High-throughput confocal imaging of differentiated 3D liver-like spheroid cellular stress response reporters for identification of drug-induced liver injury liability

Hiemstra, S.; Ramaiahgari, S.C.; Wink, S.; Callegaro, G.; Coonen, M.; Meerman, J.; ... ; Water, B. van de

IN VITRO SYSTEMS

High-throughput confocal imaging of differentiated 3D liver-like spheroid cellular stress response reporters for identification of drug-induced liver injury liability

Steven Hiemstra[1] · Sreenivasa C. Ramaiahgari[1,4] · Steven Wink[1] · Giulia Callegaro[1] · Maarten Coonen[3] · John Meerman[1] · Danyel Jennen[3] · Karen van den Nieuwendijk[1] · Anita Dankers[2] · Jan Snoeys[2] · Hans de Bont[1] · Leo Price[1] · Bob van de Water[1]

Abstract

Adaptive stress response pathways play a key role in the switch between adaptation and adversity, and are important in drug-induced liver injury. Previously, we have established an HepG2 fluorescent protein reporter platform to monitor adaptive stress response activation following drug treatment. HepG2 cells are often used in high-throughput primary toxicity screening, but metabolizing capacity in these cells is low and repeated dose toxicity testing inherently difficult. Here, we applied our bacterial artificial chromosome-based GFP reporter cell lines representing Nrf2 activation (Srxn1-GFP and NQO1-GFP), unfolded protein response (BiP-GFP and Chop-GFP), and DNA damage response (p21-GFP and Btg2-GFP) as long-term differentiated 3D liver-like spheroid cultures. All HepG2 GFP reporter lines differentiated into 3D spheroids similar to wild-type HepG2 cells. We systematically optimized the automated imaging and quantification of GFP reporter activity in individual spheroids using high-throughput confocal microscopy with a reference set of DILI compounds that activate these three stress response pathways at the transcriptional level in primary human hepatocytes. A panel of 33 compounds with established DILI liability was further tested in these six 3D GFP reporters in single 48 h treatment or 6 day daily repeated treatment. Strongest stress response activation was observed after 6-day repeated treatment, with the BiP and Srxn1-GFP reporters being most responsive and identified particular severe-DILI-onset compounds. Compounds that showed no GFP reporter activation in two-dimensional (2D) monolayer demonstrated GFP reporter stress response activation in 3D spheroids. Our data indicate that the application of BAC-GFP HepG2 cellular stress reporters in differentiated 3D spheroids is a promising strategy for mechanism-based identification of compounds with liability for DILI.

Keywords High-throughput imaging · HepG2 spheroids · Drug-induced liver injury · BAC-reporter cells · Cellular stress response · Liver transcription factors

---

Steven Hiemstra and Sreenivasa C. Ramaiahgari have contributed equally.

✉ Bob van de Water
  b.water@lacdr.leidenuniv.nl

1 Division of Drug Discovery and Safety, Leiden Academic Centre for Drug Research (LACDR), Leiden University, Leiden, The Netherlands

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

3 Department of Toxicogenomics, GROW School of Oncology and Development Biology, Maastricht University, Maastricht, The Netherlands

4 Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

Introduction

Drug-induced liver injury (DILI) is a major problem in the clinic, as 50% of all liver failures are caused by drugs (Ostapowicz et al. 2002). In 13% of these cases, liver failure

is caused by idiosyncratic drug reactions. Idiosyncratic DILI occurs only in rare cases, in approximately 0.1–0.01% 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 make it complicated to predict which drug will eventually cause idiosyncratic DILI in the clinic. 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 (Wilke et al. 2007).

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 candidate chemical entity. This is valuable information as idiosyncratic DILI only occurs in susceptible individuals and likely also involves initial cell injury to hepatocytes. Therefore, over the past decade, transcriptomic approaches have unraveled some of the mechanisms of action by which the drugs with DILI liabilities affect the cellular homeostasis (Cui and Paules 2010; Jennings et al. 2013; Jiang et al. 2015). While transcriptomics has allowed identification of 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 have prevented the clear identification of benchmark concentrations for point of departure of cellular perturbations. Furthermore, while mRNA expression levels can be indicative of cellular function, they do not necessarily represent the changes in the cellular protein landscape after drug exposure. In addition, not only up- or down-regulation of genes affects the overall biological outcome of drug exposure, e.g., as a consequence of post-translational modification and changes in cellular location determine protein function and, hence, the cell biologically. 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 (Wink et al. 2016). Using bacterial artificial chromosome recombineering technology (Poser et al. 2008), we tagged individual components of three different stress response pathways with green fluorescent protein (GFP): oxidative stress, unfolded protein response and DNA damage (Wink et al. 2016). 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 (Wink et al. 2016, 2018).

