High throughput microscopy of mechanism-based reporters in druginduced liver injury

Hiemstra, S.W.

Chapter 4

A 3D HepG2 GFP reporter platform to screen for drug-induced liver injury liabilities

Steven Hiemstra[1], Steven Wink[1], Karen van den Nieuwendijk[1], Sreenivasa Ramaiaghari[1], Anita Dankers[2], Hans de Bont[1] and Bob van de Water[1]

[1]Division of Toxicology, Leiden Academic Centre for Drug Research (LACDR), Leiden University, The Netherlands

[2]Janssen Pharmaceutical Companies of Johnson &Johnson, Department of Pharmacokinetics, Dynamics and Metabolism, Beerse, Belgium

Manuscript in preparation

Abstract

Drug-induced liver injury is a major problem in the clinic and in drug development. Adaptive stress response pathways play a key role in the switch between adaptation and adversity, thus are important players in cell death and in liver injury. Previously, we have established a HepG2 GFP reporter platform to monitor early adaptive stress response activation following drug treatment. HepG2 cells are often used in primary toxicity screening. However, the disadvantage of HepG2 cells is their low metabolizing capacity. Therefore, we cultured our GFP reporter cell lines representing Nrf2 activation (Srxn1 and NQO1), unfolded protein response (BiP and CHOP) and DNA damage response (p21 and Btg2) as 3D spheroids. We confirmed that these six GFP reporter lines differentiated in a similar manner as wild type HepG2 cells. We developed a method to assess and quantitatively analyze GFP reporter intensity by using high throughput confocal imaging and an image analysis pipeline. We optimized the application of the 3D spheroid reporters by evaluating a compound set that activates these stress response pathways at the transcriptional level in primary human hepatocytes. Next, a large set of drugs with known DILI liability were tested in these six 3D GFP reporters in single 48 hr treatment or 6 day repeated treatment at five different concentration. Most stress response activation was observed after 6 day repeated treatment, with the BiP and Srxn1 reporter most responsive. Drugs with most-DILI-concern clustered together. Several compounds that did not show GFP reporter activation in 2D monolayer, demonstrated clear stress response activation in 3D spheroids. Together these six GFP-reporters showed a high sensitivity and high specificity for determining DILI liability.

Introduction

Drug Induced Liver-injury (DILI) is a major problem in the clinic, as 50% of all liver failures are caused by drugs[7]. The same study shows in 13% of the cases that liver failure is caused by idiosyncratic drug reactions. Idiosyncratic DILI occurs only in rare cases, in approximately 0,1% to 0,01% of patients administrating the drug and has a variable latency time, ranging from a few days up to a few years after first administration of the drug. These properties makes it complicated to predict which drug will eventually cause idiosyncratic DILI in the clinic in which patients. Therefore, idiosyncratic DILI onset is often missed in preclinical toxicity assessment during drug development, hence DILI is the leading cause for drug withdrawal from the market[8].

Traditionally, *in vitro* tools have been used during preclinical toxicity to screen for cell viability in simple end-point measurements. However, these assays do not reflect the early cell state changes occurring when cells adapt to the insult induced by the drug. This is valuable information as idiosyncratic DILI only occurs in susceptible individuals. Therefore, over the past decade transcriptomics approaches have unraveled some of the mechanisms of action by which the drugs with DILI liabilities affect the cellular homeostasis[48,51,52]. While transcriptomics has

allowed the identification of the global cellular changes at the transcriptional level following drug exposure, it has deemed less suitable in the evaluation of pharmaceutical safety. Limitations in concentration and time courses has prevented the clear identification of benchmark concentrations for point of departure of cellular perturbations. Furthermore, while mRNA expression levels can be indicative for cellular function, they do not necessarily represent the changes in the cellular protein landscape after drug exposure. Also, not only up- or down-regulation of genes affects the overall biological outcome of drug exposure, e.g. as a consequence also post-translational modification and changes in cellular location determine protein function and hence the cell biological. To overcome some of these limitations, we have built a fluorescent based proteomic platform to screen for critical early cell state changes after drug exposure[145]. Using bacterial artificial chromosome recombineering technology[96], we tagged individual components of three different stress response pathways with green fluorescent protein (GFP): oxidative stress, unfolded protein response and DNA damage[145]. Using high content live cell confocal microscopy we can quantitatively evaluate the dynamics of the activation of individual stress response pathway components at the individual cell level.

Different *in vitro* liver cell lines are used in toxicity screening. Primary human hepatocytes (PHH) are viewed the 'golden standard' in *in vitro* liver toxicity screening, as they have the best representation of the human liver in terms of liver functional properties[163]. Major shortcomings of the use of primary human hepatocytes are the limited batch size, donor variability and high costs[4]. The liver carcinoma cell line HepG2 is a cheap, easy to use cell line with unlimited lifespan, which makes it highly suitable for high throughput toxicity screening. We have used the HepG2 cell line to establish the BAC-GFP reporter platform. These reporters show excellent performance in detecting mode of action of stress response activation. A caveat of HepG2 cells cultured in a conventional 2D monolayer are the high proliferative capacity and dedifferentiated phenotype. Thus, limited drug metabolism capacity or liver properties like bile duct formation and albumin expression are present. In addition, 2D cultured HepG2 are not optimally suited for repeated dose exposures, due to the high proliferative capacity.

Previously, we reported an advanced HepG2 3D spheroid model with a highly differentiated liver phenotype, that is suitable for high throughput screening[27]. By culturing HepG2 cells in a hydrogel, cells stop proliferating and start differentiating with neonatal-hepatocytes-like properties. Within 21 days HepG2 spheroids are formed which start to increase the expression of phase I and II metabolizing enzymes and phase III transporters. Furthermore, liver specific functions as bile duct formation and glycogen storage are increased. Given the stability of these HepG2 spheroid cultures, repeated treatment over a prolonged period of time is feasible.

The goal of the current study was to evaluate the application of the HepG2 stress response reporters in 3D spheroid systems in combination with high throughput microscopy. The data demonstrate that the GFP reporter platform can be used effectively and efficiently for repeated

dose exposure for the detection of stress response activation caused by DILI compounds.

Material and methods

Reagents and antibodies

All compounds used were obtained from Sigma (Zwijndrecht, The Netherlands), except for: acetaminophen, amiodarone, bosentan, clozapine, diclofenac, entacapone, fialuridine, metformin, perhexilline, pioglitazone, tolcapone, troglitazone and ximelagatran (MIP DILI consortium), bendazac (kind gift from Dr. Dankers, Janssen Pharmaceuticals), adefovir (Shanghai PI chemicals), azathioprine, bicalutamide, buspirone, busulfan, colchicine, clotrimazole, danazole, furosemide, meclizine, methimazole, mexiletine, nitrofurantoin, nitrofurazone, primidone, procyclidine, propylthiouracil, sulindac, tacrine, thioridazine, trazodone, zafirlukast (kind gift from Dr. Weida Tong, NCTR-FDA, all vendor are listed in supplemental table 3). All compounds were dissolved in DMSO, except for metformin, venlafaxine, methapyrilene, fluphenazine, buthionine sulfoximine, lomustine and isoniazid (all PBS), acetaminophen and phenobarbital (DMEM).

Anti-human albumin (A80-229F, polycolonal), was obtained from Bethyl Laboratories Inc (Texas, USA), anti-β-Catenin antibody (610153, monoclonal) was obtained from BD Transduction Laboratories (Breda, The Netherlands) and anti-MRP2 (M2III-6) (AB3373, monoclonal) was obtained from Abcam (Cambridge, UK). Goat anti-mouse Alexa-488 (A11001) and goat anti-rabbit Alexa-488 (A11008) were obtained from Molecular Probes (Breda, The Netherlands). Hoechst 33342 (Thermo Scientific, Leiden, The Netherlands) 200 ng/ml was used to stain nuclei of live cells. Hoechst 33258 (Sigma, Zwijndrecht, The Netherlands) 2 µg/ml was used to stain nuclei of fixed spheroids. Rhodamine-phalloidin was obtained from Sigma (Zwijndrecht, The Netherlands).

Cell lines

Human hepatoma cell line HepG2 was obtained from the American type Tissue Culture Collection (ATCC, Wesel, Germany). For experiments HepG2 cells were cultured in phenol red free Dublecco's modified Eagles medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS), 25 U/ml penicillin and 25 µg streptomycin (PSA, Invitrogen).

GFP BAC reporter generation

Human *SRXN1*, *NQO1*, *CDKN1A* (p21), *BTG2*, *DDIT3* (Chop) and *HSPA5* (BiP) were selected and tagged with enhanced-GFP as described previously using the bacterial artificial chromosome (BAC) recombineering technique[96] and stably introduced into HepG2 cells by transfection and 500 µg/ml G418 selection as reported previously[145]. At least 20 of the monoclonal BAC transfected HepG2 colonies were separately grown out and GFP-positive clones with

homogenous expression and detection levels that meet the imaging criteria were selected to complement the BAC-GFP stress response reporter platform[145].

