fatty acid metabolism	
HMGCS2	3-hydroxy-3-methylglutaryl-CoA synthase 2	204607_at	mitochondrial enzyme; catalyzes HMG-CoA formation	bile acid/ cholesterol metabolism
AKR1D1	aldo-keto reductase family 1, member D1	207102_at	cytosolic aldoketo reductase; reduction of several steroid hormones and bile acid intermediates	
FABP1	fatty acid-binding protein 1	205892_s_at	cytosolic protein; regulates intracellular lipid transport, required for cholesterol synthesis and metabolism	lipid transport
MGEA5[1]	meningioma expressed antigen 5 (hyaluronidase)	223494_at 235868_at	glycosidase; removes O-GlcNAc modifications; further regulates DNA metabolic processes	DNA metabolic processes

Category: Inflammatory Response and Immune System

C2CD4A	C2 calcium-dependent domain containing 4A	241031_at	nuclear protein; involved in inflammatory processes, cell architecture and adhesion	inflammatory response
IL33[1]	Interleukin 33	209821_at	inflammatory cytokine; dual function as cytokine and nuclear factor with transcriptional regulatory properties	
NDST1[1]	N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 1	1554010_at	bifunctional enzyme; involved in heparan sulfate biosynthesis and the inflammatory response	
HAMP	hepcidin antimicrobial peptide	220491_at	secreted signaling molecule; maintenance of iron homeostasis; disrupts the cell membranes of cellular pathogens	immune response
IFIT2	interferon-induced protein with tetratricopeptide repeats 2	226757_at	cytoplasmic, antiviral protein; immune response, inhibits expression of viral messenger RNAs, can promote apoptosis	
RSAD2[2]	radical S-adenosyl methionine domain containing 2	242625_at	interferone inducible membrane protein; inhibits DNA and RNA viruses, T-cells activation and differentiation	
CEBPA	CAAT/enhancer binding protein (C/EBP), alpha	204039_at	transcription factor; differentiation of granulocytes and myeloid cells, further role in inhibition of proliferation	
UNC5B[1]	unc-5 homolog B	226899_at	membrane recptor; immune cell adhesion, migration, inflammatory response, further role in apoptosis	
CLEC2B	C-type lectin domain family 2, member B	209732_at	myeloid-specific activating membrane receptor; involved in cross-talk between myeloid cells and NK cells	
TSC22D3	TSC22 domain family, member 3	208763_s_at	peptide with role in anti-inflammatory and immunesuppressive effects, acts as transcriptional regulator	

Category: Development and Differentiation

EFNB2	ephrin-B2	202668_at	cell surface transmembrane ligand; involved in migration, adhesion, is further involved in angiogenesis	signal transducing enzymes and proteins
EPC1	enhancer of polycomb homolog 1	238633_at	nuclear protein; growth regulation, differentiation and DNA repair	
GAREM	GRB2 associated, regulator of MAPK1	228115_at	adapter protein that plays a role in intracellular signaling cascades and proliferation	
GEM	GTP binding protein overexpressed in skeletal muscle	204472_at	regulatory membrane protein; functions in signal transduction, is induced by mitogens	
LATS2[1]	large tumor suppressor kinase 2	230348_at	serine/threonine kinase; regulation of cytokinesis, cell proliferation, apoptosis, component in the Hippo signaling pathway	
NEDD9[1]	neural precursor cell expressed, developmentally down-regulated 9	1569020_at	docking protein for tyrosine-kinase-based signaling; cell adhesion, growth control and proliferation	
S100P[1,2]	S100 calcium binding protein P	204351_at	signaling molecule; functions as calcium sensor, stimulates proliferation in an autokrine manner	
SPATA5	spermatogenesis associated 5	241546_at	ATPase; differentiation, functions as chaperone	
WNK1[1]	WNK lysine deficient protein kinase 1	1555068_at	cytosolic enzyme; regulation of sodium and chloride ion transport, cell signaling, survival and proliferation	
ADM	adrenomedullin	202912_at	secreted hormone; regulates hormone secretion, growth modulation, angiogenesis and antimicrobial activity	secreted proteins and hormones
DKK3	dickkopf WNT signaling pathway inhibitor 3	230508_at	secreted protein; antagonizes canonical Wnt signaling	
FIBIN[1]	fin bud initiation factor homolog	226769_at	secreted growth factor; function in limb development	
HBEGF[1]	heparin-binding EGF-like growth factor	38037_at 203821_at	secreted growth factor; mediates its effects via EGFR, involved in macrophage-mediated cellular proliferation	

TSKU	tsukushi, small leucine rich proteoglycan	218245_at	secreted protein; involved in intracellular transport and extracellular secretion	
NOV1	nephroblastoma overexpressed gene	214321_at	secreted protein; role in cell growth regulation, proliferation and differentiation	
FOS[1]	FBJ murine osteosarcoma viral oncogene homolog	209189_at	proto-oncogenic transcription factor; signal transduction, cell proliferation and differentiation	transcription factors
RBPMS[1]	RNA binding protein with multiple splicing	1557223_at	nuclear protein acting as a coactivator of transcriptional activity; interaction with SMAD proteins	transcription factors

Category: Protein Modification and Degradation

DNAJC12	DnaJ (Hsp40) homolog, subfamily C, member 12	223722_at	molecular chaperone of HSP70 proteins; involved in protein folding and export	protein folding
HSPD1	heat shock 60kDa protein 1 (chaperonin)	241716_at	mitochondrial chaperone; protein folding, mitochondrial protein import and macromolecular assembly	protein folding
HSPA6	heat shock 70kDa protein 6 (HSP70B')	117_at 213418_at	chaperone; function in protein stabilization and folding	protein folding
KLHL20	kelch-like family member 20	210634_at	adaptor for ubiquitin ligase complex, interferon response and Golgi to endosome transport, negative regulator of apoptosis	ubiquitinylation and degradation
MKRN1	makorin ring finger protein 1	209845_at	enzyme; catalyzes ubiquitinylation of substrate proteins for proteasomal degradation	ubiquitinylation and degradation
UBE2D3	ubiquitin-conjugating enzyme E2D 3	240383_at	membrane protein; ubiquitinylation and proteasomal degradation, involved in DNA damage repair and apoptosis	ubiquitinylation and degradation
MMP1	matrix metallopeptidase 1	204475_at	secreted protease; degrades extracellular matrix proteins	ubiquitinylation and degradation

Category: RNA processing, DNA Repair and Recombination

EIF4G3[1]	eukaryotic translation initiation factor 4 gamma, 3	1554309_at	scaffold protein for transcription factors; recognition of the mRNA cap and recruitment of mRNA to the ribosome	RNA processing
ZFC3H1[1]	zinc finger, C3H1-type containing	1553736_at	metal ion binding enzyme in RNA processing	RNA processing
EID3	EP300 interacting inhibitor of differentiation 3	231292_at	nuclear protein; repair of DNA double-strand breaks, repressor of nuclear receptor-dependent transcription	DNA repair/ recombination

Category: Metabolism, Xenobiotics

ADH1B	alcohol dehydrogenase 1B (class I), beta polypeptide	209612_s_at 209613_s_at	cytosolic enzyme; ethanol metabolism	phase I enzyme
CYP1A2	cytochrome P450, subfamily 1A, polypeptide 2	207608_x_at	membrane bound metabolic enzyme	phase I enzyme
CYP1B1	cytochrome P450, subfamily 1B, polypeptide 1	202437_s_at	membrane bound metabolic enzyme	phase I enzyme
CYP3A 4	cytochrome P450, sub-fam. 3A, polypeptide4	231704_at	membrane bound metabolic enzyme	phase I enzyme
CYP2C8	cytochrome P450, sub-fam. 2C, polypeptide 8	208147_s_at	membrane bound metabolic enzyme	phase I enzyme
TAT	tyrosine aminotransferase	214413_at	cytosolic enzyme; involved in tyrosine and phenylalanin catabolism	amino acid catabolism

MTHFD2	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyl-tetrahydrofolate cyclohydrolase	201761_at	mitochondrial enzyme; methylenetetrahydrofolate dehydrogenase and methenyltetrahydrofolate cyclohydrolase activity	one carbon metabolism
SULT1C2	sulfotransferase family, cytosolic, 1C, member 2	211470_s_at	cytosolic enzyme; catalyzes sulfonation	phase II enzyme
ARG1	arginase 1	206177_s_at	final enzyme of the urea cycle	urea cycle
CPS1	carbamoyl-phosphate synthase 1	204920_at 217564_s_at	mitochondrial protein; important role in the urea cycle by removing excess ammonia from the cell	

Category: Stress Response and Apoptosis (ER stress, cell cycle arrest)

BCL10	B-cell CLL/lymphoma 10	1557257_at	cytosolic protein; promotes apoptosis	apoptosis
BEX2	brain expressed X-linked 2	224367_at	cytosolic protein; regulator of mitochondrial apoptosis and cell cycle	
DEDD2	death effector domain containing 2	225434_at	nuclear protein; mediates apoptosis, cell cycle regulation and inhibits mitosis	
GULP1[1]	GULP, engulfment adaptor PTB domain containing 1	204237_at	adapter protein; phagocytosis of apoptotic cells; glycol-sphingolipid and cholesterol transport, endosomal trafficking	
STK17B[1]	serine/threonine kinase 17b	205214_at	calmodulin-dependent kinase; functions as positive regulator of apoptosis	
DDIT3	DNA-damage-inducible transcript 3	209383_at	transcription factor; induces cell cycle arrest and apoptosis in response to ER stress	stress response
ERN1	endoplasmic reticulum to nucleus signaling 1	235745_at	kinase in the ER membrane; sensor for unfolded protein inside the ER, triggers growth arrest and apoptosis	
FAM129A[1]	family with sequence similarity 129, member A	217967_s_at	cytosolic protein; involved in apoptosis, survival and ER stress response	
MT1G	metallothionein 1G	210472_at	havy metal binding protein, role in protective stress responses	
PPP1R15A	protein phosphatase 1, regulatory subunit 15A	202014_at 37028_at	phosphatase; mediates growth arrest and apoptosis in response to DNA damage ; transcriptional activities	
SESN3[1,2]	sestrin 3	235683_at	stress-induced protein; reduces intracellular oxygen species, role in regulation of blood glucose and lipid storage	
ZFAND2A[1]	zinc finger, AN1-type domain 2A	226650_at	zinc ion binding protein; proteasomal degradation during cell stress	

Category: Transporter

BLZF1	basic leucine zipper nuclear factor 1	210462_at	protein in the Golgi lumen; protein transport from the ER through the Golgi apparatus to the cell surface	Protein/ RNA transport
XPO1[1]	exportin 1	235927_at	nuclear protein; nuclear export of RNAs and proteins (cargos), role in proliferation and chromosome region maintenance	
SLC20A1	solute carrier family 20 (phosphate transporter), member 1	230494_at	membrane protein with role in phosphate transport	ion and amino acid transport
SLC30A1	solute carrier family 30 (zinc transporter), member 1	228181_at	membrane protein; involved in zinc transport out of the cell	
SLC7A11	solute carrier family 7 member 11	209921_at	membrane anchored transport protein; cystine and glutamate transporter	
STC2[1]	stanniocalcin 2	203438_at	secreted glycoprotein hormone; regulation of calcium and phosphate transport and homeostasis	

Category: Cytoskeleton and Cell Cycle

MYLIP[1]	myosin regulatory light chain interacting protein	228098_s_at	transmembrane protein; links actin to membrane-bound proteins at the cell surface, inhibitor of cholesterol uptake	cytoskeleton/ trafficking
SH3BP4[1]	SH3-domain binding protein 4	231468_at	membrane protein involved in endocytosis; regulates cell growth, proliferation and autophagy	
SMEK2[1]	SMEK homolog 2, suppressor of mek1	1568627_at	cytosolic enzyme; regulates microtubule organization, role in cell cycle and cytoskeleton	
SPTBN1[1]	spectrin, beta, non-erythrocytic 1	226342_at	cytoskeletal, calmodulin binding protein; cell shape, organization of organelles and molecular traffic	
SERTAD1	SERTA domain containing 1	223394_at	nuclear protein; cell cycle regulation by activation and formation of CDK4 complexes	cell cycle
WRAP53	WD repeat containing, antisense to TP53	224185_at	nuclear protein; replication of chromosome termini	

Category: Other

ANKRD33	ankyrin repeat domain 33	242209_at	protein motif, ankyrin repeat proteins are composed of tandem repeats of a basic structural motif
HMOX1	heme oxygenase (decycling) 1	203665_at	oxygenase for heme degradation; functions in apoptosis and vascularization
KANSL3[1]	KAT8 regulatory NSL complex subunit 3	1558652_at	nuclear protein; role in regulation of transcription
NPTX2	neuronal pentraxin II	213479_at	secreted protein with biochemical properties of a Ca-dependent lectin; modifies properties underlying longterm plasticity

[3]1553133_at, [3]1556602_at, [3]1557459_at, [3]1569403_at, [3]200800_s_at, [3]202581_at, [3]204760_s_at, [3]208180_s_at, [3]210387_at, [3]214469_at, [3]214472_at, [3]215078_at, [3]215779_s_at, [3]218541_s_at, [3]227099_s_at, [3]232035_at, [3]235102_x_at, [3]235456_at, [3]236898_at, [3]239203_at, [3]239845_at, [3]242981_at, [3]243631_at, [3]243918_at, [3]244677_at, [3]AFFX-M27830_5_at

Down regulated SV 3 genes overlap in large parts with down regulated SV 20 genes that are associated with cell cycle progression. Further SV 3 genes belong to biological categories, such as differentiation, endogenous and xenobiotics metabolism, cytoskeletal organization, immune response, transporters, energy and lipid metabolism, and apoptosis (Table 3.8 B).

B. Down regulated SV 3 genes. The 60 probe sets (47 genes) with the highest fold changes are given below. 90 probe sets (66 genes) are overlapping with downregulated genes in SV20 and are not listed here.

Symbol	Gene	Probe set	Function of the Gene Product	
Category: Development and Differentiation				
AXL[1,2]	AXL receptor tyrosine kinase	202686_s_at	receptor tyrosine kinase; signal transduction regulating survival, proliferation, migration and differentiation	signal transducing enzymes and proteins
ATAD2	ATPase family, AAA domain containing 2	222740_at	ATPase; involved in genome regulation for proliferation, cell growth, differentiation and apoptosis	
PTPN14	protein tyrosine phosphatase, non-receptor type 14	226282_at	non-receptor tyrosine phosphatase; involved in adhesion, migration, cell growth and proliferation	
TRIM55	tripartite motif containing 55	236175_at	RING finger zinc containing protein; signal transduction, development, transcription repression and ubiquitination	
NREP[2]	neuronal regeneration related protein	201310_s_at	cytosolic protein with roles in in cellular differentiation	
ARID5B[1]	AT rich interactive domain 5B (MRF1-like)	212614_at	transcription coactivator; key role in liver development, regulation of adipogenic genes	transcription factors
CUX2[1]	cut-like homeobox 2	213920_at	transcription factor; STAT5 dependent and GH regulated, controls proliferation and differentiation	
SOX6[1]	SRY (sex determining region Y)-box 6	227498_at	transcription factor; role in several developmental process-es	
EGR1[1,2]	early growth response 1	227404_s_at	transcription factor; role in mitogenesis and differentiation; directly controls TGF-β expression	
CTGF[1,2]	connective tissue growth factor	209101_at	secreted extracellular matrix protein; proliferation, migra-tion, adhesion, survival, differentiation, induces EMT	secreted proteins
SEMA3C	semaphorin 3C	203789_s_at	secreted protein; regulation of developmental processes	
Category: Metabolism, Xenobiotics				
ADH1B[1,2]	alcohol dehydrogenase 1B (class I), beta polypeptide	209613_s_at	cytosolic enzyme; ethanol metabolism	phase I enzymes
AKR1B10	aldo-keto reductase family 1, member B10 (aldose reductase)	206561_s_at	secreted enzyme; reduction and detoxification of aliphatic and aromatic aldehydes	
CHST9	carbohydrate (N-acetylgalactosamine 4-0) sulfotransferase 9	223737_x_at 224400_s_at	enzyme that sulfates carbohydrates and glycolipids	
CYP4A11[1,2]	cytochrome P450, subfamily 4A, polypeptide 11	207407_x_at	membrane bound phase I metabolic enzyme	
CYP8B1[1]	cytochrome P450, subfamily 8B polypeptide 1	232494_at		
UGT2A3	UDP glucuronosyltransferase 2 family, polypeptide A3	219948_x_at	membrane protein; conjugates lipophilic substrates with glucuronic acid	phase II enzyme
UGT2B15[1,2]	UDP glucuronosyltransferase 2 family, polypeptide B15	207392_x_at 216687_x_at		
GBA3[1,2]	glucosidase, beta, acid 3	222943_at	cytosolic enzyme; involved in carbohydrate metabolic processes	glycoside metabol.

Category: Cytoskeleton and Trafficking

COTL1[1,2]	coactosin-like F-actin binding protein 1	224583_at	actin binding protein; regulates the actin cytoskeleton	cytoskeletal organization
KRT19	keratin 19	201650_at	intermediate filament protein; involved in the organization of myofibers	
MICAL2	microtubule associated monooxygenase, calponin and LIM domain containing 2	212473_s_at	nuclear monooxygenase; promotes F-actin depolymerization	
SORBS2[1]	sorbin and SH3 domain containing 2	225728_at	cytoskeletal adapter protein; assembles signaling complexes, promotes ubiquitination and proteosomal degradation	
FLRT3[1]	fibronectin leucine rich transmembrane protein 3	219250_s_at 222853_at	membrane protein; function in cell adhesion and receptor signaling	adhesion
RAB27B	RAB27B, member RAS oncogene family	228708_at	prenylated, membrane-bound protein involved in vesicular fusion and trafficking	vesicular trafficking
ANXA1	annexin A1	201012_at	calcium/phospholipid-binding protein; promotes membrane fusion and is involved in endocytosis	

Category: Immune Response

CXCL1	chemokine (C-X-C motif) ligand 1	204470_at	secreted protein with chemotactic activity for neutrophils; role in inflammation and development	Cytokines/ chemokines
CXCL2[1,2]	chemokine (C-X-C motif) ligand 2	209774_x_at		
IL18	interleukin 18	206295_at	secreted, proinflammatory cytokine; augments natural killer and T-cell activity	
TRIM22[1]	tripartite motif containing 22	213293_s_at	interferon-induced antiviral protein; involved in cell innate immunity, blocks viral transcription and replication	other
UBASH3B	ubiquitin associated, SH3 domain containing B domain-containing protein B	238462_at	protein tyrosine phosphatase; regulates receptor mediated signaling in T-cells, ubiquitin ligand for t-cells	

Category: Transporter

ATP2B4	ATPase, Ca++ transporting, plasma membrane 4	212136_at	membrane bound enzyme; catalyzes the hydrolysis of ATP coupled with the transport of calcium out of the cell	ion transport
SLC26A2[1]	solute carrier family 26 member 2	205097_at	transmembrane glycoprotein; sulfate transporter	
TRPM8[1,2]	transient receptor potential cation channel, subfamily M, member 8	243483_at	membrane located, Ca(2+)-permeable cation channel; activated by temperature or pH changes	
SLC2A10[1,2]	solute carrier family 2 member 10	221024_s_at	membrane integrated glucose transporter, role in maintaining glucose homeostasis	other
SLC38A4[1,2]	solute carrier family 38, member 4	220786_s_at	membrane bound, sodium-dependent amino acid transporter	

Category: Energy and Lipid Metabolism

PCK1[1,2]	phosphoenolpyruvate carboxykinase 1 (soluble)	208383_s_at	gluconeogenic enzyme; production of glucose from lactate and other precursors derived from the citric acid cycle	glucose homeo-stasis
SLC2A2[2]	solute carrier family 2, member 2	206535_at	glucose transporter; transfer of glucose across the plasma membrane	
GPAM[1]	glycerol-3-phosphate acyltransferase	225424_at	mitochondrial membrane enzyme; glycerolipid biosynthesis, regulates of triacylglycerol and phospholipid levels	lipid synthesis

Category: Cell Cycle and Cytoskeleton

MCM6	minichromosome maintenance complex component 6	201930_at	replicative helicase; essential for 'once per cell cycle' DNA replication
MKI67	marker of proliferation Ki-67	212022_s_at	nuclear protein; required for maintaining cell proliferation
NUF2	NUF2, NDC80 kinetochore complex component	223381_at	nuclear protein; chromosome segregation and spindle checkpoint activity

Category: Apoptosis

BRI3BP	BRI3 binding protein	225716_at	membrane protein; negative regulator of the p53 tumor suppressor, involved in apoptosis mediated by TNF
CASP1[1]	caspase 1, apoptosis-related cysteine peptidase	211367_s_at 211368_s_at	cytosolic protease; promotes apoptosis, regulates inflammatory processes by proteolytical cleavage of IL proteins

Category: Other

ART4[1]	ADP-ribosyltransferase 4 (Dombrock blood group)	242496_at	lipid anchor protein; covalent transfer of ADP-ribose residue from NAD^+ to amino acids in target proteins
CLRN3[1]	clarin 3	229777_at	transmembrane protein with homology to the tetraspanin family
HNMT[1,2]	histamine N-methyltransferase	228772_at	cytosolic protein; inactivates histamine by N-methylation, role in degrading histamine

[3]206963_s_at, [3]210289_at, [3]214069_at, [3]227794_at, [3]230554_at, [3]241914_s_at, [3]1562049_at, [3]230577_at, [3]233604_at

The overlap between unstable baseline genes and stereotypically altered SV 20 genes, as well as rather compound specific SV 3 genes, is illustrated in Figure 3.14. The Venn diagrams show how many and, for the SV 20 genes, which genes are altered by the hepatocyte isolation and cultivation procedures. These gene expression changes are possibly not related to compound-induced effects.

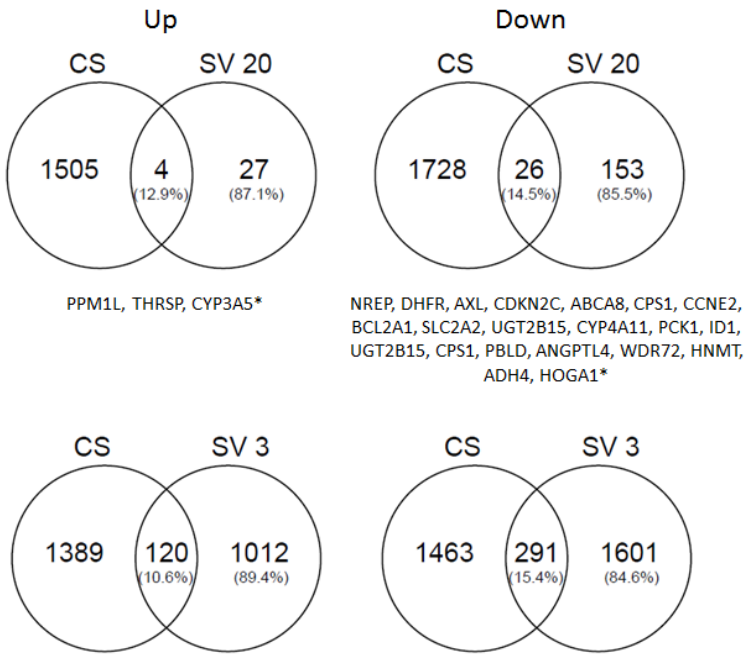


Figure 3.14: Overlap between 'unstable baseline genes' (CS) and the SV 20 (SV 3) genes. The uniquely annotated genes in the overlap of the SV20 genes are listed below the corresponding Venn diagrams (the asterisk refers to probe sets that are not annotated).

3.1.10 Over representative gene ontology groups and transcription factor binding sides

In addition to the manual classification of the genes, gene ontology analysis was performed to assign the stereotypic response genes to biological categories (Table 3.9 A). This computational analysis revealed up regulation of genes involved in metabolism of xenobiotics and endogenous compounds as predominant biological function and down regulation of cell cycle and proliferation associated genes. Thus, it confirmed the manually obtained results.

Furthermore, transcription factor binding sides (TFBS) were analyzed by PRIMA software (Table 3.9 B). Among the up regulated motifs are binding sides for transcription factors like hepatocyte nuclear factor 4 (HNF 4) and further transcription factors involved in developmental and differentiation processes, TCF-4, Nkx and GATA. HNF4 represents a well characterized transcription factor with a pivotal role in regulation of liver function and differentiation (Kamiya et al. 2003; Watt et al. 2003). It influences the expression of large sets of genes controlling liver functions, such as xenobiotics detoxification, energy metabolism, bile acid synthesis and serum protein production (Hayhurst et al. 2001; Inoue et al. 2002; Li et al. 2000; Stoffel and Duncan 1997; Tirona et al. 2003).

Table 3.9: A Overrepresented GO groups for sv20 genes (unadjusted p-value ≤0.01, in total 13 up-regulated, here are all listed, in total 88 down-regulated, here only the top 15 are listed). **B** Overrepresented TFBS (unadjusted p-value ≤ 0.01).

A. Overrepresented GO groups
Up
xenobiotic metabolic process (GO:0006805); exogenous drug catabolic process (GO:0042738); negative regulation of transcription from RNA polymerase II promoter (GO:0000122); oxidative demethylation (GO:0070989); phosphatidylethanolamine biosynthetic process (GO:0006646); monoterpenoid metabolic process (GO:0016098); alkaloid catabolic process (GO:0009822); response to unfolded protein (GO:0006986); negative regulation of endoplasmic reticulum unfolded protein response (GO:1900102); glucose 6-phosphate metabolic process (GO:0051156); steroid metabolic process (GO:0008202); nucleosome assembly (GO:0006334); glucosamine biosynthetic process (GO:0006042); triglyceride metabolic process (GO:0006641); response to sucrose stimulus (GO:0009744); negative regulation of apoptotic process (GO:0043066); negative regulation of fatty acid biosynthetic process (GO:0045717); bile acid catabolic process (GO:0030573); endocardial cushion to mesenchymal transition involved in heart valve formation (GO:0003199); negative regulation of fat cell differentiation (GO:0045599)
Down
mitotic prometaphase (GO:0000236); cell division (GO:0051301); cell cycle checkpoint (GO:0000075); DNA strand elongation involved in DNA replication (GO:0006271); regulation of transcription involved in G1/S phase of mitotic cell cycle (GO:0000083); DNA replication initiation (GO:0006270); S phase of mitotic cell cycle (GO:0000084); G1/S transition of mitotic cell cycle (GO:0000082); M/G1 transition of mitotic cell cycle (GO:0000216); DNA replication (GO:0006260); response to progesterone stimulus (GO:0032570); response to drug (GO:0042493); mitotic chromosome condensation (GO:0007076); DNA unwinding involved in replication (GO:0006268); CENP-A containing nucleosome assembly at centromere (GO:0034080)

B. Overrepresented TFBS
Up
N-Myc (M00055); HNF4 (M01032); Helios_A (M01004); GATA-1 (M00127); AHRHIF (M00976); MEF-2 (M00006); Oct-1 (M00162); ETF (M00695); NF-AT (M00302); COMP1 (M00057); MZF1 (M00084); RSRFC4 (M00026); ZF5 (M00333); Freac-3 (M00291); AIRE (M00999); NF-kappaB_(p65) (M00052); Nkx2-5 (M00240); TATA (M00216); SRY (M00148); GCM (M00634)
Down
HNF4 (M01032); Helios_A (M01004); POU6F1 (M00465); Cdc5 (M00478); Pax-4 (M00377); POU1F1 (M00744); MEF-2 (M00006); Nkx6-2 (M00489); Oct-1 (M00162); TEF-1 (M00704); E2F (M00024); Oct-1 (M00137); ETF (M00695); NKX3A (M00451); HFH-1 (M00129)

Down regulated transcription factor binding motifs comprise, for example, E2F and activating transcription factors (ATF), which are associated with hepatocyte proliferation. These TFBS results are consistent with the GO analysis and the manual classification of stereotypically altered genes.

In addition, the SV 3 genes were verified with computational analysis. An overview of GO groups and TFBS of genes that are deregulated as a consequence of rather compound specific effects is depicted in Table 3.10. GO analysis of deregulated SV 3 genes revealed a broad spectrum of biological functions, including also the stereotypic gene expression responses. SV 3 up regulated genes belong to categories, such as xenobiotic and energy and lipid me-

tabolism, and exogenous drug catabolism, but also include diverse biosynthetic processes, stress response and many more effects. Again, the manual classification of gene expression alterations was confirmed. Similar results were obtained with the TFBS analysis, which identified motifs for transcription factors involved in various biological functions. Up regulated TFBS were determined for several transcription factors involved in development (myocyte enhancer factor 2 (MEF2), GATA-binding factor 1, the n-Myc proto-oncogenic transcription factor); furthermore, binding sites for Helios A were found to be altered, a well-established player in immune cell activation (Akimova et al. 2011). In addition, TFBS for nuclear factor kappa B (NF-kB) are induced, a transcription factor playing a role in the regulation of various biological processes, such as immune response, proliferation, or cell death. In conclusion, the expression response observed for SV 3 genes is much more diverse than for the stereotypic SV 20 genes.

Table 3.10: A Overrepresented GO groups for SV 3 genes (unadjusted p-value ≤0.01, in total 129 up-regulated, here only the top 20 are listed, in total 135 down-regulated, here only the top 15 are listed). **B** Top 20 of the overrepresented TFBS for up regulated genes and top 15 of the overrepresented TFBS for down regulated genes (unadjusted p-value ≤ 0.01).

A. Overrepresented GO groups
Up
xenobiotic metabolic process (GO:0006805); exogenous drug catabolic process (GO:0042738); negative regulation of transcription from RNA polymerase II promoter (GO:0000122); oxidative demethylation (GO:0070989); phosphatidylethanolamine biosynthetic process (GO:0006646); monoterpenoid metabolic process (GO:0016098); alkaloid catabolic process (GO:0009822); response to unfolded protein (GO:0006986); negative regulation of endoplasmic reticulum unfolded protein response (GO:1900102); glucose 6-phosphate metabolic process (GO:0051156); steroid metabolic process (GO:0008202); nucleosome assembly (GO:0006334); glucosamine biosynthetic process (GO:0006042); triglyceride metabolic process (GO:0006641); response to sucrose stimulus (GO:0009744); negative regulation of apoptotic process (GO:0043066); negative regulation of fatty acid biosynthetic process (GO:0045717); bile acid catabolic process (GO:0030573); endocardial cushion to mesenchymal transition involved in heart valve formation (GO:0003199); negative regulation of fat cell differentiation (GO:0045599)
Down
mitotic prometaphase (GO:0000236); cell division (GO:0051301); cell cycle checkpoint (GO:0000075); DNA strand elongation involved in DNA replication (GO:0006271); regulation of transcription involved in G1/S phase of mitotic cell cycle (GO:0000083); DNA replication initiation (GO:0006270); S phase of mitotic cell cycle (GO:0000084); G1/S transition of mitotic cell cycle (GO:0000082); M/G1 transition of mitotic cell cycle (GO:0000216); DNA replication (GO:0006260); response to progesterone stimulus (GO:0032570); response to drug (GO:0042493); mitotic chromosome condensation (GO:0007076); DNA unwinding involved in replication (GO:0006268); CENP-A containing nucleosome assembly at centromere (GO:0034080)
B. Overrepresented TFBS
Up
N-Myc (M00055); HNF4 (M01032); Helios_A (M01004); GATA-1 (M00127); AHRHIF (M00976); MEF-2 (M00006); Oct-1 (M00162); ETF (M00695); NF-AT (M00302); COMP1 (M00057); MZF1 (M00084); RSRFC4 (M00026); ZF5 (M00333); Freac-3 (M00291); AIRE (M00999); NF-kappaB_(p65) (M00052); Nkx2-5 (M00240); TATA (M00216); SRY (M00148); GCM (M00634)
Down
HNF4 (M01032); Helios_A (M01004); POU6F1 (M00465); Cdc5 (M00478); Pax-4 (M00377); POU1F1 (M00744); MEF-2 (M00006); Nkx6-2 (M00489); Oct-1 (M00162); TEF-1 (M00704); E2F (M00024); Oct-1 (M00137); ETF (M00695); NKX3A (M00451); HFH-1 (M00129)

3.1.11 Overlap of chemical-induced gene expression alterations and gene expression changes in liver diseases

Transcriptome analysis of compound-exposed primary human hepatocytes *in vitro* offers a new approach to identify biomarkers of toxicity in humans, but a direct comparison to possible effects *in vivo* is for ethical reasons not possible. It remains challenging to assess whether genes which are found to be deregulated *in vitro* would respond similarly under the same conditions of *in vivo* exposure. Moreover, human *in vivo* data is restricted to tissue from intoxicated patients or patients with liver diseases undergoing surgery, and the conditions of exposure are not precisely defined.