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 (Soldatow et al. 2013). Major shortcomings of the use of primary human hepatocytes are the limited batch size, donor variability, and high costs (LeCluyse et al. 2012). The liver carcinoma cell line HepG2 is an economical, 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 (Wink et al. 2018). A caveat of HepG2 cells cultured in a conventional 2D monolayer is the high proliferative capacity and a dedifferentiated phenotype (LeCluyse et al. 2012). 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-like phenotype, which is suitable for high-throughput screening (Ramaiahgari et al. 2014). By culturing HepG2 cells in a matrigel matrix, 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 such 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.

Materials 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, perhexiline, 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, danazol, furosemide, meclizine, methimazole, mexiletine, nitrofurantoin, nitrofurazone, primidone, procyclidine, propylthiouracil, sulindac, tacrine, thioridazine, trazodone, and zafirlukast (kind gift from Dr. Weida Tong, NCTR-FDA; all

vendors are listed in Supplemental Table 6). 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, polyclonal), 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). HepG2 GFP BAC reporters for human *SRXN1*, *NQO1*, *CDKN1A* (p21), *BTG2*, *DDIT3* (Chop), and *HSPA5* (BiP) have been generated previously (Wink et al. 2016). For experiments, HepG2 wt and GFP-BAC cell lines were cultured in phenol red-free Dubelco's modified Eagles medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS), 25 U/ml penicillin, and 25 μg streptomycin (PSA, Invitrogen).

3D cell culturing

Matrigel (cat. No: 354230) 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 as described (Ramaiahgari et al. 2014). 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 h after exposure. LC–MS analysis of metabolites was performed by Janssen Pharmaceutics (Beerse, Belgium) in a similar way to what has been described earlier (Tanaka et al. 2014).

Cell treatment and viability

Compound exposures were performed in three different ways: 48-h single exposure, 6 days repeated, and 11 days repeated exposure. For 6 or 11 days repeated exposure, every 24 h, the medium was aspirated from the cells and freshly diluted compound in fresh medium was added. 24 h after the last exposure, the imaging was started. For 48 h single exposure and 6 days repeated exposure, the imaging was started at day 28 after seeding the cells in 384 wells. For 11 day 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 (Ramaiahgari et al. 2014). Cell death was calculated as a percentage of viable cells after compound exposure compared to its vehicle control, assuming that 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 Fig. 2). 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 individual identified spheroids and 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 (Wink et al. 2016). Raw CEL files can be downloaded at: http://dbarchive.biosciencedbc.jp/en/open-tggates/desc.html.

RNA isolation and microarray analysis of 2D and 3D HepG2 cells: RNA was extracted from 3-day cultured 2D HepG2 cells and 3D HepG2 cells cultured for 3, 7, 14, 21, and 28 days. Total RNA was extracted from 2D/3D cultured cells using Trizol reagent (Invitrogen, Carlsbad, USA) followed by clean up using RNeasy mini kit (Qiagen, Hilden, Germany). RNA quality and integrity were determined using the Agilent bioanalyzer (Agilent Technologies Inc, Santa Clara, USA). Biotinylated cRNA was prepared using the Affymetrix 3′ IVT-Express Labeling Kit (Affymetrix, Santa Clara, USA) and hybridization steps were performed by Service XS B.V (Leiden, The Netherlands) on Affymetrix HT Human Genome U133 plus PM plate. Array plates were scanned using the Affymetrix GeneTitan scanner. Raw cel files are submitted to GEO (Number: GSE128763). BRB Array Tools software (developed by Dr. Richard Simon and BRB-ArrayTools Development Team) was used to normalize the cel files data using Robust Multichip Average (RMA) method. Differentially expressed genes between the various experimental conditions were identified with an ANOVA test followed by calculation of the false discovery rate according to Benjamini and Hochberg on Partek Genomics Suite 7.0 (Partek, Missouri, USA). Principal component analyses (PCA) of basal 2D/3D HepG2 gene expression were performed on Partek Genomics Suite 7.0 (Partek, Missouri, USA). Classification of the selected genes according to their biological and toxicological functions was performed using the Ingenuity Pathway Analysis IPA® software (Ingenuity Systems, Redwood, USA). Heatmap representations and hierarchical clustering (using Pearson correlation) were performed using the Multi Experiment Viewer software.