3D cell culturing

Matrigel (cat. No: 354230, Lot: 3164773, 4321005, 3115921 and 3032578) was obtained from BD biosciences (Erembedegem, Belgium). Matrigel batches vary in total protein concentration; a stock of matrigel was diluted to 5 mg protein/ml and used to culture the spheroids. Cells were incubated at 37°C with 5% CO_2. Culture medium was refreshed every 3-4 days. Greiner Bio-one (Alphen aan de Rijn, The Netherlands) µclear 384 well plates were used for culturing and confocal imaging of the spheroids.

Determination of metabolic activity of CYP450 enzymes

A cocktail of five different chemicals (25 µM diclofenac (CYP2C9), 10 µM midazolam (CYP3A4), 100 µM phenacetin (CYP1A2), 500 µM bupropion (CYP2B6) and 15 µM dextromethorphan (CYP2D6)) in methanol was applied to 2D and 3D cultures. Cell culture supernatant was collected at 0, 4, 24 and 48 hours after exposure. LC-MS analysis of metabolites was performed by Janssen Pharmaceutics (Beerse, Belgium) in a similar way to what has been described earlier[164].

Cell treatment and viability

Compound exposures were performed in three different ways: 48 hours single exposure, six days repeated and eleven days repeated exposure. For six or eleven days repeated exposure, every 24 hours the medium was aspirated from the cells and freshly diluted compound in fresh medium was added. 24 hours after the last exposure the imaging was started. For 48 hours single exposure and six days repeated exposure the imaging was started at day 28 after seeding the cells in 384 wells. For eleven days repeated exposure the imaging started at 33 days after seeding. Cell viability was measured in HepG2 wild type cultures, using the ATP-lite luminescence assay kit (Perkin Elmer) as described previously[27]. Cell death was calculated as a percentage of viable cells after compound exposure compared to its vehicle control, assuming all wells have an equal amount of spheroids. All experiments were performed in triplicate independent experiments.

Confocal microscopy image acquisition.

The GFP intensity levels of reporter cell lines were determined using a Nikon TiE2000 confocal laser microscope equipped with an automated stage, perfect focus system and live cell control to ensure 37°C and CO_2 conditions during imaging. Images were acquired with 20X objective (NA = 0.75, Violet Corrected) and 2X zoom. Lasers 405 detecting Hoechst 33342 and 488 detecting GFP were used. An in-house build NisElements-based macro enabled automated

image acquisition of spheroids (supplemental figure 1). First, the macro acquired multiple images, which were stitched together to form a large image almost covering the entire well; software identified the spheroids in the large image based on Hoechst 33342 intensity. Next, the microscope zoomed in on one of the identified spheroids and started generating z-stacks with step size of 15 µm; this allowed the identification of the z-stack with maximal intensity of Hoechst 33342. As a last step the software acquired one image containing Hoechst 33342 and GFP intensity values. A maximum of seven spheroids were imaged per well per replicate treatment.

Gene expression analysis

Open TG-GATES database "Toxicogenomics Project and Toxicogenomics Informatics Project under CC Attribution-Share Alike 2.1 Japan" was used to extract primary human hepatocyte gene expression data as reported previously[145].

Quantitative image analysis and statistics

Image quantification was performed with CellProfiler 2.1.1 (Broad Institute, Cambridge, USA), HDF5 (version 2.10.1), R (version 3.2.2) and R studio (version 0.97.551) (Wink *et al*, manuscript in preparation). Based on the Hoechst image CellProfiler defines two objects: a mask of the entire spheroid and a mask of the single nuclei in the spheroid. The integrated GFP intensity values were determined with the mask of the entire spheroid and normalized for spheroid size by dividing by the mask of single nuclei. Next, possible outliers were removed based on the two-tailed Grubb's test in vehicle DMSO treatment per imaged plate. Each treatment/dose combination was subsequently normalized by the DMSO vehicle control. Final values were obtained by removing possible outliers in obtained fold change values. Fold changes were considered significant as $p<0,05$ in a one-tailed students t-test.

Results

GFP reporter HepG2 cell lines differentiate towards liver-like spheroids in a similar way as wild type HepG2

For generating the stress response reporter platform, BAC-GFP construct were transfected in HepG2 cells followed by selection using G418[145]. Surviving single cell-derived clones were ultimately obtained, thereby possibly introducing variability in the phenotypic characteristics of individual BAC-GFP HepG2 reporters. We reasoned that the capacity to differentiate into 3D HepG2 spheroids with liver-like properties would be a unique feature of a HepG2 cell behavior. In order to verify whether the various GFP reporter cell lines could differentiate into 3D liver-like spheroids in a similar manner as wild type HepG2, we cultured Srxn1-GFP, Nqo1-GFP, Chop-GFP, BiP-GFP, Btg2-GFP and p21-GFP in 3D matrigel. After 28 days all wells were stained for nuclei and F-actin as well as for liver-like properties that are lost in 2D HepG2 monocultures;

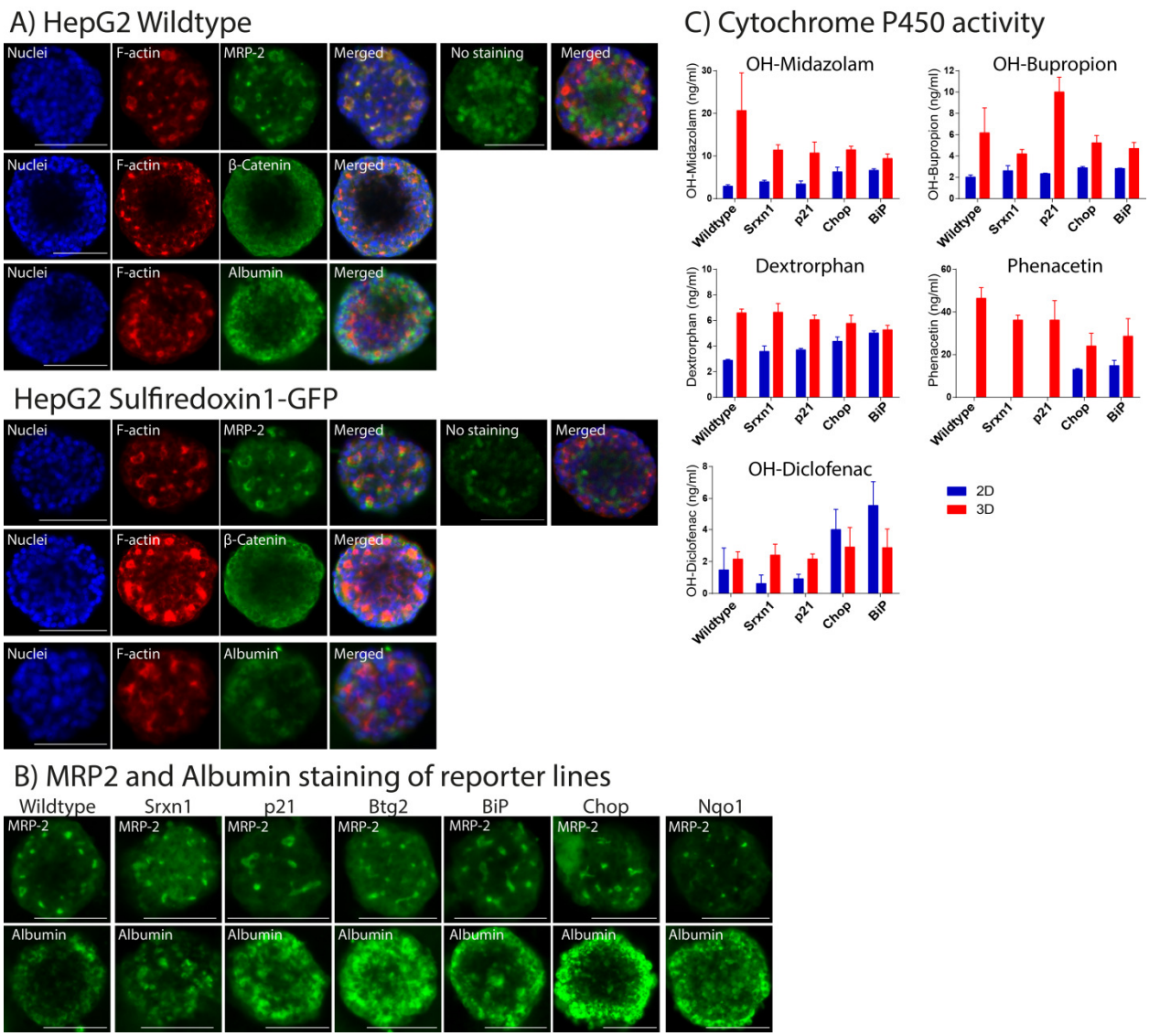


Figure 1: HepG2 GFP reporter cell lines differentiate in 3D spheroids with liver-like properties. A) Wildtype (upper panel) and Srxn1-GFP (lower panel) spheroids are stained with nuclear staining Hoechst$_{33258}$ (blue), F-actin staining rhodamine-phalloidin (red). Immunofluorescent labeling with MRP2 (upper row), β-catenin (middle row) and albumin (lower row) are shown in green. Merged images contain all three stainings (Hoechst, F-actin and GFP labeling). 'No staining' shows basal GFP background signal. White scale bar indicates 100 µm scale. B) GFP labeling of wildtype HepG2, Srxn1, p21, Chop and BiP reporter lines are shown for MRP2 (upper row) and albumin expression (lower row). White bar indicates 100 µm scale. C) Five different metabolites were measured in 2D monolayers and 3D spheroids for 48 hours. 1-OH-midazolam was measured as a measurement for CYP3A4 function, dextrorphan as a measurement for CYP2D6 function, phenacetin as a measurement of CYP1A2 function, OH-bupropion as a measurement for CYP2B6 function and 4-OH-diclofenac as a measurement for CYP2C9 function. Graphs show averages of three technical replicates and is normalized to the average amount of cells in 3D and 2D.