To bridge the gap between gene expression alterations *in vitro* and a possible relevance for a gene *in vivo*, publically available whole transcriptome data sets from liver tissue samples of liver disease patients were used. Liver disease tissue transcriptomics data from patients suffering from liver diseases, such as hepatocellular carcinoma (HCC), cirrhosis or non-alcoholic steatohepatitis (NASH) was compared to healthy or non-tumor tissue. Overlapping genes that are altered in liver disease tissue, as well as by chemicals *in vitro*, may hint at a possible relevance for a gene *in vivo* and decreases the probability that the effect observed *in vitro* is just an *in vitro* artifact.

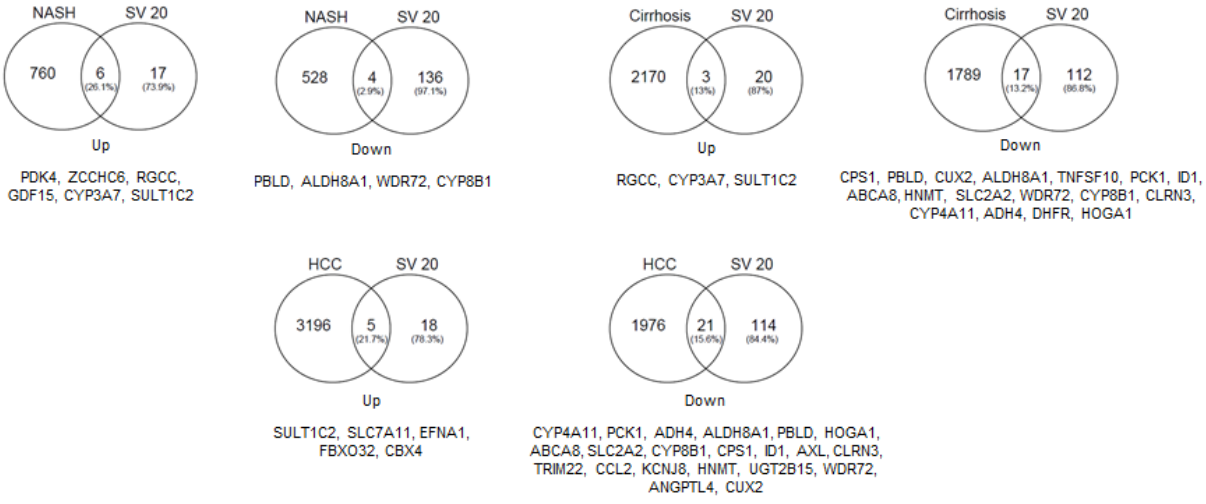


Figure 3.15: Overlap of SV 20 genes altered by chemicals and genes deregulated in the human liver diseases non-alcoholic steatohepatitis (NASH), cirrhosis and hepatocellular carcinoma (HCC). The genes in the overlap are listed below the corresponding Venn diagrams.

The overlap between liver disease and stereotypic response SV 20 genes is depicted in Figure 3.15. 13 – 21.7 % of the up regulated SV 20 genes were found to be also induced in cirrhotic and hepatocellular carcinoma tissue, and the overlap among down regulated genes was 13.2 - 15.6 %. 26.1 % of the SV 20 genes were also found to be induced by NASH, but the fraction of overlapping down regulated genes is 4 % only. One of the genes that are altered as a stereotypical response to chemically-induced stress, as well as in human liver diseases, is, for

example, CYP3A7. CYP3A7 is a phase I enzyme involved in metabolism of endogenous and xenobiotic compounds. It represents the predominant cytochrome P450 enzyme in human fetal liver and accounts for 30–50% of the total CYP in fetal liver and 87–100% of total fetal hepatic CYP3A content (Pang et al. 2012). In addition, SULT1C2 was found to be induced. This phase II metabolic enzyme belongs to the family of sulfotransferases and probably represents the major detoxification enzyme system expressed in the human fetus (Stanley et al. 2005). Regulator of cell cycle (RGCC) is another gene which is up regulated *in vitro* after chemical-exposure and in liver diseases. RGCC is reported to be induced by p53 and modulates the activity of cell cycle specific kinases in response to DNA damage (Huang et al. 2009; Saigusa et al. 2007).

Among the down regulated genes in the overlap of liver disease genes and chemically altered SV 20 genes are the aldehyde dehydrogenase family members ALDH8A1 and ALDH4, the gluconeogenesis regulating enzyme PCK1, the sterol and fatty acid metabolizing cytochrome P450 isoenzymes CYP8B1 and CYP4A11, the ATP dependent lipophilic drug transporter ABCA8, the urea cycle enzyme CPS1 and the glucose transporter SLC2A2.

Similarly, the SV 3 genes revealed a strong overlap with genes altered in human liver diseases (Figure 3.16). 12.9 - 18.6 % of the SV 3 induced genes were altered in the same direction of expression in NASH, HCC and cirrhosis.

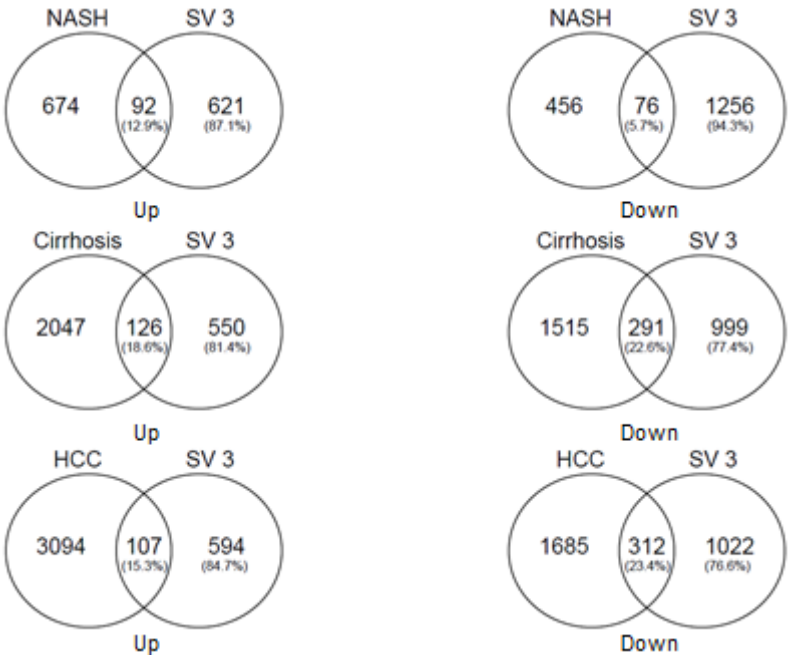


Figure 3.16: Overlap of SV 3 genes altered by chemicals and genes deregulated in the human liver diseases non-alcoholic steatohepatitis (NASH), cirrhosis and hepatocellular carcinoma (HCC). The genes in the overlap are listed in Supplemental Table 4

Down regulated SV 3 genes overlap about 23 % with genes that are down regulated in cirrhosis and HCC, but only 5.7 % of down regulated SV 3 genes are also altered in NASH. The small overlap between liver disease genes and SV 20 and SV 3 genes, respectively, may be explained by the small data set, which is available from patients suffering from NASH. A detailed overview of individual SV 3 genes overlapping with human liver disease genes is given in Supplemental Table 4.

3.2 Application of the toxicogenomics directory: Identification of biomarker candidate genes and their potential to predict human hepatotoxic blood concentrations.

The second part of this thesis focusses on the identification of biomarker genes based on the toxicogenomics directory and the applicability of these genes in predicting blood concentrations that are associated with hepatotoxicity *in vivo*. Ideally, a biomarker gene differentiates between hepatotoxic and non-hepatotoxic compounds and changes its expression only in case of a hepatotoxic or cytotoxic effect. Unsupervised cluster analysis of the Open TG-GATEs data (Figure 3.11, 24h exposure at a slightly cytotoxic concentration) did not differentiate between hepatotoxic and non-hepatotoxic compounds. However, the compounds were tested in a slightly cytotoxic, but not an *in vivo* relevant concentration range, so that cytotoxic effects may have covered differences in gene expression between the two groups of compounds. To test whether a compound is associated with an increased risk of hepatotoxicity *in vivo*, it is recommendable to analyze biomarker expression within an *in vivo* relevant concentration range, which covers also the plasma concentration of a therapeutic dose. The following sections of this thesis will describe, step by step, how suitable biomarkers of toxicity were extracted from the curated data base and how these genes were analyzed *in vitro* and applied to predict human hepatotoxic blood concentrations. HepG2 cells and primary human hepatocytes were chosen as *in vitro* systems.

3.2.1 Selection of compounds

Two sets of compounds were defined, namely a set of compounds that are associated with a high risk of hepatotoxicity at therapeutic doses and a set of compounds for which no hepatotoxic effects are reported (Table 3.11). Additionally, the drug acetaminophen was included in the set of hepatotoxic compounds. Acetaminophen is not associated with increased risk for hepatotoxicity at therapeutic doses, but the toxic blood concentration is well documented in literature. Liver toxicity due to intoxication from acetaminophen overdose represents a pervasive problem in society.
The selection of compounds is based on a manual literature search in Pubmed [1] and data from the database Livertox [2], which delivers information on the hepatotoxic potential of

numerous compounds. Compounds were considered as hepatotoxic if a relatively high number of patients developed any kind of liver injury after taking a therapeutic dose.

Table 3.11: Selection of compounds with increased risk of hepatotoxicity and negative control compounds without reported liver toxic effects. Information on hepatotoxicity was obtained by searching in Pubmed [1] and the Livertox database [2].

Hepatotoxic compounds		Non-hepatotoxic compounds	
APAP	Acetaminophen	**BPR**	Buspirone
ASP	Aspirin	**CHL**	Chlorpheniramine
CBZ	Carbamazepine	**CLO**	Clonidine
DFNa	Diclofenac	**FAM**	Famotidine
INAH	Isoniazid	**HYZ**	Hydroxyzine
KC	Ketoconazole	**LEV**	Levofloxacin
LBT	Labetalol	**MEL**	Melatonin
NIM	Nimesulide	**PMZ**	Promethazine
NFT	Nitrofurantoin	**PPL**	Propranolol
PhB	Phenylbutazone		
RIF	Rifampicin		
VPA	Valproic acid		

Compounds were considered as non-hepatotoxic if the following criteria were fulfilled: (i) the compound was not listed as hepatotoxic in the livertox database and (ii) Pubmed search based on the terms 'compound' AND hepatotoxicity OR liver toxicity' was performed but no evidence for hepatotoxicity in humans at therapeutic doses was obtained. Compounds which were difficult to interpret are listed in Table 3.12. These compounds were mainly found to be harmless, but single cases of acute liver injury were reported.

Table 3.12: Compounds which are questionable regarding their hepatotoxic potential. For these compounds single cases of hepatotoxicity in different contexts were reported, but evidence for a direct liver toxic effect is missing.

Compounds	Hepatotoxic effect	Reference
Famotidine	rare cases of clinically apparent liver injury, cases varied in the time to onset and pattern of injury; in the reported cases famotidine was combined with other drugs	Gupta et al. 2009 Ament et al. 1994 Hashimoto et al. 1994
Chlorpheniramine	clinically apparent liver injury exceedingly rare, few cases reported in the literature but it is not considered to be a hepatotoxic drug	Mignot et al. 2000 Pagani et al. 1987 Farrell et al. 1994 Stricker et al. 1995
Levofloxacin	rare instances of clinically apparent hepatic injury marked by a short latency period and a hepatocellular pattern of enzyme elevations; used as non-hepatotoxic control compound by Cosgrove et al. (2009)	Karim et al. 2001 Gulen et al. 2015

For the set of hepatotoxic compounds, a more detailed overview of hepatotoxic effects, possible mechanisms of toxicity and frequencies of liver injury when administered at therapeutic doses were elaborated (Table 3.13 and Table 3.14).

Table 3.13: Medication, phenotype and frequency of liver injury observed for the selected hepatotoxic compounds.

Compound	Medication	Phenotype of hepatotoxicity	Frequency of liver injury	References
Acetaminophen	Non-steroidal anti-inflammatory drug, analgesic and anti-pyretic agent	necrosis	Not hepatotoxic at therapeutic doses; but toxic doses account for approx. 50% of all acute liver failure cases in the USA and many Western countries	Lee 2012
Aspirin	Non-steroidal anti-inflammatory drug, analgesic and anti-pyretic agent	hepatitis, steatosis, Reye syndrome (lactic acidosis, microvesicular fat, hepatic dysfunction)	At least 6 cases of severe liver injury have been reported.	Musumba et al. 2004 Kanada et al. 1978 Laster and Satoskar 2014 Chen et al. 2001
Carbamazepine	Anticonvulsant, used for the treatment of epilepsy and psychatric disorders	cholestasis hepatitis	In 165 cases of CBZ hypersensitivity up to 1998, 47% of the cases were associated with liver injury. CBZ-induced liver injury is estimated to occur in 16 per 100,000 patients per year.	Pirmohamed et al. 2013 Feldmann et al. 2015
Diclofenac	Non-steroidal anti-inflammatory drug, treatment of mild to moderate pain, treatment of arthritis	hepatitis, necrosis cholestasis (rare)	Published cases of severe hepatotoxicity amount to approximately 250 reports, with a case fatality rate of approximately 10%. DFN-induced liver injury occurs with a frequency of 1-5 cases per 100,000 patients.	Lewis et al. 2003 Garcia Rodriguez et al. 1994 Chitturi and George 2002
Isoniazid	Antibiotic agent, used for treatment of tuberculosis	hepatitis	In a study with 2,321 men treated with isoniazid, 19 patients showed clincial signs of liver disease and 2 patients died. Another study revealed 1% of 14,000 individuals treated with INAH developed hepatitis. The frequency of INAH-induced liver injury is estimated to occur in 1.6 % of all patients, in 2.55 % when combined therapy with rifampicin.	Kopanoff et al. 1978 Garibaldi et al. 1972 Saukkonen et al. 2006 Steele et al. 1991
Ketoconazole	Treatment of fungal infections	hepatitis, phospholipidosis in mice	Up to 1987 there were 82 reports of possible hepatotoxicity in patients taking oral ketoconazole, including five deaths. The frequency of KC-induced liver injury is estimated to occur in 0.1 - 1% of all patients	Lake-Bakaar et al. 1987 Rodriguez and Acosta 1997
Labetalol	Anti-hypertensive agent, used for the treatment of high blood pressure	hepatitis, necrosis	At least 5 cases of severe liver injury have been reported.	Long et al. 2007 Marinella 2002 Stronkhorst et al. 1992 Douglas et al. 1989 Clark et al. 1990

Compound	Medication	Phenotype of hepatotoxicity	Frequency of liver injury	References
Nimesulide	Non-steroidal anti-inflammatory drug, analgesic and antipyretic agent	hepatitis	A number of severe cases of NIM-induced hepatotoxicity have been reported. The frequency of NIM-induced liver injury is estimated to be 0.1 of 100,000 patients.	Tan et al. 2007 Merlani et al. 2001 Chatterjee et al. 2008 Bessone 2010 Boelsterli 2002
Nitrofurantoin	Antibiotic agent, treatment of bladder infections	granulomatous, cholestatic or chronic hepatitis	A number of severe cases of NFT-induced hepatotoxicity have been reported. The frequency of NFT-induced liver injury is estimated to be 0.00003%	Kiang et al. 2011 Appleyard et al. 2010 Amit et al. 2002 Sherigar et al. 2012 Moseley 2013
Phenylbutazone	Non-steroidal anti-inflammatory drug, analgesic and antipyretic agent for short term treatment in animals	necrosis, hepatitis	Severe hepatotoxic effects have been observed in a large number of patients. Several studies revealed severe liver injury in 1-5 % of all patients at therapeutic doses of 400mg/day.	Benjamin et al. 1981 Feldmann et al. 2015
Rifampicin	Antibiotic agent, used for treatment of tuberculosis, leprosy and legionella	necrosis, hepatitis, cholestasis	Although RIF-induced hepatotoxicity is especially associated with elevated transaminase levels, a number of severe cases of liver injury have been reported. Hepatotoxicity occurs in up to 1 % of all patients.	Prince et al. 2002 van Hest et al. 2004 Steele et al. 1991
Valproic acid	Anticonvulsant, used for the treatment of epilepsy	microvesicular steatosis	In one study, 1197 patients were monitored. 42 cases of severe hepatitis, 3 cases Reye's like syndrome and 22 instances of hyperammonemia were observed. The risk of fatal hepatotoxicity by VPA is estimated to affect 1/500 children below the age of 2 years, 1/12,000 in adults used in polytherapy and 1/37,000 in adults used in monotherapy.	Powell-Jackson et al. 1984 Ahmed and Siddiqi 2006

Table 3.14: Suggested mechanisms and possible explanations underlying the hepatotoxic effect of the selected compounds**.**

Compound	Suggested mechanisms of hepatotoxicity	References
Acetamino-phen	Metabolism via CYP2E1, 3A4 and 1A2 produces the highly reactive metabolite N-acetyl-p-benzoquinone imine (NAPQI). NAPQI is detoxified by glutathione and eliminated into the urine. With increased doses of APAP, NAPQI covalentely binds to proteins and depletes the glutathione stores, resulting in oxidative stress.	Daly et al. 2008 Dart et al. 2006 James et al. 2003
Aspirin	ASP mediated hepatotoxicity occurs only at high doses. The mechanism is not fully elucidated but mitochondrial dysfunction is reported to play an important role. Inhibition of the β-oxidation and the delivery of metabolites to the electron transport chain; the mitochondrial fuel supply and energy flux are reduced. ASP further causes intracellular ATP decrease, which leads to hepatocellular injury mediated by lipid peroxidation.	Fromenty and Pessayre 1995 Doi and Horie 2010
Carbamaze-pine	The mechanism of CBZ hepatotoxicity appears to be hypersensitivity or an immunological response to a metabolically generated drug-protein complex. Metabolization via CYP3A4 leads to the formation of reactive metabolites, which further involve the immune system and results in tissue injury.	Pandit et al. 2012 Forbes et al. 1992 Mitchell et al. 1981
Diclofenac	Extensive metabolization in the liver via cytochrome P450 enzymes CYP2C 9 and CYP2C8. The toxic products, acyl glucuronide and benzoquinone imines modify proteins covalently. Accumulation of reactive metabolites generates oxidative stress and is accompanied by mitochondrial impairment.	Pandit et al. 2012 Ponsoda et al. 1995 Bort et al. 1999 Chitturi and George 2002
Isoniazid	Metabolism via the N-acetyltransferase 2 to monoacetyl hydrazine, which is further metabolized via CyP2E1 to toxic metabolites that covalently bind to hepatic macromolecules. Hepatotoxicity seems to be an idiosyncratic response and is dependent on the CYP2E1 genotype. INAH has an inhibitory effect on CYP1A2, 2A6, 2C19 and 3A4 activity. It can induce its own toxicity, probably by the induction or inhibition of these enzymes.	Pandit et al. 2012 Tostmann et al. 2008 Lauterburg et al. 1985 Steele et al. 1991
Ketoconazole	Biotransformation to N-desacetyl ketoconazole (DAK) and further metabolization of DAK by flavin-containing mono-oxygenases results in covalent binding to hepatic proteins and glutathione depletion.	Greenblatt and Greenblatt 2014 Rodriguez and Acosta 1997 Rodriguez et al. 1999 Rodriguez and Buckholz 2003
Labetalol	The mechanism of LAB-induced hepatotoxicity is unknown. Histological patterns of inflammation suggest an immune-mediated response.	Halegoua-De Marzio and Navarro 2013 Clark et al. 1990 [3]
Nimesulide	NIM has been associated with idiosyncratic hepatotoxicity in susceptible patients. The molecular mechanisms of NIM-induced hepatotoxicity have not yet been fully elucidated. It has been suggested that NIM undergoes bio-reductive metabolism and forms reactive metabolites, which have been implicated in oxidative stress, covalent binding to hepatic proteins and mitochondrial injury.	Chitturi and George 2002 Tripathi et al. 2010 Singh et al. 2012 Bessone et al. 2010 Boelsterli 2002
Nitro-furantoin	The mechanism of NFT-mediated hepatotoxicity is poorly understood and presumed to be the result of an immunologic process or a direct cytotoxic reaction. Its nitro-reductive metabolism produces oxidative free radicals, which result in hepatocyte damage.	Sakaan et al. 2014 Moseley et al. 2013

Compound	Suggested mechanisms of hepatotoxicity	References
Phenyl-butazone	PhB is extensively metabolized in the liver to oxyphenbutazone, gamma-hydroxyphenylbutazone and p, gamma-dihydroxyphenyl-butazone. Involved enzymes and the underlying mechanism of hepatotoxicity are not well understood.	Dieterle et al. 1976 Aarbakke et al. 1977 Aronson 2009
Rifampicin	RIF is a potent inducer of the hepatic CYP450 system, such as CYP 2C and 3A, thereby increasing the metabolism of other compounds. The detailed mechanism of its hepatotoxicity is not well understood. RIF is often combined with isoniazid, leading to an increased risk of hepatotoxicity. The cause of injury is most likely due to idiosyncratic metabolic products that are either directly toxic or induce an immunologic reaction.	Grange et al. 1994 Tostmann et al. 2008 Pandit et al. 2012 Steele et al. 1991
Valproic acid	Metabolism via glucuronidation and mitochondrial b-oxidation. VPA enters the mitochondria via the long chain fatty acid transport system, which uses carnitine as a co-factor. VPA is first attached to coenzyme A to form VPA-CoA. VPA-CoA is then esterified with L-carnitine to form VPA-carnitine ester, which is subsequently transported into the mitochondrial matrix by carnitine translocase in exchange for free carnitine. Conjugation of VPA to carnitines results in carnitine depletion, which inhibits β-oxidation of endogenous lipids and results in microvesicular steatosis and mitochondrial dysfunction. Further reports show that VPA is associated with the formation of increased reactive oxygen species.	Sztajnkrycer 2002 Pandit et al. 2012 Pourahmad et al. 2012 Begriche et al. 2011

3.2.1 Identification of peak plasma concentrations and selection of a concentration range

To evaluate a risk of hepatotoxicity *in vivo*, concentrations for all compounds used in this study were selected based on *in vivo* relevant concentrations. For all 21 compounds, a literature search was performed to identify peak plasma concentrations of therapeutic doses in patients. These results are summarized in Table 3.15 and Table 3.16, which deliver information on therapeutic doses of a compound and the resulting peak plasma concentration. The study is based on free, meaning non-protein bound, concentrations.

Table 3.15: Overview of peak plasma concentrations of compounds which are associated with a high risk of hepatotoxicity at therapeutic doses. The table is based on literature search and delivers information on recommended doses and resulting plasma levels of a drug. PPB = plasma protein binding, values are from drugbank.ca; except for VPA (reference O'Brien et al. 2006).

Compound		Therapeutic dose	Dose and route of application	Peak plasma concentration	Plasma levels [M]	PPB	Reference
APAP	**Acetaminophen**	660 - 1000 mg every 4 - 6 h max. 3 g /day	a) c) 20-40mg/kg rectal 3-5h after administr. b) 2x 235mg oral dose 60-85 min after administr.	a) c) 10-20ug/ml (40mg/kg) <10µg/ml (20 mg/kg) b) 6.93 / 7.72µg/ml d) 10-20 µg/ml	a) c) 66.15 - 165.39µM b) 45.85µM/ 51.07µM d) 66.16 - 132.31 µM	25%	a) Beck et al.2000 b) Albert et al. 1974 c) Stocker and Montgomery 2001 d) Winek et al. 2001
ASP	**Aspirin**	330 - 660 mg every 4 - 6 h	b) 1200mg oral dose peak 10-20 min after administration	a) 0.1-2mM b) 17-40 µg/ml	a) 0.1 - 2 mM b) 94-222 µM	99.50%	a) Frantz et al. 1995 b) Seymor et al.1984
CBZ	**Carbamazepine**	200 mg 2x daily	a) ~ 200 mg oral dose mean concentration	a) 5.4+/-2.5 µg/ml b) 4-12 µg/ml	a) 12.27 -33.44 µM b) 16.93 - 50.79 µM	76%	a) Eichelbaum et al. 1976 b) McMillin et al. 2010
DFN	**Diclofenac**	for chronic arthritis 50 mg 3x daily	a) 50 mg oral dose b) concentration associated with efficacy	a) 1.7 µg/ml	a) 5.8 µM b) 4.2 µM	99%	a) Kircheiner et al. 2003 b) O'Brien et al. 2006
INAH	**Isoniazid**	300 mg daily or 900 mg/day 2-3x/week	b) 5-10 mg/kg/day in children c) 5.15 mg/kg/day daily and 12.8 mg/kg/day 2x weekly	a) 0.6-20 µg/ml b) 3.2 - 8.11 µg/ml c) 2.5 - 3 µg/ml and 8-10 µg/ml	a) 4.38 - 145.84 µM b) 21.9-36.5 µM	0-10 %	a) Winek et al. 2001 b) Thee et al. 2011 c) Requena-Méndez et al.(2014
KC	**Ketoconazole**	fungal infections 200-400 mg daily prostate cancer 400 mg 3x daily	a) 200mg oral dose 1h after administr. b) 400-2000mg oral dose 4-6h after administr.	a) 3+/- 0.5 µg/ml b) 7-17 µg/ml	a) 4.7 - 6.59 µM b) 13- 32 µM	99%	a) Schäfer-Korting et al. 1984 b) Sugar et al 1987
LBT	**Labetalol**	200-400 mg daily	a) 200mg oral dose 2h after administr. b) 100-1000 mg dose 3x daily up to 6 weeks c) 100-400 mg oral dose	a) 275 +/- 99 ng/ml b) 37-510 ng/ml c) 100mg: 32 ng/ml; 200mg: 83 ng/ml; 400mg: 165 ng/ml	a) 482.4 nM - 1.025 µM b) 101.4 nM - 1.4 µM c) 87.7 nM; 227.5 nM; 452.2 nM	50%	a) Lalonde et al. 1990 b) Sanders al. 1979 c) Richards et al. 1977
NIM	**Nimesulide**	200 -400 mg/day	a) 100mg oral dose 30 min after administr. b) 100 mg oral dose	a) 2.0 - 2.5 µg/ml b) 1.95 ± 0.67 µg/ml	a) 6.49 - 8.11µM b) 4.15 - 8.5 µM	>97.5%	a) Bianchi et al. 2006 b) Macpherson et al. 2013

Compound		Therapeutic dose	Dose and route of application	Plasma peak concentration	Plasma levels [M]	PPB	Reference
NFT	**Nitrofurantoin**	50-100 mg 4 x daily	a) 100 mg oral dose b) 100mg oral dose 2-2.3 h after administr.	a) 1-1.5µg/ml b) 0.87-1.1 µg/ml c) 1.8 µg/ml	a) 6.3 µM b) 3.65 - 4.62 µM c) 7.56 µM	20-60 %	a) Albert et al. 1974 b) Adkison et al. 2008 c) Winek et al. 2001
PhB	**Phenylbutazone**	50 - 300 mg/day	a) 50, 100, 200 or 300 mg/day oral dose	a) ~25-100µg/ml b) 16-150 µg/ml	a) 81.07 -324.29 µM b) 51.89 - 486.43 µM	up to 95%	a) Orme et al. 1976 b) Winek et al. 2001
RIF	**Rifampicin**	600 mg daily (~10 mg/kg)	a) b) 450 mg oral dose c) 600mg, 900mg oral dose	a) 6-9µg/ml b) 4-32 µg/ml c) 1-15µg/ml	a) 10.94 µM b) 4.86 - 38.89µM c) 1.215 - 18.22 µM	89%	a) Ellenhorn and Barceloux 1988 b) Mandel and Sande 1985 c) Mehta et al. 2001
VPA	**Valproic acid**	10 - 15 mg/kg/ day	a) 1200-1600mg iv	a) 50-100µg/ml b) 50-100 µg/ml	a) 346.72 -693.43 µM b) 346.72 -693.43 µM	93%	a) De Turck et al. 1998 b) Winek et al. 2001

Table 3.16: Overview of peak plasma concentrations of control compounds at therapeutic doses. The table is based on literature search and delivers information on recommended doses and resulting plasma levels of a drug. PPB =plasma protein binding, values are from drugbank.ca; except for MEL (reference Cardinali et al. 1972).

Compound		Therapeutic dose	Dose and route of application	Peak plasma concentrations	Plasma levels [M]	PPB	Reference
BPR	**Buspirone**	15-30 mg daily in devided doses	a) 20 mg oral dose b) 20mg oral dose c) 30mg oral dose	a) 2.5ng/ml b) 1.15 +/- 0.77ng/ml c)6.6 ± 3.7 ng/ml	a) 5.9 nM b) 0.9 - 4.55 nM c) 15.6 - 24.4 nM	95%	a) Mahmood and Sahajwalla 1999 b) Dalhoff et al. 1987 c) Lamberg 1998
CHL	**Chlorpheniramine**	2-4 mg 3-4 x daily	a) 4-10 mg oral dose b) 2-6mg oral dose	a) 11.9-35.6 ng/ml b) 2mg: 3.4-7.4 ng/ml 6mg: 2-14.3 ng/ml	a) 30.4 nM b) 8.7 nM; 5.1 nM	72%	a) Huang et al. 1982 b) Tagawa et al. 2002
CLO	**Clonidine**	0.1 - 0.6 mg daily in 2-3 doses	a) 150µg iv b) 75 - 300µg oral dose	a) 0.846+/-0.288 ng/ml b) 300µg: 1.17+/-0.12 ng/ml 75µg dose continous treatment steady state level of 0.3-0.35 ng/ml	a) 2.1 - 4.3 nM b) 3.9 - 4.8 nM; 1.1 - 1.3 nM	20-40%	a) Klein et al. 2013 b) Keränen et al. 1978
FAM	**Famotidine**	20 mg 2x/day	b) 40mg oral dose 40 mg dose	a) 17 - 139 ng/ml b) 4-137 ng/ml for 20mg 15-358 ng/ml for 40mg	a) 50.3 - 411.9 nM b) 45 - 1060 nM	15-20%	a) Morgan et al. 1990 b) Chremos 1987
HYZ	**Hydroxyzine**	25-100 mg 3-4 x daily	b) 50mg oral dose (439 ± 66 mg = 07mg/kg)	a) 0.022-0.08 µg/ml b) 116.5+/-60.6 ng/ml	a) 49.1 - 178.6 nM b) 124.8 - 395.5 nM	93%	a) Winek et al. 2001 b) Simons et al. 1989
LEV	**Levofloxacin**	250-750 mg daily	a) 50-1000mg oral dose b) 250-500mg iv c) 500mg oral dose	a) 0.6-9.4 µg/ml b) 1.2-7.7 µg/ml c) 6.34±1.42 µg/ml	a) 1.66 - 26 µM b) 3.32 - 21.3µM c) 17.5 - 21.5 µM	24-38%	a) Fish and Chow 1997 B) Malone et al. 2001 c) Chien et al. 1997
MEL	**Melatonin**	10 mg daily	a) 2 mg dose	a) 10-23 pg/ml	a) 43 - 99 pM	61–78%	a) Aldhous et al. 1985
PMZ	**Promethazine**	12.5 - 25 mg every 4-6 h	a) 50 mg oral dose b) 50mg oral dose c) 50 mg dose oral or rectal	a) 19.3 ng/ml b) 11-23 ng/ml 4h c) 12.1-17.3 ng/ml	a) 60.1 nM b) 34.3 - 71.7 nM c) 37.7 - 53.9 nM	93%	a) Strenkoski-Nix et al. 2000 b) Wallace et al. 1981 c) Schwinghammer et al. 1984
PPL	**Propranolol**	60-240 mg daily devided in 2 doses	a) 40 mg dose peroral and sublingual b) 20mg oral dose c) 160 mg oral dose	a) Peroral 41 ± 12 ng/ml Sublingual 147 ± 72 ng/ml b) 24-28.5 ng/ml c) 38-194 ng/ml	a) 98 nM - 740 nM b) 81.1 nM - 96.3 nM c) 129 - 656 nM	> 90%	a) Mansu et al. 1988 b) Sharoky et al. 1988 c) Aro et al. 1982

HepG2 cells and primary human hepatocytes were exposed for 24 hours and each compound was tested in a concentration range covering the peak plasma concentration of a therapeutic dose and increasing up to a slightly cytotoxic concentration (concentrations C1 – C5, Table 3.17). The highest concentration (C5) was supposed to be slightly cytotoxic, in the range of the IC_{10} – IC_{20} (for cytotoxicity tests see Supplemental Figure 6 and Supplemental Figure 7). Usually a dilution factor of 5 was chosen. Due to cell killing events, a smaller dilution factor was chosen for the highest concentration of the hepatotoxic compound ketoconazole. For some chemicals, especially the non-hepatotoxic compounds, even higher dilution factors were applied for the highest concentration. For this purpose dilution factors of up to 5,000 (for melatonin) were included.