Re-annotation, normalization, and data filtering for comparative gene expression profiling across various in vitro and in vivo models

To compare basal gene expression data from 2D and 3D HepG2 cells with other liver tissue/cell models, different data sources were combined. Raw data files from untreated HepaRG and HepG2 cells were obtained from the department of Toxicogenomics, Maastricht University, and GEO database (3D HepG2—GSE128763; HepaRG—GSE109511; 2D HepG2—GSE109513). Primary cryopreserved human hepatocyte data were downloaded from TG-GATEs and GEO (PHH1—GSE53399;

PHH2—GSE122660); Postmortem liver data were from GEO (GSE13471,GSE3526, and GSE107037). Raw data files were loaded into R version 2.15.2 for Windows (64-bit), re-annotated to EntrezGene using Brainarray's custom CDF version 15.1.0. The R-packages used were obtained from BioConductor version 2.11. Since the combined data set was originating from the Affymetrix Human Genome U133 Plus 2.0 and Affymetrix HT Human Genome U133 plus PM (GeneTitan) platforms, normalization was performed in a multi-step procedure. Data from different chip types were merged based on overlapping probe identifiers, followed by scaling of the data and quantile normalization on the merged set (Quackenbush 2002). Principal Component Analysis (PCA) was applied to identify data patterns and to highlight data similarity and differences between the treated and untreated cell lines at different time points. PCA analysis was performed on whole-genome expression and filtered gene sets of hepatocyte-specific canonical pathways obtained from pathway annotation of Ingenuity Pathway Analysis (IPA). Differentially expressed genes ($p < 0.05$; FDR < 0.05, fold change ± 1.5) of 2D/3D HepG2 and HepaRG compared to human liver and primary human hepatocyte expression were calculated using Partek Genomics Suite 7.0 and provided in supplemental data. PCA data were also visualized using Tibco Silver Spotfire (Paulo Alto, CA, USA).

Transcription factor analysis

To further characterize the functional status of transcriptional regulatory circuits, prediction of transcription factor (TF) activities based on TF-target relationship was assessed. TF activities were estimated following the pipeline proposed in the Dorothea tool version 2 (https://saezlab.github.io/DoRothEA/, (Garcia-alonso et al. 2018)). Differentially expressed genes calculated for each pair 2D 3 days—3D other days ($p_{val} < 0.001$, |FC| > 1.5). In case of multiple probe-gene annotation, the probe with highest FC was included. To estimate TF rank for each treatment, we used the manually curated human regulon of Dorothea v2 (Garcia-alonso et al. 2018). Version 2 of DoRothEA provides updated TF regulons derived from a broader collection of resources and strategies. The new TF regulons are signed (to account for activation/repression), when possible, and each TF-target interaction has been assigned a confidence score, ranging from A to E, being A the most confident interactions. Confidence set ABC was chosen for this analysis. We estimated TF activities as a proxy of the expression levels of the targeted genes using the analytic Rank-based Enrichment Analysis (aREA) method from the viper R package 2, a statistical test based on the average ranks of the targets. The approach assumes that the activity of a TF can be estimated

from the mRNA levels of its direct target genes. TF scores have been z-scored TF-wise for all the samples.

Quantitative image analysis

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) as reported previously (Wink et al. 2016). Based on the Hoechst33342 image, CellProfiler defines two objects: a mask of the entire spheroid and a mask of all the single nuclei combined 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, resulting in values for normalized GFP intensities.

Data processing and statistics

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 using the two-tailed Grubb's test. Fold changes were considered significant as $p < 0.05$ in a one-tailed Student's t test.

Results

Transcriptome analysis of 3D spheroid HepG2 differentiation

Previously, we reported differentiation of HepG2 cells when cultured as 3D spheroids with increased liver-like properties as a result (Ramaiahgari et al. 2014). Our objective in the current study was to assess the applicability of HepG2 fluorescent cellular stress response reporters in 3D differentiated screening systems. As a first step, we further characterized the 3D differentiation in more detail using detailed whole-genome transcriptome analysis in comparison with other liver model systems. First, we evaluated the gene expression profile of HepG2 cultured in 2D monolayers with HepG2 cultured as 3D spheroids (Fig. 1a–d; Supplemental Fig. 1). Gene expression profiles of 2D monolayer at 3 days after seeding as well as 3D spheroid samples at 3, 7, 14, 21, and 28 days after seeding were retrieved. A change in the gene expression of 3D cultured genes was observed with a strong increase at day 14 (Supplemental Fig. 1A). This shift in gene expression changes after day 14 can also be observed in a principal component analysis (PCA) (Fig. 1b). Genes representing differentiated liver