e.g. bile duct formation (MRP2), albumin expression and basal-lateral polarization (β-catenin). All BAC-GFP HepG2 reporter cell lines analyzed acquired the same type of differentiation as HepG2 wildtype (Figure 1A and 1B). 3D HepG2 spheroids have an improved metabolizing capacity compared to 2D monolayers. Therefore, next we evaluated the activity of some

CYP450 enzymes in a subset of the reporter lines (Srxn1-GFP, p21-GFP, Chop-GFP and BiP-GFP). For this we measured the CYP450 metabolic capacity of the cells in 2D and 3D cultures using a mixture of five different substrates and LC-MS analysis. The compounds specific CYP450 enzyme activity allowed the exact determination of the CYP450 enzyme activity. Figure 1C shows that an enhanced metabolizing capacity of CYP3A4, CYP2C9, CYP2B6, CYP1A2 and CYP2D6 in 3D cultures was observed for all reporter cell lines tested, which similar to wild type HepG2 spheroids. Taken together, these data shows a differentiated liver-like phenotype of all GFP reporter cell lines.

Detection of specific GFP reporter activity in 3D spheroids

We previously determine the specific activation of GFP stress response reporter activity acquisition over time in 2D monolayers[145]. To determine the functionality of our reporters in 3D spheroids, we used well-established reference chemicals that specifically activate the stress response pathways: tert-butylhydroquinone (tBHQ) for the oxidative stress response pathway (Srxn1-GFP and Nqo1-GFP); tunicamycin for the UPR pathway (BiP-GFP and Chop-GFP) and aflatoxin B1 for the DNA damage response pathway (p21-GFP and Btg2-GFP). The effect of the reporter activity was evaluated using confocal microscopy and analyzed using our in-house built analysis pipeline[145]. All six reporter cell lines demonstrated strong and significant upregulation of GFP intensity after stimulation with reference inducers (figure 2A and 2B). This demonstrates that all the BAC-GFP reporter cell lines remain their sensing properties under 3D culture conditions which can be detected and quantified using our confocal microscopy settings.

TG-GATES selected compounds with target gene specific responses demonstrate high sensitivity of reporter cell lines

As a next step we systematically compared the reporter responses of the 3D spheroid system to primary human hepatocytes gene expression data from the TG-GATEs database. The TG-GATEs database contains gene expression data of ~150 compounds exposed to primary human hepatocytes in three different concentrations and three different time points. For each of the six GFP reporter genes we ranked the TG-GATES compounds (at high dose, 24 hours treatment) based on fold change. For each reporter ten compounds were selected (except for Chop-GFP; only nine compounds were used) originating from the top 12 ranked compounds that caused upregulation of mRNA levels of the respective genes in PHH. Since a major advantage of 3D HepG2 spheroid cultures is the possibility to perform repeated treatments[27], we also assessed which exposure regimen corresponded best with the PHH TG-GATEs transcript data: 48 hours single exposure, six days repeated exposure or eleven days repeated exposure treatment conditions. For direct comparison we used the same three concentrations used in TG-GATEs (supplemental table 1). To apply the 3D GFP reporter system for pharmaceutical screening automated imaging is preferred. Therefore, we first established an automated imaging of 3D

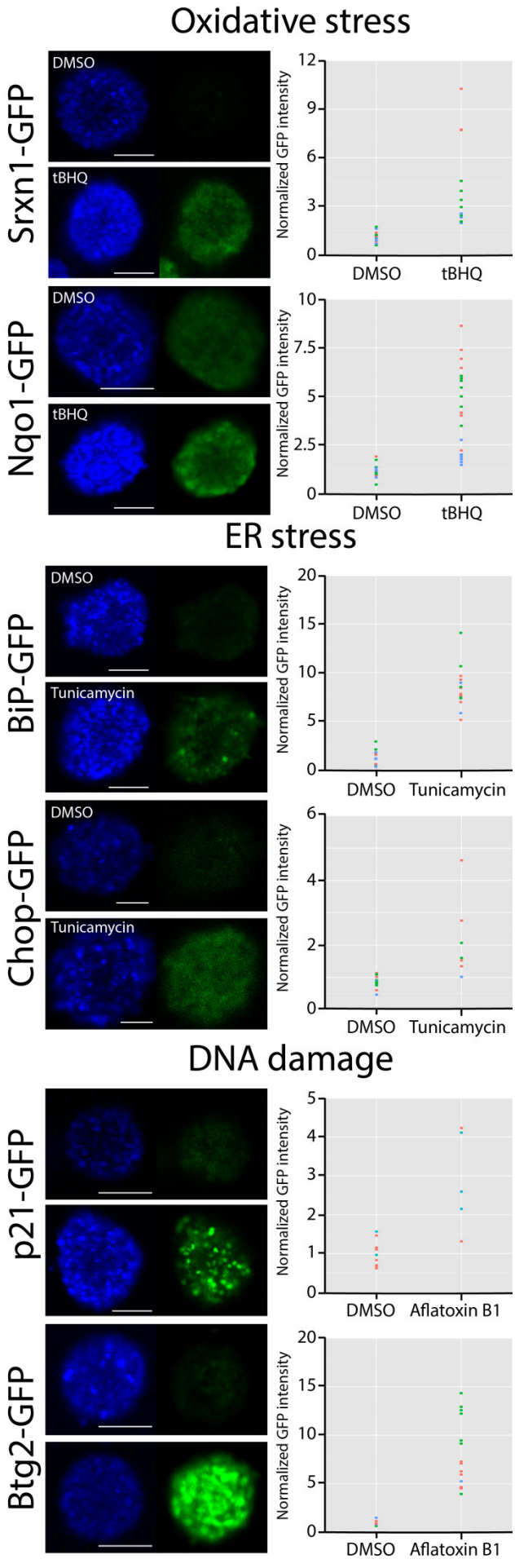


Figure 2: Confocal microscopy can detect differences in GFP intensity for control compounds. Representative images of single 3D spheroids for Srxn1-GFP and Nqo1-GFP (induced by 0,15 mM t-BHQ); BiP-GFP and Chop-GFP (12 µM tunicamycin); p21-GFP (5 µM aflatoxin B1) and Btg2-GFP (25 µM aflatoxin B1) compared to DMSO control. The graphs show dots which represent individual spheroids as a fold change to DMSO average. Different colors represent different biological replicates.

spheroids. 3D spheroids grow in different layers and at random x,y positions through-out the well, making it impossible to use the perfect focus system. An in-house built macro, which identifies spheroids' x,y positions by making a large image and subsequently determines z positions by zoomed z-stacking, does enable automated imaging (supplemental figure 1). All reference control compounds were positive in their corresponding reporters; for some maximal response was observed after 48 hours (Aflatoxin B1: Btg2-GFP and p21-GFP reporters); for others this was primarily at 6 day repeated exposure (tBHQ: Srxn1-GFP and Nqo1-GFP; and thapsigargin: BiP-GFP). Chop-GFP responses were low, but in the range of what was expected based on experience in 2D (figure 3)[145]. For the selected TG-GATEs compounds most significant reporter activation was observed for the six and eleven days repeated exposure