Table 3.17: Selection of concentrations for each compound. Five concentrations (C1-C5) were defined, spanning from sub therapeutic doses and increasing up to slightly cytotoxic concentrations (C5). Peak plasma concentrations are marked by bold letters.

Compound	C1	C2	C3	C4	C5	Solvent	Stock solution
Acetaminophen	8 µM	**40 µM**	**200 µM**	1 mM	5 mM	Medium	-
Aspirin	8 µM	**40 µM**	**200 µM**	**1mM**	5mM	0.2% DMSO	2.5 M
Carbamazepine	1.6 µM	8 µM	**40 µM**	200 µM	1 mM	0.5% DMSO	200 mM
Diclofenac	64 nM	320 nM	**1.6 µM**	**8 µM**	400 µM	0.2% DMSO	200 mM
Isoniazid	8 µM	**40 µM**	200 µM	1 mM	10 mM	Medium	-
Ketoconazole	0.32 µM	**1.6 µM**	**8 µM**	40 µM	100 µM	0.1% DMSO	100 mM
Labetalol	64 nM	**0.32 µM**	**1.6 µM**	8 µM	40 µM	Medium	-
Nimesulide	0.32 µM	**1.6 µM**	**8 µM**	40 µM	330 µM	0.2% DMSO	165 mM
Nitrofurantoin	0.32 µM	**1.6 µM**	**8 µM**	40 µM	200 µM	0.1% DMSO	200 mM
Phenylbutazone	1.6 µM	8 µM	**40 µM**	**200 µM**	1 mM	0.2% DMSO	500 mM
Rifampicin	0.32 µM	1.6 µM	**8 µM**	**40 µM**	200 µM	0.1% DMSO	200 mM
Valproic acid	8 µM	40 µM	**200 µM**	**1 mM**	5 mM	Medium	-
Buspirone	0.51 nM	**2.6 nM**	**12.8 nM**	64 nM	30 µM	1 % H_2O	3 mM
Chlorpheniramine	2.6 nM	**12.8 nM**	64 nM	0.32 µM	90 µM	1 % H_2O	9 mM
Clonidine	0.51 nM	2.6 nM	**12.8 nM**	64 nM	1mM	Medium	-
Famotidine	12.8 nM	**64 nM**	320 nM	1.6 µM	700 µM	0.2% DMSO	350 mM
Hydroxyzine	12.8 nM	**64 nM**	**320 nM**	1.6 µM	40 µM	1 % H_2O	4 mM
Levofloxacin	0.32 µM	**1.6 µM**	**8 µM**	40 µM	200 µM	1 % H_2O	20 mM
Melatonin	20 pM	**0.1 nM**	0.51 nM	2.6 nM	100 nM	0.1% DMSO	100 µM
Promethazine	12.8 nM	**64 nM**	320 nM	1.6 µM	35 µM	1 % H_2O	3.5 mM
Propranolol	64 nM	**320 nM**	**1.6 µM**	8 µM	40 µM	1 % H_2O	10 mM

3.2.2 Identification of biomarker candidate genes according to the toxicogenomics directory

In a next step, potential biomarker candidate genes were identified by using genome-wide expression data from the Open TG GATEs data base. With the exception of levofloxacin, clonidine and melatonin, genome-wide expression data is available for 18 out of the 21 selected compounds. To focus on genes which are strongly altered by chemicals, the top ten genes with the highest fold change of induction were characterized for all 18 compounds (Table 3.18). Based on literature search, all genes were manually assigned to biological categories. Next to the biological function of the gene, this list provides information on the overlap with liver disease genes, and by how many compounds the expression is influenced.

Table 3.18: List of genes which are up regulated with the highest fold change among all 18 compounds where gene array data was available. For each compound the top ten genes are listed and characterized according to their function and the marked whether the expression is also altered in liver diseases or because of the culture conditions (CS). SV up and SV down show, by how many compounds the expression of the appropriate gene is altered in which direction. 1 probe set not annotated.

Symbol	Gene	Probesets	NASH	Cirrhosis	HCC	CS	SV ↑ (FC3)	SV ↓ (FC3)	Function of the gene product	Category
Category: Metabolism, Xenobiotics										
ADH1B	alcohol dehydrogenase 1B (class I), beta polypeptide	209612_s_at 209613_s_at	0	0	↓	↓	11	7	cytosolic enzyme; ethanol metabolism	phase I enzymes
CYP1A1	cytochrome P450, subfamily 1A, polypeptide 1	205749_at	0	↓	0	↓	35	0	metabolic enzyme in the ER	
CYP1A2	cytochrome P450, subfamily. 1A, polypeptide 2	207608_x_at 207609_s_at	↓	↓	↓	↓	18	1		
CYP1B1	cytochrome P450, subfamily. 1B, polypeptide 1	202437_s_at	↑	↑	0	0	18	1		
CYP2C9	cytochrome P450, subfamily 2C, polypeptide 9	217558_at	0	↓	↓	0	32	4		
CYP3A4	cytochrome P450, subfamily 3A, polypeptide 4	205998_x_at 208367_x_at 231704_at	0	↓	↓	↓	7	0		
CYP3A5	cytochrome P450, subfamily 3A, polypeptide 5	214235_at 243015_at	0	↓	0	↓	23	0		
CYP3A7	cytochrome P450, subfamily 3A, polypeptide 7	205939_at 211843_x_at	↑	↑	0	↓	39	0		
SULT1C2	sulfotransferase 1C2	205342_s_at 205343_at 211470_s_at	↑	↑	↑	0	22	1	cytosolic enzyme; catalyzes sulfonation	phase II enzymes
SULT2A1	sulfotransferase 2A1	206292_s_at 206293_at	0	↓	↓	↓	25	1		

Category: Development, differentiation and signal transduction

CTSB	cathepsin B	227961_at	↑	↑	↓	↓	11	1	protease; degradation and turnover of proteins active in tumorigenesis, angiogenesis, invasion and metastasis	development and differentiation
ENO2	enolase 2	201313_at	0	0	0	0	10	0	extracellular enzyme; neuronal development	
KISS1R	KISS1 receptor	242517_at	0	0	0	0	0	12	G-protein coupled receptor; metastasis suppressor protein, regulation of endocrine function and the onset of puberty	
LOX	lysyl oxidase	213640_s_at 215446_s_at	0	0	↑	0	0	0	extra cellular matrix protein; cross-linking of extracellular matrix proteins	development and differentiation
TMPRSS2	transmembrane protease, serine 2	211689_s_at	0	↓	↓	↓	10	0	serin protease; putative role in angiogenesis and development	
TNFSF15	tumor necrosis factor (ligand) superfamily, member 15	229242_at	0	0	0	0	2	0	membrane receptor protein; cell proliferation, immune regulation, inflammation, apoptosis	
ADRB1	adrenoceptor beta 1	229309_at	0	0	↓	↓	3	1	G-protein coupled receptor	secreted proteins and hormone signaling
TSKU	tsukushi, small leucine rich proteoglycan	218245_at	0	0	↓	0	17	0	secreted protein; intracellular transport and extracellular secretion	
FGF21	fibroblast growth factor 21	221433_at	0	0	0	0	28	2	secreted growth factor; mitosis and survival	growth factors
FIBIN	fin bud initiation factor homolog	226769_at	0	↑	0	0	13	0	secreted growth factor; function in limb development	

ASCL1	achaete-scute family bHLH transcription factor 1	209988_s_at	0	0	0	↓	1	0	transcription factor; neuronal differentiation and development	**transcription factors**
CEBPA	CAAT/enhancer binding protein (C/EBP), alpha	204039_at	0	↓	↑	0	17	0	transcription factor; differentiation of granulocytes and myeloid cells, inhibition of proliferation	
IRF6	interferon regulatory factor 6	1552478_a_at	0	0	0	↑	4	0	transcription factor; role in proliferation and differentiation	

Category: Energy and lipid metabolism

ACSS2	acyl-CoA synthetase short-chain family member 2	234312_s_at	↓	↓	0	↑	6	0	cytoplasmic protein; activation of acetate for use in lipid synthesis and energy generation	lipid synthesis
PLA1A	phospholipase A1 member A	219584_at	0	↓	0	0	8	0	secreted enzyme, hydrolyzation of phospholipids into fatty acids	
G6PD	glucose-6-phosphate dehydrogenase	202275_at	0	0	↑	0	7	0	enzyme in pentose phosphate pathway	fatty acid synthesis
PPM1L	protein phosphatase, Mg2+/Mn2+ dependent, 1L	228108_at 229506_at	0	0	0	↑	23	0	membrane bound enzyme; regulation of blood-glucose	glucose regulation
INSIG1	insulin induced gene 1	201625_s_at 201627_s_at	0	↓	0	↑	12	0	ER membrane protein; control of cholesterol synthesis, may play a role in growth and differentiation of tissues involved in metabolic control	**cholesterol and triglyceride synthesis**
THRSP	thyroid hormone responsive SPOT 14	229476_s_at 229477_at 1553583_a_at	0	↓	↓	↑	21	8	cytosolic protein; involved in lipogenesis and biosynthesis of triglycerides	

Category: Cytoskeletal organization and cell cycle

KIF5C	kinesin family member 5C	203130_s_at	0	0	0	↑	1	0	microtubule motor protein; intracellular transport of organelles	microtubule motor and stabilization proteins
TUBB2B	tubulin, beta 2B class lib	214023_x_at	0	↑	0	0	5	0	major constituent of microtubules; mitosis and intra-cellular transport	
SRPX	sushi-repeat containing protein, X-linked	204955_at	0	↑	↓	↓	13	0	secreted cell surface protein; involved in cell migration and adhesion	migration and adhesion
PLXNC1	plexin C1	213241_at	↑	0	↑	↑	4	0	receptor for semaphorines which are involved in cytoskeletal rearangement, signal transduction, cell adhesion, immune response	cytoskeletal rearangement and adhesion
TEX14	testis expressed 14	221035_s_at	0	0	0	0	2	0	cytoplasmic protein; role in mitosis and meiosis, formation of cell junctions	junction formation
ATF3	activating transcription factor3	202672_s_at 1554420_at	0	0	↓	0	35	0	transcription factor; stress response, further involved in cell cycle regulation, DNA repair, apoptosis	cell cycle arrest
RGCC	regulator of cell cycle	218723_s_at	↑	↑	0	↓	25	0	cytosolic protein; induced by p53 modulates the activity of cell cycle specific kinases in response to DNA damage	

Category: Transport

SLC16A14	solute carrier family 16, member 14	238029_s_at	0	0	0	↓	5	0	membrane protein for mono-carboxylate transport	ion and small molecule transport
SLCO4C1	solute carrier organic anion transporter family, member 4C1	222071_s_at	↓	0	↓	0	5	1	membrane protein; organic anion transporter	
BSPRY	B-box and SPRY domain containing	222746_s_at	0	0	0	0	1	0	membrane protein; involved in calcium transport	

STXBP1	syntaxin binding protein 1	202260_s_at	0	0	↑	0	6	0	cytoplasmic protein; vesicular trafficking	vesicular trafficking
BLZF1	basic leucine zipper nuclear factor 1	210462_at	0	0	0	0	10	0	protein in the Golgi lumen; protein transport from the ER through the Golgi apparatus to the cell surface	protein and RNA transport

Category: Protein stabilization and degradation

CCT4	chaperonin containing TCP1, subunit 4 (delta)	227171_at	0	↑	↑	↑	0	0	chaperone; function in protein stabilization and folding	protein stabilization and transport
HSPA6	heat shock 70kDa protein 6	117_at 213418_at	0	0	0	↓	16	4	chaperone; function in protein stabilization and folding	
SCG5	secretogranin V	203889_at	0	↓	0	0	12	0	secreted chaperone involved in intracellular protein transport	
FBXL16	F-box and leucine-rich repeat protein 16	227641_at	0	0	↑	0	1	0	protein-ubiquitin ligase, ubiquitination and proteasomal degradation	protein degradation
FBXO32	F-box protein 32	225803_at	0	0	↑	0	29	0	cytosolic protein; ubiquitination and proteasomal degradation	
KLHL24	kelch-like family member 24	221986_s_at	0	↓	0	↓	30	0	cytosolic protein; role in protein degradation	

Category: Apoptosis and ER stress response

FAM129A	family with sequence similarity 129, member A	217967_s_at	↑	↑	0	0	10	2	cytosolic protein; involved in apoptosis, survival and ER stress response
BEX2	brain expressed X-linked 2	224367_at	0	0	↑	↑	13	0	cytosolic protein; regulator of mitochondrial apoptosis and cell cycle

Category: Other

ANKRD33	ankyrin repeat domain 33	242209_at	0	0	↑	0	14	0	protein motif, ankyrin repeat proteins are composed of tandem repeats of a basic structural motif
BTBD11	BTB (POZ) domain containing 11	238692_at	0	0	0	0	2	0	membrane protein
HMOX1	heme oxygenase 1	203665_at	0	↑	0	↓	14	0	oxygenase for heme degradation; functions in apoptosis and vascularization
KNG1	kininogen 1	217512_at	0	↓	↓	↓	3	0	secreted protein; gene product can be processed to 2 isoforms; high molecular weight form is involved in blood coagulation, inflammatory response
LAMP3	lysosomal-associated membrane protein 3	205569_at	0	0	0	0	3	0	lysosomal membrane protein; role in dendritic cell function and in adaptive immunity
NPTX2	neuronal pentraxin II	213479_at	0	0	0	0	8	0	secreted protein with biochemical properties of a Ca-depen-dent lectin; modifies properties underlying longterm plasticity
PKIB	protein kinase inhibitor beta	231120_x_at	0	0	↑	↑	6	7	inhibitor of cAMP-dependent protein kinase activity
PSG9	pregnancy specific beta-1-glycoprotein 9	209594_x_at	0	0	0	0	4	0	secreted, pregnancy related signaling protein
RUSC1-AS1	RUSC1 antisense RNA 1	230256_at	0	0	↑	↑	4	5	
TRIM73	tripartite motif containing 73	1554250_s_at	0	0	0	0	3	0	cytosolic ubiquitin ligase
TTC9	tetratricopeptide repeat domain 9	213172_at	0	↑	↑	0	2	0	protein containing three tetratricopeptide repeats, gene is hormonally regulated
ZCCHC6	zinc finger, CCHC domain containing 6	242776_at	↑	↓	↓	0	23	0	enzyme involved in RNA processing

[1]200800_s_at (14/0); [1]202581_at (19/0); [1]205122_at (2/0); [1]208180_s_at (6/0); [1]208575_at (2/0); [1]214469_at (12/0); [1]215078_at (14/4); [1]215779_s_at (9/0); [1]227062_at (13/0); [1]235102_x_at (5/0); [1]236542_at (21/0); [1]237031_at (37/0); [1]242981_at (14/0); [1]243489_at (7/0); [1]242313_at (0/2)

The selection of potential biomarker genes was based on the following criteria: (i) To cover a broad set of chemicals, the biomarker candidate gene should be strongly altered by many compounds. (ii) The gene should also be deregulated in human liver diseases. The overlap with human liver disease genes implies relevance for the gene *in vivo* which makes it less probable that the chemically-induced response is just an *in vitro* artifact. (iii) The gene is not altered in the same direction of expression alteration by the culture conditions or the isolation procedure and (iv) the most relevant toxic mechanisms should be covered. Seven biomarker candidate genes were selected for further analysis (Table 3.19). Among these genes are the aforementioned SV 20 genes CYP3A7, RGCC, SULT1C2 and the stress response protein, Fbxo32 that polyubiquitinates proteins for proteasomal degradation (Cleveland and Evenhuis 2010). In addition, CYP 1B1 was included for further analysis, another isoenzyme of the Cytochrome P450 family. This enzyme metabolizes a variety of environmental and xenobiotic toxicants and is transcriptionally regulated via the aryl hydrocarbon receptor. However, it shows low expression in healthy human liver (Beedanagari et al. 2010; Bhattacharyya et al. 1995). From the category 'energy and lipid metabolism,' the enzyme glucose-6-phosphate dehydrogenase (G6PD) was selected as a potential biomarker. It catalyzes the rate-limiting step of the oxidative pentose-phosphate pathway and provides reducing power (NADPH) and pentose phosphates for fatty acid and nucleic acid synthesis. Interestingly, G6PD expression levels were found to be directly correlated to hepatoma cell migration and invasion (Hu et al. 2014).

Table 3.19: Potential biomarker candidate genes for further analysis. The selected genes cover a wide range of biological functions, are up regulated in different human liver diseases and are not induced due to cultivation conditions or the isolation procedure (CS).

Symbol	Liver disease	CS	SV ↑ (FC3)	SV ↓ (FC3)	Gene Function	Category
CYP1B1	up	0	18	1	metabolic enzyme in the ER (phase I enzyme)	Metabolism Xenobiotics
CYP3A7	up	down	39	0		
SULT1C2	up	0	22	1	cytosolic enzyme; catalyzes sulfonation (phase II enzyme)	
G6PD	up	0	7	0	enzyme in pentose phosphate pathway → fatty acid synthesis	Energy and Lipid metabolism
TUBB2B	up	0	5	0	major constituent of microtubules; functions in mitosis and intracellular transport	Cytoskeleton Cell cycle
RGCC	up	down	25	0	cytosolic protein; induced by p53 modulates the activity of cell cycle specific kinases in response to DNA damage	
FBXO32	up	0	29	0	cytosolic protein; ubiquitination and proteasomal degradation	Protein degradation

3.2.3 Prediction of hepatotoxic blood concentrations *in vivo*

Based on the available biomarkers so far, an *in vitro* system was established to predict human hepatotoxic blood concentrations. HepG2 cells and cultivated primary human hepatocytes were exposed to the two sets of chemicals for 24 h for the defined concentration range. In both cell systems, two read outs were considered: First, the expression of the selected biomarkers was analyzed. Ideally, biomarkers differentiate between hepatotoxic and non-hepatotoxic compounds at therapeutic concentrations. However, not all of the compounds may be covered with the selected biomarker genes, but may exhibit other cytotoxic effects. For this reason, cytotoxicity tests were included as a second readout to identify the lowest cytotoxic concentration *in vitro*. Based on cytotoxicity and biomarker induction, the lowest observed effect concentration *in vitro* was determined and compared to peak plasma concentrations of therapeutic doses.

3.2.3.1 Prediction of hepatotoxic blood concentrations in HepG2 cells

For the biomarker analysis, gene expression in compound-exposed hepatocytes was quantified by qRT-PCR. Relative expression values of the analyzed genes in HepG2 cells and primary human hepatocytes for all compounds are listed in detail in the Supplemental Table 8 and Supplemental Table 9. For each compound, the lowest concentration of biomarker induction was identified. This *in vitro* alert concentration was defined as the lowest concentration that causes a significant increase of at least 2.5 fold induction of at least one of the marker genes. Figure 3.17, as an example, shows how valproic acid alters the biomarker expression in HepG2 cells. Corresponding diagrams for all compounds are given in Supplemental Figure 4 and Supplemental Figure 5. The fold changes represent mean values of three independent experiments. The lowest alert concentration for valproic acid in HepG2 cells is 1mM. Five of the selected marker genes respond at similar, but not identical concentrations.

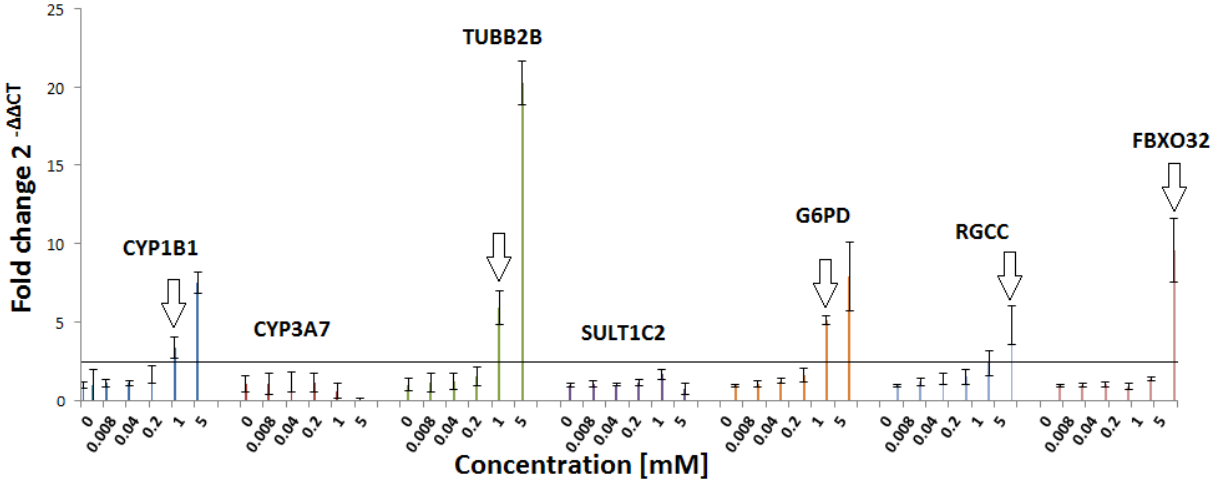


Figure 3.17: Valproic acid induced biomarker expression in HepG2 cells. A gene was considered to be up regulated when crossing the threshold line, which illustrates a significant increase of at least 2.5 fold change induction. Arrows indicate the lowest alert concentration for biomarker induction in vitro. The error bars illustrate the standard deviation of three independent experiments.

The lowest concentration of biomarker induction was determined for each compound and compared to peak plasma concentrations of therapeutic doses *in vivo*. Figure 3.18 illustrates to which degree the biomarker-based *in vitro* system predicts human hepatotoxic blood concentrations. Red symbols represent compounds which are associated with an increased risk of hepatotoxicity when administered at therapeutic doses, whereas green symbols show compounds that are considered harmless. Already in HepG2 cells the two groups of compounds cluster mostly apart from each other. Each compound was tested in three independent experiments. The x-axis shows the lowest concentration of biomarker induction *in vitro*; dots and dashed lines represent differences in the lowest alert concentration of three independent experiments with cells from different donors. The median of the three replicates is highlighted by enlarged symbols. On the y-axis, the peak plasma concentration of a therapeutic dose *in vivo* is shown and for each compound a concentration range is given. The highlighted concentration represents the mean value of peak plasma concentrations that were identified in different studies. The diagonal line indicates identical concentrations of the lowest biomarker inducing concentration *in vitro* and the peak plasma concentration which has a therapeutic effect *in vivo*. If a compound is located on this '*in vivo* line', at least one of the selected marker genes was induced at a concentration, which corresponds to a dose with a therapeutic effect *in vivo*. Compounds which cluster below the line induced the biomarkers at concentrations which are higher than a therapeutic dose *in vivo*, meaning the genes were induced at a cytotoxic concentration only. Compounds which cluster at the very right did not induce any of the marker genes within the tested concentration range. Ideally, compounds that are associated with a high risk of hepatotoxicity at therapeutic doses cluster close to the '*in vivo* line', whereas harmless compounds cluster to the lower right. In HepG2 cells, the two groups of compounds cluster mostly apart from each other, but the majority of hepatotoxic compounds cluster below the line, indicating that hepatotoxic effects *in vitro* are observed at concentrations which are higher than critical concentrations *in vivo*. However, the hepatotoxic compounds cluster closer to the line than the non-hepatotoxic compounds.

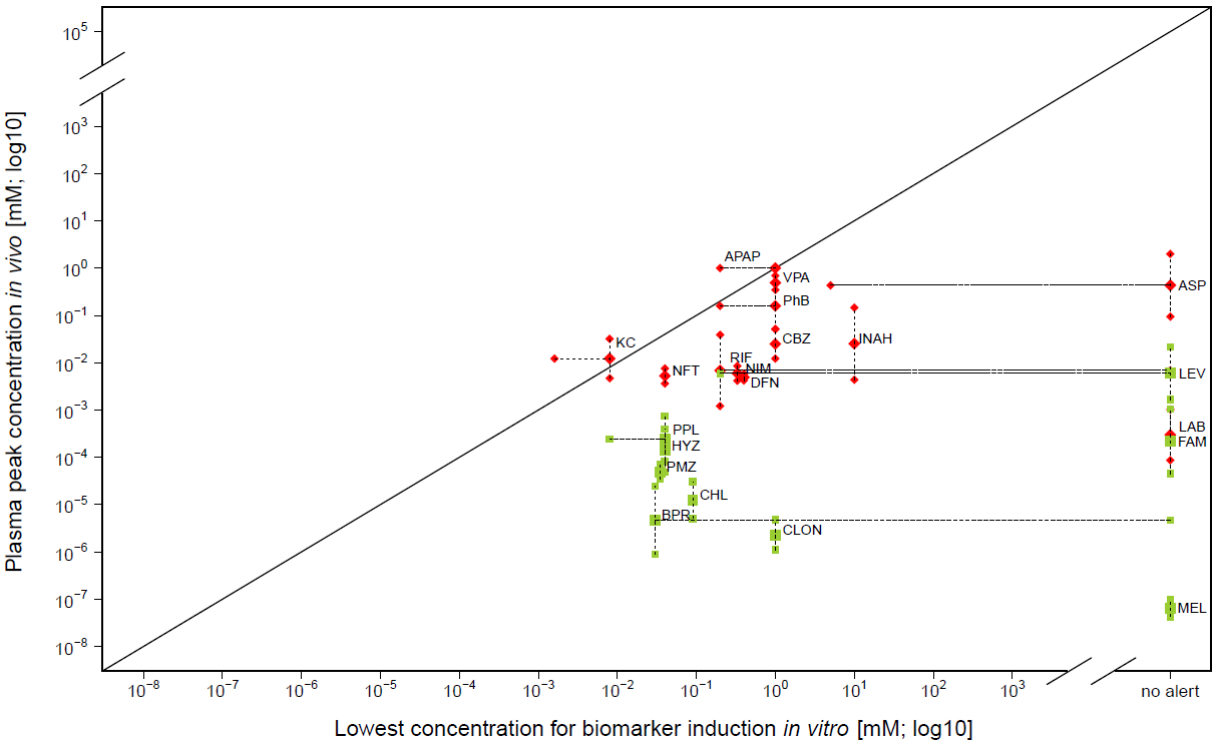


Figure 3.18: The lowest alert concentrations of biomarker induction in exposed HepG2 cells are shown in relation to peak plasma concentrations that have a therapeutic effect *in vivo*. In red: compounds which are associated with increased risk of hepatotoxicity at therapeutic doses; in green: non-hepatotoxic compounds. The x-axis shows the lowest concentration of biomarker induction *in vitro* whereas the y-axis gives the peak plasma concentration of therapeutic doses in vivo. The peak plasma concentration of each compound is shown as a concentration range. Values on the x-axis represent lowest alert concentrations of 3 individual experiments for each compound; the median is highlighted by enlarged symbols. The line indicates identical concentrations of the biomarker inducing concentration and the in vivo relevant concentration.

With a few exceptions, all hepatotoxic compounds cluster close to the '*in vivo* line'. These compounds induced the biomarker genes *in vitro* at therapeutic or close to therapeutic concentrations that are associated with a high risk of hepatotoxicity *in vivo*. In contrast, all non-hepatotoxic compounds either did not up regulate these genes or induced them at concentrations which lie far above critical concentrations *in vivo*.

A close correlation is observed for the hepatotoxic compounds valproic acid (VPA), ketoconazole (KC), phenylbutazone (PhB) and acetaminophen (APAP). These compounds show hepatotoxic effects *in vitro* at concentrations which correspond to peak plasma concentrations of critical concentrations *in vivo*. Thus, the biomarker based *in vitro* model is able to precisely predict human hepatotoxic blood concentrations for these compounds. Acetaminophen is the only compound among the set of hepatotoxic chemicals which is not associated with hepatotoxicity at therapeutic doses. For this reason, *in vitro* alert concentrations were not compared to a dose with a therapeutic effect, but to the well documented toxic plasma concentration of 1mM.

Isoniazid (INAH) clusters relatively far away from the other hepatotoxic compounds; a significant induction of the biomarkers in HepG2 cells was observed only at a concentration that is a factor of about 250 fold higher than the dose with a high risk of hepatotoxicity *in vivo*.

However, isoniazid metabolism requires N-acetyl transferase 2, an enzyme which is only marginally expressed in HepG2 cells (Husain et al. 2007; Saukkonen et al. 2006). Due to the reduced metabolic capacity of this cell line, isoniazid-induced hepatotoxicity is probably attenuated in these cells. Labetalol (LAB) and aspirin (ASP) represent outliers that cluster outside the group of hepatotoxic compounds, because none of the analyzed biomarkers was induced up to the tested concentration range. In contrast to hepatotoxic compounds, none of the biomarker genes were induced at therapeutic doses of non-hepatotoxic drugs. For non-hepatotoxic compounds the predicted blood concentration of hepatotoxicity corresponds to cytotoxic concentrations only. Ideally, the distance between *in vitro* alert concentrations of hepatotoxic compounds and non-hepatotoxic compounds should be larger (Figure 3.18). Nevertheless, non-hepatotoxic chemicals cluster apart from the '*in vivo* line' because the factor between an *in vivo* relevant dose and a cytotoxic concentration is much higher than for compounds which are associated with an increased risk of hepatotoxicity. In general, the selected biomarkergenes seem to be suitable to predict hepatotoxic blood concentrations in HepG2 cells. Keeping in mind that the genes were selected based on data of cultivated primary human hepatocytes the result obtained in compound-exposed HepG2 cells is very promising. With a few exceptions, the set of biomarkers differentiates between hepatotoxic and non-hepatotoxic compounds at therapeutic doses. However, not all compounds were captured with the available biomarkers so far, such as labetalol (LAB) and aspirin (ASP). Other compounds, such as isoniazid (INAH), diclofenac (DFN), nimesulide (NIM) and nitrofurantoin (NFT) did not induce the biomarker genes at therapeutic, but at slightly cytotoxic concentrations only, although these drugs exhibit also a high risk of hepatotoxicity in the range therapeutic doses.

To optimize the prediction of human hepatotoxic blood concentrations, cytotoxicity tests were performed as a second readout. For all compounds, alert concentrations of cytotoxicity *in vitro* were identified by using the Cell Titer Blue cytotoxicity test. HepG2 cells were exposed for 48 h and dose response curves were compiled to identify the lowest cytotoxic concentration for each compound *in vitro* (Supplemental Figure 6). The lowest cytotoxic concentration was defined as 20% loss of viability after 48h of exposure. The clustering of the two groups of compounds based on cytotoxicity is shown in Figure 3.19. Again, the *in vivo* relevant concentration for acetaminophen (APAP) corresponds to the toxic blood concentration of 1mM.