functions associated with fetal hepatocyte differentiation were strongly up-regulated (Fig. 1c) (Takebe et al. 2013). To further examine these changes, regulatory transcription factors were extracted in the comparison between 2D and 3D samples. Top 50 transcription factors revealed major changes in liver-specific transcription factors (HNF4α, C/EBP, EGR1, and FOXA1) (Fig. 1d; Supplemental Table 2) (Guerquin et al. 2013; Guzman-lepe et al. 2018). For the top five TFs, we observed large down-regulated regulons for E2F4 and ETS1 (mainly involved in cell cycle) (Hsu and Sage 2016; Fry and Inoue 2018) with suppressed activity over time, and large up-regulated regulons for HNF4α, C/EBP, and STAT1 (mainly involved in differentiation) with increased activation over time (Fig. 1d; Supplemental Fig. 1B, C) with various transcripts in common for these differentiation TFs. These results support the improved liver-like properties in 28 day 3D spheroid cultured HepG2 cells (Ramaiahgari et al. 2014). To investigate whether these gene expression changes also correlate with gene expression profile in other liver systems, we compared 2D HepG2 monolayer cultures and 3D HepG2 spheroids (at 28 days) with primary human hepatocytes, HepaRG cells, and human liver. Important functional pathways of hepatocytes were selected and compared in the different liver models using a PCA plot (Fig. 1e). For all pathways selected, a shift is observed of 3D HepG2 spheroids from 2D HepG2 towards human liver. This supports the differentiation of HepG2 cells towards more liver-like cells when cultured in a 3D matrigel environment. We further systematically compared gene expression profiles (fold change ± 1.5; $p < 0.001$) of HepG2, primary human hepatocytes (PHH), HepaRG, and human liver (see Supplemental Table 1). Of particular interest, delta-like 1 homolog (DLK1) was > 100-fold higher in 3D HepG2 spheroids compared to PHH and HepaRG. In particular, we observed increased expression of factors involved in differentiation. Thus, DLK1 is associated with Notch signaling pathway, and was shown to have a role in development and differentiation of various tissue types such as liver, lung, adipose tissue, and skeletal muscle (Falix et al. 2012). Sex Determining Region Y-Box 9 (SOX9), which has a role in embryonic development of various tissues, was 20-fold and ninefold higher in 3D HepG2 compared to PHH and HepaRG, respectively. Similarly, Forkhead Box Q1 (FOXQ1) is 12- to 14-fold higher in 3D HepG2 spheroids (Bieller et al. 2001). As a likely consequence of this enhanced differentiation, the terminal hepatocyte differentiation marker Glucose-6-phosphatase (G6PC) gene was 4- to 11-fold higher in 3D HepG2 spheroids compared to PHH and HepaRG. Cell cycle-related genes such as CCND2, CCND3, CDK9, CDKN1A, CFL1, ECT2, KIF11, MDM2, etc. were down-regulated in 3D HepG2 spheroids compared to PHH, and this effect was much more pronounced in comparisons with HepaRG cells, indicative of the non-proliferating differentiated

A

day 3   day 7   day 14   day 21   day 28

HepG2 cells in 3D culture

B

Attribute
- 2D day 3
- 3D day 3
- 3D day 7
- 3D day 14
- 3D day 21
- 3D day 28

C

D

E

Cell cycle regulation

Xenobiotic metabolism

Liver proliferation

Bile acid biosynthesis

Complement system

Liver metabolism

- Human liver
- PHH
- HepaRG
- 3D HepG2
- 2D HepG2

◄**Fig. 1** Transcriptome profile of HepG2 cells cultured as 3D spheroids. **a** Phase contrast images of HepG2 cells in 3D Matrigel cultures over a period of 28 days. Scale bars 100 μm. **b** Principal Component Analysis plot of basal-level whole-genome gene expression of 2D HepG2 cells at day 3 and HepG2 spheroids at various timepoints in 3D culture. **c** Fold change expression of genes involved in differentiation and development of the liver (see gene list Takebe et al.) in HepG2 spheroids at various time periods in 3D culture compared to 2D monolayer cultures, data are average (four experiments) fold change compared to 2D HepG2 gene expression. **d** Transcription factor (TF) rank plot: TF (*y*-axis) are ranked for the number of the associated differentially regulated target genes (*x*-axis) in the different 2D vs 3D time point comparisons. Colors represent the different 2D vs 3D time point comparisons. TFs are ranked for the comparison 2D vs 2D-day28 (violet). Only top 50 TFs are shown. **d** Top right panel. Activity of top 5 TFs plot: TF activity (*y*-axis) is shown for the 5 TF that have largest regulon differentially expressed (colors) for the different 2D vs 3D time points comparisons (*x*-axis). TF scores have been *z*-scored TF-wise for all the comparisons. **e** Principal component analysis of pathways related to hepatocyte function and differentiation: cell cycle regulation, xenobiotic metabolism, bile acid biosynthesis, complement system, liver proliferation, and liver metabolism, in 2D/3D HepG2 cells, primary human hepatocytes, HepaRG cells, and human liver (color figure online)