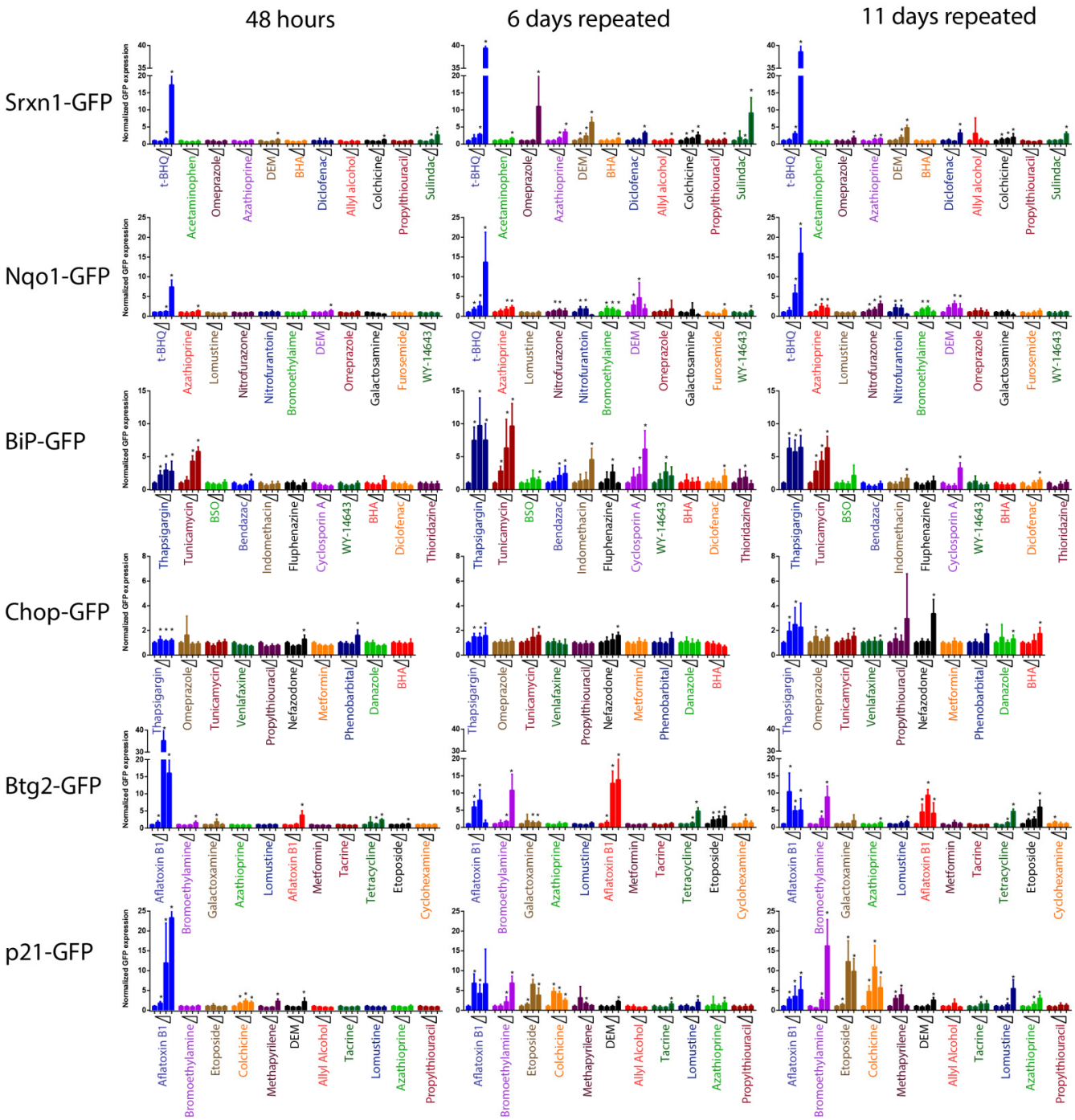


Figure 3: Activation of GFP stress response pathways after stimulation with DILI relevant inducers. Srxn1-GFP, Nqo1-GFP, BiP-GFP, Chop-GFP, p21-GFP and Btg2-GFP exposed to ten compounds from top twelve inducers in the TG-GATES 24 hours high dose primary human hepatocyte dataset. 48 hours single exposure (left), six days repeated exposure (middle) and eleven days repeated exposure (right) are depicted. Graphs are the fold-change of DMSO average and are displayed as mean values with 95% confidence interval over all spheroids tested as error bars. Concentrations are the same as tested in the TG-GATES dataset (Supplemental table 1). Compounds are ordered based on their ranking in TG GATES dataset. Significance was based on a one-tailed Student's t-test, * = $p < 0,05$.

compared to 48 hours single exposure. In particular the Srxn1-GFP and BiP-GFP reporters were most representative for the TG-GATEs PHH transcriptomics responses (figure 3). For all six reporter lines the average combined sensitivity was 86%, with Srxn1-GFP as a single reporter even a 100% sensitivity (table 1). Thus, the 3D repeated exposures reporter activity showed high resemblance with gene expression data of primary human hepatocytes. This notion was supported comparing the fold changes of the TG-GATEs data set with the maximal fold change obtained from our 3D reporter system (supplemental table 1). Taken together, we conclude that our 3D BAC-GFP stress response reporter system shows high resemblance with responses observed in PHH.

Sensitivity	Srxn1	Nqo1	BiP	Chop	p21	Btg2
48 hours	30%	20%	20%	22%	30%	50%
6 days repeated	100%	70%	90%	22%	80%	60%
11 days repeated	60%	50%	40%	89%	80%	70%
Combined	100%	70%	90%	89%	80%	80%

Table 1: Sensitivity of GFP responses induced by TG-GATES selected compounds. Sensitivity was calculated as percentage of compounds significantly activated in one of the tested concentrations in total amount of compounds. For combined sensitivity all time points are included.

Most-DILI-concern drugs are over represented in activated clusters of 3D exposures

Next, to study the applicability of the 3D BAC-GFP reporter cell lines in drug safety assessment screening, a systemic analysis of in total 33 drugs was performed with known DILI liabilities: 21 drugs categorized as most-DILI-concern drugs, 2 drugs with less-DILI-concern and 10 drugs with no-DILI-concern drugs were selected (supplemental table 2). The evaluation was conducted with the six GFP reporter cell lines: Srxn1-GFP, Nqo1-GFP, p21-GFP, Btg2-GFP, BIP-GFP and Chop-GFP. 3D spheroids were treated for 48 hr single exposure or six days repeated exposure using concentrations that were based on peak plasma levels measured after administrations to patients during clinical trials (Cmax): 1, 5, 10, 25, 50 and 100 times Cmax [53,54,165,166]. Six positive reference control compounds were also included: diethyl maleate (DEM) and t-BHQ (oxidative stress reporters); etoposide and aflatoxin B1 (DNA damage reporters); tunicamycin and thapsigargin (UPR reporters). As a first step the effect on viability was determined with ATP-lite measurements in wildtype HepG2 spheroids at the same concentrations (supplemental figure 2). Some DILI compounds already caused loss of viability after 48 hours at the highest concentrations, which was further increased after 6 days. In the no-DILI-concern group, entacapone and pioglitazone resulted in some cell death in 48 hours and in 6 days repeated exposure.

Next, we carefully assessed the GFP-intensity of individual spheroids. Clear increases in GFP positive spheroids was observed for various DILI compounds, also those for which responses in