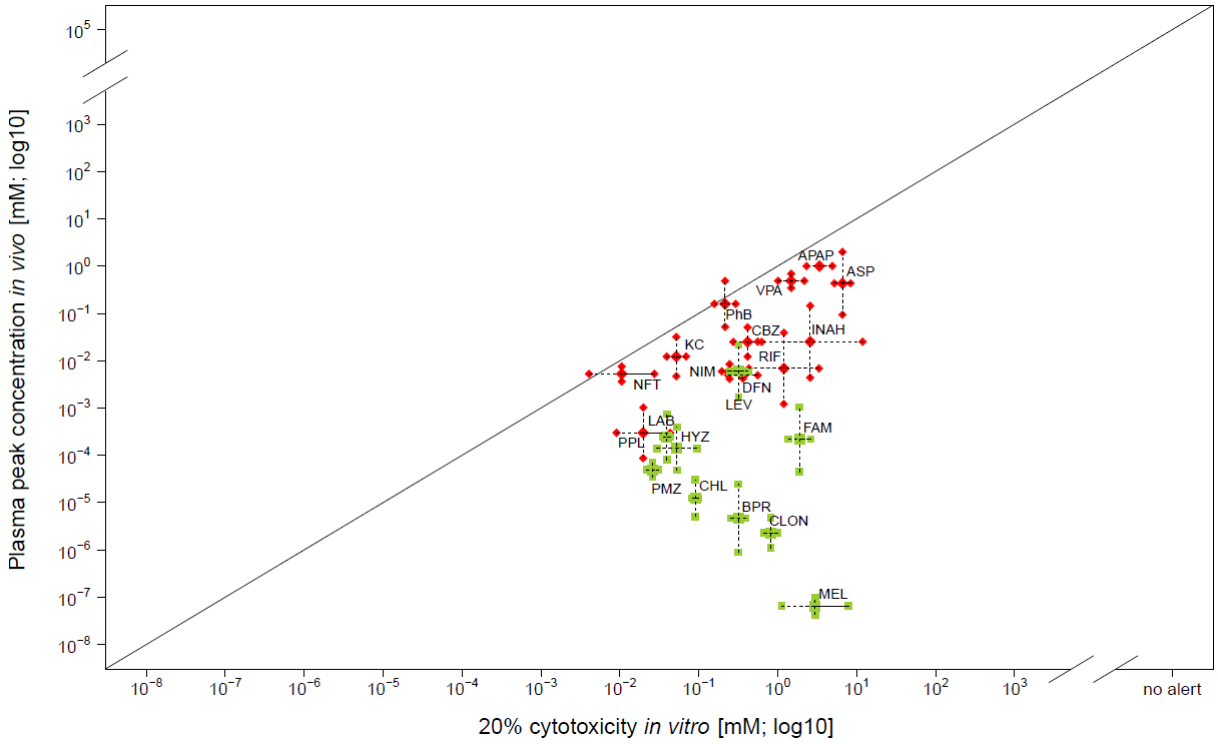


Figure 3.19: Prediction of toxic blood concentrations in HepG2 cells. Based on cell titer blue cytotoxicity data, the lowest cytotoxic concentration, representing 20 % loss of cell viability after 48 h of compound exposure, was determined. In red: compounds that are associated with increased risk of hepatotoxicity at therapeutic doses; in green: non-hepatotoxic compounds. Each compound was tested in 3 individual experiments. The x-axis shows the concentration at which the viability decreased by 20 % (IC20) *in vitro* whereas the y-axis gives the peak plasma concentration of therapeutic doses *in vivo*. The peak plasma concentration of each compound is shown as a concentration range. Values on the x-axis represent the IC20 values of 3 individual experiments, the mean value by enlarged symbols with the estimated confidence intervals. The line indicates identical concentrations of the IC20 *in vitro* and the therapeutic range *in vivo*.

The line represents identical concentrations of doses with a therapeutic effect *in vivo* and 20 % loss of viability *in vitro*. Ideally, compounds which are associated with a high risk of hepatotoxicity at therapeutic doses cluster close to the line. This is, for example, the case for valproic acid (VPA), nitrofurantoin (NFT) and phenylbutazone (PhB). The three compounds show cytotoxic effects *in vitro* at concentrations which correspond to critical concentrations *in vivo*, indicating that the cytotoxicity-based *in vitro* system precisely predicts hepatotoxic blood concentrations.

The majority of hepatotoxic compounds cluster close to the line, whereas the non-hepatotoxic compounds cluster apart. A clear cluster formation between the two groups of compounds was not observed. A few candidates, such as levofloxacin (LEV), labetalol (LAB) or propranolol (PPL) overlap among the two clusters. However, for some compounds, the cytotoxicity-based prediction system improves. Carbamazepine (CBZ), phenylbutazone (PhB), isoniazid (INAH), aspirin (ASP) and labetalol (LAB) of the hepatotoxic compounds show lower effect concentrations *in vitro* and shift closer to the '*in vivo* line'. Aspirin and labetalol were not captured by the biomarker analysis alone, but based on cytotoxicity data they cluster closer to the other hepatotoxic compounds. Similarly, the system was more sensitive for the

non-hepatotoxic compounds levofloxacin (LEV), famotidine (FAM), buspirone (BPR), melatonin (MEL) and chlorpheniramine (CHL). These compounds induced cytotoxic effects at lower concentrations at which the biomarkers were not induced, but these cytotoxic concentrations are still far away from doses with a therapeutic effect. In contrast, the lowest cytotoxic concentration *in vitro* of valproic acid (VPA), ketoconazole (KC) and acetaminophen (APAP) is much higher than the biomarker-inducing concentration. Thus, the biomarker-based prediction system is more sensitive for these hepatotoxic compounds.

In summary, cytotoxicity data alone is not sufficient to predict human hepatotoxic blood concentrations. The two prediction systems based on biomarker expression and cytotoxicity data exhibit individual advantages for a subset of compounds. Some compounds show lower alert concentrations based on biomarker expression, others based on cytotoxicity data. To capture all compounds and to identify the lowest hepatotoxic concentration *in vitro*, the two readouts were combined. To improve the sensitivity of the prediction system, the *in vitro* alert concentration from the readout that gives the lower concentration was considered for further analysis. Based on either cytotoxicity data or biomarker induction, the lowest observed effect concentration (LOEC) *in vitro* was identified. The LOEC *in vitro* for all compounds in relation to peak plasma concentrations of a therapeutic dose *in vivo* is presented in Figure 3.20. The combination of the two readouts distinctly improves the prediction quality of the system. Hepatotoxic compounds cluster closer to the line, indicating that hepatotoxic effects *in vitro* are observed at concentrations that are close to critical concentrations *in vivo*. Thus, distance between the two sets of compounds increases and a better clustering is achieved.

A few candidates, such as levofloxacin (LEV), propranolol (PPL), labetalol (LAB) and hydroxyzine (HYZ) remain in the overlap of the two clusters. For labetalol, the available biomarkers so far, as well as the cytotoxicity data, were not sufficient to predict hepatotoxicity in the range of therapeutic doses. Follow-up studies will focus on the data driven identification of further biomarkers, which allow a more precise prediction and a clustering of labetalol within the cluster of hepatotoxic compounds. However, the predominant clustering of the two groups of compounds demonstrates that the biomarker-based system with HepG2 cells represents already a promising tool to predict hepatotoxicity *in vivo*.

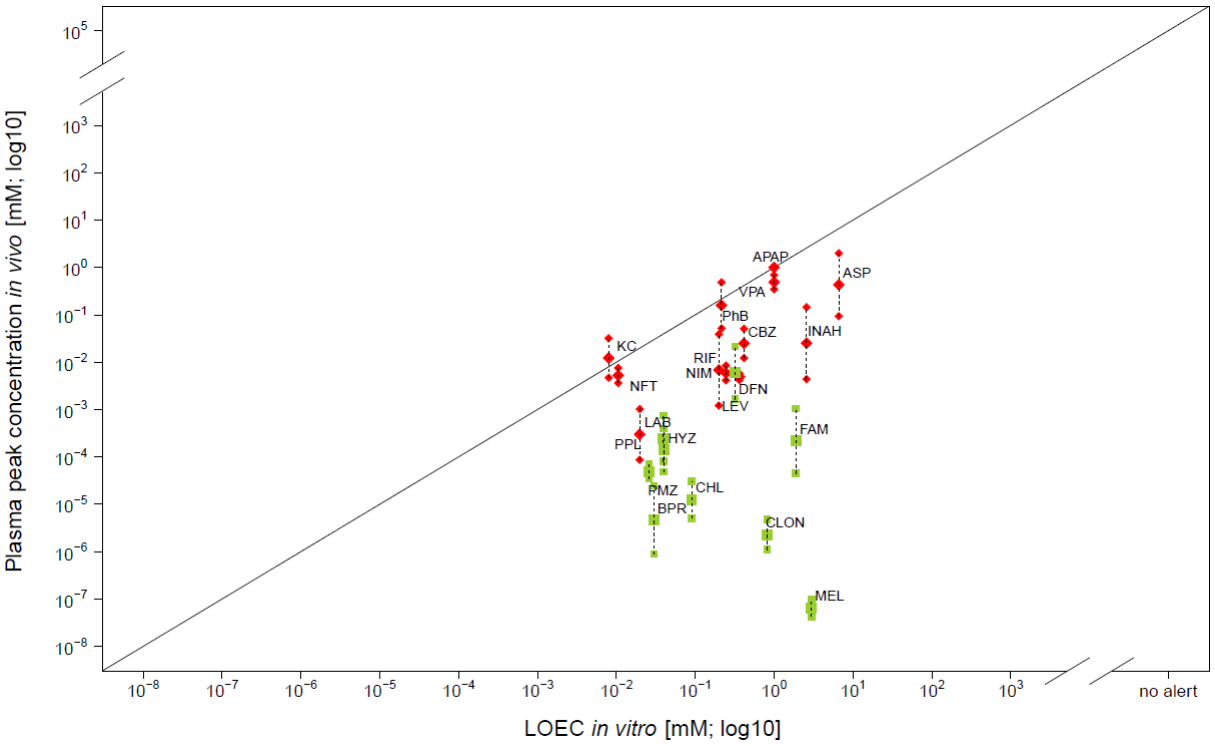


Figure 3.20: Prediction of hepatotoxic blood concentrations based on the lowest alert concentration (LOEC) in HepG2 cells. The LOEC corresponds either to the concentration where at least one of the selected marker genes was induced or to the concentration which was associated with a 20 % decrease of cell viability after 48 h of compound exposure. In red: compounds associated with high risk for hepatotoxicity. In green: non-hepatotoxic compounds. The line indicates identical concentration of the LOEC *in vitro* and the peak plasma concentration *in vivo*.

3.2.3.2 Prediction of hepatotoxic blood concentrations in primary human hepatocytes

In a next step, primary human hepatocytes were used for the biomarker-based *in vitro* model to predict human hepatotoxic blood concentrations. In contrast to the strongly dedifferentiated HepG2 cells, the metabolic capacity in primary hepatocytes is predominantly maintained. Therefore, primary human hepatocytes have become the "gold standard" for evaluating hepatotoxicity of drugs.

The expression of the selected biomarker genes was analyzed in compound-exposed primary human hepatocytes and for each compound the lowest concentration of biomarker induction was identified. This alert concentration *in vitro* represents a significant increase of at least 2.5 fold change induction of at least one biomarker gene. Primilary results with cells from 1-3 individual donors revealed to which degree the biomarker-based *in vitro* system with primary human hepatocytes is able to predict human hepatotoxic blood concentrations (**Figure 3.21**).

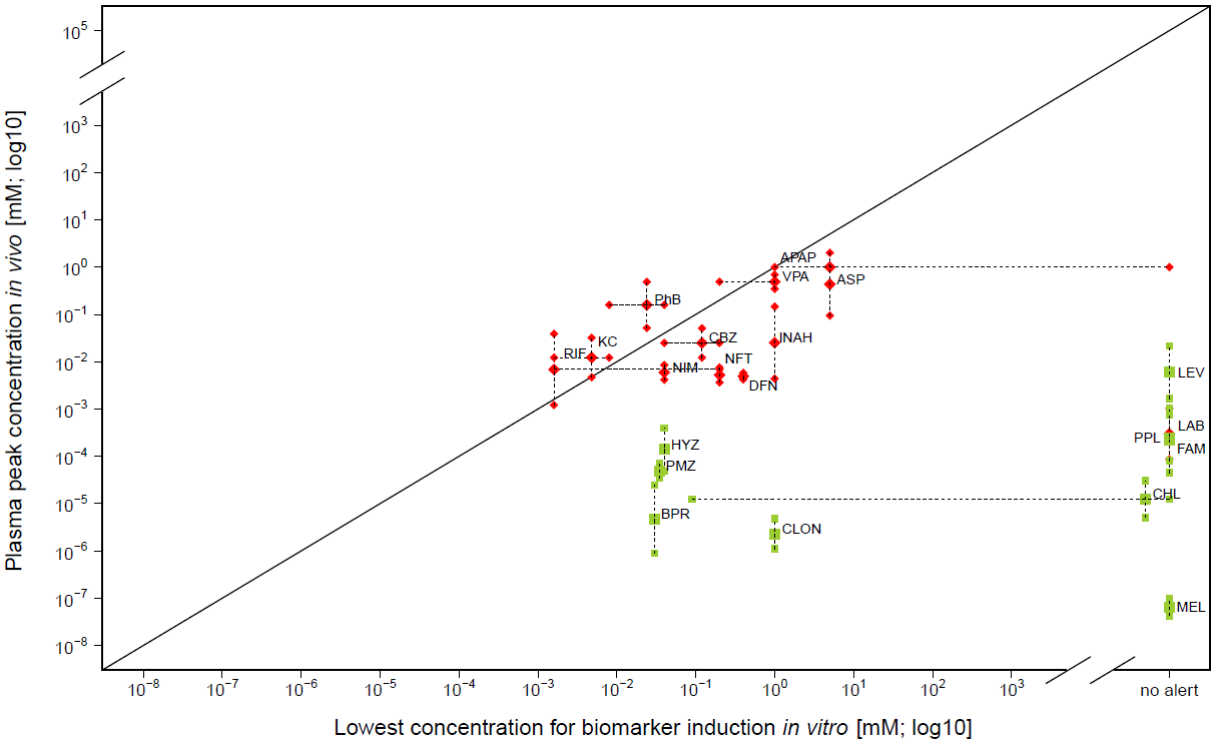


Figure 3.21: The lowest alert concentrations of biomarker induction in exposed primary human hepatocytes are shown in relation to peak plasma concentrations of therapeutic doses *in vivo*. In red: compounds which are associated with increased risk of hepatotoxicity at therapeutic doses; in green: non-hepatotoxic compounds. The x-axis shows the lowest concentration of biomarker induction *in vitro* whereas the y-axis gives the peak plasma concentration of therapeutic doses *in vivo*. The peak plasma concentration of each compound is shown as a concentration range. Values on the x-axis represent lowest alert concentrations of 1-3 individual experiments with cells from different donors; the median is highlighted in enlarged symbols. The line indicates identical concentrations of the biomarker inducing concentration and the *in vivo* relevant concentration. (n=1 for NIM, LAB, PPL, HYZ and MEL, n=2 for ASP, DFN, KC, NFT, PhB, CBZ, INAH, BPR, FAM, PMZ, CHL, CLON and LEV, n=3 for APAP, RIF and VPA).

First results show that primary human hepatocytes represent a more sensitive *in vitro* system because the compounds show hepatotoxic effects at lower concentrations. The majority of compounds cluster closer to the '*in vivo* line' than it was observed in HepG2 cells. The biomarkers of hepatotoxicity were induced at concentrations that have a therapeutic effect. In this *in vitro* model, aspirin (ASP) and isoniazid (INAH) were also captured with the selected biomarkers and therefore cluster together with the other hepatotoxic compounds. Labetalol (LAB) still represents an outlier and is not captured with the set of biomarkers. Therefore, more suitable biomarkers have to be identified.

As a second read out, Cell Titer Blue cytotoxicity tests were performed in primary human hepatocytes. The lowest cytotoxic concentration, representing 20 % loss of viability, was determined for each compound, in order to assess whether a clustering of the two sets of compounds is also possible based on cytotoxicity data (Supplemental Figure 7). So far, only a subset of compounds was tested in cells from 1-2 different donors, including acetaminophen (APAP), ketoconazole (KC), nitrofurantoin (NFT), diclofenac (DFN), rifampicin (RIF), valproic acid (VPA), nimesulide (NIM), labetalol (LAB), hydroxyzine (HYZ), famotidine (FAM), buspirone (BPR) and melatonin (MEL) (Figure 3.22). The lowest cytotoxic concentration in cultivated primary hepatocytes was determined and compared to the concentration, which has a

therapeutic effect *in vivo*. In primary human hepatocytes, a clear separation and clustering of the two sets of compounds based on cytotoxicity data alone is not achieved, as well. Hepatotoxic compounds cluster closer to the '*in vivo* line' than non-hepatotoxic compounds and show cytotoxic effects at concentrations that are close to therapeutic doses. The prediction efficiency improves for nitrofurantoin (NFT), labetalol (LAB), diclofenac (DFN) and acetaminophen (APAP). For these compounds, cytotoxic alert concentrations *in vitro* were observed at concentrations that are lower than biomarker-inducing concentrations. Similarly, the cytotoxicity-based prediction system is more sensitive for the non-hepatotoxic compounds famotidine (FAM), propranolol (PPL) and melatonin (MEL). A risk of hepatotoxicity is detected at lower concentrations *in vitro*, but these alert concentrations are still much higher than concentrations that have a therapeutic effect *in vivo*.

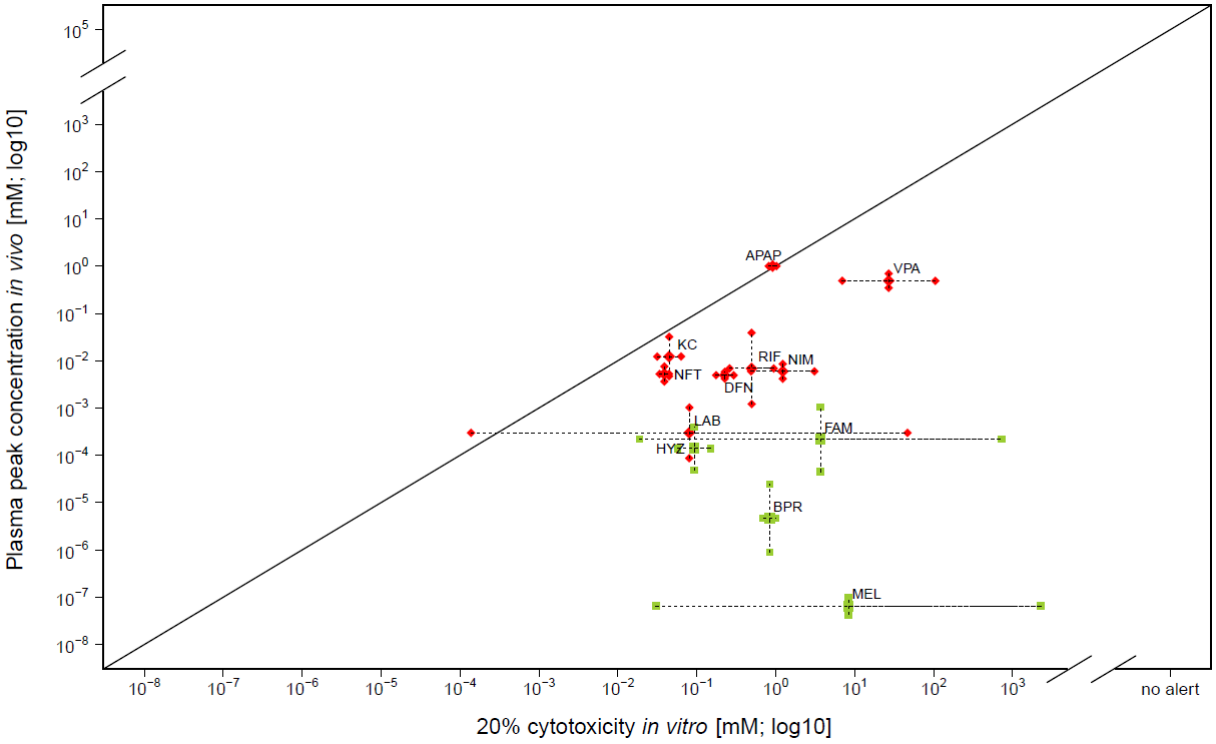


Figure 3.22: Prediction of toxic blood concentrations in primary human hepatocytes. Based on cell titer blue cytotoxicity data, the lowest cytotoxic concentration, representing 20 % decrease of cell viability after 48 h of compound exposure, was determined. In red: compounds which are associated with increased risk of hepatotoxicity at therapeutic doses; in green: non-hepatotoxic compounds. The x-axis shows the concentration of 20% loss of viability *in vitro* whereas the y-axis gives the peak plasma concentration of therapeutic doses *in vivo*. The peak plasma concentration of each compound is shown as a concentration range. Values on the x-axis represent the IC20 values of 1-2 individual experiments, the mean value in enlarged symbols with the estimated confidence interval. The line indicates identical concentrations of the IC20 *in vitro* and the therapeutic range *in vivo*.

In contrast to the biomarker-based prediction model, the cytotoxicity-based system is less sensitive for rifampicin (RIF), ketoconazole (KC), acetaminophen (APAP) and valproic acid (VPA). These compounds showed no cytotoxic effects *in vitro* at concentrations that resemble critical concentrations *in vivo*. Cytotoxic effects occur at higher concentrations only

which are already in a slightly cytotoxic range. Similarly as observed in HepG2 cells, the prediction system becomes more sensitive when the two readouts were combined, both biomarker expression and cytotoxicity assays (Figure 3.23). With the lowest observed effect concentrations (LOECs), hepatotoxic compounds cluster closer to the '*in vivo* line' and the prediction efficiency improves. The majority of hepatotoxic compounds showed hepatotoxic effects *in vitro* at concentrations that correspond to *in vivo* relevant concentrations with an increased risk of hepatotoxicity. For these compounds, the test system predicts a risk of hepatotoxicity *in vivo*. However, the data represent preliminary results, because cytotoxicity data is not yet available for all compounds and biomarker expression, as well as cytotoxicity has not been analyzed in three independent experiments yet. In case of aspirin (ASP), phenylbutazone (PhB), carbamazepine (CBZ), isoniazid (INAH), chlorpheniramine (CHL), clonidine (CLON), levofloxacin (LEV), propranolol (PPL) and promethazine (PMZ) the LOEC corresponds to the biomarker inducing concentration. Levofloxacin is not recognized by the biomarkers up to the tested concentration, therefore it clusters apart.

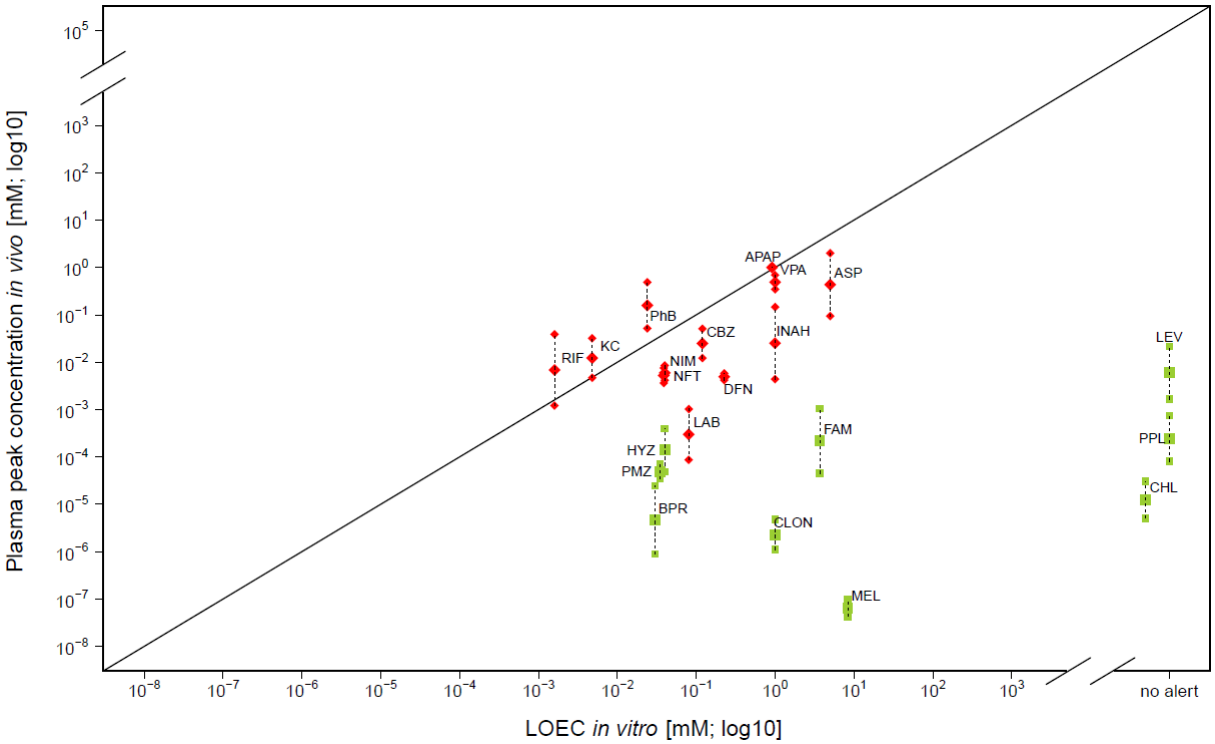


Figure 3.23: Prediction of hepatotoxic blood concentrations based on the lowest alert concentration *in vitro* (LOEC) in primary human hepatocytes. The LOEC corresponds either to the concentration where at least one of the selected marker genes was induced or to the concentration which was associated with loss of 20 % cell viability after 48 h of compound exposure. In red: compounds associated with high risk for hepatotoxicity. In green: non-hepatotoxic compounds. The line indicates identical concentration of the LOEC *in vitro* and the peak plasma concentration *in vivo*. For ASP, PhB, CBZ, INAH, CHL, CLON, LEV, PPL and PMZ cytotoxicity data is still in progress, here the LOEC represent the biomarker inducing concentration.

Compared to HepG2 cells (Figure 3.20), primary human hepatocytes represent the more sensitive *in vitro* model to predict hepatotoxic blood concentrations. The compounds show lower LOECs, indicating that a risk of hepatotoxicity would be detected at lower concentrations. This is especially the case for the compounds ketoconazole (KC), rifampicin (RIF), phe-

nylbutazone (PhB), valproic acid (VPA) and carbamazepine (CBZ), which cluster directly on the '*in vivo* line'. For these compounds, the biomarker- and cytotoxicity-based *in vitro* model represents a promising tool that precisely predicts blood concentrations, which are associated with an increased risk of hepatotoxicity *in vivo*. Similarly, nimesulide (NIM), isoniazid (INAH), aspirin (ASP) and diclofenac (DFN) cluster closer to the line in primary human hepatocytes than in HepG2 cells. The prediction of hepatotoxic blood concentrations for these compounds fluctuates within an error range of factor 5-25. Despite hepatotoxic blood concentrations of non-hepatotoxic compounds are far above therapeutic doses, the sensitivity of the prediction system also improves for some non-hepatotoxic compounds, such as buspirone (BPR) and clonidine (CLON).

Considering biomarker expression and cytotoxicity data individually, the so far available biomarkers capture all hepatotoxic compounds except for labetalol (LAB). A precise prediction of hepatotoxic blood concentration based on the biomarkers alone is achieved for rifampicin, ketoconazole, valproic acid, acetaminophen, nimesulide and carbamazepine, whereas the prediction for the other hepatotoxic compounds fluctuates within an error range of factor 100-1000. For these compounds, the situation improves when including the cytotoxicity data. However, cytotoxicity data alone does not achieve a clear cluster formation of the two groups of compounds. The combination of the two readouts improves the prediction sensitivity for the hepatotoxic, as well as non-hepatotoxic compounds and a risk of hepatotoxicity is detected at lower concentrations. Combining the biomarker and cytotoxicity-based prediction system, the two groups of compounds cluster mainly apart from each other in HepG2 cells (Figure 3.20) and even more so in primary human hepatocytes (Figure 3.23). Even some idiosyncratic hepatotoxic compounds, such as nimesulide, were identified and differentiated from non-hepatotoxic compounds, such as clonidine and buspirone. Although the *in vitro* prediction system presented is still in a developmental stage, preliminary results indicate that both systems are suitable to predict human hepatotoxic blood concentrations, at least within a certain error range.

4 Discussion

4.1 Establishment of a toxicogenomics directory for compound exposed hepatocytes

An overall goal for safety evaluation and human health risk assessment is the prediction of hepatotoxicity *in vivo* based on *in vitro* data. In recent years, numerous research groups have focused on the identification and development of biomarkers of hepatotoxicity, which can be applied in *in vitro* systems. In this context, the use of a genomics approach to identify patterns of changes in mRNA transcripts, referred to as toxicogenomics, has gained in popularity (Yang et al. 2012). Genomic biomarkers may be more reliable and sensitive than conventional morphological or serum markers, making it possible to detect hepatotoxicity at low doses and during the early stages of drug exposure (Pfannkuch et al., 2014). Transcriptomics data from rodent liver and cultivated hepatocytes are frequently used to identify novel candidate genes for further evaluation, and to elucidate the molecular mechanisms of drug-induced liver injury (Mendrick and Schnackenberg 2009; Shi et al. 2010). Several emerging data bases, such as DrugMatrix, diXa and Open TG-GATEs comprise genome wide expression data from *in vivo* animal studies, as well as from cultivated hepatic cell lines and primary hepatocytes, where hundreds of chemicals have been tested (Jiang et al. 2015). However, a comprehensive analysis summarizing the key features of chemically-influenced gene expression has not yet been performed. To establish a systematic strategy for the identification of potential biomarker genes, the first part of this thesis focuses on the definition of key principles of global gene expression alterations, by utilizing genome wide expression data from the Open TG-GATES database. The database comprises Affymetrix files from the analysis of cultivated, primary human hepatocytes that were exposed to 158 chemicals for different time points and different concentrations (Igarashi et al. 2015). Furthermore, microarray data sets investigating global gene expression changes in human liver diseases were acquired from public data repositories. Before analyzing the structure of chemically-induced gene expression alterations, a set of curation steps was performed to improve the reliability of the underlying data. Based on the optimized dataset, a toxicotranscriptomics directory was developed and is now publically available (http://wiki.toxbank.net/toxico-genomics-map/). For each gene, the following questions were addressed:

(i) Are there alterations in gene expression by chemicals, and if yes, which class of compound and how many compounds induced a change in expression?
(ii) Is the gene also altered in human liver diseases, which implies a potential function *in vivo*?
(iii) Is the change in gene expression influenced by the hepatocyte isolation and cultivation procedures?

(iv) What is the function of the gene? Is there available evidence of its involvement in some toxic mechanism?

In the following paragraphs, the stepwise development of the toxicogenomics directory, as well as the strategy for biomarker identification will be discussed.

4.1.1 Stereotypic versus compound specific gene expression alterations and detection of biological motifs

When developing predictive biomarkers of hepatotoxicity derived from gene expression profiles of compound-exposed hepatocytes, it is necessary to know whether the expression of the gene is frequently altered by any kind of hepatocellular stress or whether the observed deregulation can be attributed to a very specific compound. Chemically-induced stress by exposure of cells to compounds at a cytotoxic concentration results in the differential expression of a certain amount of genes as a stereotypical response. This response is independent of the toxic mode of action of the compounds and should be differentiated from compound-specific effects. A first overview of stereotypical versus compound-specific gene expression effects was obtained using an unsupervised cluster analysis, which was performed based on the 100 most deregulated probe sets across all compounds. This analysis identified clusters of genes that are deregulated by many compounds (Figure 3.11). Clusters of up regulated genes were for example, associated with cellular stress or metabolic activity via cytochrome P450 enzymes; whereas, a cluster of proliferation-associated genes was found to be strongly down regulated. Similarly, chemically-induced patterns of gene expression were identified by cluster analysis based on the 100 genes with the highest variability (Krug et al. 2013; Waldmann et al. 2014). In addition to providing an overview of the biological motifs altered by chemical exposure, this kind of analysis helps to identify interesting candidate genes for further evaluation.

In order to systematically distinguish between stereotypic and compound-specific gene expression alterations, the selection value (SV) concept was introduced. For example, SV (x), delivered a ranked list of genes that were at least three fold up or down regulated by at least x compounds. For the identification of stereotypical gene expression alterations, the SV 20 genes were considered. The expression of these genes was at least three fold up or down regulated by at least twenty compounds. Analysis of the proportion of compounds contributing to the deregulation of the 100 most up and down regulated genes revealed that only 32 compounds were responsible for the 100 most up regulated genes and only 23 compounds for the 100 strongest down regulations (Figure 3.5 and Figure 3.6). For this reason, twenty compounds represent a relatively large fraction to define stereotypical gene expression responses. Twenty-two annotated genes (31 probe sets) were identified for the SV 20 up regulated genes (Table 3.7 A). These genes are associated with biological functions such as phase I and II metabolism of xenobiotics, energy and lipid metabolism, development and

differentiation, protein processing, as well as stress response. One hundred and seventy nine probe sets (74 annotated genes) were down regulated upon chemically-induced stress (Table 3.7 B). Most of these genes are associated with cell cycle progression; however, others belonged to biological motifs such as DNA repair and synthesis, immune response, cytoskeleton, metabolism and intracellular trafficking. Since these SV 20 genes were found to be deregulated by many compounds and cover a broad range of biological motifs, they represent interesting candidates for biomarker genes. Although specific toxic mode of actions might not be covered by these genes, they capture a wide spectrum of compounds and are indicative of many types of cellular stress.