phenotype of HepG2 spheroids. On the contrary, relative gene expression levels of cytochrome P450 enzymes were generally higher in PHH and HepaRG cells, indicating that the enhanced differentiation did not lead to equal levels of metabolism enzymes. Analogous to previous findings (Limonciel et al. 2018), basal-level expression of GAGE genes which are associated with cancer were significantly higher in HepaRG cells with an 80- to 100-fold change increase. Apolipoprotein C2 (APOC2), Inter-Alpha-Trypsin Inhibitor Heavy Chain 3 (ITIH3), Protein C Inactivator Of Coagulation Factors Va And VIIIa (PROC), DnaJ Heat Shock Protein Family (Hsp40) Member C15 (DNAJC15), and Thyroid Hormone Responsive (THRSP) were top down-regulated genes in HepaRG compared to PHH (Supplemental Table 1).

Liver-like spheroid differentiation of fluorescent reporter HepG2 cell lines

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 not present in 2D HepG2 monocultures; e.g., bile duct formation (MRP2), albumin expression, and basal-lateral polarization (β-catenin) (Ramaiahgari et al. 2014). All BAC-GFP HepG2 reporter cell lines analyzed acquired the similar type of differentiation as HepG2 wild type (Fig. 2a, b).

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 compound-specific CYP450 enzyme activity allowed the exact determination of the CYP450 enzyme activity. Enhanced metabolizing capacity of CYP3A4, CYP2C9, CYP2B6, CYP1A2, and CYP2D6 was observed in 3D cultures for all reporter cell lines tested, to a similar extend as wild-type HepG2 spheroids. These combined data demonstrate a differentiated liver-like phenotype of all GFP reporter cell lines.

Confocal imaging-based detection of GFP reporter activity in 3D liver spheroids

We previously determined the specific activation of GFP stress response reporter activation dynamics in 2D monolayer reporter cultures (Wink et al. 2016). To determine the functionality of our reporters in 3D spheroids, we applied 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 followed by quantitative image analysis (Wink et al. 2016). All six reporter cell lines demonstrated strong and significant up-regulation of GFP intensity after stimulation with reference inducers (Fig. 3a, b). 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.

3D spheroid reporter responses represent drug-induced gene activation in primary human hepatocytes

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 (Supplemental Table 3). For each of the six GFP reporter genes, we ranked the TG-GATES compounds (at high dose, 24 h 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 up-regulation of
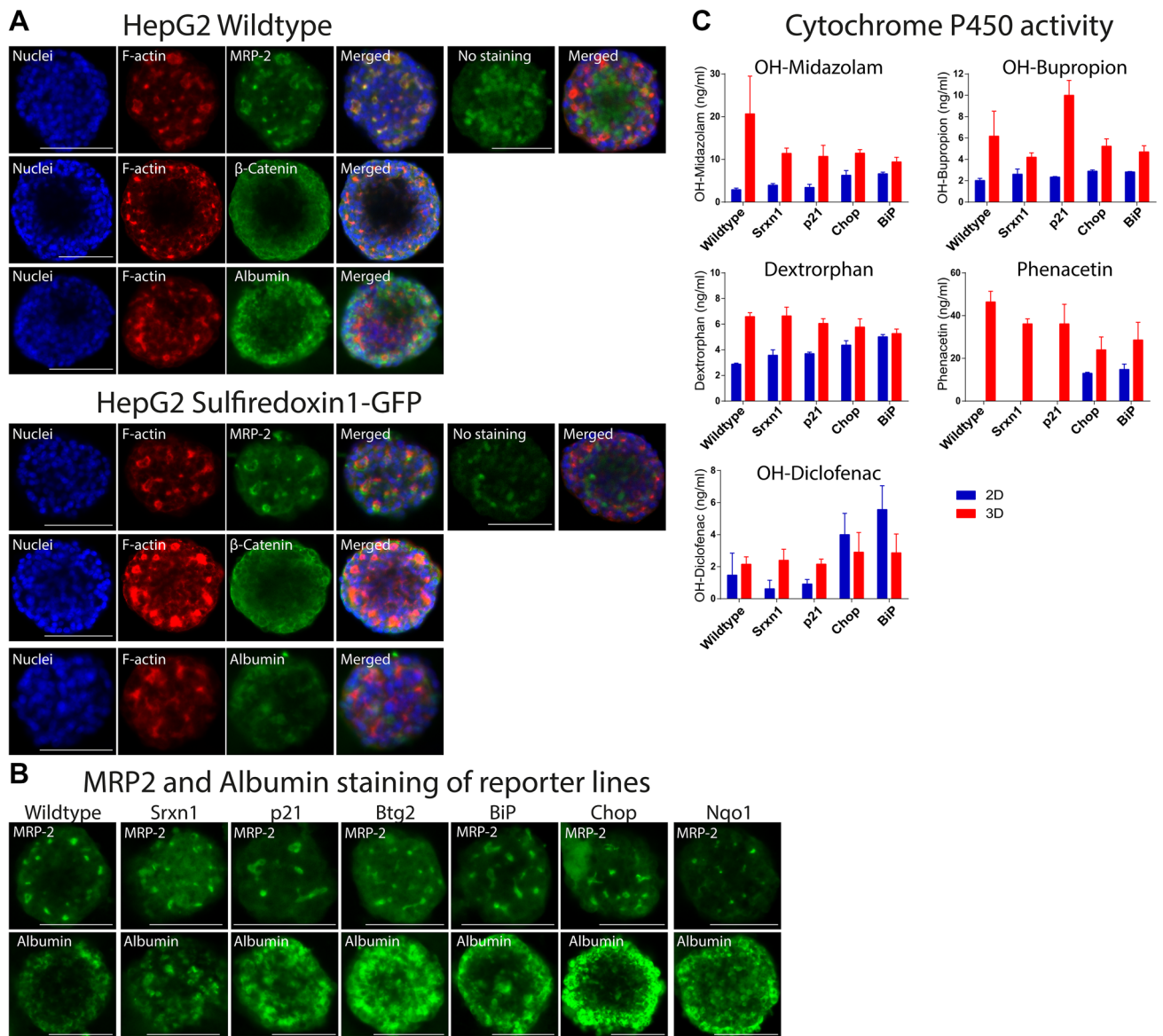
Fig. 2 Characterization of the differentiated phenotype of 3D liver-like spheroid HepG2 GFP reporter cell lines. **a** Wild-type (upper panel) and Srxn1-GFP (lower panel) spheroids are stained with nuclear staining Hoechst$_{33258}$ (blue), F-actin staining rhodamine phalloidin (red). Immunofluorescent 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 wild-type 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 h. 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 as represented by the average ATP content of 80 parallel 384-well 2D and 3D cultures (color figure online)