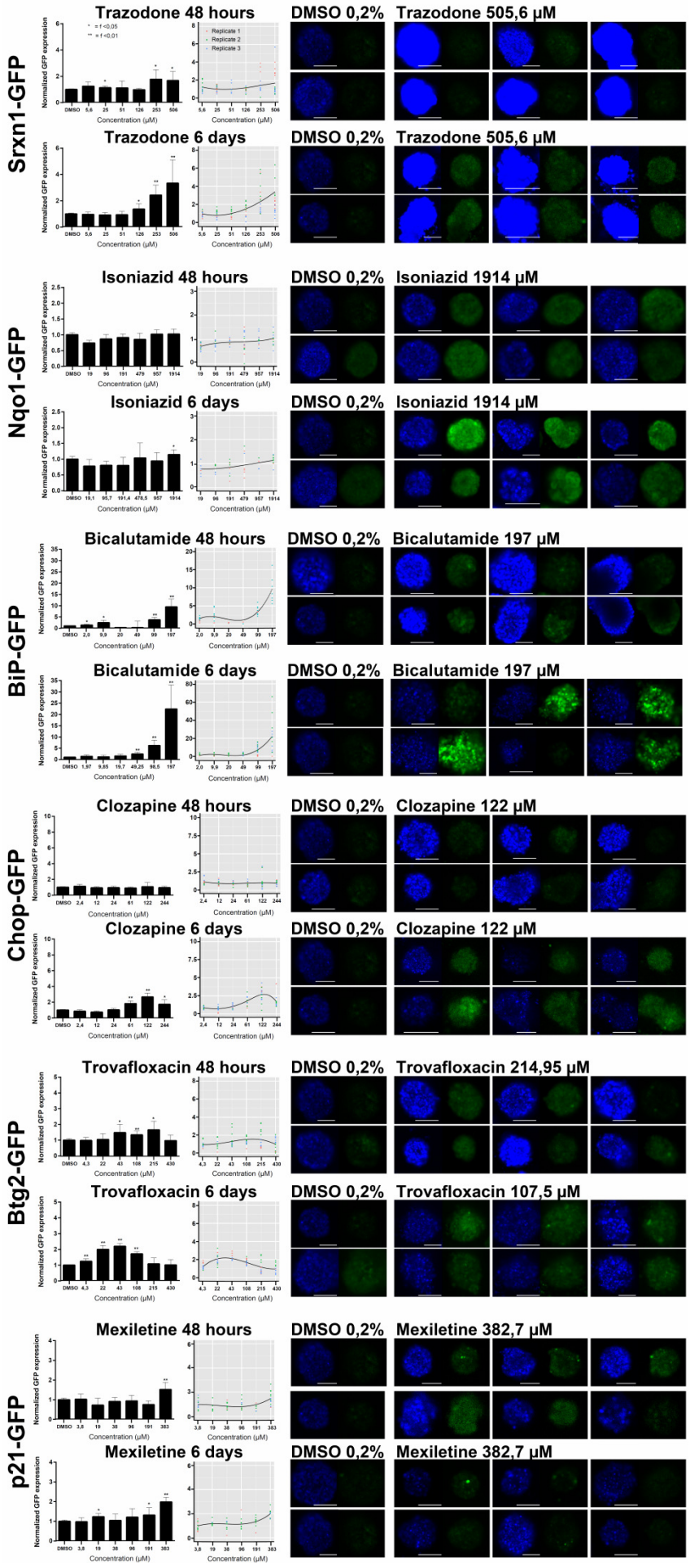


Figure 4: Activation of different stress response pathway by most-DILI-concern drugs. One most-DILI-concern drug for each Srxn1-GFP, Nqo1-GFP, BiP-GFP, Chop-GFP, p21-GFP and Btg2-GFP is shown in concentration response graphs for 48 hours single exposure and six days repeated exposure. The left graph shows mean fold change with 95% confidence interval error bars. Significance is tested with a one-tailed student's t-test with * = p<0,05 and ** = p<0,01. The right graph shows a dose response with dots for individual spheroids (colors represent biological replicates). The line is fitted with a b-spline fit with 3 degrees of freedom. Two example images are shown for DMSO control and six example images are shown for one concentration of a significantly induced most-DILI-concern drug.

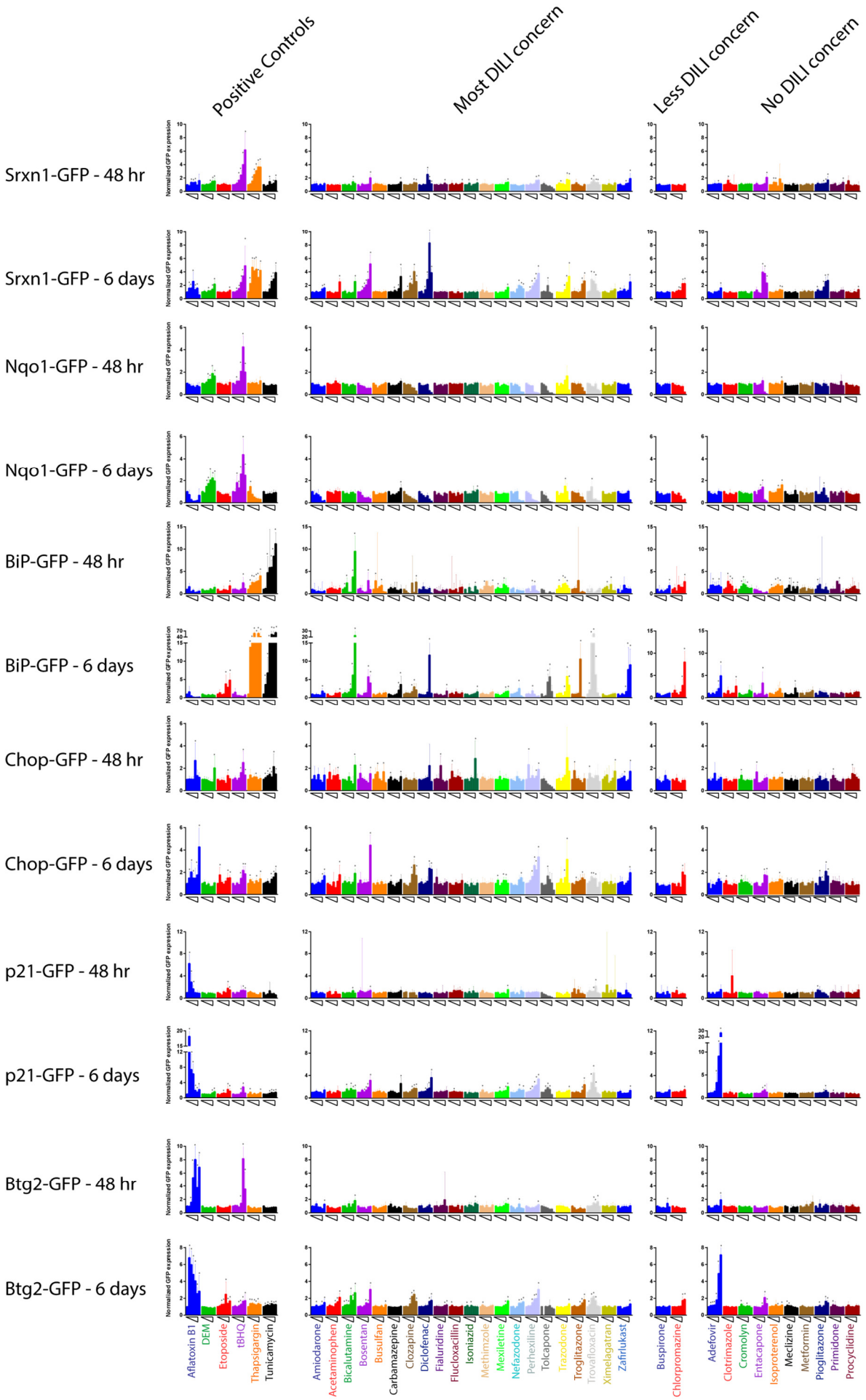

Positive Controls
Most DILI concern
Less DILI concern
No DILI concern
Srxn1-GFP - 48 hr
Srxn1-GFP - 6 days
Nqo1-GFP - 48 hr
Nqo1-GFP - 6 days
BiP-GFP - 48 hr
BiP-GFP - 6 days
Chop-GFP - 48 hr
Chop-GFP - 6 days
p21-GFP - 48 hr
p21-GFP - 6 days
Btg2-GFP - 48 hr
Btg2-GFP - 6 days
Normalized GFP expression
Aflatoxin B1
DEM
Etoposide
tBHQ
Thapsigargin
Tunicamycin
Amiodarone
Acetaminophen
Bicalutamine
Bosentan
Busulfan
Carbamazepine
Clozapine
Diclofenac
Fialuridine
Flucloxacillin
Isoniazid
Methimzole
Mexiletine
Nefazodone
Perhexiline
Tolcapone
Trazodone
Troglitazone
Trovafloxacin
Ximelagatran
Zafirlukast
Buspirone
Chlorpromazine
Adefovir
Clotrimazole
Cromolyn
Entacapone
Isoproterenol
Meclizine
Metformin
Pioglitazone
Primidone
Procyclidine

Figure 5: Activation of GFP stress response pathways after stimulation with drugs with known DILI hazard. Srxn1-GFP, Nqo1-GFP, BiP-GFP, Chop-GFP, p21-GFP and Btg2-GFP exposed to six positive control compounds, 21 most-DILI-concern drugs, two less-DILI-concern drugs and ten no-DILI-concern drugs shown for 48 hours single exposure and six days repeated exposure. Graphs are a fold change of DMSO average and are displayed as mean values with 95% confidence interval over all spheroids tested as error bars. Concentrations are based on maximal plasma concentrations after administration to patients (Cmax). Dose range tested is 1, 5, 10, 25, 50 and 100x Cmax. Significance calculated with a one-tailed Student's t-test, * = $p < 0,05$.