In contrast to these stereotypic gene expression alterations, compound-specific effects and defined toxic mechanisms were identified with the SV 3 genes, which are genes showing at least a threefold up or down regulation in expression by at least three chemicals. Although more individual effects could potentially be obtained by considering up or down regulated genes of single compounds (SV 1 genes), the risk of false positive results is relatively high due to the low number of available replicates. A compromise between reliability and individuality is given when considering gene expression alterations observed by at least three compounds. The SV 3 genes can be attributed to various biological functions, such as energy and lipid metabolism, the inflammatory and immune response, development and differentiation, protein modification and degradation, transcriptional regulation, endogenous and xenobiotics metabolism, cytoskeleton, transport, stress response and apoptosis (Table 3.8 A and B). SV 3 genes are included in the list of SV 20 genes, but cover an even more diverse pattern of biological motifs. These genes might be helpful in identifying specific toxic mechanisms that underlie the hepatotoxicity of a particular compound.

4.1.2 Overlap with human liver disease genes

As previously reported one challenge, when extracting biomarker candidates from transcriptomics data, is that compound-induced effects observed in *in vitro* systems are not automatically representative for the situation in the human liver (Kienhui*s et al.* 2009). Therefore, the toxicogenomics directory also considers whether a gene deregulated by chemicals *in vitro*, is also indicative of a disturbed situation in the human liver *in vivo*. For this purpose, genome wide expression data from liver tissue samples of patients suffering from liver diseases such as hepatocellular carcinoma (HCC), cirrhosis or non-alcoholic steatohepatitis (NASH) were used to identify genes that overlap with chemically-induced expression alterations *in vitro*. Such an overlap would indicate that the gene has some *in vivo* relevance, and decreases the probability that the chemically-induced effect observed *in vitro* is just an *in vitro* artifact. The directory also provides information as to whether the gene is altered in the same direction in human liver diseases. SV 20 genes which are altered by many compounds *in vitro* as well as in human liver diseases, represent the most interesting candidates for biomarkers. Overlapping candidates between the two sets of up regulated genes include the regulator of cell cycle (RGCC) and sulfotransferase 1C2 (SULT1C2). The major function of the

latter phase II enzyme is the sulfonation of endo- and xenobiotic compounds, thus facilitating their excretion (Alnouti and Klaassen 2006). SULT1C2 is not expressed in the adult human liver, but is predominantly found in the developing fetus, implicating a role in developmental physiology (Stanley et al. 2005). Therefore, SULT1C2 expression might reactivate a fetal expression pattern as a stereotypical response *in vivo* and *in vitro*. Sulfonation protects against numerous potentially toxic drugs and chemicals; therefore the activation of this detoxification system by the up regulation of SULT1C2 expression may be a direct consequence of chemical exposure (Blanchard et al. 2004; Coughtrie 2002). In addition, sulfonation can also yield unstable electrophilic species, which lead to the formation of protein and DNA adducts, or which bind to other macromolecules, resulting in features that are reminiscent of a diseased liver (Glatt et al. 1998).

RGCC is also up regulated by chemicals *in vitro* and in liver tissue from patients suffering from NASH and cirrhosis. This gene is a direct target of p53 in different human cells and is induced in response to DNA damage (Saigusa et al. 2007). RGCC regulates cell cycle by modulating the expression and activation of cyclins and cyclin dependent kinases (An et al. 2009; Badea et al. 2002). Previous studies reported on its induction by hypoxia, showing that p53 is activated upon any form of cellular stress (An et al. 2009).

A strong overlap was observed for a number of down regulated genes between chemically deregulated SV 20 genes and liver disease genes. This list of overlapping genes comprises for example the aldehyde dehydrogenase (ALDH) family members, ALDH8A1 and ALDH4, phosphoenolpyruvate carboxykinase 1 (PCK1), and carbamoyl phosphate synthase 1 (CPS1). ALDH8A1 encodes a retinal dehydrogenase isoenzyme, which converts all-trans and 9-cis retinol to all-trans and 9-cis retinoic acid (Lin and Napoli 2000). The latter is the major ligand of the retinoic acid receptor (RAR) and retinoic X receptor (RXR). RXR represents one of the key transcription factors in hepatocellular gene expression and plays a pivotal role in the metabolism of xenobiotics *in vivo* (Cai et al. 2002). ALDH8A1 is down regulated in NASH, cirrhosis and HCC, indicating possible pathophysiological relevance. In contrast, the second aldehyde dehydrogenase family member among the down regulated SV 20 and liver disease genes, ALDH4, is a mitochondrial matrix enzyme with various substrates from numerous endogenous and exogenous precursors (Kimura et al. 2009). ALDH4 was reported to be a negative regulator of p53-induced apoptosis (Yoon et al. 2004), and its down regulation in chemically-treated cells, as well as during liver diseases, may indicate that there is a shift in the balance from cell survival to death.

Phosphoenolpyruvate carboxykinase 1 (PCK1) catalyzes the conversion of oxaloacetate to phosphoenol pyruvate, the rate limiting step in gluconeogenesis that produces glucose from lactate and other precursors derived from the citric acid cycle (Beale et al. 2007). A decrease in PCK1 activity due to chemical exposure or liver disease may result in reduced gluconeogenesis. The susceptibility of this gene to cellular stress was, for example, shown in severe sepsis, where PCK1 activity significantly decreased leading to deregulated gluconeogenesis (Feingold et al. 2012). Previous studies have shown that inflammation in several tissues is also associated with decreased PCK1 expression (Feingold et al. 2012).

CPS1 (carbamoyl phosphate synthase 1) is a mitochondrial, liver specific, rate-limiting enzyme in the urea cycle and has a pivotal role in ammonia detoxification (Weerasinghe et al. 2014). The encoding gene is down regulated in liver diseases, as well as by a large number of chemicals. Previous studies reported that CPS1 protein is most readily secreted from stressed hepatocytes, for example, in the case of acetaminophen intoxication, under septic conditions, or from apoptotic hepatocytes (Weerasinghe et al. 2014). Therefore, CPS1 is a potential serum marker for detecting mitochondrial injury of the liver. CPS1 is decreased in stressed hepatocytes, due to both a down regulation at the transcriptional level, as well as secretion of the protein itself. However, a mechanistic link between this outward transfer and liver injury phenotype is not known.

Altered gene expression – either as a stereotypical response to chemicals or in human liver diseases, may serve as markers of an unhealthy liver. These deregulations are not compound specific but can result from numerous forms of liver injury. Approximately 20 % of the chemically-induced gene expression alterations in hepatocytes *in vitro* overlap with the fraction of genes that is deregulated in liver diseases. These genes are promising biomarker candidates, because a certain *in vivo* relevance can be assumed. The remaining 80 % are not altered in liver diseases, and may rather be representative of more chemically-specific mechanisms of toxicity. Further studies will be required to the functional relevance of these genes *in vivo*.

4.1.3 Unstable baseline genes

One limitation when using cultivated primary hepatocytes for toxicogenomics studies is that the hepatocyte isolation and cultivation procedures already induce strong gene expression changes. Genome wide analysis of cultivated primary hepatocytes revealed that the liver cells undergo massive gene expression alterations within the first 24h of cultivation (Zellmer et al. 2010). This “noisy background” may interfere with chemically-induced gene expression alterations. Genes, which are altered in the same direction by the hepatocyte isolation and cultivation procedure, as well as a consequence of chemically-induced stress, should be considered with caution, as they tend to be unsuitable as biomarker candidates. For this reason, the toxicogenomics directory also provides information on whether a gene is deregulated by the hepatocyte isolation and cultivation procedures. These unstable baseline genes may be useful biomarkers, but time-matched untreated controls are then crucial to avoid misinterpretation of the data.

4.2 Application of the toxicogenomics directory: Identification of biomarker candidate genes and their potential to predict human hepatotoxic blood concentrations.

Having established a toxicogenomics directory which summarizes key features of compound-exposed hepatocytes, the second part of this thesis focusses on the identification of predictive biomarkers and on their suitability to predict human hepatotoxicity *in vivo*. In a pilot study, a biomarker-based *in vitro* system was developed which discriminates between hepatotoxic and non-hepatotoxic compounds, and predicts blood concentrations that are associated with an increased risk of hepatotoxicity in humans. The novelty and strength of this study are based on two aspects: First, the prediction model represents a human based test system, which not only allows for toxicity evaluation *in vitro*, but also forges a link to the human *in vivo* situation. The model can also be used to analyze the expression of genes that are induced as a consequence of chemically-induced stress *in vitro*, in addition to those that are deregulated in human liver diseases, such as NASH, cirrhosis and HCC. Second, the model facilitates the prediction of hepatotoxic blood concentrations in humans *in vivo* by comparing critical concentrations *in vitro* to *in vivo* blood concentrations where a therapeutic effect was observed. If a compound induces hepatotoxic effects *in vitro* within the range of *in vivo* therapeutic concentrations, the risk of hepatotoxicity *in vivo* is increased. For the development of the prediction model the following strategy was applied:

1) Identification of predictive biomarkers of toxicity based on the previously described toxicogenomics directory.
2) Selection of two sets of compounds for testing the predictability of the biomarkers: a) set of hepatotoxic compounds, which are associated with an increased risk of hepatotoxicity when administered at therapeutic doses, and b) a set of non-hepatotoxic compounds. Ideally, the selected biomarker genes discriminate between hepatotoxic and non-hepatotoxic compounds.
3) Identification of *in vivo* relevant concentrations. For all compounds, a literature search was performed to identify peak plasma concentrations at therapeutic doses. If critical concentrations *in vitro* resemble concentrations with a therapeutic effect *in vivo*, the compound has a high risk of hepatotoxicity.
4) Identification of critical concentrations *in vitro*. Compound-induced biomarker expression in HepG2 cells and primary human hepatocytes was analyzed using a range of concentrations covering the peak plasma concentration *in vivo* and increasing up to slightly cytotoxic concentrations. For all compounds, the lowest observed effect concentrations that induced expression of the biomarker genes, in addition to those that caused cytotoxic effects were identified.
5) Prediction of hepatotoxic blood concentrations. Human *in vitro* as well as *in vivo* data were integrated into a model that estimates non protein bound plasma concentrations associated with an increased risk of human hepatotoxicity.

The development of predictive biomarkers of hepatotoxicity derived from gene expression profiles of compound-exposed hepatocytes is challenging, and a number of criteria need to be fulfilled. Ideally, biomarker genes of toxicity are indicative of many types of cellular stress, and therefore strongly respond to a large number of different chemicals when tested at a relatively high concentration. Nevertheless, different compounds may induce different forms of liver injury and act via various mechanisms of toxicity, which in turn could result in unique gene expression profiles. In order to cover a broad spectrum of chemicals, the focus should not be on single genes, but rather on a set of biomarker genes that are involved in different biological functions, which cover the most relevant toxic mechanisms of action.

In the present study, potential biomarker candidate genes were identified with the help of the previously established toxicogenomics directory. According to the aforementioned strategy, seven biomarker candidate genes were selected to predict human hepatotoxicity (Table 3.19): the cytochrome P450 isoenzymes CYP1B1 and CYP3A7, the phase II metabolism enzyme sulfotransferase SULT1C2, glucose-6-phosphate dehydrogenase (G6PD), the cytoskeleton component tubulin 2b (TUBB2B), regulator of cell cycle (RGCC), and the proteasomal degradation factor, Fbxo32. Enzymes belonging to the cytochrome P450 family metabolize a variety of environmental and xenobiotic toxicants, thereby playing important roles in the detoxification and clearance of toxic compounds. For this reason, the expression levels of these enzymes are often used to predict potential problems with compound metabolism or drug-drug interactions (Cheng et al. 2011). In this study, the expression of the two isoenzyme biomarker candidate genes, CYP3A7 and CYP1B1 was deregulated after treatment of five out of ten hepatotoxicants (namely rifampicin, ketoconazole, valproic acid, phenylbutazone and carbamazepine), and could therefore be used to discriminate between hepatotoxic and non-hepatotoxic compounds. Since the hepatotoxicity observed *in vitro* occurred at concentrations that correspond to critical concentrations *in vivo*, it could be concluded that these two biomarkers are able to correctly predict blood concentrations that are associated with an increased risk of hepatotoxicity. The CYP3A subfamily accounts for as much as 30 % of total liver cytochrome P450 content, and metabolizes at least 50 % of all drugs (Burk et al. 2002). CYP3A7 is predominantly expressed in the fetal human liver; therefore, its induction by chemicals, as well as in human liver diseases suggests that under these conditions, the liver regresses to a fetal expression pattern (Burk et al. 2002; Pang et al. 2012). Like CYP3A7, CYP1B1 is also expressed at low levels in healthy human liver. It is transcriptionally regulated by the aryl hydrocarbon receptor, which regulates biological responses to a variety of chemicals. It was also found to be important for fetal liver development (Hakkola et al. 1997; Lahvis et al. 2000).

Glucose-6-phosphate dehydrogenase (G6PD) is an enzyme involved in energy and lipid metabolism in the liver. It catalyzes the rate-limiting step of the oxidative pentose-phosphate pathway and provides reducing power in the form of NADPH, and pentose phosphates for fatty acid and nucleic acid synthesis (Hu et al. 2014; Wang et al. 2014). Due to its essential role in the oxidative stress response by producing the main intracellular reductant NADPH, G6PD is considered a guardian of cellular redox potential during oxidative stress (Filosa et al. 2003). Its expression is also strongly induced in many cancers, as well as in compound-

exposed hepatocytes. Consequently, G6PD is considered as a relevant biomarker for the identification of compounds that interfere with hepatic energy homeostasis, as well as those that disturb the redox potential of the cell. Another protein involved in the cellular stress response is the gene product of the suggested biomarker gene Fbxo32. Fbxo32 is a direct target of the transcription factor, FOXO3, which upregulates genes involved in the ubiquitin-proteasome system (Webb and Brunet 2014). Proteins, which might have accumulated in response to hepatocellular injury, are polyubiquitinated by Fbxo32 for proteasomal degradation (Cleveland and Evenhuis 2010). Fbxo32 was further described as a novel cell death regulator, as a mediator of drug-induced apoptosis, and a disrupter of the pro-survival PI3K/Akt pathway (Li et al. 2007; Stitt et al. 2004; Tan et al. 2007).

Further studies have shown that there are additional classes of genes that are highly predictive of hepatotoxicity, such as inflammatory genes or genes encoding proteins involved in DNA repair mechanisms (Cheng et al. 2011). The seven biomarkers identified in the current study cover the toxicological motifs energy and lipid metabolism, metabolism of xenobiotics, cytoskeleton, cell cycle and protein degradation. In order to test, whether these genes are able to predict human hepatotoxicity, two sets of compounds were applied: a set of hepatotoxic compounds that pose a high risk of hepatotoxicity when administered at therapeutic doses, and a set of non-hepatotoxic compounds, which are considered harmless at doses providing a therapeutic effect. A summary of the hepatotoxic compounds and their suggested hepatotoxic mechanism of action are presented in Table 3.14.

All selected candidate genes represent a set of highly predictive biomarkers, which are able to discriminate between hepatotoxic and non-hepatotoxic compounds at given concentrations. In the case of hepatotoxic compounds, the biomarker based *in vitro* system identifies hepatotoxic effects at concentrations within in the range of therapeutic blood concentrations that induced a hepatotoxic effect *in vivo*. In contrast, non-hepatotoxic compounds exhibit hepatotoxic effects only at concentrations that are already within a toxic range and several factors higher than an *in vivo* therapeutic dose.

In primary human hepatocytes, the biomarker and cytotoxicity data based discrimination of the two sets of compounds has been successfully done (Figure 3.23). The *in vitro* model is applicable for a number of hepatotoxic compounds, such as the anticonvulsant valproic acid, and the analgesic acetaminophen, as it precisely predicts the blood concentration that is associated with human hepatotoxicity. Valproic acid is an anti-epileptic agent with a well-characterized toxicity profile. In approximately 0.1 % of all patients, concentrations within a range of its therapeutic dose have been associated with increased risk of hepatotoxicity (Engel et al. 2007). VPA is primarily metabolized in the liver via glucuronidation and β-oxidation (Sztajnkrycer 2002). Its hepatotoxic effect is caused by its interference of the β-oxidation of endogenous lipids. VPA enters the mitochondria via the long chain fatty acid transport system, which uses carnitine as a co-factor. VPA is first attached to coenzyme A (CoA) to form VPA-CoA. VPA-CoA is then esterified with L-carnitine to form VPA-carnitine ester, which is subsequently transported into the mitochondrial matrix by carnitine translocase in exchange for free carnitine. Conjugation of VPA to carnitines results in carnitine depletion, which in-

hibits β-oxidation of endogenous lipids. This results in the accumulation of fat, manifesting itself as microvesicular steatosis, and leads to mitochondrial dysfunction (Begriche et al. 2011; Chitturi and George 2002; Sztajnkrycer 2002). Furthermore, VPA cytotoxicity is associated with increased formation of reactive oxygen species (Tong et al. 2005). With the biomarker based *in vitro* model, VPA-induced hepatotoxicity is observed with *in vitro* concentrations that correspond to hepatotoxic concentrations *in vivo*. Similarly, the model correctly predicts toxic blood concentrations of acetaminophen. Acetaminophen is a classical dose dependent hepatotoxin and not hepatotoxic when administered at therapeutic doses. However, intoxication with this drug is responsible for almost 50 % of all acute liver failure cases in the Western world (Lee 2012). In humans, the risk of hepatotoxicity increases when blood concentrations exceed 1 mM (Winek et al. 2001). Based on the selected biomarkers of hepatotoxicity, the described *in vitro* model predicts exactly this concentration for human hepatotoxicity *in vivo*.

The prediction model is sufficiently sensitive for the antibiotic agent rifampicin, the antifungal drug ketoconazole, as well as for the non-steroidal anti-inflammatory drug (NSAID), phenylbutazone. It predicts hepatotoxicity at already very low concentrations, which are even lower than therapeutic doses. However, it cannot be excluded that hepatotoxic effects in humans *in vivo* already occur at lower concentrations. Lowest observed effect levels *in vivo* are only available for animal models, and cannot be determined for the human system.

Correct evaluations of hepatotoxicity were further obtained for a number of non-hepatotoxic compounds, such as the antihypertensive drug, clonidine or melatonin, which is administered to treat sleep disorders. With the biomarker based *in vitro* system, these compounds are evaluated as harmless. Hepatotoxic effects could only be predicted for very high concentrations, which are several orders of magnitudes higher than a dose with a therapeutic effect. The model is indeed promising as a prediction tool for many of the tested compounds; however, it is not yet applicable to all analyzed compounds. It underestimates the risk of several hepatotoxic compounds, such as diclofenac or labetalol. Diclofenac is a NSAID drug that is used to treat arthritis, and moderate to acute pain. It was previously reported to induce severe idiosyncratic liver injury in approximately 1-5 of 100,000 patients (Chitturi and George 2002). The mechanism of hepatotoxicity is not fully understood, but the formation of toxic metabolites, as well as covalent binding to hepatic proteins leading to oxidative stress and mitochondrial impairment have been proposed (Bort et al. 1999; Ponsoda et al. 1995; Pandit et al. 2012). Labetalol-induced hepatotoxicity is exceedingly rare and the mechanism of hepatotoxicity is unknown. It is primarily metabolized in the liver, ant its metabolic products are thought to induce idiosyncratic drug reactions [3]. Therapeutic doses of diclofenac as well as labetalol are associated with an increased risk of liver injury, but the *in vitro* model only detects hepatotoxic effects at concentrations higher than those associated with hepatotoxicity *in vivo*. Whereas hepatotoxic blood concentrations of diclofenac are in the range of 4-5 µM, the *in vitro* model does not detect hepatotoxicity below a concentration of 400 µM (Kirchheiner et al. 2003; O'Brien et al. 2006). In case of labetalol, hepatotoxicity *in vitro* is detected at a concentration that is 1000-fold higher than the dose associated with a hepatotoxic effect *in vivo*. One possible reason why the prediction model underestimates the risk of

labetalol hepatotoxicity may be that the expression of the selected biomarker candidate genes is not influenced by labetalol *in vitro*. Labetalol is the only compound in this study that did not induce any of the marker genes at any tested concentration in HepG2 cells, as well as primary human hepatocytes. The lowest observed effect concentration *in vitro* corresponded to a cytotoxic concentration with 20 % loss of viability. A logical assumption is that gene expression alterations always occur prior to cytotoxic effects. Therefore, it is highly probable that the so far available biomarkers cannot account for possible mechanisms of toxicity induced by labetalol. To enable a more precise prediction and to detect labetalol-induced hepatotoxicity at lower concentrations *in vitro*, a more sensitive set of biomarkers is required.

Assuming that gene expression alterations always occur prior to cell killing events, the *in vitro* based model for the prediction of hepatotoxic blood concentrations would be most aptly based solely on biomarker expression. Although the seven biomarkers of hepatotoxicity identified to date already cover the majority of known toxic modes of action, not all hepatotoxic compounds could be ascertained. Based on biomarker expression alone (**Figure 3.21**), the system is already very sensitive for the hepatotoxic compounds valproic acid, rifampicin, ketoconazole, acetaminophen and phenylbutazone. Nimesulide- and carbamazepine-induced hepatotoxicity *in vitro* is detected in a concentration range that is tenfold higher than the dose that causes hepatotoxicity *in vivo*. Because of inter-individual susceptibilities among the different donors, predicting a hepatotoxic blood concentration that varies within a factor of 10 is still within an acceptable range. However, except for labetalol, hepatotoxic blood concentrations could not be precisely predicted for nitrofurantoin, diclofenac, isoniazid and aspirin. Biomarker inducing concentrations *in vitro* are a factor of 100-1000 higher than doses associated with hepatotoxicity *in vivo*. A more accurate prediction for these compounds is partially achieved using cytotoxicity data. However, although drug-induced cytotoxicity tests *in vitro* qualitatively support the potential for human toxicity *in vivo*, they were not quantitatively predictive. Prediction of hepatotoxicity based on the lowest cytotoxic concentration alone is much less sensitive than biomarker expression, and does not discriminate between hepatotoxic and non-hepatotoxic compounds (Figure 3.22). Because the so far available biomarkers do not capture all hepatotoxic compounds, the lowest observed effect level from either biomarker induction or 20 % loss of viability are required to discriminate between the two sets of compounds. Follow up studies aim to identify further biomarkers that are more sensitive and which can more precisely predict hepatotoxic blood concentrations *in vivo*. Data driven test system improvements will be performed to establish a set of suitable biomarkers covering all hepatotoxic compounds that will discriminate them from harmless drugs.

The biomarker based *in vitro* system established in HepG2 cell line also discriminates between hepatotoxic and non-hepatotoxic compounds, albeit with less sensitivity than primary human hepatocytes (Figure 3.20). For a number of compounds, these cells predict hepatotoxic alert concentrations in a concentration range that is much higher than doses associated with hepatotoxicity *in vivo*. The model underestimates the risk of hepatotoxicity for phenyl-

butazone, isoniazid, rifampicin, nitrofurantoin and carbamazepine, compared to human primary hepatocytes, which more accurately predict toxicity. The reduced sensitivity in HepG2 cells may be due to the limited metabolic capacity of these cells. HepG2 cells are highly dedifferentiated, but display many features of normal liver cells (Sassa et al. 1987). They express most of the liver specific enzymes (phase I and II metabolism) as well as a number of transcription factors, which are essential for drug metabolism and toxicity responses (Adachi et al. 2007; Hewitt and Hewitt 2004; Vollmer et al. 1999). However, the expression levels of almost all of these enzymes are much lower compared to primary human hepatocytes. Furthermore, the number of differentially expressed genes in this cell line is comparatively low (Gerets et al. 2012; Guo et al. 2011; Westerink and Schoonen 2007). The metabolic competence of HepG2 cells limits the production of reactive metabolites, thus explaining possible differences in hepatotoxic effects observed in HepG2 cells and primary human hepatocytes. Differences in the metabolic capacity might for instance explain why prediction of hepatotoxic blood concentrations were less sensitive for isoniazid in HepG2 cells compared to primary human hepatocytes. Isoniazid (INAH) is an antibiotic agent used to treat tuberculosis and is associated with idiosyncratic hepatotoxicity. About 1.6 % of patients taking INAH develop liver injury, ranging from asymptomatic elevation of serum transaminases to hepatic failure requiring liver transplantation (Steele et al. 1991; Pandit et al. 2012). Isoniazid is cleared mostly by the liver, particularly by acetylation by the N-acetyltransferase 2 (NAT-2) (Saukkonen et al. 2006). NAT-2 acetylated INAH is further hydrolyzed via Cyp2E1 to acetylhydrazine, which together with hydrazine, participates in reactions that generate oxidative stress [4] (Huang et al. 2003). Previous studies have shown that the key enzyme in INAH metabolism, NAT-2, is not expressed in HepG2 cells, but in primary human hepatocytes and liver tissue (Guo et al. 2011; Husain et al. 2007). Reduced production of the toxic metabolite, hydrazine may explain why the HepG2-based *in vitro* model overestimates the hepatotoxic blood concentrations.

A general limitation when using *in vitro* models to predict hepatotoxicity is that *in vitro* observed effects might be attenuated or aggravated under conditions of *in vivo* exposure. Toxicities occurring at the organ level, such as cholestasis, cannot efficiently be predicted. This is also true for toxicities that arise due to the interaction of organs with the systemic circulation, such as the immune and inflammatory response (O'Brien et al. 2006). In addition, drug properties, such as protein binding, transport and pharmacokinetic properties (ADME) are not fully modelled in *in vitro* systems. These limitations may provide a further explanation why there was no sensitive prediction of the hepatotoxic blood concentration for all hepatotoxic compounds. Although the selected biomarkers cover a broad spectrum of toxicological motifs and indicate general hepatocellular stress, specific mechanisms of action might not have been captured. The *in vitro* system does not fully reflect the *in vivo* situation - for example, interactions at the organ level or with the systemic circulation are missing. Therefore, the prediction model may not be able to accurately predict many different mechanisms of toxicity to the same degree. Nevertheless, previous studies have shown that *in vitro* systems are indeed useful to identify hepatotoxic compounds and enable a biomarker based prediction of hepatotoxicity (Cheng et al. 2011; Fischer et al. 2015; Gomez-Lechon et al. 2010).

Likewise, the presented pilot study already revealed promising results and captured many compounds with different toxic mode of actions and different phenotypes of liver injury. In HepG2 cells, and even more sensitive in primary human hepatocytes, the biomarker based *in vitro* model discriminates between hepatotoxic and non-hepatotoxic compounds. In both cell systems, hepatotoxic effects *in vitro* are detected at concentrations that correspond to doses with a high risk of hepatotoxicity *in vivo*. This novel strategy allows the *in vitro* based prediction of human hepatotoxic blood concentrations for a large number of compounds – at least within a certain margin of error. Even some idiosyncratic hepatotoxic compounds, such as diclofenac and nimesulide were identified and distinguished from non-hepatotoxic compounds, such as buspirone and clonidine. This is especially remarkable, because most of the so far available test and prediction systems do not cover idiosyncratic hepatotoxic compounds.

With the selected biomarker set, the prediction model allows a rough estimation of whether a therapeutic dose of a novel compound would be associated with a high or a low risk of hepatotoxicity *in vivo*. The clustering within the set of hepatotoxic or the non-hepatotoxic compounds provides valuable knowledge, e.g. for ranking and prioritization of compounds in early drug development; and therefore, provides a potentially promising tool to assess a putative risk of hepatotoxicity for unknown compounds. In conclusion, the presented model provides a proof of concept for the use of an *in vitro* system to evaluate hepatotoxicity. However, further optimization steps are required because the prediction model still underestimates the risk of hepatotoxicity for some compounds. Follow-up studies will focus on the identification of further biomarkers, which enhances prediction and improves the prediction model. Validation with independent sets of compounds will include hepatotoxic agents, which are usually administered at comparatively low doses, as well as high-dosed non-hepatotoxic compounds. Altogether, this approach will provide additional evidence that the prediction model works independently of concentration effects.

5 References

Aarbakke J, Bakke OM, Milde EJ, Davies DS (1977) Disposition and oxidative metabolism of phenylbutazone in man. European journal of clinical pharmacology 11(5):359-66

Adachi T, Nakagawa H, Chung I, et al. (2007) Nrf2-dependent and -independent induction of ABC transporters ABCC1, ABCC2, and ABCG2 in HepG2 cells under oxidative stress. Journal of experimental therapeutics & oncology 6(4):335-48

Adkison KK, Vaidya SS, Lee DY, et al. (2008) The ABCG2 C421A polymorphism does not affect oral nitrofurantoin pharmacokinetics in healthy Chinese male subjects. British journal of clinical pharmacology 66(2):233-9 doi:10.1111/j.1365-2125.2008.03184.x

Ahmed SN, Siddiqi ZA (2006) Antiepileptic drugs and liver disease. Seizure 15(3):156-64 doi:10.1016/j.seizure.2005.12.009

Akimova T, Beier UH, Wang L, Levine MH, Hancock WW (2011) Helios expression is a marker of T cell activation and proliferation. PloS one 6(8):e24226 doi:10.1371/journal.pone.0024226

Albert KS, Sedman AJ, Wilkinson P, Stoll RG, Murray WJ, Wagner JG (1974) Bioavailability studies of acetaminophen and nitrofurantoin. Journal of clinical pharmacology 14(5-6):264-70

Aldhous M, Franey C, Wright J, Arendt J (1985) Plasma concentrations of melatonin in man following oral absorption of different preparations. British journal of clinical pharmacology 19(4):517-21

Alexa A, Rahnenführer J (2010) topGO: topGO: enrichment analysis for gene ontology. R package version 2.12.0

Alnouti Y, Klaassen CD (2006) Tissue distribution and ontogeny of sulfotransferase enzymes in mice. Toxicological sciences : an official journal of the Society of Toxicology 93(2):242-55 doi:10.1093/toxsci/kfl050

Ament PW, Roth JD, Fox CJ (1994) Famotidine-induced mixed hepatocellular jaundice. The Annals of pharmacotherapy 28(1):40-2

Amit G, Cohen P, Ackerman Z (2002) Nitrofurantoin-induced chronic active hepatitis. The Israel Medical Association journal : IMAJ 4(3):184-6

An X, Jin Y, Guo H, et al. (2009) Response gene to complement 32, a novel hypoxia-regulated angiogenic inhibitor. Circulation 120(7):617-27 doi:10.1161/CIRCULATIONAHA.108.841502

Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome biology 11(10):R106 doi:10.1186/gb-2010-11-10-r106

Appleyard S, Saraswati R, Gorard DA (2010) Autoimmune hepatitis triggered by nitrofurantoin: a case series. Journal of medical case reports 4:311 doi:10.1186/1752-1947-4-311

Aro A, Anttila M, Korhonen T, Sundquist H (1982) Pharmacokinetics of propranolol and sotalol in hyperthyroidism. European journal of clinical pharmacology 21(5):373-7

Aronson JK (2008) Phenylbutazone. In: Aronson JK, editor. Meyler's side effects of analgesics and anti-inflammatory drugs. Boston, MA: Elsevier.