mRNA levels of the respective genes in PHH. Since a major advantage of 3D HepG2 spheroid cultures is the ability to perform repeated treatments (Ramaiahgari et al. 2014), we also assessed which exposure regimen did correspond best with the PHH TG-GATEs transcript data: 48 h single exposure, 6 day repeated exposure, or 11 day repeated exposure

treatment conditions. For direct comparison, we used the same three concentrations used in TG-GATEs (Supplemental Table 4). To apply the 3D GFP reporter system for pharmaceutical screening automated imaging, we established an automated imaging strategy (see Supplemental Fig. 2). All 3D GFP reporter lines show significant increases in

Fig. 3 Application of 3D spheroid HepG2 GFP reporters in automated confocal microscopy. Representative images of single 3D spheroids for Srxn1-GFP and Nqo1-GFP (150 µM *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 0.1% DMSO control. The graphs show response of individual spheroids as a fold change over average DMSO intensity; colors represent different biological replicates (color figure online) ▶

the TG-GATES (Fig. 4; Supplemental Fig. 3). In addition, all reference control compounds were positive in their corresponding reporters; for some maximal response was observed after 48 h (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 2D reporter activity (Supplemental Fig. 3) (Wink et al. 2016). For the selected TG-GATEs compounds, most significant reporter activation was observed for the 6 and 11 day repeated exposure compared to 48 h single exposure. In particular, the Srxn1-GFP and BiP-GFP reporters were most representative for the TG-GATEs' PHH transcriptomics responses (Supplemental Fig. 3). For all six reporter lines, the average combined sensitivity was 86%, with Srxn1-GFP as a single reporter even a 100% sensitivity (Supplemental Table 4). Thus, the 3D repeated exposure reporter activity showed high resemblance with gene expression data of primary human hepatocytes. This notion was further supported when comparing the fold changes of the TG-GATEs data set with the maximal fold change obtained from our 3D reporter system (Supplemental Table 5). Taken together, we conclude that our 3D BAC-GFP stress response reporter system shows high resemblance with responses observed in PHH.