2D HepG2 have so far not been observed (Wink *et al*, manuscript in preparation). Some variability in responses between spheroids was evident for different reporters (figure 4). Overall stress response activation was stronger after six days repeated exposure. We then summarized the reporter data for all compound treatments (figure 5). The positive control group showed significant concentration dependent upregulation of the selective pathway reporters. Concentration dependent responses were observed for most DILI compounds, however some compounds (e.g. trovafloxacin, tolcapone and clozapine) displayed a decrease in activity at the highest concentrations, especially in 6 days repeated exposure. This correlates with a strong increase in overall cytotoxicity (supplemental figure 2 and supplemental table 2). Typically, all six GFP reporter lines demonstrated more significantly upregulated responses after six days

		Srxn1	Nqo1	BiP	Chop	p21	Btg2	Combined	Clustering
≤ 100 Cmax	Sensitivity	85%	19%	95%	95%	66%	81%	100%	72%
	Specificity	30%	70%	20%	20%	60%	30%	10%	60%
≤ 50 Cmax	Sensitivity	66%	9%	81%	85%	57%	71%	100%	NA
	Specificity	50%	70%	20%	20%	60%	30%	10%	NA

Table 2: Sensitivity and specificity for DILI compounds. Compounds were counted positive as one concentration in one time point showed a significant increase. Sensitivity is calculated as percentage positive most-DILI-concern drugs over all most-DILI-concern drugs. Specificity is calculated as percentage negative no-DILI-concern drugs over total no-DILI-concern drugs. Sensitivity and specificity are shown for all the GFP reporters separate and combined. Combined means a positive reaction in at least one reporter. Last column indicates sensitivity and specificity based on hierarchical clustering in figure 6.

repeated exposure than after 48 hours single exposure. BiP-GFP was most sensitive at 6 days repeated exposure, followed by Srxn1-GFP, but no direct overlap in responses was observed. All GFP reporter lines, except Nqo1-GFP, had at least 14 out of 21 (66% sensitivity) most-DILI-concern compounds where at least one concentration significantly increased the reporter activity (figure 5 and table 2). The negatives included fialuridine, xilamagatran, flucloxacillin, busulfan and isoniazid. For no-DILI concern compounds 4 out of 10 compounds show significant upregulation; these included entacapone, pioglitazone, adefovir and isoproterenol. BiP-GFP and Srxn1-GFP were most sensitive for the majority of severe-DILI-concern compounds; Nqo1-GFP was not significantly upregulated by any of the compounds, likely due to high basal levels of the Nqo1-GFP under control conditions. Interestingly, trovafloxacin induced all three different

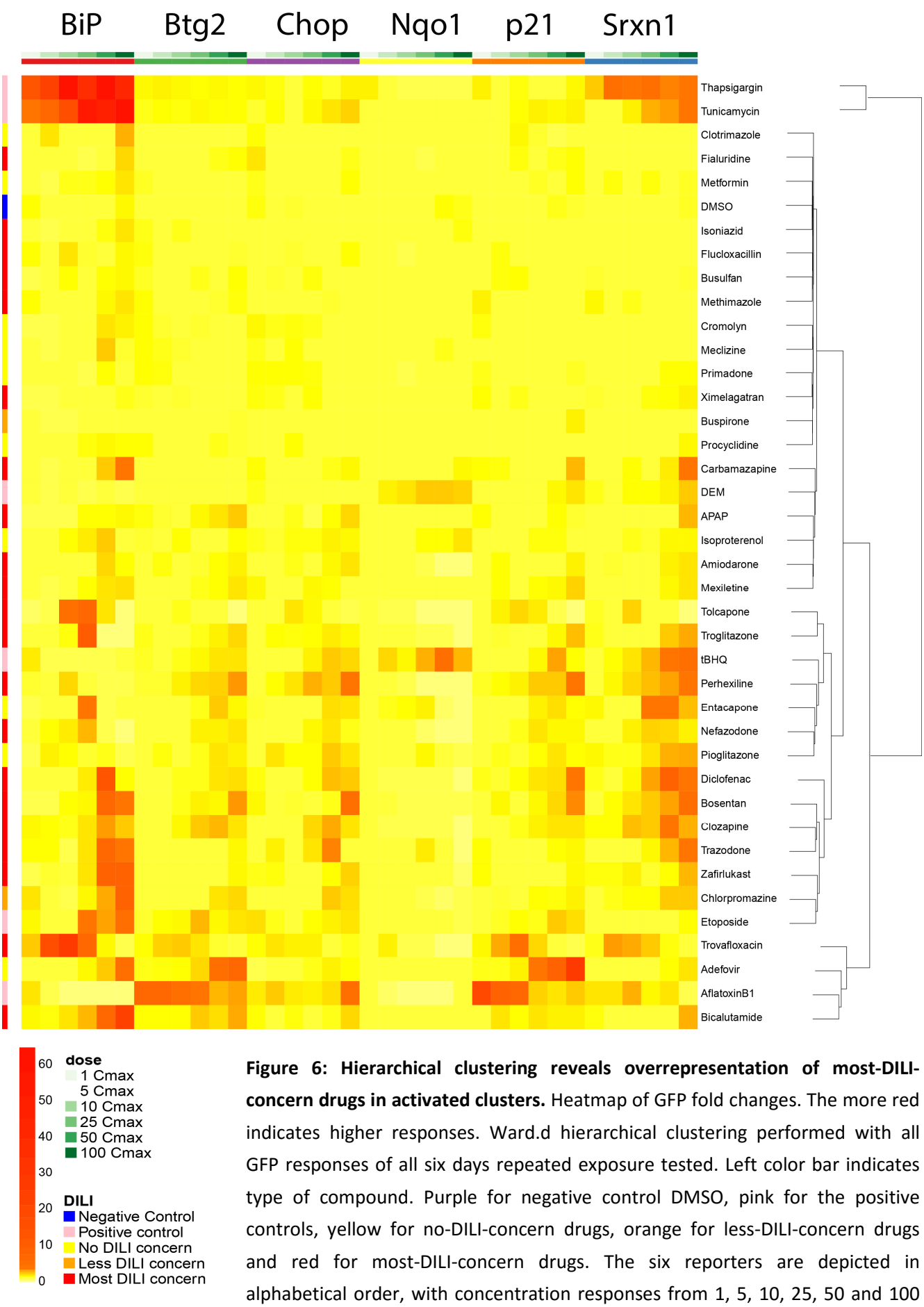


Figure 6: Hierarchical clustering reveals overrepresentation of most-DILI-concern drugs in activated clusters. Heatmap of GFP fold changes. The more red indicates higher responses. Ward.d hierarchical clustering performed with all GFP responses of all six days repeated exposure tested. Left color bar indicates type of compound. Purple for negative control DMSO, pink for the positive controls, yellow for no-DILI-concern drugs, orange for less-DILI-concern drugs and red for most-DILI-concern drugs. The six reporters are depicted in alphabetical order, with concentration responses from 1, 5, 10, 25, 50 and 100 Cmax (from left to right).

pathways; remarkably, the non-DILI-concern compound adefovir activated the BiP-GFP, Btg2-GFP and p21-GFP reporter, questioning the correctness of DILI labeling.

Next, we performed hierarchal clustering to evaluate compound responses. Because 6 days repeated exposure showed the most significant responses clustering was performed for all concentrations and all compounds (figure 6). Six different clusters were distinguished (clusters I-VI): five clusters (clusters I, III-VI) with activation of multiple pathways and one cluster (cluster II) with minor activation in one pathway (mainly BiP-GFP). Cluster II also contained DMSO control as well as 6 out of 10 no-DILI-concern compounds; also 6 out of 21 most-DILI-concern compounds. Thus, hierarchical clustering resulted in an overrepresentation of most-DILI-concern compounds in activated clusters, suggesting 72% sensitivity and 60% specificity.

Discussion

In the current study we established that our HepG2 BAC-GFP reporters in cultured as 3D spheroids have an enhanced correct identification of DILI liability compared to 2D monolayer cultures. In particular BiP-GFP and Srnx1-GFP reporter cell lines seems most promising as a combined set of BAC-GFP reporter cells. The Nqo1-GFP reporters seems less valuable in 3D spheroids.

In this study, we combined our GFP reporter platform and the enhanced liver-like properties of HepG2 3D spheroid cultures. We demonstrated that GFP reporter cell lines differentiate in a comparable manner as the HepG2 wild type cells, as evidenced by various differentiation markers at the morphological (i.e. bile canaliculi formation) as well as protein expression level (i.e. albumin expression). We established a method to automatically acquire GFP intensity in 3D spheroids and to quantify these GFP intensities. Well-established positive reference compounds showed robust upregulation of GFP intensity of all six reporter lines validating the image acquisition and quantification method. Moreover, based on TG-GATES selected compound screening we were able to verify that our 3D system shows very high concordance with primary human hepatocyte responses, as sensitivity ranges from 70% up to 100%.

After establishing the 3D HepG2 GFP reporter platform as an imaging-based high throughput liver toxicity screening method, we evaluated the application for the identification of DILI liability. Our combined screening resulted in an overrepresentation of most-DILI-concern drugs in compounds that together did activate various reporters. Of particular interest was the observations that several DILI drugs, such as bicalutamide and trazodone, did not show any response in 2D monolayers, showed significant concentration-dependent activation in the 3D spheroids. This is possibly related to the more differentiated phenotype, included drug metabolizing capacity, of the 3D spheroids, although we can not exclude that this is solely due to the repeated dosing condition. Regardless, we are confident that our 3D HepG2 spheroids GFP reporter systems contributes to a more refined assessment of DILI liabilities.