Aubrecht J, Schomaker SJ, Amacher DE (2013) Emerging hepatotoxicity biomarkers and their potential to improve understanding and management of drug-induced liver injury. Genome medicine 5(9):85 doi:10.1186/gm489

Badea T, Niculescu F, Soane L, et al. (2002) RGC-32 increases p34CDC2 kinase activity and entry of aortic smooth muscle cells into S-phase. The Journal of biological chemistry 277(1):502-8 doi:10.1074/jbc.M109354200

Bandara LR, Kennedy S (2002) Toxicoproteomics -- a new preclinical tool. Drug discovery today 7(7):411-8

Bauer A, Schumann A, Gilbert M, et al. (2009) Evaluation of carbon tetrachloride-induced stress on rat hepatocytes by 31P NMR and MALDI-TOF mass spectrometry: lysophosphatidylcholine generation from unsaturated phosphatidylcholines. Chemistry and physics of lipids 159(1):21-9 doi:10.1016/j.chemphyslip.2009.02.006

Beale EG, Harvey BJ, Forest C (2007) PCK1 and PCK2 as candidate diabetes and obesity genes. Cell biochemistry and biophysics 48(2-3):89-95

Beck DH, Schenk MR, Hagemann K, Doepfmer UR, Kox WJ (2000) The pharmacokinetics and analgesic efficacy of larger dose rectal acetaminophen (40 mg/kg) in adults: a double-blinded, randomized study. Anesthesia and analgesia 90(2):431-6

Beedanagari SR, Taylor RT, Bui P, Wang F, Nickerson DW, Hankinson O (2010) Role of epigenetic mechanisms in differential regulation of the dioxin-inducible human CYP1A1 and CYP1B1 genes. Molecular pharmacology 78(4):608-16 doi:10.1124/mol.110.064899

Begriche K, Massart J, Robin MA, Borgne-Sanchez A, Fromenty B (2011) Drug-induced toxicity on mitochondria and lipid metabolism: mechanistic diversity and deleterious consequences for the liver. Journal of hepatology 54(4):773-94 doi:10.1016/j.jhep.2010.11.006

Benjamin SB, Ishak KG, Zimmerman HJ, Grushka A (1981) Phenylbutazone liver injury: a clinical-pathologic survey of 23 cases and review of the literature. Hepatology 1(3):255-63

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met 57(1):289-300

Bessone F (2010) Non-steroidal anti-inflammatory drugs: What is the actual risk of liver damage? World journal of gastroenterology 16(45):5651-61

Bessone F, Colombato L, Fassio E, Reggiardo MV, Vorobioff J, Tanno H. (2010) The spectrum of nimesulide-induced-hepatotoxicity. An overview. Anti-Inflammatory & Anti-Allergy Agents in Medicinal Chemistry 9(4):355-365. DOI: 10.2174/1871523011009040355

Bhattacharyya KK, Brake PB, Eltom SE, Otto SA, Jefcoate CR (1995) Identification of a Rat Adrenal Cytochrome-P450 Active in Polycyclic-Hydrocarbon Metabolism as Rat Cyp1b1 - Demonstration of a Unique Tissue-Specific Pattern of Hormonal and Aryl-Hydrocarbon Receptor-Linked Regulation. Journal of Biological Chemistry 270(19):11595-11602 doi:DOI 10.1074/jbc.270.19.11595

Bianchi M, Ferrario P, Balzarini P, Broggini M (2006) Plasma and synovial fluid concentrations of nimesulide and its main metabolite after a single or repeated oral administration in patients with knee osteoarthritis. The Journal of international medical research 34(4):348-54

Blanchard RL, Freimuth RR, Buck J, Weinshilboum RM, Coughtrie MWH (2004) A proposed nomenclature system for the cytosolic sulfotransferase (SULT) superfamily. Pharmacogenetics 14(3):199-211 doi:10.1097/00008571-200403000-00009

Boelsterli UA (2002) Mechanisms of NSAID-induced hepatotoxicity: focus on nimesulide. Drug safety 25(9):633-48

Bort R, Ponsoda X, Jover R, Gomez-Lechon MJ, Castell JV (1999) Diclofenac toxicity to hepatocytes: a role for drug metabolism in cell toxicity. The Journal of pharmacology and experimental therapeutics 288(1):65-72

Burk O, Tegude H, Koch I, et al. (2002) Molecular mechanisms of polymorphic CYP3A7 expression in adult human liver and intestine. The Journal of biological chemistry 277(27):24280-8 doi:10.1074/jbc.M202345200

Cai Y, Konishi T, Han G, Campwala KH, French SW, Wan YJ (2002) The role of hepatocyte RXR alpha in xenobiotic-sensing nuclear receptor-mediated pathways. European journal of pharmaceutical sciences : official journal of the European Federation for Pharmaceutical Sciences 15(1):89-96

Cha HJ, Ko MJ, Ahn SM, et al. (2010) Identification of classifier genes for hepatotoxicity prediction in non steroidal anti inflammatory drugs. Mol Cell Toxicol 6(3):247-253 doi:10.1007/s13273-010-0034-1

Chatterjee S, Pal J, Biswas N (2008) Nimesulide-induced hepatitis and toxic epidermal necrolysis. Journal of postgraduate medicine 54(2):150-1

Chen M, Zhang M, Borlak J, Tong W (2012) A decade of toxicogenomic research and its contribution to toxicological science. Toxicological sciences : an official journal of the Society of Toxicology 130(2):217-28 doi:10.1093/toxsci/kfs223

Chen TC, Ng KF, Jeng LB, Yeh TS, Chen CM (2001) Aspirin-related hepatotoxicity in a child after liver transplant. Digestive diseases and sciences 46(3):486-8

Cheng F, Theodorescu D, Schulman IG, Lee JK (2011) In vitro transcriptomic prediction of hepatotoxicity for early drug discovery. Journal of theoretical biology 290:27-36 doi:10.1016/j.jtbi.2011.08.009

Chien SC, Rogge MC, Gisclon LG, et al. (1997) Pharmacokinetic profile of levofloxacin following once-daily 500-milligram oral or intravenous doses. Antimicrobial agents and chemotherapy 41(10):2256-60

Chitturi S, George J (2002) Hepatotoxicity of commonly used drugs: nonsteroidal anti-inflammatory drugs, antihypertensives, antidiabetic agents, anticonvulsants, lipid-lowering agents, psychotropic drugs. Seminars in liver disease 22(2):169-83 doi:10.1055/s-2002-30102

Chremos AN (1987) Clinical pharmacology of famotidine: a summary. Journal of clinical gastroenterology 9 Suppl 2:7-12

Clark JA, Zimmerman HJ, Tanner LA (1990) Labetalol hepatotoxicity. Annals of internal medicine 113(3):210-3

Cleveland BM, Evenhuis JP (2010) Molecular characterization of atrogin-1/F-box protein-32 (FBXO32) and F-box protein-25 (FBXO25) in rainbow trout (Oncorhynchus mykiss): Expression across tissues in response to feed deprivation. Comparative biochemistry and physiology Part B, Biochemistry & molecular biology 157(3):248-57 doi:10.1016/j.cbpb.2010.06.010

Cosgrove BD, King BM, Hasan MA, et al. (2009) Synergistic drug-cytokine induction of hepatocellular death as an in vitro approach for the study of inflammation-associated idiosyncratic drug hepatotoxicity. Toxicology and applied pharmacology 237(3):317-330 doi:10.1016/j.taap.2009.04.002

Coughtrie MW (2002) Sulfation through the looking glass--recent advances in sulfotransferase research for the curious. The pharmacogenomics journal 2(5):297-308 doi:10.1038/sj.tpj.6500117

Dalhoff K, Poulsen HE, Garred P, et al. (1987) Buspirone pharmacokinetics in patients with cirrhosis. British journal of clinical pharmacology 24(4):547-50

Daly FF, Fountain JS, Murray L, et al. (2008) Guidelines for the management of paracetamol poisoning in Australia and New Zealand--explanation and elaboration. A consensus statement from clinical toxicologists consulting to the Australasian poisons information centres. The Medical journal of Australia 188(5):296-301

Dart RC, Erdman AR, Olson KR, et al. (2006) Acetaminophen poisoning: an evidence-based consensus guideline for out-of-hospital management. Clinical toxicology 44(1):1-18

De Turck BJ, Diltoer MW, Cornelis PJ, et al. (1998) Lowering of plasma valproic acid concentrations during concomitant therapy with meropenem and amikacin. The Journal of antimicrobial chemotherapy 42(4):563-4

Dieterle W, Faigle JW, Fruh F, et al. (1976) Metabolism of phenylbutazone in man. Arzneimittel-Forschung 26(4):572-7

Doi H, Horie T (2010) Salicylic acid-induced hepatotoxicity triggered by oxidative stress. Chemico-biological interactions 183(3):363-8 doi:10.1016/j.cbi.2009.11.024

Douglas DD, Yang RD, Jensen P, Thiele DL (1989) Fatal labetalol-induced hepatic injury. The American journal of medicine 87(2):235-6

Draghici S, Khatri P, Eklund AC, Szallasi Z (2006) Reliability and reproducibility issues in DNA microarray measurements. Trends in genetics : TIG 22(2):101-9 doi:10.1016/j.tig.2005.12.005

Eichelbaum M, Bertilsson L, Lund L, Palmer L, Sjoqvist F (1976) Plasma levels of carbamazepine and carbamazepine-10,11-epoxide during treatment of epilepsy. European journal of clinical pharmacology 09(5-6):417-21

Elkon R, Linhart C, Sharan R, Shamir R, Shiloh Y (2003) Genome-wide in silico identification of transcriptional regulators controlling the cell cycle in human cells. Genome research 13(5):773-80 doi:10.1101/gr.947203

Ellenhorn MJ, Barceloux DG (1988) Anti-infective Drugs. Medical Toxicology: Diagnosis and Treatment of Human Poisoning. Amsterdam, Elsevier

Ellinger-Ziegelbauer H, Gmuender H, Bandenburg A, Ahr HJ (2008) Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term in vivo studies. Mutation research 637(1-2):23-39 doi:10.1016/j.mrfmmm.2007.06.010

Engel J, Pedley TA, Aicardi J: Epilepsie (2007): A comprehensive textbook, Lippincott Williams & Wilkins

Farrell GC (1994) Drug-induced liver disease. Edinburgh: Churchill Livingstone, 69

Feldman M, Friedman LS, Brandt LJ (2015) Sleisenger and Fordtran's Gastrointestinal and Liver Disease: Pathophysiology, Diagnosis, Management Elsevier Health Sciences

Feingold KR, Moser A, Shigenaga JK, Grunfeld C (2012) Inflammation inhibits the expression of phosphoenolpyruvate carboxykinase in liver and adipose tissue. Innate immunity 18(2):231-40 doi:10.1177/1753425911398678

Filosa S, Fico A, Paglialunga F, et al. (2003) Failure to increase glucose consumption through the pentose-phosphate pathway results in the death of glucose-6-phosphate dehydrogenase gene-deleted mouse embryonic stem cells subjected to oxidative stress. The Biochemical journal 370(Pt 3):935-43 doi:10.1042/BJ20021614

Fischer BM, Neumann D, Piberger AL, Risnes SF, Koberle B, Hartwig A (2015) Use of high-throughput RT-qPCR to assess modulations of gene expression profiles related to genomic stability and interactions by cadmium. Archives of toxicology doi:10.1007/s00204-015-1621-7

Fish DN, Chow AT (1997) The clinical pharmacokinetics of levofloxacin. Clinical pharmacokinetics 32(2):101-19 doi:10.2165/00003088-199732020-00002

Forbes GM, Jeffrey GP, Shilkin KB, Reed WD (1992) Carbamazepine hepatotoxicity: another cause of the vanishing bile duct syndrome. Gastroenterology 102(4 Pt 1):1385-8

Frantz B, Oneill EA (1995) The Effect of Sodium-Salicylate and Aspirin on Nf-Kappa-B. Science 270(5244):2017-2018 doi:DOI 10.1126/science.270.5244.2017

Fromenty B, Pessayre D (1995) Inhibition of mitochondrial beta-oxidation as a mechanism of hepatotoxicity. Pharmacology & therapeutics 67(1):101-54

Garcia Rodriguez LA, Williams R, Derby LE, Dean AD, Jick H (1994) Acute liver injury associated with nonsteroidal anti-inflammatory drugs and the role of risk factors. Archives of internal medicine 154(3):311-6

Garibaldi RA, Drusin RE, Ferebee SH, Gregg MB (1972) Isoniazid-associated hepatitis. Report of an outbreak. The American review of respiratory disease 106(3):357-65 doi:10.1164/arrd.1972.106.3.357

Gerets HH, Tilmant K, Gerin B, et al. (2012) Characterization of primary human hepatocytes, HepG2 cells, and HepaRG cells at the mRNA level and CYP activity in response to inducers and their predictivity for the detection of human hepatotoxins. Cell biology and toxicology 28(2):69-87 doi:10.1007/s10565-011-9208-4

Giannini EG, Testa R, Savarino V (2005) Liver enzyme alteration: a guide for clinicians. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne 172(3):367-79 doi:10.1503/cmaj.1040752

Gingell D, Garrod DR, Palmer JF (1970). In “Calcium and Cellular Function”; (Cuthbert AW, ed.); 59-64. Macmillan, London

Glatt H, Davis W, Meinl W, Hermersdorfer H, Venitt S, Phillips DH (1998) Rat, but not human, sulfotransferase activates a tamoxifen metabolite to produce DNA adducts and gene mutations in bacteria and mammalian cells in culture. Carcinogenesis 19(10):1709-13

Godoy P, Hengstler JG, Ilkavets I, et al. (2009) Extracellular matrix modulates sensitivity of hepatocytes to fibroblastoid dedifferentiation and transforming growth factor beta-induced apoptosis. Hepatology 49(6):2031-43 doi:10.1002/hep.22880

Godoy P, Hewitt NJ, Albrecht U, et al. (2013) Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. Archives of toxicology 87(8):1315-530 doi:10.1007/s00204-013-1078-5

Godoy P, Lakkapamu S, Schug M, et al. (2010a) Dexamethasone-dependent versus -independent markers of epithelial to mesenchymal transition in primary hepatocytes. Biological chemistry 391(1):73-83 doi:10.1515/BC.2010.010

Godoy P, Schug M, Bauer A, Hengstler JG (2010b) Reversible manipulation of apoptosis sensitivity in cultured hepatocytes by matrix-mediated manipulation of signaling activities. Methods in molecular biology 640:139-55 doi:10.1007/978-1-60761-688-7_7

Goidin D, Mamessier A, Staquet MJ, Schmitt D, Berthier-Vergnes O (2001) Ribosomal 18S RNA prevails over glyceraldehyde-3-phosphate dehydrogenase and beta-actin genes as internal standard for quantitative comparison of mRNA levels in invasive and noninvasive human melanoma cell subpopulations. Analytical biochemistry 295(1):17-21 doi:10.1006/abio.2001.5171

Gomez-Lechon MJ, Lahoz A, Gombau L, Castell JV, Donato MT (2010) In vitro evaluation of potential hepatotoxicity induced by drugs. Current pharmaceutical design 16(17):1963-77

Grange JM, Winstanley PA, Davies PD (1994) Clinically significant drug interactions with antituberculosis agents. Drug safety 11(4):242-51

Greenblatt HK, Greenblatt DJ (2014) Liver injury associated with ketoconazole: review of the published evidence. Journal of clinical pharmacology 54(12):1321-9 doi:10.1002/jcph.400

Grinberg M, Stober RM, Edlund K, et al. (2014) Toxicogenomics directory of chemically exposed human hepatocytes. Archives of toxicology 88(12):2261-87 doi:10.1007/s00204-014-1400-x

Gulen M, Ay MO, Avci A, Acikalin A, Icme F (2015) Levofloxacin-induced hepatotoxicity and death. American journal of therapeutics 22(3):e93-6 doi:10.1097/MJT.0b013e3182a44055

Guo L, Dial S, Shi L, et al. (2011) Similarities and differences in the expression of drug-metabolizing enzymes between human hepatic cell lines and primary human hepatocytes. Drug metabolism and disposition: the biological fate of chemicals 39(3):528-38 doi:10.1124/dmd.110.035873

Gupta N, Patel C, Panda M (2009) Hepatitis following famotidine: a case report. Cases journal 2(1):89 doi:10.1186/1757-1626-2-89

Hakkola J, Pasanen M, Pelkonen O, et al. (1997) Expression of CYP1B1 in human adult and fetal tissues and differential inducibility of CYP1B1 and CYP1A1 by Ah receptor ligands in human placenta and cultured cells. Carcinogenesis 18(2):391-7

Halegoua-De Marzio D, Navarro VJ (2013) Hepatotoxicity of cardiovascular and antidiabetic drugs. In: Kaplowitz N, DeLeve LD, eds. Drug-Induced Liver Disease. 3rd ed. Amsterdam: Elsevier; 2013: 522–6.

Hamadeh HK, Bushel PR, Jayadev S, et al. (2002a) Prediction of compound signature using high density gene expression profiling. Toxicological sciences : an official journal of the Society of Toxicology 67(2):232-40

Hamadeh HK, Bushel PR, Jayadev S, et al. (2002b) Gene expression analysis reveals chemical-specific profiles. Toxicological sciences : an official journal of the Society of Toxicology 67(2):219-31

Harbron C, Chang KM, South MC (2007) RefPlus: an R package extending the RMA Algorithm. Bioinformatics 23(18):2493-4 doi:10.1093/bioinformatics/btm357

Hashimoto F, Davis RL, Egli D (1994) Hepatitis following treatments with famotidine and then cimetidine. The Annals of pharmacotherapy 28(1):37-9

Hayhurst GP, Lee YH, Lambert G, Ward JM, Gonzalez FJ (2001) Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. Molecular and cellular biology 21(4):1393-403 doi:10.1128/MCB.21.4.1393-1403.2001

Heise T, Schug M, Storm D, et al. (2012) In vitro - in vivo correlation of gene expression alterations induced by liver carcinogens. Current medicinal chemistry 19(11):1721-30

Hendrickx DM, Aerts HJ, Caiment F, et al. (2015) diXa: a data infrastructure for chemical safety assessment. Bioinformatics 31(9):1505-7 doi:10.1093/bioinformatics/btu827

Hewitt NJ, Hewitt P (2004) Phase I and II enzyme characterization of two sources of HepG2 cell lines. Xenobiotica; the fate of foreign compounds in biological systems 34(3):243-56 doi:10.1080/00498250310001657568

Hoehme S, Brulport M, Bauer A, et al. (2010) Prediction and validation of cell alignment along microvessels as order principle to restore tissue architecture in liver regeneration. Proceedings of the National Academy of Sciences of the United States of America 107(23):10371-6 doi:10.1073/pnas.0909374107

Holland PM, Abramson RD, Watson R, Gelfand DH (1991) Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. Proceedings of the National Academy of Sciences of the United States of America 88(16):7276-80

Holzapfel B, Wickert L (2007): Die quantitative Real-Time-PCR. Biol. Unserer Zeit 37(2): 120-126.

Hu H, Ding X, Yang Y, et al. (2014) Changes in glucose-6-phosphate dehydrogenase expression results in altered behavior of HBV-associated liver cancer cells. American

journal of physiology Gastrointestinal and liver physiology 307(6):G611-22 doi:10.1152/ajpgi.00160.2014

Huang SM, Athanikar NK, Sridhar K, Huang YC, Chiou WL (1982) Pharmacokinetics of chlorpheniramine after intravenous and oral administration in normal adults. European journal of clinical pharmacology 22(4):359-65

Huang WY, Li ZG, Rus H, Wang X, Jose PA, Chen SY (2009) RGC-32 mediates transforming growth factor-beta-induced epithelial-mesenchymal transition in human renal proximal tubular cells. The Journal of biological chemistry 284(14):9426-32 doi:10.1074/jbc.M900039200

Huang YS, Chern HD, Su WJ, et al. (2003) Cytochrome P450 2E1 genotype and the susceptibility to antituberculosis drug-induced hepatitis. Hepatology 37(4):924-30 doi:10.1053/jhep.2003.50144

Husain A, Zhang X, Doll MA, States JC, Barker DF, Hein DW (2007) Identification of N-acetyltransferase 2 (NAT2) transcription start sites and quantitation of NAT2-specific mRNA in human tissues. Drug metabolism and disposition: the biological fate of chemicals 35(5):721-7 doi:10.1124/dmd.106.014621

Igarashi Y, Nakatsu N, Yamashita T, et al. (2015) Open TG-GATEs: a large-scale toxicogenomics database. Nucleic acids research 43(Database issue):D921-7 doi:10.1093/nar/gku955

Inoue Y, Hayhurst GP, Inoue J, Mori M, Gonzalez FJ (2002) Defective ureagenesis in mice carrying a liver-specific disruption of he-datocvte nuclear factor 4 alpha (HNF4 alpha) - HNF4 alpha regulates ornithine transcarbamylase in vivo. Journal of Biological Chemistry 277(28):25257-25265 doi:10.1074/jbc.M203126200

Jaeschke H, Gores GJ, Cederbaum AI, Hinson JA, Pessayre D, Lemasters JJ (2002) Forum - Mechanisms of hepatotoxicity. Toxicological Sciences 65(2):166-176 doi:DOI 10.1093/toxsci/65.2.166

James LP, Mayeux PR, Hinson JA (2003) Acetaminophen-induced hepatotoxicity. Drug Metabolism and Disposition 31(12):1499-1506 doi:DOI 10.1124/dmd.31.12.1499

Jiang J, Wolters JE, van Breda SG, Kleinjans JC, de Kok TM (2015) Development of novel tools for the in vitro investigation of drug-induced liver injury. Expert opinion on drug metabolism & toxicology 11(10):1523-37 doi:10.1517/17425255.2015.1065814

Kamiya A, Inoue Y, Gonzalez FJ (2003) Role of the hepatocyte nuclear factor 4alpha in control of the pregnane X receptor during fetal liver development. Hepatology 37(6):1375-84 doi:10.1053/jhep.2003.50212

Kanada SA, Kolling WM, Hindin BI (1978) Aspirin hepatotoxicity. American journal of hospital pharmacy 35(3):330-6

Kaplowitz N, DeLeve LD (2013) Drug-Induced Liver Disease; Academic Press, 18.04.2013

Karim A, Ahmed S, Rossoff LJ, Siddiqui RK, Steinberg HN (2001) Possible levofloxacin-induced acute hepatocellular injury in a patient with chronic obstructive lung disease. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 33(12):2088-90 doi:10.1086/338156

Keranen A, Nykanen S, Taskinen J (1978) Pharmacokinetics and side-effects of clonidine. European journal of clinical pharmacology 13(2):97-101

Kiang TK, Ford JA, Yoshida EM, Partovi N (2011) Nitrofurantoin-associated lung and liver toxicity leading to liver transplantation in a middle-aged patient. The Canadian journal of hospital pharmacy 64(4):262-70

Kienhuis AS, van de Poll MC, Wortelboer H, et al. (2009) Parallelogram approach using rat-human in vitro and rat in vivo toxicogenomics predicts acetaminophen-induced hepatotoxicity in humans. Toxicological sciences : an official journal of the Society of Toxicology 107(2):544-52 doi:10.1093/toxsci/kfn237

Kimura Y, Nishimura FT, Abe S, Fukunaga T, Tanii H, Saijoh K (2009) Polymorphisms in the promoter region of the human class II alcohol dehydrogenase (ADH4) gene affect both transcriptional activity and ethanol metabolism in Japanese subjects. The Journal of toxicological sciences 34(1):89-97

Kirchheiner J, Meineke I, Steinbach N, Meisel C, Roots I, Brockmoller J (2003) Pharmacokinetics of diclofenac and inhibition of cyclooxygenases 1 and 2: no relationship to the CYP2C9 genetic polymorphism in humans. British journal of clinical pharmacology 55(1):51-61

Kis E, Ioja E, Rajnai Z, et al. (2012) BSEP inhibition: in vitro screens to assess cholestatic potential of drugs. Toxicology in vitro : an international journal published in association with BIBRA 26(8):1294-9 doi:10.1016/j.tiv.2011.11.002

Klein RH, Alvarez-Jimenez R, Sukhai RN, et al. (2013) Pharmacokinetics and pharmacodynamics of orally administered clonidine: a model-based approach. Hormone research in paediatrics 79(5):300-9 doi:10.1159/000350819

Kopanoff DE, Snider DE, Jr., Caras GJ (1978) Isoniazid-related hepatitis: a U.S. Public Health Service cooperative surveillance study. The American review of respiratory disease 117(6):991-1001 doi:10.1164/arrd.1978.117.6.991

Kothapalli R, Yoder SJ, Mane S, Loughran TP (2002) Microarray results: how accurate are they? Bmc Bioinformatics 3 doi:Artn 22 Doi 10.1186/1471-2105-3-22

Krug AK, Kolde R, Gaspar JA, et al. (2013) Human embryonic stem cell-derived test systems for developmental neurotoxicity: a transcriptomics approach. Archives of toxicology 87(1):123-43 doi:10.1007/s00204-012-0967-3

Lahvis GP, Lindell SL, Thomas RS, et al. (2000) Portosystemic shunting and persistent fetal vascular structures in aryl hydrocarbon receptor-deficient mice. Proceedings of the National Academy of Sciences of the United States of America 97(19):10442-7 doi:10.1073/pnas.190256997

Lake-Bakaar G, Scheuer PJ, Sherlock S (1987) Hepatic reactions associated with ketoconazole in the United Kingdom. Br Med J (Clin Res Ed) 294(6569):419-22

Lake AD, Novak P, Fisher CD, et al. (2011) Analysis of global and absorption, distribution, metabolism, and elimination gene expression in the progressive stages of human nonalcoholic fatty liver disease. Drug metabolism and disposition: the biological fate of chemicals 39(10):1954-60 doi:10.1124/dmd.111.040592

Lalonde RL, O'Rear TL, Wainer IW, Drda KD, Herring VL, Bottorff MB (1990) Labetalol pharmacokinetics and pharmacodynamics: evidence of stereoselective disposition. Clinical pharmacology and therapeutics 48(5):509-19

Lamberg TS, Kivisto KT, Neuvonen PJ (1998) Concentrations and effects of buspirone are considerably reduced by rifampicin. British journal of clinical pharmacology 45(4):381-5

Laster J, Satoskar R (2014) Aspirin-Induced Acute Liver Injury. ACG case reports journal 2(1):48-9 doi:10.14309/crj.2014.81

Lauterburg BH, Smith CV, Todd EL, Mitchell JR (1985) Pharmacokinetics of the toxic hydrazino metabolites formed from isoniazid in humans. The Journal of pharmacology and experimental therapeutics 235(3):566-70

Lee JH, Kim EJ, Kim DK, et al. (2012) Hypoxia induces PDK4 gene expression through induction of the orphan nuclear receptor ERRgamma. PloS one 7(9):e46324 doi:10.1371/journal.pone.0046324

Lee WM (2012) Acute Liver Failure. Semin Resp Crit Care 33(1):36-45 doi:10.1055/s-0032-1301733

Lewis JH (2003) Nonsteroidal anti-inflammatory drugs: pathology and clinical presentation of hepatotoxicity. In Kaplowitz N, DeLeve DL eds. Drug-Induced Liver Disease, Marcel Dekker, New York, 377–404

Li HH, Willis MS, Lockyer P, et al. (2007) Atrogin-1 inhibits Akt-dependent cardiac hypertrophy in mice via ubiquitin-dependent coactivation of Forkhead proteins. The Journal of clinical investigation 117(11):3211-23 doi:10.1172/JCI31757

Li J, Ning G, Duncan SA (2000) Mammalian hepatocyte differentiation requires the transcription factor HNF-4alpha. Genes & development 14(4):464-74

Lin M, Napoli JL (2000) cDNA cloning and expression of a human aldehyde dehydrogenase (ALDH) active with 9-cis-retinal and identification of a rat ortholog, ALDH12. The Journal of biological chemistry 275(51):40106-12 doi:10.1074/jbc.M008027200

Long RC, Wofford MR, Harkins KG, Minor DS (2007) Hepatocellular necrosis associated with labetalol. Journal of clinical hypertension 9(4):287-90

Macpherson D, Best SA, Gedik L, Hewson AT, Rainsford KD, Parisi S (2013) The Biotransformation and Pharmacokinetics of 14 C-Nimesulide in Humans Following a Single Dose Oral Administration. J Drug Metab Toxicol 4: 140.

Mahmood I, Sahajwalla C (1999) Clinical pharmacokinetics and pharmacodynamics of buspirone, an anxiolytic drug. Clinical pharmacokinetics 36(4):277-87 doi:10.2165/00003088-199936040-00003

Malone RS, Fish DN, Abraham E, Teitelbaum I (2001) Pharmacokinetics of levofloxacin and ciprofloxacin during continuous renal replacement therapy in critically ill patients. Antimicrobial agents and chemotherapy 45(10):2949-54 doi:10.1128/AAC.45.10.2949-2954.2001

Mandel GL, Sande MA (1985) Antimicrobial Agents. Goodman & Gilman's the pharma- co-logic Basis of Therapeutics, 7th ed. MacMillan Publish Co.

Mansur AP, Avakian SD, Paula RS, Donzella H, Santos SR, Ramires JA (1998) Pharmacokinetics and pharmacodynamics of propranolol in hypertensive patients after sublingual administration: systemic availability. Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas / Sociedade Brasileira de Biofisica [et al] 31(5):691-6

Marinella MA (2002) Labetalol-induced hepatitis in a patient with chronic hepatitis B infection. Journal of clinical hypertension 4(2):120-1

McMillin GA, Juenke JM, Tso G, Dasgupta A (2010) Estimation of carbamazepine and carbamazepine-10,11-epoxide concentrations in plasma using mathematical equations generated with two carbamazepine immunoassays. American journal of clinical pathology 133(5):728-36 doi:10.1309/AJCPFAHVB26VVVTE

Mehta JB, Shantaveerapa H, Byrd RP, Jr., Morton SE, Fountain F, Roy TM (2001) Utility of rifampin blood levels in the treatment and follow-up of active pulmonary tuberculosis in patients who were slow to respond to routine directly observed therapy. Chest 120(5):1520-4

Mendrick DL, Schnackenberg L (2009) Genomic and metabolomic advances in the identification of disease and adverse event biomarkers. Biomarkers in medicine 3(5):605-15 doi:10.2217/bmm.09.43

Merlani G, Fox M, Oehen HP, et al. (2001) Fatal hepatoxicity secondary to nimesulide. European journal of clinical pharmacology 57(4):321-6

Mignot G. (2000) Gastrointestinal drugs. In: Dukes MNG, Aronson JK, editors. Meyler's Side Effects of Drugs. 14. Amsterdam: Elsevier; 1246–337

Mitchell MC, Boitnott JK, Arregui A, Maddrey WC (1981) Granulomatous hepatitis associated with carbamazepine therapy. The American journal of medicine 71(4):733-5

Morgan MY, Stambuk D, Cottrell J, Mann SG (1990) Pharmacokinetics of famotidine in normal subjects and in patients with chronic liver disease. Alimentary pharmacology & therapeutics 4(1):83-96

Moscona A, Trowell OA and Willmer EN (1965). In “Cells and Tissues in Culture” (Willmer EN, ed.); Academic Press New York; Vol 1, 19-98

Moseley RH. (2013) Antifungal agents. Antibacterial and antifungal agents. In, Kaplowitz N, DeLeve LD, eds. Drug-induced liver disease. 3rd ed. Amsterdam: Elsevier, 470-3.

Musumba CO, Pamba AO, Sasi PA, English M, Maitland K (2004) Salicylate poisoning in children: report of three cases. East African medical journal 81(3):159-63

Navarro VJ, Senior JR (2006) Drug-related hepatotoxicity. The New England journal of medicine 354(7):731-9 doi:10.1056/NEJMra052270

Nibourg GA, Chamuleau RA, van Gulik TM, Hoekstra R (2012) Proliferative human cell sources applied as biocomponent in bioartificial livers: a review. Expert opinion on biological therapy 12(7):905-21 doi:10.1517/14712598.2012.685714

O'Brien PJ, Irwin W, Diaz D, et al. (2006) High concordance of drug-induced human hepatotoxicity with in vitro cytotoxicity measured in a novel cell-based model using high content screening. Archives of toxicology 80(9):580-604 doi:10.1007/s00204-006-0091-3

Orme M, Holt PJ, Hughes GR, et al. (1976) Plasma concentration of phenylbutazone and its therapeutic effect-studies in patients with rheumatoid arthritis. British journal of clinical pharmacology 3(1):185-91

Ostapowicz G, Fontana RJ, Schiodt FV, et al. (2002) Results of a prospective study of acute liver failure at 17 tertiary care centers in the United States. Annals of internal medicine 137(12):947-54

Pagani A, Rizzetto M. (1987) Chlorpheniramine hepatotoxicity. It J Gastroenterol.;19:179–83

Pandit A, Sachdeva T, Bafna P. (2012) Drug induced hepatotoxicity: A Review. Journal of Applied Pharmaceutical Science 2(5):233-243. DOI: 10.7324/JAPS.2012.2541

Pang XY, Cheng J, Kim JH, Matsubara T, Krausz KW, Gonzalez FJ (2012a) Expression and regulation of human fetal-specific CYP3A7 in mice. Endocrinology 153(3):1453-63 doi:10.1210/en.2011-1020

Pang XY, Cheng J, Kim JH, Matsubara T, Krausz KW, Gonzalez FJ (2012b) Expression and Regulation of Human Fetal-Specific CYP3A7 in Mice. Endocrinology 153(3):1453-1463 doi:10.1210/en.2011-1020

Pfannkuch F, Suter-Dick L Mannhold R, Kubinyi H, Folkers G (2014) Predictive Toxicology: From Vision to Reality, Volume 64. Wiley Verlag; ISBN: 978-3-527-33608-1

Pirmohamed M, Leeder SJ. Anticonvulsant agents. (2013) In, Kaplowitz N, DeLeve LD, eds. Drug- induced liver disease. 3rd ed. Amsterdam: Elsevier, 2013: pp 423-41.