Drugs with clinical DILI concern activate stress response reporters in 3D spheroids

To further 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: 20 drugs categorized as most-DILI-concern drugs, 7 drugs with less-DILI concern, and 6 drugs with no-DILI-concern drugs were selected (Supplemental Table 6). These compounds were selected based on limited sensitivity in our 2D HepG2 reporter study (Wink et al. 2018). The evaluation was conducted with six GFP reporter cell lines: Srxn1-GFP, Nqo1-GFP, p21-GFP, Btg2-GFP, BIP-GFP, and Chop-GFP. 3D spheroids were treated for 48 h single exposure or 6 day repeated exposure using concentrations that were based on peak plasma levels measured after administrations to patients during clinical trials (C_{max}): 1, 5, 10, 25, 50, and 100 times C_{max} (Dawson et al. 2012; Khetani et al. 2013; Jenkins et al. 2009; Keisu and
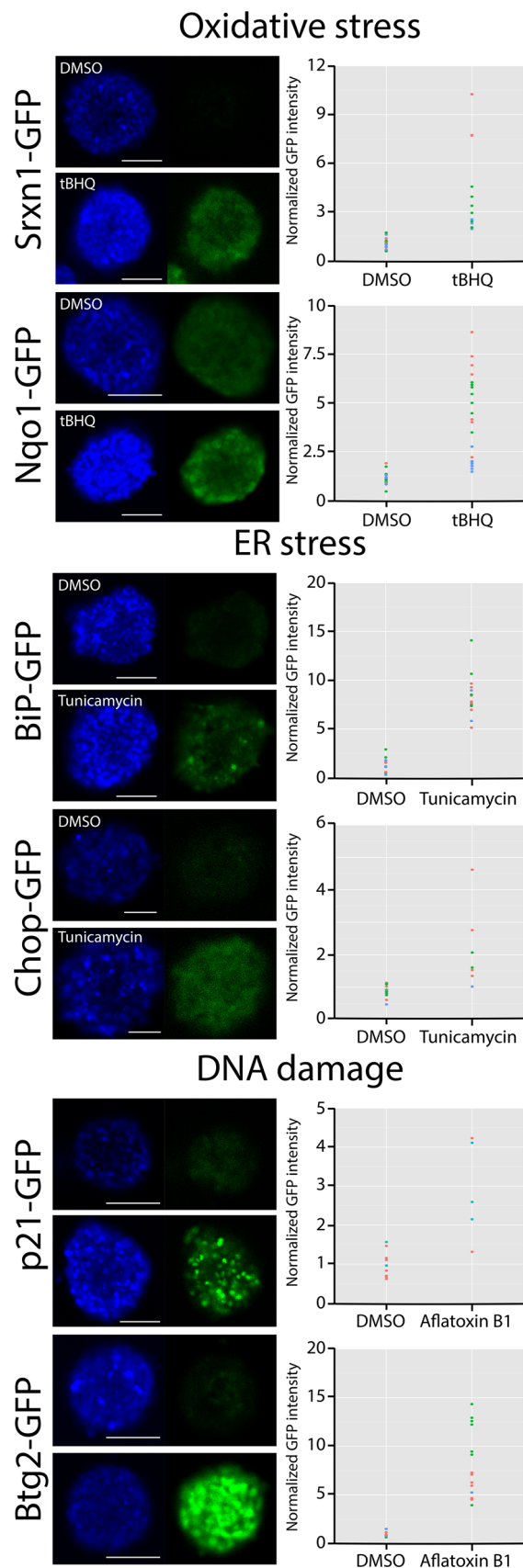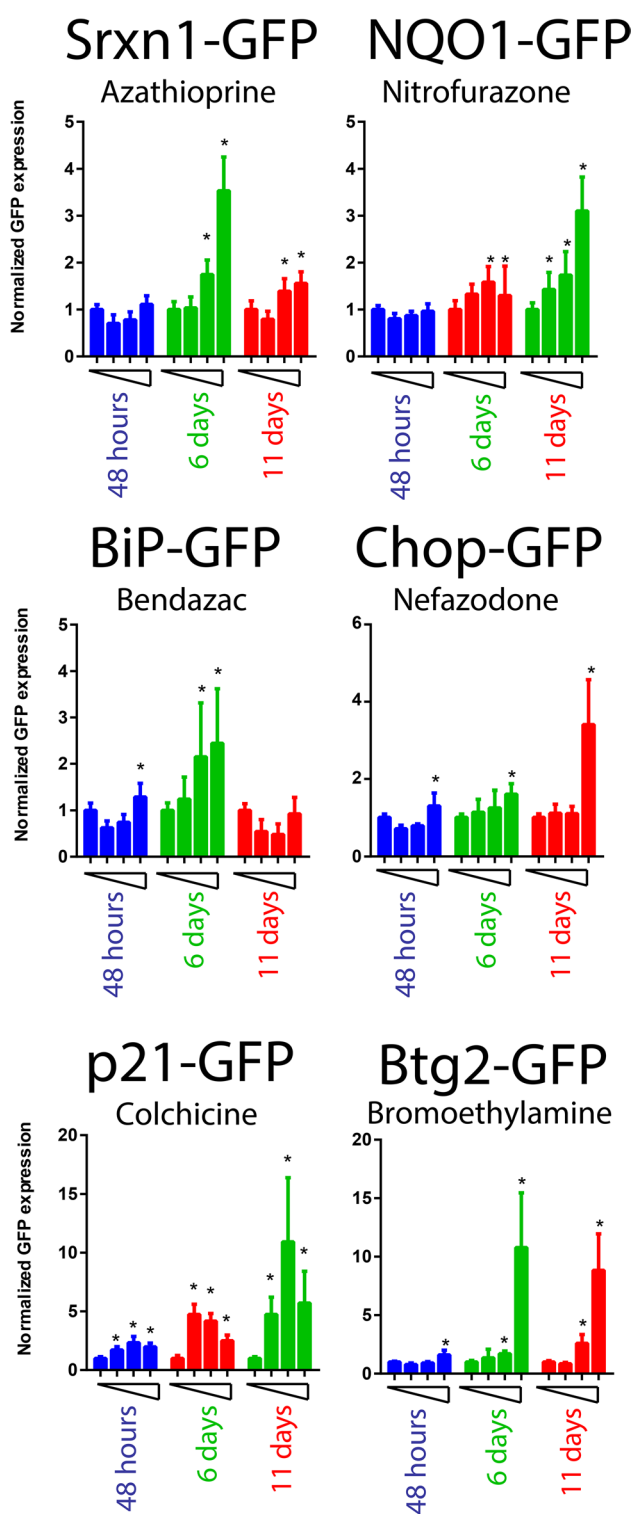
Fig. 4 Example of cellular stress response GFP reporter activation by reference DILI compounds in 3D spheroids. Srxn1-GFP, Nqo1-GFP, BiP-GFP, Chop-GFP, p21-GFP, and Btg2-GFP were exposed to ten different compounds from top 12 inducers in the TG-GATES primary human hepatocyte data set. For each reporter, one example compound is shown for 48 h single exposure (blue), 6 day repeated exposure (green), and 11 day repeated exposure (red). Graphs are the fold-change of average DMSO response 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 data set (see Supplemental Table 5) (color figure online)