Our hierarchical clustering analysis showed six most-DILI-concern compounds which were inactive in our GFP reporter system. These were busulfan, flucloxacillin, ximelagatran, fialuridine, isoniazid and methimazole. Four of the inactive most-DILI-concern drugs, flucloxacillin, ximelagatran, methimazole and isoniazid, are associated with immune related liver injury[167–171]. Flucloxacillin, ximelagatran and methimazole are more specifically linked to single nucleotide polymorphisms in the human leukocyte antigen (HLA) allele, typically associated with iDILI responses. Daly *et al* performed a genome wide association study (GWAS) to identify single nucleotide polymorphisms involved in flucloxacillin mediated liver injury[168]. The HLA-B*5701 appeared to be significantly different in patients suffering from flucloxacillin mediated liver injury compared to control patients. In 2008, a retrospective study was performed for ximelagatran, where a strong association was found between elevated ALAT levels and single nucleotide polymorphisms in DRB1*07 and DQA1*02[169]. More recent, methimazole patients were found to be at higher risk for agranulocytosis when possessing the HLA-B*38:02:01 allele[167]. While the past decade received much debate about the mechanism of injury of isoniazid, more recent literature points towards immune mediated mechanisms of liver injury[171]. The mechanism of fialuridine induced liver injury has been suggested to be mitochondrial related[172] due to inhibition of the mitochondrial DNA replication; typically this response takes several weeks, and might be missed in the 6 day repeat exposure, which may not have culminated in the activation of cellular stress responses. For future studies we might require integration of Immune-related and mitochondrial related toxicity in our reporter platform. This could be simply mitochondrial mass detection using GFP-BAC constructs for mitochondrial markers. For immune-mediated signaling this could involve signaling components downstream of cytokine signaling[112], although more complex adaptive immune response signaling would not allow integration in reporter systems.

Four no-DILI-concern compounds clustered in activated clusters: entacapone, pioglitazone, adefovir and isoproterenol. Entacapone is structurally related to tolcapone and have both been used as catechol-O-methyltransferase (COMT) inhibitor to treat Parkinson's disease. Also, both drugs have been associated with mitochondrial dysfunction and bile salt export protein (BSEP) inhibitors[173]. Previously, we also observed stress response activation in 2D of entacapone (Wink *et al*., manuscript in preparation). This indicates both tolcapone and entacapone induce similar stress response pathways. The reason why entacapone does not lead to DILI and why tolcapone does, remains unknown, but is possibly dose dependent. Pioglitazone is known as mitochondrial toxic and affects cellular ATP levels[174], which could explain the observed activation Srxn1-GFP in 3D (our work) and 2D (Wink *et al*., manuscript in preparation). Adefovir is well known to induce nephrotoxicity, which seems due to depletion of mitochondrial DNA[175]. We demonstrate that adefovir activates both BiP-GFP as well as p21-GFP and Btg2-GFP, the latter suggests an involvement of DNA-damage. There is currently debate on the DILI liability of adefovir. Isoproterenol is a β-adrenalreceptor agonist which has been classified with no-DILI-

concern. In rats isoproterenol induces TNF-α and IL-6 elevation indicating induction of an innate immune response [176]. This could be the reason why mild stress response activation is observed in our assay.

Next to the adaptive stress response activation we also measured viability. When we determine the lowest observed effective level of the stress response activation, it was in all cases lower than the lowest concentration to affect the maximal loss of cell viability. This emphasizes that adaptive stress responses are an earlier marker for cellular perturbations than cytotoxicity, supporting our paradigm to use adaptive stress response activation as an early and advanced marker for drug-induced hazard assessment.

The enhanced liver like properties present in 3D spheroids compared to 2D monolayer culturing not only provides the advantage of enhanced metabolizing capacity, it also enables repeated dose exposures. After approximately seven to fourteen days 3D spheroids stop proliferating and start differentiating [27]. This means that from this moment, the spheroids will not overgrow and can be maintained for several weeks afterwards. This allows the possibility of repeated dose exposure. In many cases of idiosyncratic DILI, latency time varies from the first week to two years after administration. During this period, patients obtain at least a daily dose of the drug. By applying fresh drug every day for six or eleven days, we resemble closer the human drug administration situation, which seems pivotal in hepatotoxicity screening. This is recognized by much higher sensitivity values in both six and eleven days repeated exposure compared to 48 hour single exposure (Table 1). Currently we lack the information on the intracellular concentrations and drug accumulation in our 3D spheroid system. The current concentration ranges were based on human peak plasma levels (Cmax). For *in vitro* repeated dose studies it has been shown drugs can accumulate to high intracellular concentrations [177]. For the same drug, postmortem studies showed 53-fold increase compared to Cmax levels [178]. Both studies indicate intracellular accumulation of drugs *in vitro* and in human tissue. Therefore, we are convinced that repeated dose exposure in a broader concentration range that extends above the (predicted) Cmax concentration is essential in DILI liability assessment.

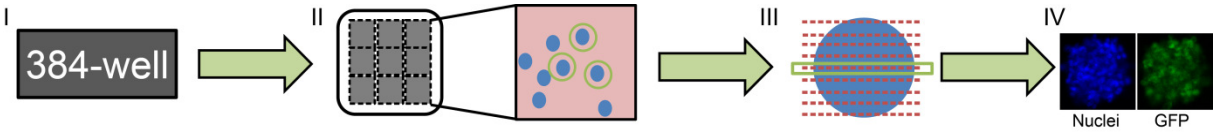
The current study provides suboptimal sensitivity and specificity levels to accurately discriminate most-DILI-concern from no-DILI-concern drugs. We see several solutions to improve this. First, we need to extend the number of stress response endpoints and likely include immune and mitochondrial pathways to increase predictivity. Second, we will integrate bioinformatics approaches to identify a classifier that will rank drugs for DILI liability based on pathway activation, as well as on multiple defined criteria, including combinations of activated stress pathways, fold change induction cut-off, and TC50 values. Such a classifier will likely increase the sensitivity and specificity determinants. Ultimately we foresee the integration of 2D monolayer GFP reporter activation screening as a tier 1 (Wink *et al.*, manuscript in preperation) and 3D spheroid screening as a tier 2.

Taken together, we have established a HepG2 3D spheroid GFP reporter platform with enhanced liver-like features and improved predictivity which is suited for imaging-based high throughput screening for DILI liabilities.

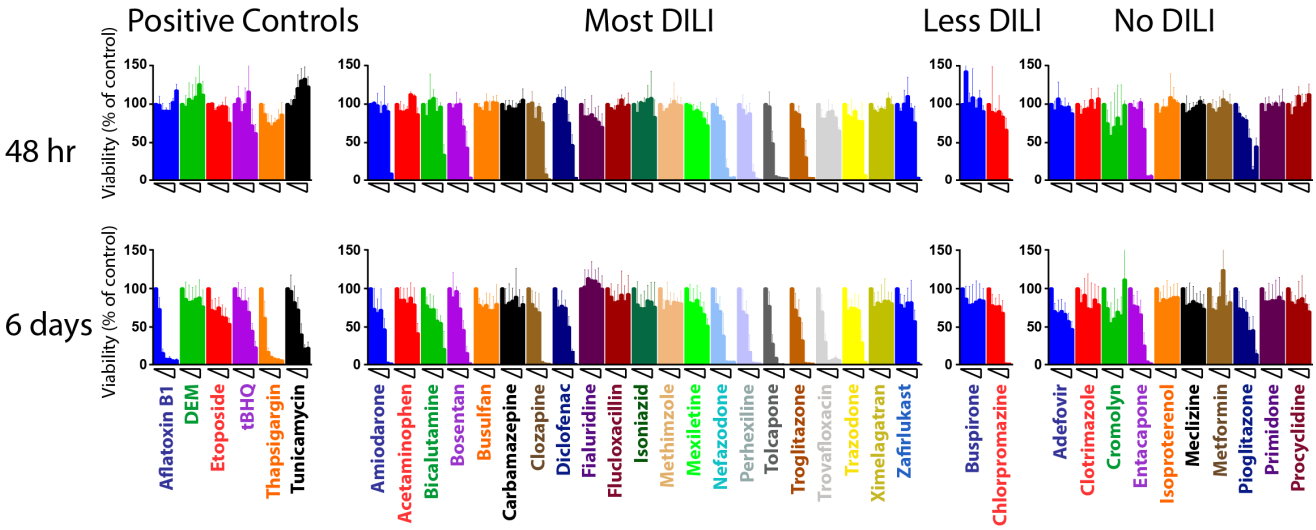
Acknowledgements

This work was supported by the IMI MIP DILI project (grant agreement number 115336) and the EU FP7 Seurat-1 Detective project (grant agreement number 266838).

Supplementals



Supplemental figure 1: Set-up for automated high throughput confocal imaging of 3D spheroids. I) Spheroids are cultured in a hydrogel in a 384 well plate. II) A large image is stitched out of multiple images to visualize the whole well. The macro identifies spheroid objects and zooms in one by one. III) the macro determines by z-stacking the plane with the highest intensity of nuclear staining Hoechst$_{33342}$. IV) It takes one image in the plane with highest Hoechst$_{33342}$ intensity. The image contains nuclear and GFP intensity.