Ponsoda X, Bort R, Jover R, Gomez-Lechon MJ, Castell JV (1995) Molecular mechanism of diclofenac hepatotoxicity: Association of cell injury with oxidative metabolism and decrease in ATP levels. Toxicology in vitro : an international journal published in association with BIBRA 9(4):439-44

Pourahmad J, Eskandari MR, Kaghazi A, Shaki F, Shahraki J, Fard JK (2012) A new approach on valproic acid induced hepatotoxicity: involvement of lysosomal membrane leakiness and cellular proteolysis. Toxicology in vitro : an international journal published in association with BIBRA 26(4):545-51 doi:10.1016/j.tiv.2012.01.020

Powell-Jackson PR, Tredger JM, Williams R (1984) Hepatotoxicity to sodium valproate: a review. Gut 25(6):673-81

Prince MI, Burt AD, Jones DE (2002) Hepatitis and liver dysfunction with rifampicin therapy for pruritus in primary biliary cirrhosis. Gut 50(3):436-9

Requena-Mendez A, Davies G, Waterhouse D, et al. (2014) Effects of dosage, comorbidities, and food on isoniazid pharmacokinetics in Peruvian tuberculosis patients. Antimicrobial agents and chemotherapy 58(12):7164-70 doi:10.1128/AAC.03258-14

Richards DA, Maconochie JG, Bland RE, Hopkins R, Woodings EP, Martin LE (1977) Relationship between plasma concentrations and pharmacological effects of labetalol. European journal of clinical pharmacology 11(2):85-90

Rodriguez RJ, Acosta D, Jr. (1997) N-deacetyl ketoconazole-induced hepatotoxicity in a primary culture system of rat hepatocytes. Toxicology 117(2-3):123-31

Rodriguez RJ, Buckholz CJ (2003) Hepatotoxicity of ketoconazole in Sprague-Dawley rats: glutathione depletion, flavincontaining monooxygenases-mediated bioactivation and

hepatic covalent binding. Xenobiotica; the fate of foreign compounds in biological systems 33(4):429-441 doi:10.1080/0049825031000072243

Rodriguez RJ, Proteau PJ, Marquez BL, Hetherington CL, Buckholz CJ, O'Connell KL (1999) Flavin-containing monooxygenase-mediated metabolism of N-deacetyl ketoconazole by rat hepatic microsomes. Drug Metabolism and Disposition 27(8):880-886

Roth A, Boess F, Landes C, et al. (2011) Gene expression-based in vivo and in vitro prediction of liver toxicity allows compound selection at an early stage of drug development. Journal of biochemical and molecular toxicology 25(3):183-94 doi:10.1002/jbt.20375

Saigusa K, Imoto I, Tanikawa C, et al. (2007) RGC32, a novel p53-inducible gene, is located on centrosomes during mitosis and results in G2/M arrest. Oncogene 26(8):1110-21 doi:10.1038/sj.onc.1210148

Sakaan SA, Twilla JD, Usery JB, Winton JC, Self TH (2014) Nitrofurantoin-induced hepatotoxicity: a rare yet serious complication. Southern medical journal 107(2):107-13 doi:10.1097/SMJ.0000000000000059

Sanders GL, Routledge PA, Ward A, Davies DM, Rawlins MD (1979) Mean steady-state plasma concentrations of labetalol in patients undergoing antihypertensive therapy. British journal of clinical pharmacology 8(Suppl 2):153S-155S

Sassa S, Sugita O, Galbraith RA, Kappas A (1987) Drug metabolism by the human hepatoma cell, Hep G2. Biochemical and biophysical research communications 143(1):52-7

Saukkonen JJ, Cohn DL, Jasmer RM, et al. (2006) An official ATS statement: hepatotoxicity of antituberculosis therapy. American journal of respiratory and critical care medicine 174(8):935-52 doi:10.1164/rccm.200510-1666ST

Schafer-Korting M, Korting HC, Mutschler E (1985) Human plasma and skin blister fluid levels of griseofulvin following a single oral dose. European journal of clinical pharmacology 29(1):109-13

Schug M, Stober R, Heise T, et al. (2013) Pharmacokinetics explain in vivo/in vitro discrepancies of carcinogen-induced gene expression alterations in rat liver and cultivated hepatocytes. Archives of toxicology 87(2):337-45 doi:10.1007/s00204-012-0999-8

Schwinghammer TL, Juhl RP, Dittert LW, Melethil SK, Kroboth FJ, Chung VS (1984) Comparison of the bioavailability of oral, rectal and intramuscular promethazine. Biopharmaceutics & drug disposition 5(2):185-94

Seglen PO (1976) Preparation of isolated rat liver cells. Methods in cell biology 13:29-83

Seymour RA, Williams FM, Ward A, Rawlins MD (1984) Aspirin metabolism and efficacy in postoperative dental pain. British journal of clinical pharmacology 17(6):697-701

Sharoky M, Perkal M, Turner R, Lesko LJ (1988) Steady state relative bioavailability and pharmacokinetics of oral propranolol in black and white North Americans. Biopharmaceutics & drug disposition 9(5):447-56

Sherigar JM, Fazio R, Zuang M, Arsura E (2012) Autoimmune hepatitis induced by nitrofurantoin. The importance of the autoantibodies for an early diagnosis of immune disease. Clinics and practice 2(4):e83 doi:10.4081/cp.2012.e83

Shi Q, Hong H, Senior J, Tong W (2010) Biomarkers for drug-induced liver injury. Expert review of gastroenterology & hepatology 4(2):225-34 doi:10.1586/egh.10.8

Shinde V, Stober R, Nemade H, et al. (2015) Transcriptomics of Hepatocytes Treated with Toxicants for Investigating Molecular Mechanisms Underlying Hepatotoxicity. Methods in molecular biology 1250:225-40 doi:10.1007/978-1-4939-2074-7_16

Simons FE, Watson WT, Chen XY, Minuk GY, Simons KJ (1989) The pharmacokinetics and pharmacodynamics of hydroxyzine in patients with primary biliary cirrhosis. Journal of clinical pharmacology 29(9):809-15

Singh BK, Tripathi M, Chaudhari BP, Pandey PK, Kakkar P (2012) Natural terpenes prevent mitochondrial dysfunction, oxidative stress and release of apoptotic proteins during nimesulide-hepatotoxicity in rats. PloS one 7(4):e34200 doi:10.1371/journal.pone.0034200

Stanley EL, Hume R, Coughtrie MW (2005) Expression profiling of human fetal cytosolic sulfotransferases involved in steroid and thyroid hormone metabolism and in detoxification. Molecular and cellular endocrinology 240(1-2):32-42 doi:10.1016/j.mce.2005.06.003

Steele MA, Burk RF, DesPrez RM (1991) Toxic hepatitis with isoniazid and rifampin. A meta-analysis. Chest 99(2):465-71

Stitt TN, Drujan D, Clarke BA, et al. (2004) The IGF-1/PI3K/Akt pathway prevents expression of muscle atrophy-induced ubiquitin ligases by inhibiting FOXO transcription factors. Molecular cell 14(3):395-403

Stocker ME, Montgomery JE (2001) Serum paracetamol concentrations in adult volunteers following rectal administration. British journal of anaesthesia 87(4):638-40

Stoffel M, Duncan SA (1997) The maturity-onset diabetes of the young (MODY1) transcription factor HNF4alpha regulates expression of genes required for glucose transport and metabolism. Proceedings of the National Academy of Sciences of the United States of America 94(24):13209-14

Strenkoski-Nix LC, Ermer J, DeCleene S, Cevallos W, Mayer PR (2000) Pharmacokinetics of promethazine hydrochloride after administration of rectal suppositories and oral syrup to healthy subjects. American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists 57(16):1499-505

Stricker BHC, Biour M, Wilson JHP. (1995) Individual agents: drugs. In: Stricker B, editor.Drug-induced hepatic injury. Drug-induced disorders. Amsterdam: Elsevier; 5:71–524

Stronkhorst A, Bosma A, van Leeuwen DJ (1992) A case of labetalol-induced hepatitis. The Netherlands journal of medicine 40(3-4):200-2

Sugar AM, Alsip SG, Galgiani JN, et al. (1987) Pharmacology and toxicity of high-dose ketoconazole. Antimicrobial agents and chemotherapy 31(12):1874-8

Sun J, Slavov S, Schnackenberg LK, et al. (2014) Identification of a metabolic biomarker panel in rats for prediction of acute and idiosyncratic hepatotoxicity. Computational and structural biotechnology journal 10(17):78-89 doi:10.1016/j.csbj.2014.08.001

Suter L, Babiss LE, Wheeldon EB (2004) Toxicogenomics in predictive toxicology in drug development. Chemistry & biology 11(2):161-71 doi:10.1016/j.chembiol.2004.02.003

Sztajnkrycer MD (2002) Valproic acid toxicity: overview and management. Journal of toxicology Clinical toxicology 40(6):789-801

Tagawa M, Kano M, Okamura N, et al. (2002) Differential cognitive effects of ebastine and (+)-chlorpheniramine in healthy subjects: correlation between cognitive impairment and plasma drug concentration. British journal of clinical pharmacology 53(3):296-304

Tan HH, Ong WM, Lai SH, Chow WC (2007a) Nimesulide-induced hepatotoxicity and fatal hepatic failure. Singapore medical journal 48(6):582-5

Tan J, Yang X, Zhuang L, et al. (2007b) Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. Genes & development 21(9):1050-63 doi:10.1101/gad.1524107

Thee S, Seddon JA, Donald PR, et al. (2011) Pharmacokinetics of isoniazid, rifampin, and pyrazinamide in children younger than two years of age with tuberculosis: evidence for implementation of revised World Health Organization recommendations. Antimicrobial agents and chemotherapy 55(12):5560-7 doi:10.1128/AAC.05429-11

Tirona RG, Lee W, Leake BF, et al. (2003) The orphan nuclear receptor HNF4alpha determines PXR- and CAR-mediated xenobiotic induction of CYP3A4. Nature medicine 9(2):220-4 doi:10.1038/nm815

Tong V, Teng XW, Chang TK, Abbott FS (2005) Valproic acid II: effects on oxidative stress, mitochondrial membrane potential, and cytotoxicity in glutathione-depleted rat hepatocytes. Toxicological sciences : an official journal of the Society of Toxicology 86(2):436-43 doi:10.1093/toxsci/kfi185

Tostmann A, Boeree MJ, Aarnoutse RE, de Lange WC, van der Ven AJ, Dekhuijzen R (2008) Antituberculosis drug-induced hepatotoxicity: concise up-to-date review. Journal of gastroenterology and hepatology 23(2):192-202 doi:10.1111/j.1440-1746.2007.05207.x

Tripathi M, Singh BK, Mishra C, Raisuddin S, Kakkar P (2010) Involvement of mitochondria mediated pathways in hepatoprotection conferred by Fumaria parviflora Lam. extract against nimesulide induced apoptosis in vitro. Toxicology in vitro : an international journal published in association with BIBRA 24(2):495-508 doi:10.1016/j.tiv.2009.09.011

Tung EK, Mak CK, Fatima S, et al. (2011) Clinicopathological and prognostic significance of serum and tissue Dickkopf-1 levels in human hepatocellular carcinoma. Liver international : official journal of the International Association for the Study of the Liver 31(10):1494-504 doi:10.1111/j.1478-3231.2011.02597.x

Tuschl G, Lauer B, Mueller SO (2008) Primary hepatocytes as a model to analyze species-specific toxicity and drug metabolism. Expert opinion on drug metabolism & toxicology 4(7):855-70 doi:10.1517/17425255.4.7.855

Ulitsky I, Maron-Katz A, Shavit S, et al. (2010) Expander: from expression microarrays to networks and functions. Nature protocols 5(2):303-22 doi:10.1038/nprot.2009.230

van Hest R, Baars H, Kik S, et al. (2004) Hepatotoxicity of rifampin-pyrazinamide and isoniazid preventive therapy and tuberculosis treatment. Clinical infectious diseases : an official

publication of the Infectious Diseases Society of America 39(4):488-96 doi:10.1086/422645

Varemo L, Nielsen J, Nookaew I (2013) Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. Nucleic acids research 41(8):4378-91 doi:10.1093/nar/gkt111

Vollmer CM, Ribas A, Butterfield LH, et al. (1999) p53 selective and nonselective replication of an E1B-deleted adenovirus in hepatocellular carcinoma. Cancer research 59(17):4369-74

Waldmann T, Rempel E, Balmer NV, et al. (2014) Design principles of concentration-dependent transcriptome deviations in drug-exposed differentiating stem cells. Chemical research in toxicology 27(3):408-20 doi:10.1021/tx400402j

Wallace JE, Shimek EL, Jr., Harris SC, Stavchansky S (1981) Determination of promethazine in serum by liquid chromatography. Clinical chemistry 27(2):253-5

Wang YP, Zhou LS, Zhao YZ, et al. (2014) Regulation of G6PD acetylation by SIRT2 and KAT9 modulates NADPH homeostasis and cell survival during oxidative stress. The EMBO journal 33(12):1304-20 doi:10.1002/embj.201387224

Watt AJ, Garrison WD, Duncan SA (2003) HNF4: a central regulator of hepatocyte differentiation and function. Hepatology 37(6):1249-53 doi:10.1053/jhep.2003.50273

Webb AE, Brunet A (2014) FOXO transcription factors: key regulators of cellular quality control. Trends in biochemical sciences 39(4):159-69 doi:10.1016/j.tibs.2014.02.003

Weber LW, Boll M, Stampfl A (2003) Hepatotoxicity and mechanism of action of haloalkanes: carbon tetrachloride as a toxicological model. Critical reviews in toxicology 33(2):105-36 doi:10.1080/713611034

Weerasinghe SV, Jang YJ, Fontana RJ, Omary MB (2014) Carbamoyl phosphate synthetase-1 is a rapid turnover biomarker in mouse and human acute liver injury. American journal of physiology Gastrointestinal and liver physiology 307(3):G355-64 doi:10.1152/ajpgi.00303.2013

Westerink WM, Schoonen WG (2007) Cytochrome P450 enzyme levels in HepG2 cells and cryopreserved primary human hepatocytes and their induction in HepG2 cells. Toxicology in vitro : an international journal published in association with BIBRA 21(8):1581-91 doi:10.1016/j.tiv.2007.05.014

Wilke RA, Lin DW, Roden DM, et al. (2007) Identifying genetic risk factors for serious adverse drug reactions: current progress and challenges. Nature reviews Drug discovery 6(11):904-16 doi:10.1038/nrd2423

Winek CL, Wahba WW, Winek CL, Jr., Balzer TW (2001) Drug and chemical blood-level data 2001. Forensic science international 122(2-3):107-23

Wong ML, Medrano JF (2005) Real-time PCR for mRNA quantitation. BioTechniques 39(1):75-85

Wu J, Wang C, Li S, et al. (2013) Thyroid hormone-responsive SPOT 14 homolog promotes hepatic lipogenesis, and its expression is regulated by liver X receptor alpha through a sterol regulatory element-binding protein 1c-dependent mechanism in mice. Hepatology 58(2):617-28 doi:10.1002/hep.26272

Yang Y, Salminen WF, Schnackenberg LK (2012) Current and emerging biomarkers of hepatotoxicity. Curr Biomarker Find; 2:43-55

Yoon KA, Nakamura Y, Arakawa H (2004) Identification of ALDH4 as a p53-inducible gene and its protective role in cellular stresses. Journal of human genetics 49(3):134-40 doi:10.1007/s10038-003-0122-3

Zellmer S, Schmidt-Heck W, Godoy P, et al. (2010) Transcription factors ETF, E2F, and SP-1 are involved in cytokine-independent proliferation of murine hepatocytes. Hepatology 52(6):2127-36 doi:10.1002/hep.23930

Zhu Q, Mariash A, Margosian MR, et al. (2001) Spot 14 gene deletion increases hepatic de novo lipogenesis. Endocrinology 142(10):4363-70 doi:10.1210/endo.142.10.8431

Zidek N, Hellmann J, Kramer PJ, Hewitt PG (2007) Acute hepatotoxicity: a predictive model based on focused illumina microarrays. Toxicological sciences : an official journal of the Society of Toxicology 99(1):289-302 doi:10.1093/toxsci/kfm131

Further references (12th of January 2015):

[1] http://www.ncbi.nlm.nih.gov/pubmed

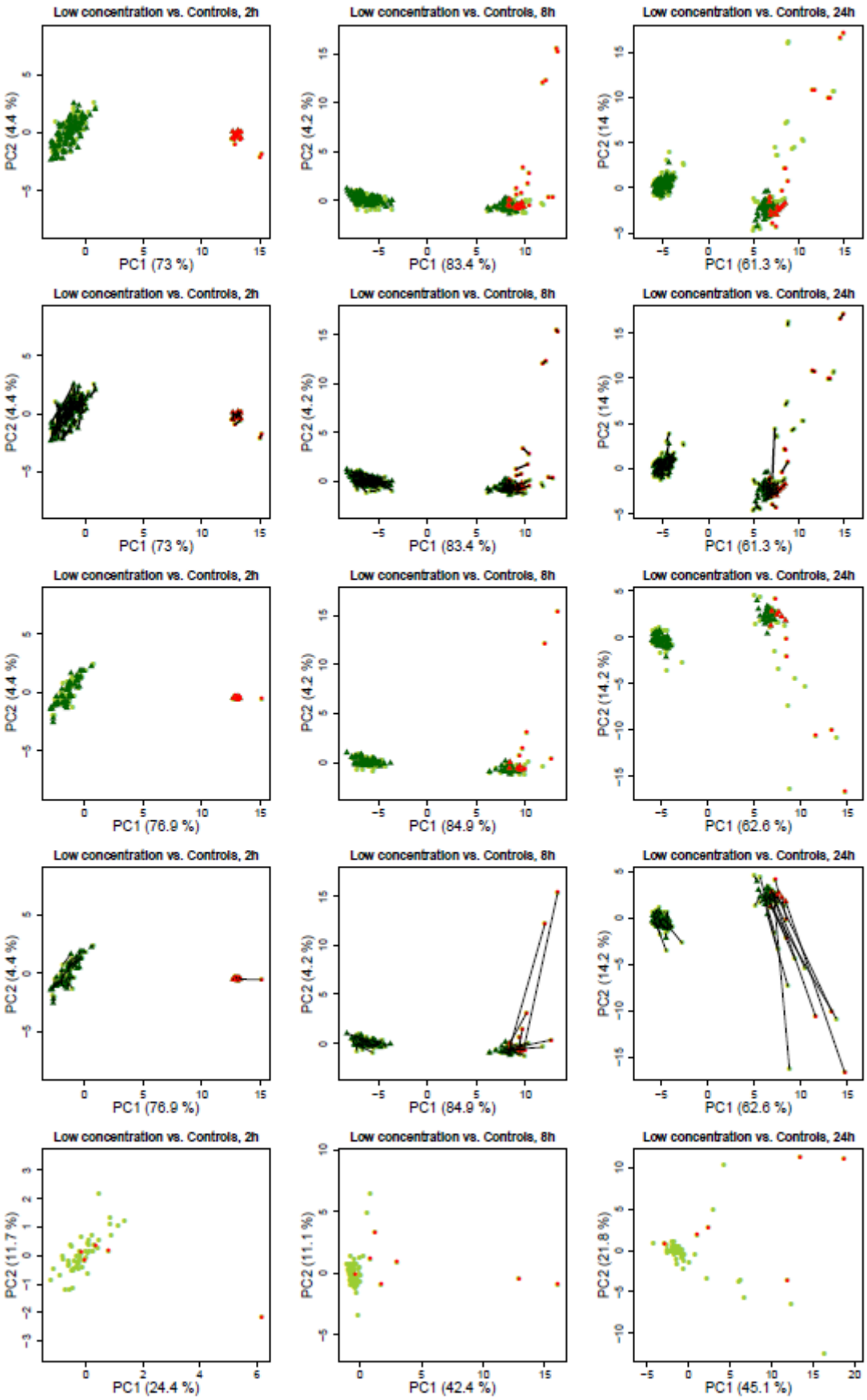
[2] http://livertox.nih.gov/

[3] http://livertox.nlm.nih.gov/Labetalol.htm

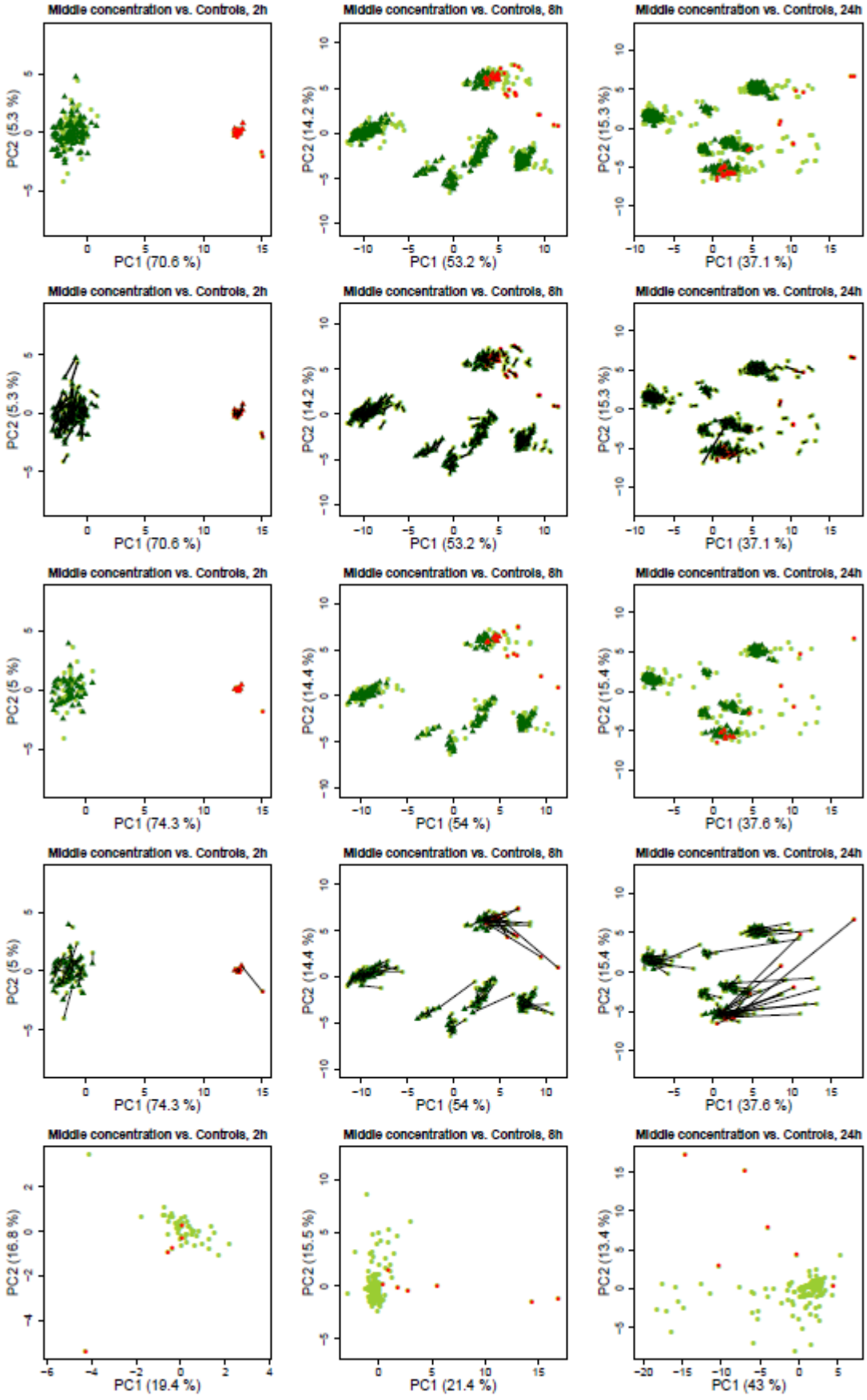
[4] http://www.goldfrankstoxicology.com/cases/GTE_INH-Associated_Hepatotoxicity.pdf

6 Appendix

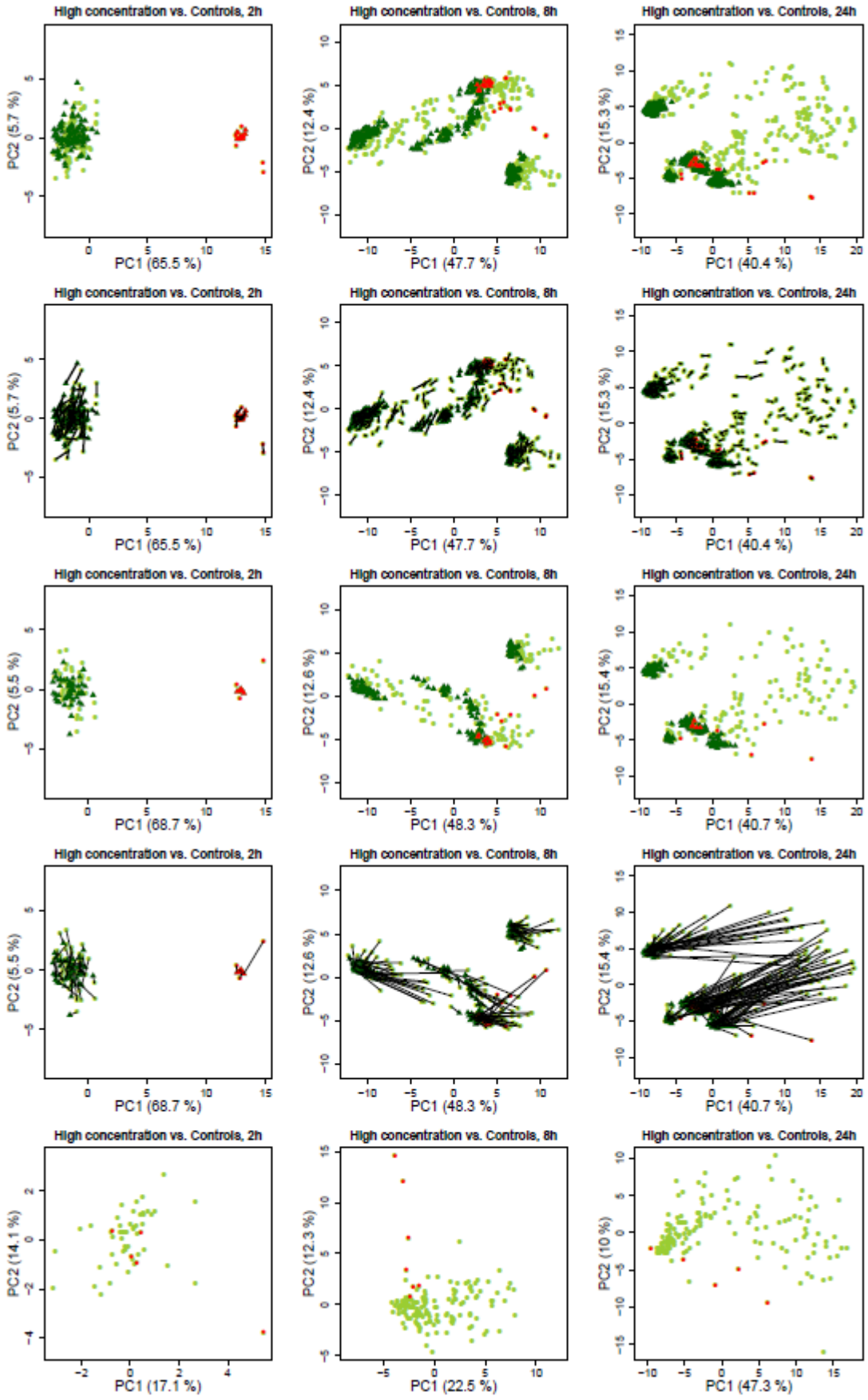
6.1 Supplemental figures:



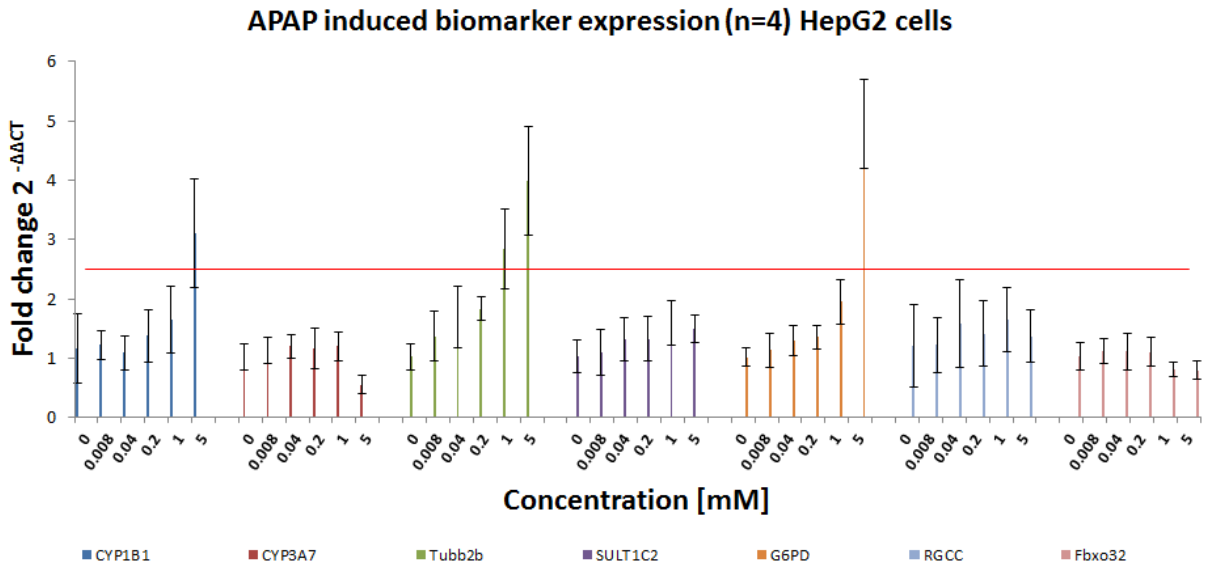
Supplemental Figure 1: Corresponding data to Figure 3.1 summarizing all further incubation conditions besides the high concentration and 24h exposure already shown in Fig. 1. NA: A sample for this time point and concentration was not available. Low concentration, 2h, 8h and 24h incubation**.**



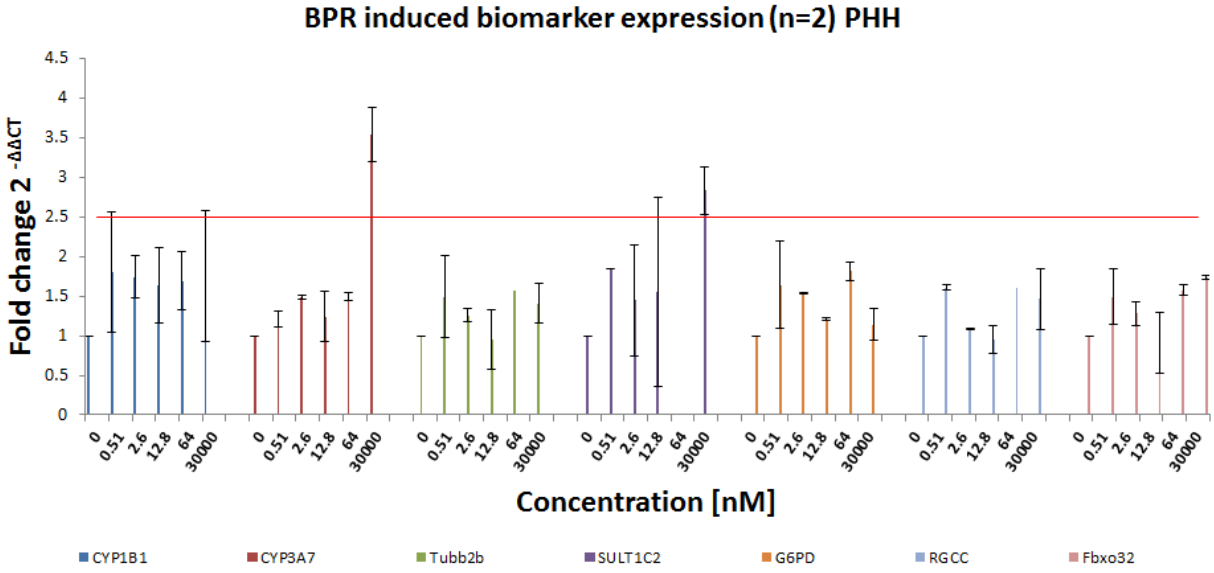
Supplemental Figure 2: Corresponding data to Figure 3.1 summarizing all further incubation conditions besides the high concentration and 24h exposure already shown in Fig. 1. NA: A sample for this time point and concentration was not available. Middle concentration, 2h, 8h and 24h incubation.