Supplemental figure 2: Viability of drugs with known DILI hazard. Srxn1-GFP, Nqo1-GFP, BiP-GFP, Chop-GFP, p21-GFP and Btg2-GFP exposed to six positive control compounds, 21 most-DILI-concern drugs, two less-DILI-concern drugs and ten no-DILI-concern drugs shown for 48 hours single exposure and six days repeated exposure. Graphs are a percentage of DMSO average (set to 100%) and are displayed as mean values with standard deviations as error bars over three biological replicates. Concentrations are based on maximal plasma concentrations after administration to patients (Cmax). Dose range tested is 1, 5, 10, 25, 50 and 100x Cmax.

Compounds TG-GATES	Concentratie (uM)			Srxn1		Nqo1		BiP		Chop		p21		Btg2	
	High	Middle	Low	TGP FC	3D FC	TGP FC	3D FC	TGP FC	3D FC	TGP FC	3D FC	TGP FC	3D FC	TGP FC	3D FC
Acetaminophen	5000	1000	200	4.79	1.65										
Aflatoxin B1	6	1.2	0.24											1.44	13.82
Allyl alcohol	70	14	2.8	2.52	3.16							1.66	1.74		
Azathioprine	73	14.6	2.92	3.53	3.53	2.51	2.47					1.58	3.02	1.61	1.32
Bendazac	200	40	8					2.08	2.44						
Bromoethylamine	500	100	20			1.73	2.15					2.79	16.25	3.20	10.78
Buthionine sulfoximine	10000	2000	400					2.13	2.16						
Butylated hyrdoxyaniso	200	40	8	2.77	1.66			1.60	1.45	3.61	1.74				
Colchicine	4000	800	160	2.44	2.59							2.10	10.92		
Cyclohexamine	100	20	4											1.27	1.89
CyclosporinA	6	1.2	0.24					1.70	6.15						
Danazole	35	7	1.4							3.68	1.43				
DEM	1500	300	60	3.43	6.34	1.41	4.67					1.94	2.56		
Diclofenac	400	80	16	2.73	3.26			1.53	2.09						
Etoposide	330	66	13.2									2.25	12.28	1.33	5.87
Fluphenazine	20	4	0.8					1.82	2.68						
Furosemide	2500	500	100			1.36	1.47								
Galactosamine	10000	2000	400			1.37	1.65							1.92	1.92
Indomethacin	200	40	8					1.92	4.56						
Lomustine	120	24	4.8			2.22	1.01					1.59	5.46	1.60	1.58
Metformin	1000	200	40							4.06	1.12			1.43	1.44
Methapyrilene	600	120	24									1.99	3.90		
Nefazodone	30	6	1.2							4.20	3.36				
Nitrofuratoin	125	25	5			1.85	2.16								
Nitrofurazone	50	10	2			1.98	3.10								
Omeprazole	600	120	24	4.58	11.02	1.38	1.85			13.00	1.62				
Phenobarbital	10000	2000	400							3.72	1.72				
Propylthiouracil	4000	800	160	2.42	1.38					4.44	2.96	1.46	1.26		
Sulindac	3000	600	120	2.26	9.10										
Tacrine	80	16	3.2									1.65	1.60	1.39	1.11
Tetracycline	25	5	1											1.33	4.71
Thioridazine	15	3	0.6					1.52	1.88						
Tunicamycin	10	2	0.4					5.95	9.65	8.81	1.59				
Venlafaxine	1200	240	48							6.02	1.11				
WY-14643	150	30	6			1.26	1.28	1.67	2.70						

Supplemental table 1: Fold changes of compounds in TG GATES and in our 3D GFP reporter system. For all compounds used, low, middle and high concentrations are shown. The fold changes shown for TG GATES and 3D reporters are fold change compared to DMSO average control. TG GATES fold changes taken as 24 hours high dose. 3D GFP reporter fold changes taken as maximal fold changes over all conditions tested.

	Compound	Cmax (uM)	DILI classification	Metabolism?	Maximal concentration used (uM)	Vendor/Supplier	TC50 48hr	TC50 6 days
1	Acetaminophen	139	Most-DILI-concern	YES	13900	MIP DILI consortium	NA	100 Cmax
2	Adefovir	0.085	No-DILI-concern	NO	8.5	Shanghai PI chemicals	NA	100 Cmax
3	AflatoxinB1	1	Control	NA	100	Sigma Aldrich	NA	5 Cmax
4	Amiodarone	0.807	Most-DILI-concern	NO	80.7	MIP DILI consortium	100 Cmax	25 Cmax
5	Bicalutamide	1.97	Most-DILI-concern	YES	197	USP	100 Cmax	100 Cmax
6	Bosentan	7.4	Most-DILI-concern	YES	740	MIP DILI consortium	50 Cmax	25 Cmax
7	Buspirone	0.016	Less-DILI-concern	YES	1.6	Sigma Aldrich	NA	NA
8	Busulfan	0.276	Most-DILI-concern	YES	27.6	Sigma Aldrich	NA	NA
9	Carbamazepine	50.79	Most-DILI-concern	NO	800	Sigma Aldrich	NA	NA
10	Chlorpromazine	0.94	Less-DILI-concern	YES	94	Sigma Aldrich	100 Cmax	50 Cmax
11	Clotrimazole	0.087	No-DILI-concern	NO	8.7	Sigma Aldrich	NA	NA
12	Clozapine	2.44	Most-DILI-concern	YES	244	Sigma Aldrich	50 Cmax	25 Cmax
13	Cromolyn	0.58	No-DILI-concern	NO	58	Sigma Aldrich	NA	NA
14	Diethyl maleate	NT	Control	NA	200	Sigma Aldrich	NA	NA
15	Diclofenac	10.1	Most-DILI-concern	YES	1010	MIP DILI consortium	50 Cmax	50 Cmax
16	Entacapone	3.93	No-DILI-concern	YES	393	MIP DILI consortium	50 Cmax	25 Cmax
17	Etoposide	NT	Control	NA	50	Sigma Aldrich	NA	NA
18	Fialuridine	1	Most-DILI-concern	NO	100	MIP DILI consortium	NA	NA
19	Flucloxacillin	31.9	Most-DILI-concern	NO	1595	Sigma Aldrich	NA	NA
20	Isoniazid	76.57	Most-DILI-concern	YES	1914	Sigma Aldrich	NA	NA
21	Isoproterenol	2.02	No-DILI-concern	NO	202	Sigma Aldrich	NA	NA
22	Meclizine	0.026	No-DILI-concern	NO	2.6	USP	NA	NA
23	Metformin	7.78	No-DILI-concern	YES	778	MIP DILI consortium	NA	NA
24	Methimazole	2.619	Most-DILI-concern	YES	261.9	Sigma Aldrich	NA	NA
25	Mexiletine	3.827	Most-DILI-concern	YES	382.7	Sigma Aldrich	NA	NA
26	Nefazodone	3.95	Most-DILI-concern	YES	395	MIP DILI consortium	25 Cmax	10 Cmax
27	Perhexilline	2.162	Most-DILI-concern	NO	216.2	MIP DILI consortium	25 Cmax	10 Cmax
28	Pioglitazone	2.946	No-DILI-concern	NO	294.6	MIP DILI consortium	50 Cmax	25 Cmax
29	Primidone	4.673	No-DILI-concern	YES	467.3	Sigma Aldrich	NA	NA
30	Procyclidine	0.404	No-DILI-concern	NO	40.4	Sigma Aldrich	NA	NA
31	tert-Butylhydroquinone	NT	Control	NA	316	Sigma Aldrich	NA	50 Cmax
32	Thapsigargin	NT	Control	NA	2	Sigma Aldrich	NA	5 Cmax
33	Tolcapone	21.985	Most-DILI-concern	YES	2198.5	MIP DILI consortium	5 Cmax	5 Cmax
34	Trazodone	5.056	Most-DILI-concern	YES	505.6	Sigma Aldrich	100 Cmax	5 Cmax
35	Troglitazone	6.39	Most-DILI-concern	YES	639	MIP DILI consortium	25 Cmax	10 Cmax
36	Trovafloxacin	4.299	Most-DILI-concern	NO	429.9	Sigma Aldrich	NA	5 Cmax
37	Tunicamycin	NT	Control	NA	24	Sigma Aldrich	NA	25 Cmax
38	Ximelagatran	0.3	Most-DILI-concern	NO	30	MIP DILI consortium	NA	NA
39	Zafirlukast	1.211	Most-DILI-concern	YES	121.1	Cayman	100 Cmax	100 Cmax

Supplemental table 2: List of drugs tested with known DILI hazard. List of drugs tested with known DILI hazard, including Cmax concentrations, DILI classification, expected metabolism and maximal concentration used. NT means not treated in patients, NA means not applicable. TC50 value means the first Cmax concentration the viability is lower than 50%.