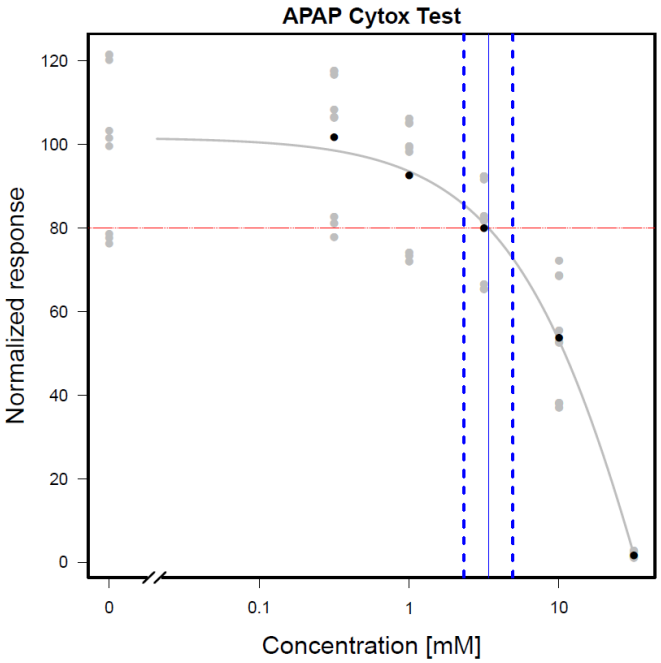
Supplemental Figure 3: Corresponding data to Figure 3.1summarizing all further incubation conditions besides the high concentration and 24h exposure already shown in Fig. 1. NA: A sample for this time point and concentration was not available. High concentration, 2h, 8h and 24h incubation.



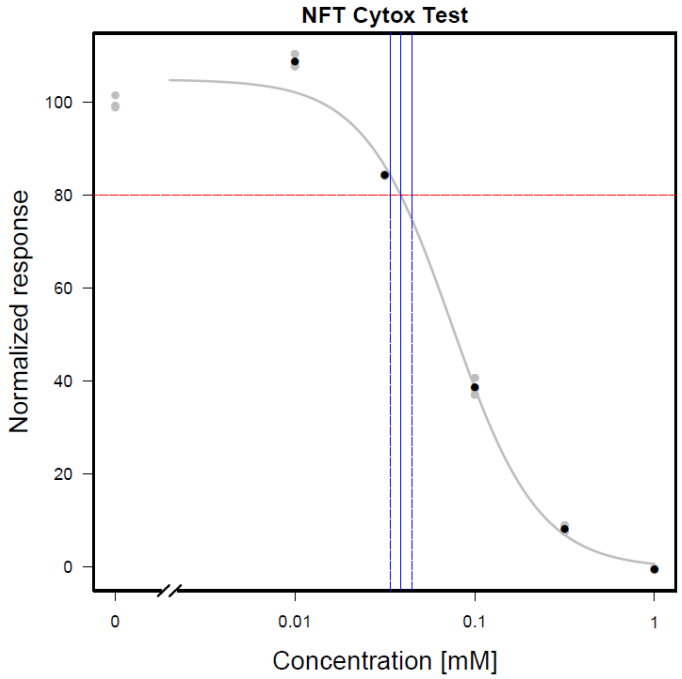
Supplemental Figure 4: Compound induced biomarker expression in HepG2 cells. The presented values (relative expression fold changes) represent mean values that were calculated from 3-5 independent experiments. The error bars illustrate the standard deviation of the independent experiments. A gene was considered to be up regulated when crossing the threshold line, which illustrates a significant increase of at least 2.5 fold change induction. The presented figure shows exemplary the relative expression values for acetaminophen. The diagrams for all compounds are given in digital form on the attached CD.



Supplemental Figure 5: Compound induced biomarker expression in primary human hepatocytes. The presented values (relative expression fold changes) represent mean values that were calculated from 1-3 independent experiments. The error bars illustrate the standard deviation of the independent experiments. A gene was considered to be up regulated when crossing the threshold line, which illustrates a significant increase of at least 2.5 fold change induction. The presented figure shows exemplary the relative expression values for buspirone. The diagrams for all compounds are given in digital form on the attached CD.



Supplemental Figure 6: Cell Titer blue cytotoxicity data for all compounds in HepG2 cells after 48 h of compound exposure. The presented dose-response curves represent data of three independent experiments with three technical replicates each. The cell viability for each concentration is presented as normalized response data, representing the percentage of untreated controls. Gray symbols represent the viability values (fluorescence measurements) for each technical replicate normalized to untreated controls whereas black symbols show the mean values of all technical replicates for each concentration. The blue line indicates the concentration which causes 20 % loss of viability (red line). The dashed blue lines give the 95 % confidence intervals for this concentration. Estimated concentration values for 80 % viability are given in Supplemental Table 6. The presented diagram shows the cytotoxicity data for acetaminophen. Corresponding figures for all compounds are given in a digital form on the enclosed CD.



Supplemental Figure 7: Cell Titer blue cytotoxicity data for all compounds in primary human hepatocytes cells after 48 h of compound exposure. The presented dose-response curves represent data of 1-2 independent experiments with three technical replicates each. The cell viability for each concentration is presented as normalized response data, representing the percentage of untreated controls. Gray symbols represent the viability values (fluorescence measurements) for each technical replicate normalized to untreated controls whereas black symbols show the mean values of all technical replicates for each concentration. The blue line indicates the concentration which causes 20 % loss of viability (red line). The dashed blue line gives the 95 % confidence interval for this concentration. Estimated concentration values for 80 % viability are given in Supplemental Table 7. The presented diagram shows the cytotoxicity data for nitrofurantoin. Corresponding figures for all compounds are given in a digital form on the enclosed CD.

6.2 Supplemental tables

Supplemental Table 1: Matrix of the tested compounds. The table gives full and abbreviated compound names as well as the concentration in µM (µg/mL, µg/kg) and the number of independent replicates of gene array data available after incubation with a low, middle and high concentration for 2h, 8h and 24h. This table is available only in digital form on the enclosed CD**.**

Supplemental Table 2: List of compounds that deregulate (2-fold up or down compared to control) less than 20 genes in total (i.e. at the low, middle and high concentration). Compounds tested at only two concentrations were not considered here. **A** 2h, **B** 8h, **C** 24h. The term in brackets indicates the direction of the deregulation of the genes**.**

A. Compounds that up regulate genes after 2 hours of exposure	
Abbreviation	**Compound name**
ADP	adapin
AM	amiodarone
ANIT	naphthyl isothiocyanate
APAP	acetaminophen
APL	allopurinol
ASA	aspirin
AZP	azathioprine
BBZ	bromobenzene
CBZ	carbamazepine
CCL4	carbon tetrachloride
CFB	clofibrate
CIM	cimetidine
CMA	coumarin
CPA	cyclophosphamide
DFNa	diclofenac
DZP	diazepam
ET	ethionine
FP	fluphenazine
GBC	glibenclamide
GF	griseofulvin
GFZ	gemfibrozil
HCB	hexachlorobenzene
HPL	haloperidol
IM	indomethacin
INAH	isoniazid
KC	ketoconazole
LBT	labetalol
LS	lomustine
MTS	methyltestosterone
NFT	nitrofurantoin
PH	perhexiline
PhB	phenylbutazone
PHE	phenytoin
PTU	propylthiouracil
RIF	rifampicin
SS	sulfasalazine
TAA	thioacetamide
TC	tetracycline
TRZ	thioridazine
WY	WY-14643

Compounds that down regulate genes after 2 hours of exposure	
Abbreviation	**Compound name**
ADP	adapin
AM	amiodarone
ANIT	naphthyl isothiocyanate
ASA	aspirin
AZP	azathioprine
BBr	benzbromarone
BBZ	bromobenzene
CBZ	carbamazepine
CCL4	carbon tetrachloride
CFB	clofibrate
CIM	cimetidine
CMA	coumarin
CPA	cyclophosphamide
CPZ	chlorpromazine
DFNa	diclofenac
ET	ethionine
FP	fluphenazine
FT	flutamide
GBC	glibenclamide
GF	griseofulvin
GFZ	gemfibrozil
HCB	hexachlorobenzene
HPL	haloperidol
IM	indomethacin
KC	ketoconazole
LBT	labetalol
LS	lomustine
MP	methapyrilene
MTS	methyltestosterone
NFT	nitrofurantoin
OPZ	omeprazole
PH	perhexiline
PhB	phenylbutazone
PHE	phenytoin
PTU	propylthiouracil
RIF	rifampicin
SS	sulfasalazine
TAA	thioacetamide
TC	tetracycline
TRZ	thioridazine
WY	WY-14643

B. Compounds that up regulate genes after 8 hours of exposure	
Abbreviation	**Compound name**
2NF	2-nitrofluorene
AAA	acetamide
AM	amiodarone
APL	allopurinol
ASA	aspirin
BBZ	bromobenzene
BDZ	bendazac
BSO	buthionine sulfoximine
CBZ	carbamazepine
CCL4	carbon tetrachloride
CFB	clofibrate
CIM	cimetidine
CMA	coumarin
CPA	cyclophosphamide
CPZ	chlorpromazine
DEX	dexamethasone
FP	fluphenazine
GaN	galactosamine
GBC	glibenclamide
GF	griseofulvin
GFZ	gemfibrozil
HCB	hexachlorobenzene
HPL	haloperidol
IM	indomethacin
MTS	methyltestosterone
NMOR	N-nitrosomorpholine
PHE	phenytoin
RIF	rifampicin
ROT	rotenone
SS	sulfasalazine
TAA	thioacetamide
TC	tetracycline
TMD	trimethadione
TRZ	thioridazine
TZM	triazolam
WY	WY-14643

Compounds that down regulate genes after 8 hours of exposure	
Abbreviation	**Compound name**
2NF	2-nitrofluorene
AAA	acetamide
AM	amiodarone
APL	allopurinol
ASA	aspirin
BBZ	bromobenzene
BDZ	bendazac
BSO	buthionine sulfoximine
CBZ	carbamazepine
CCL4	carbon tetrachloride
CFB	clofibrate
CIM	cimetidine
CMA	coumarin
CPA	cyclophosphamide
CPZ	chlorpromazine
CSA	cyclosporine A
DEX	dexamethasone
FP	fluphenazine
GBC	glibenclamide
GF	griseofulvin
GFZ	gemfibrozil
HCB	hexachlorobenzene
HPL	haloperidol
IM	indomethacin
MTS	methyltestosterone
NMOR	N-nitrosomorpholine
PhB	phenylbutazone
PHE	phenytoin
RIF	rifampicin
ROT	rotenone
SS	sulfasalazine
TAA	thioacetamide
TC	tetracycline
TMD	trimethadione
TRZ	thioridazine
TZM	triazolam
WY	WY-14643

C. Compounds that up regulate genes after 24 hours of exposure	
Abbreviation	**Compound name**
2NF	2-nitrofluorene
AAA	acetamide
AM	amiodarone
BBZ	bromobenzene
BSO	buthionine sulfoximine
CFB	clofibrate
CIM	cimetidine
CMA	coumarin
GBC	glibenclamide
GF	griseofulvin
GFZ	gemfibrozil
HCB	hexachlorobenzene
HPL	haloperidol
LS	lomustine
NMOR	N-nitrosomorpholine
PHE	phenytoin
ROT	rotenone
SS	sulfasalazine
TRZ	thioridazine

Compounds that down regulate genes after 24 hours of exposure	
Abbreviation	**Compound name**
AAA	acetamide
AM	amiodarone
ASA	aspirin
BBZ	bromobenzene
BSO	buthionine sulfoximine
CCL4	carbon tetrachloride
CFB	clofibrate
CIM	cimetidine
CMA	coumarin
GBC	glibenclamide
GFZ	gemfibrozil
HCB	hexachlorobenzene
HPL	haloperidol
IM	indomethacin
MTS	methyltestosterone
NMOR	N-nitrosomorpholine
PHE	phenytoin
RIF	rifampicin
ROT	rotenone
SS	sulfasalazine
TAA	thioacetamide
TC	tetracycline
TZM	triazolam
WY	WY-14643

Supplemental Table 3: Progression error indices for each compound (for **A** the up- and **B** the downregulated genes) both the original as well as the modified progression profile error indicator values for the comparison of the low versus middle and middle versus high concentration for the three exposure periods 2h, 8h and 24h. The compounds that were excluded from further analyses due to their progression error profile are marked in red. This table is available only in digital form on the attached CD.

Supplemental Table 4: The overlap between 'differentially expressed liver disease genes' and chemically deregulated genes in vitro, determined by the SV3 lists of differentially expressed genes. The lists of 'differentially expressed liver disease genes' (false discovery rate (FDR) adjusted p-value ≤ 0.05 and fold-change of at least 1.3) results from the comparison of healthy human liver tissue to that of liver diseases. The SV3 (selection value 3) list includes all probe sets where the 3rd highest ranked compound has a fold change of at least 3 at the highest tested concentration for the incubation period of 24h. The sheets list for **A** NASH (Non-alcoholic steatohepatitis), **B** liver cirrhosis and **C** hepatocellular carcinoma the up- and down-regulated genes in the overlap. The genes can be identified by their Gene Symbol-ID. This table is available only in digital form on the attached CD.

Supplemental Table 5: Concentrations used for the Cell Titer Blue cytotoxicity tests for all compounds.

Compounds	C1	C2	C3	C4	C5	Solvent	Stock solution	Final DMSO conc.
Acetaminophen	316µM	1mM	3.16mM	10mM	31.6mM	Medium		
Aspirin	316µM	1mM	3.16mM	10mM	31.6mM	DMSO	6.32M	0.50%
Buspirone	100µM	316µM	1mM	3.16mM	6.32mM	Medium		
Carbamazepine	31.6µM	100µM	316µM	1mM	3.16mM	DMSO	316mM	1%
Chlorpheniramine	31.6µM	100µM	316µM	1mM	3.16mM	Medium		
Clonidine	100µM	316µM	1mM	3.16mM	10mM	Medium		
Diclofenac	31.6µM	100µM	316µM	1mM	3.16mM	DMSO	316mM	1.00%
Famotidine	100µM	316µM	1mM	3.16mM	10mM	DMSO	1 M	1%
Hydroxyzine	10µM	31.6µM	100µM	316µM	1mM	Medium		
Isoniazid	1mM	3.16mM	10mM	31.6mM	100mM	Medium		
Ketoconazole	3.16µM	10µM	31.6µM	100µM	316µM	DMSO	63.2mM	0.50%
Labetalol	3.16µM	10µM	31.6µM	100µM	316µM	Medium		
Levofloxacin	100µM	316µM	1mM	3.16mM	10mM	Medium		
Melatonin	100µM	316µM	1mM	3.16mM	10mM	DMSO	1 M	1%
Nimesulide	125µM	250µM	500µM	1mM	2mM	DMSO	200mM	1%
Nitrofurantoin	10µM	31.6µM	100µM	316µM	1mM	DMSO	100mM	1%
Phenylbutazone	31.6µM	100µM	316µM	1mM	3.16mM	DMSO	632mM	0.50%
Promethazine	10µM	31.6µM	100µM	316µM	1mM	Medium		
Propranolol	30µM	40µM	60µM	80µM	100µM	Medium		
Rifampicin	31.6µM	100µM	316µM	1mM	3.16mM	DMSO	6.32M	0.20%
Valproic acid	316µM	1mM	3.16mM	10mM	31.6mM	Medium		

Supplemental Table 6: Estimated concentrations causing 20 % loss of viability for all compounds in HepG2 cells. The values were calculated based on the fitted dose-response curves (Supplemental Figure 6).

Compound		Concentration for 80% viability [mM]	concentration lower 95 % confidence interval [mM]	concentration upper 95 % confidence interval [mM]
Acetaminophen	APAP	3.377	2.321	4.914
Aspirin	ASP	6.620	5.224	8.390
Buspirone	BPR	0.318	0.257	0.392
Carbamazepine	CBZ	0.416	0.275	0.630
Chlorpheniramine	CHL	0.090	0.083	0.099
Clonidin	CLON	0.816	0.672	0.991
Diclofenac	DFN	0.364	0.236	0.560
Famotidine	FAM	1.892	1.365	2.624
Hydroxizine	HYZ	0.053	0.029	0.095
Isoniazid	INAH	2.582	0.558	11.953
Ketoconazole	KC	0.052	0.039	0.069
Labetalol	LAB	0.020	0.009	0.043
Levofloxacine	LEV	0.322	0.247	0.421
Melatonin	MEL	2.986	1.128	7.900
Nitrofurantoin	NFT	0.011	0.004	0.027
Nimesulide	NIM	0.246	0.196	0.308
Phenylbutazone	PhB	0.216	0.158	0.294
Promethazine	PMZ	0.026	0.022	0.031
Rifampicin	RIF	1.198	0.431	3.332
Valproic acid	VPA	1.482	1.010	2.174

Supplemental Table 7: Estimated concentrations causing 20 % loss of viability for all compounds in primary human hepatocytes. The values were calculated based on the fitted dose-response curves (Supplemental Figure 7).

Compound		Concentration for 80% viability [mM]	concentration lower 95 % confidence interval [mM]	concentration upper 95 % confidence interval [mM]
Acetaminophen	APAP	0,92	0,82	1,03
Buspirone	BPR	0,84	0,70	1,01
Diclofenac	DFN	0,23	0,18	0,29
Famotidine	FAM	3,69	0,02	724,58
Hydroxizine	HYZ	0,09	0,06	0,15
Ketoconazole	KC	0,04	0,03	0,06
Labetalol	LAB	0,08	0,00	46,84
Melatonin	MEL	8,34	0,03	2283,82
Nitrofurantoin	NFT	0,04	0,03	0,04
Nimesulide	NIM	1,23	0,49	3,10
Rifampicin	RIF	0,50	0,26	0,94
Valproic acid	VPA	27,09	6,97	105,22

Supplemental Table 8: Raw data for gene expression quantification in HepG2 cells. This list gives the Ct values, which the calcualtion of expression fold changes is based on. Samples were obtained from 3-4 independent experiments; for each experiment 5 different concentrations plus untreated controls were tested. Each sample was measured in 2-4 technical replicates. Relative expression of the 7 biomarker genes was determined by normalization to the expression of the housekeeping gene GAPDH and in relation to untreated controls. The presented list in this printed version provides an insight into the Ct values of acetaminophen deregulated genes. Raw data for all compounds and all tested concentrations are shown in digital form on the attached CD.

Sample name	Concentration	Ct-Werte values for each gene after 24h of treatment with acetaminophen							
		GAPDH	Cyp1B1	Cyp3A7	Tubb2b	Sult1C2	G6PD	RGCC	Fbxo32
V03_07	0 µM	16.985	31.553	29.863	24.540	26.870	21.348	30.722	30.098
V03_07	0 µM	14.926	31.700	30.035	24.594	26.756	21.298	30.798	29.732
V03_07	0 µM	14.866							
V02_19	0 µM	15.572	32.755	30.894	24.835	27.339	21.785	30.145	29.867
V02_19	0 µM	15.772	32.560	30.980	24.804	27.295	21.756	29.917	29.919
V02_19	0 µM	16.388							
V01_13	0 µM	14.608	32.404	30.277	23.735	27.384	20.629	27.896	28.813
V01_13	0 µM	15.466	32.498	30.038	23.665	27.153	20.601	28.199	29.093
V01_13	0 µM	14.847							
V01_19	0 µM	14.528	33.461	30.305	23.439	26.959	20.842	28.661	28.771
V01_19	0 µM	14.382	33.293	30.425	23.286	26.685	20.789	28.622	28.876
V01_19	0 µM	14.587							
V03_08	8 µM	18.729	32.208	29.892	24.450	27.117	21.475	30.008	29.380
V03_08	8 µM	14.956	32.542	29.832	24.440	27.134	21.343	29.994	29.654
V03_08	8 µM	14.901							
V02_20	8 µM	15.759	32.844	30.937	24.618	27.289	21.878	29.927	30.357
V02_20	8 µM	15.934	32.861	31.178	24.716	27.179	21.838	30.366	30.344
V02_20	8 µM	17.409							
V01_14	8 µM	14.586	31.824	30.465	23.224	27.171	20.307	28.281	28.657
V01_14	8 µM	14.604	31.880	30.229	23.236	27.165	20.216	28.297	28.747
V01_14	8 µM	14.959							
V01_20	8 µM	14.358	32.303	29.710	22.788	26.817	20.691	28.636	28.815
V01_20	8 µM	14.595	32.158	29.925	22.882	26.805	20.701	28.479	28.689
V01_20	8 µM	14.720							

Supplemental Table 9: Raw data for gene expression quantification in primary human hepatocytes. This list gives the Ct values, which the calculation of expression fold changes is based on. Samples were obtained from 3-4 independent experiments; for each experiment 5 different concentrations plus untreated controls were tested. Each sample was measured in 2-4 technical replicates. Relative expression of the 7 biomarker genes was determined by normalization to the expression of the housekeeping gene GAPDH and in relation to untreated controls. The presented list in this printed version provides an insight into the Ct values of valproic acid deregulated genes. Raw data for all compounds as well as the remaining Ct values for valproic acid induced genes are shown in digital form on the attached CD.

Sample name	Concentration	Ct values GAPDH		Ct values Cyp1B1		Ct values CYP3A7		Ct values G6PD	
		Repl.1	Repl.2	Repl.1	Repl.2	Repl.1	Repl.2	Repl.1	Repl.2
V160_25	0 µM	15.873	16.500	26.332	26.784	23.631	23.570	26.756	27.064
V160_26	8 µM	17.479	17.722	27.495	28.107	24.553	25.117	28.240	27.911
V160_27	40 µM	17.130	16.483	25.775	26.108	23.131	23.056	26.845	27.508
V160_28	200 µM	16.671	16.713	25.835	26.684	22.967	23.171	27.050	26.591
V160_29	1 mM	15.730	16.480	24.708	24.961	22.793	22.215	26.136	26.052
V160_30	5 mM	16.404	16.658	26.508	26.689	24.251	24.380	26.617	26.444
V164_7	0 µM	17.833	17.783	28.355	28.330	31.314	31.036	28.379	28.295
V164_8	8 µM	17.789	17.701	28.882	28.899	32.298	32.366	28.613	28.588
V164_9	40 µM	17.428	17.591	28.287	28.482	30.964	31.231	27.768	28.199
V164_10	200 µM	16.583	16.797	27.316	27.122	30.244	30.736	27.096	27.471
V164_11	1 mM	16.949	16.692	26.624	26.944	29.588	29.266	25.011	24.979
V164_12	5 mM	17.768	17.318	25.880	25.954	30.245	29.892	22.092	21.921
V165_25	0 µM	16.319	16.410	28.642	28.678	26.523	26.456	28.058	27.869
V165_26	8 µM	17.595	17.604	28.630	28.519	26.654	26.892	27.687	27.311
V165_27	40 µM	16.234	16.344	28.854	28.745	25.060	25.668	27.503	27.481
V165_28	200 µM	16.559	16.519	28.457	28.810	24.790	24.713	26.896	26.934
V165_29	1 mM	16.342	16.392	27.706	27.981	23.609	23.524	24.620	24.740
V165_30	5 mM	16.161	16.023	27.400	27.565	26.343	26.511	22.437	22.411

7 List of figures

8 List of tables

9 Publications

9.1 Articles

Gene networks and transcription factor motifs defining the differentiation of stem cells into hepatocyte-like cells. Godoy P, Schmidt-Heck W, Natarajan K, Lucendo-Villarin B, Szkolnicka D, Asplund A, Bjorquist P, Widera A, **Stöber R**, Campos G, Hammad S, Sachinidis A, Damm G, Weiss TS, Nüssler A, Synnergren J, Edlund K, Küppers-Munther B, Hay D, Hengstler JG. **Journal of Hepatology 10/2015; 63(4):934-42**

A transcriptome-based classifier to identify developmental toxicants by stem cell testing: design, validation and optimization for histone deacetylase inhibitors. Rempel E, Hoelting L, Waldmann T, Balmer NV, Schildknecht S, Grinberg M, Das Gaspar JA, Shinde V, **Stöber R**, Marchan R, van Thriel C, Liebing J, Meisig J, Blüthgen N, Sachinidis A, Rahnenführer J, Hengstler JG, Leist M. **Archives of Toxicolgy 09/2015; 89(9):1599-618.**

Toxicogenomics directory of chemically exposed human hepatocytes. Grinberg M and **Stöber RM**, Edlund K, Rempel E, Godoy P, Reif R, Widera A, Madjar K, Schmidt-Heck W, Marchan R, Sachinidis A, Spitkovsky D, Hescheler J, Carmo H, Arbo MD, van de Water B, Wink S, Vinken M, Rogiers V, Escher S, Hardy B, Mitic D, Myatt GJ, Waldmann T, Mardinoglu A, Damm G, Seehofer D, Nüssler A, Weiss TS, Oberemm A, Lampen A , Schaap MM, Luijten M, van Steeg H, Thasler WE, Kleinjans JCS, Stierum RH, Leist M, Rahnenführer J, Hengstler JG. **Archives of Toxicology 11/2014; 88(12):2261-87**

Fluorescence-based recombination assay for sensitive and specific detection of genotoxic carcinogens in human cells. Ireno IC, Baumann C, **Stöber R**, Hengstler JG, Wiesmüller L. **Archives of Toxicology 03/2014; 88(5):1141-59**

Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. Godoy P, Hewitt NJ, Luebke Wheeler J, Gibson A, Eakins R, Goldring CEP, Naisbitt DJ, Rowe C, Park BK, Damm G, Glanemann M, Keitel V, Häussinger D, Singh B, Choi Y-J, Cho C-S, Cameron NR, Mwinyi J, Kullak-Ublick GA, Guyot C, Stieger B, Bode JG, Albrecht U, Fonsato V, Camussi G, Lutz A, Schmich K, Merfort I, Olinga P, Ramachandran A, Jaeschke H, Fraczek J, Bolleyn J, Vinken M, Vanhaecke T, Rogiers V, Burkhard B, Nüssler AK, Ito K, Sugiyama Y, Hrach J, Tetta C, Messner S, Kelm JM, Matz-Soja M, Böttger J, Gebhardt R, Pampaloni F, Ansari N, Stelzer EHK, Braeuning A, Schwarz M, Sá Ferreira K, Borner C, Hoehme S, Drasdo D, Widera A, **Stöber R**, Schelcher C, Thasler WE, Xu JJ, Hewitt P, Meyer C, Dooley S, Maltman DJ, Hayward A, Przyborski SA, Hallifax D, Houston JB, LeCluyse EL, Bhattacharya S, McMullen P, Woods CG, Yarborough KM, Pluta L, Lu P, Dong J, Pi J, Andersen ME, Budinsky RA, Rowlands JC, Dahmen U, Dirsch O; Gómez-Lechón MJ, Donato MT, Holzhütter HG, Hellerbrand C, Hengstler JG**. Archives of Toxicology 08/2013; 87(8):1315-530**

Piperazine designer drugs present cytotoxicity to primary rat hepatocytes. Dutra Arbo M, Melega S, Schug M, **Stöber R**, Hengstler JG, de Lourdes Bastos M, Carmo H. **Toxicology Letters 08/2013; 221:S157**

Pharmacokinetics explain in vivo/in vitro discrepancies of carcinogen-induced gene expression alterations in rat liver and cultivated hepatocytes. Schug M and **Stöber R**, Heise T, Mielke H, Gundert-Remy U, Godoy P, Reif R, Blaszkewicz M, Ellinger-Ziegelbauer H, Ahr H-J, Selinski S, Günther G, Marchan R, Sachinidis A, Nüssler A, Oberemm A, Hengstler JG **Archives of Toxicology 12/2012; 87(2):337-45**

Patients with liver cirrhosis demonstrate high il15 mRNA expression. Vogt R, Godoy P, Campos G, **Stöber R**, Hengstler JG, Schlagheck JMF, Schlitt HJ, Melter M, Weiss TS. **Journal of Hepatology 03/2011; 54.**

Epithelial mesenchymal Transition (EMT) associated transcription factor Goosecoid is highly expressed in human hepatocellular carcinoma (HCC). Schlagheck JM, Godoy P, Campos G, **Stöber R**, Hengstler JG, Vogt R, Melter M, Weiss TS **Zeitschrift für Gastroenterologie 01/2011; 49(01)**

IL-15 mRNA expression is associated with liver cirrhosis. Vogt R, Godoy P, Campos G, **Stöber R**, Hengstler JG, Schlitt HJ, Melter M, Weiss TS **Zeitschrift für Gastroenterologie 01/2011**

Epithelial to mesenchymal transition in liver regeneration . Widera A, Campos G, Begher-Tibbe B, Günther G, Garcia-Perez C, **Stöber R**, Hengstler JG, Godoy P **Zeitschrift für Gastroenterologie 01/2011; 49(01)**

A novel oxygenase from Pleurotus sapidus transforms valencene to nootkatone. Fraatz MA, Riemer SJL, **Stöber R**, Kaspera R, Nimtz M, Berger RG, Zorn H **Journal of Molecular Catalysis B Enzymatic 12/2009; 61(3-4-61):202-207.**

9.2 Book chapters

Transcriptomics of Hepatocytes Treated with Toxicants for Investigating Molecular Mechanisms Underlying Hepatotoxicity. Shinde V, **Stöber R**, Nemande H, Sotiriadon I, Hescheler J, Hengstler JG, Sachinidis A. **Protocols in In Vitro Hepatocyte Research, Methods in Molecular Biology, Vol. 1250** edited by Vinken M, Rogiers V (Eds, 12/2014: chapter Transcriptomics of Hepatocytes Treated with Toxicants for Investigating Molecular Mechanisms Underlying Hepatotoxicity: pages 225-241; A product of Humana Press, Springer., ISBN: 978-1-4939-2074-7

9.3 Guest editorials

Transcriptomics signature for drug induced steatosis. Stöber R EXCLI Journal 2015;14:1259-1260

Identification of carcinogens by a selected panel of DNA damage response associated genes. Stöber R EXCLI Journal 2015;14:1294-1296

Drug induced mitochondrial impairment in liver cells. Stöber R EXCLI Journal 2015;14:1297-1299

Transcriptome based differentiation of harmless, teratogenetic and cytotoxic concentration ranges of valproic acid. Stöber R EXCLI Journal 01/2014; 13:1281-1282

9.4 Contribution on congresses

Poster: **In vitro based prediction of human hepatotoxicity. Stöber R** Grinberg M, Escher S, Rahnenführer J and Hengstler JG. Safety Evaluation Ultimately Replacing Animal Testing (*S*EURAT) congress 4th of December 2015 in Brussels, Belgium

Poster: **Valproic acid case study: Detection and verification of biomarkers by using a read across approach. Stöber R**, Richarz A, Myatt GJ, Klipp A, van de Water B, Hengstler JG, Escher SE. Safety Evaluation Ultimately Replacing Animal Testing (SEURAT) congress 4th of December 2015 in Brussels, Belgium

Talk: **Biomarker study to predict hepatotoxic blood concentrations. Stöber R** DETECTIVE General Assembly No.6, 24.-25. November 2015 in Brussels, Belgium.

Talk: **In vitro based prediction of human hepatotoxicity. Stöber R**, Grinberg M, Escher SE, Rahnenführer J, Hengstler JG. European Association for Alternatives of Animal Testing (EUSAAT) 20-23 September in Linz, Austria

Poster: **The use of AOPs in risk assessment: Development of biomarker Based on a read across use case on VPA analogues in the Detective project. Stöber R**, Richarz A, Steger-Hartmann T, Myatt GJ, Hengstler JG, Escher SE. Congress of the European Societies of Toxicology (EUROTOX) 13-16 September 2015 in Porto, Portugal.

Poster: **Toxicogenomics directory of chemically exposed human hepatocytes. Stöber R**, Grinberg M, Edlund K, Rahnenführer J, Hengstler JG. Tag der Chemie 2015, 13th of February 2015 in Dortmund, Germany.

Poster: **Toxicoproteomics: Development of a robust cultivation system for primary human liver cells and establishment of proteome analysis via SAX fractionation and MS analysis**. **Stöber R**, Dietz L, Godoy P, Reif R, Damm G, Nüssler A, Hengstler JG. Safety Evaluation Ultimately Replacing Animal Testing (SEURAT) congress on 5-6th March 2013 in Lisbon, Portugal.