Andersson 2010; Regenthal et al. 1999; O'Brien et al. 2006; Xu et al. 2008; Gustafsson et al. 2014). Six positive reference control compounds, DEM, tunicamycin, and aflotoxinB1 were included on all imaging plates (Supplemental Table 7).

To verify compound potency for onset of cytotoxicity, we first compared the effect on viability between single and repeated exposure in parental HepG2 3D spheroids (Supplemental Fig. 5). Several DILI compounds caused cytotoxicity after 48 h at the highest concentrations, which was further increased with 6 day repeated exposure, in particular for compounds with severe-DILI liabilities; cell viability remained constant for compounds with no-DILI concern.

Next, we systematically determined the GFP reporter activation in individual spheroids for all DILI compounds in the six reporter cell lines. An increase of GFP reporter activity was observed for various DILI compounds, also those for which responses in 2D HepG2 have so far not been observed (Wink et al. 2018). Some variability in responses between spheroids was evident for different reporters, indicating the importance to include multiple spheroids in the analysis (Fig. 5). In general, stress response activation was strongest affected by DILI compounds after 6 day repeated exposure (Supplemental Fig. 4). The positive control group showed significant concentration-dependent up-regulation 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 day repeated exposure. This correlates with a strong increase in overall cytotoxicity (see Supplemental Fig. 5 and Supplemental Table 6). BiP-GFP was most sensitive at 6 day repeated exposure, followed by Srxn1-GFP; no direct overlap in responses was observed between BiP-GFP and Srxn1-GFP activation. BiP-GFP and Srxn1-GFP were most sensitive for the majority of most-DILI-concern compounds; Nqo1-GFP was not significantly up-regulated by any of the compounds, likely due to high basal levels of the Nqo1-GFP under control conditions. Interestingly, trovafloxacin induced all three different pathways; remarkably, the less-DILI-concern compound adefovir did activate the BiP-GFP, Btg2-GFP, and p21-GFP reporters. We performed hierarchical clustering of the 6 day repeat exposure data to compare compound responses (Fig. 6). Six clusters were distinguished (clusters I–VI): five clusters (clusters I–V) with activation of multiple pathways and one cluster (cluster VI) with minor activation in one pathway (mainly BiP-GFP). Three clusters are formed by one compound (trovafloxacin, adefovir, and bicalutamide), for their distinct activation of BiP-GFP (trovafloxacin and bicalutamide) and p21-GFP (adefovir). Clusters III and V are characterized by an activation of mainly BiP-GFP and Srxn1-GFP in the higher concentrations. 5 out of 7

Fig. 5 Concentration response activation of 3D spheroid GFP reporter 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 h single exposure and 6 day repeated exposure. The 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 $**p < 0.01$. 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
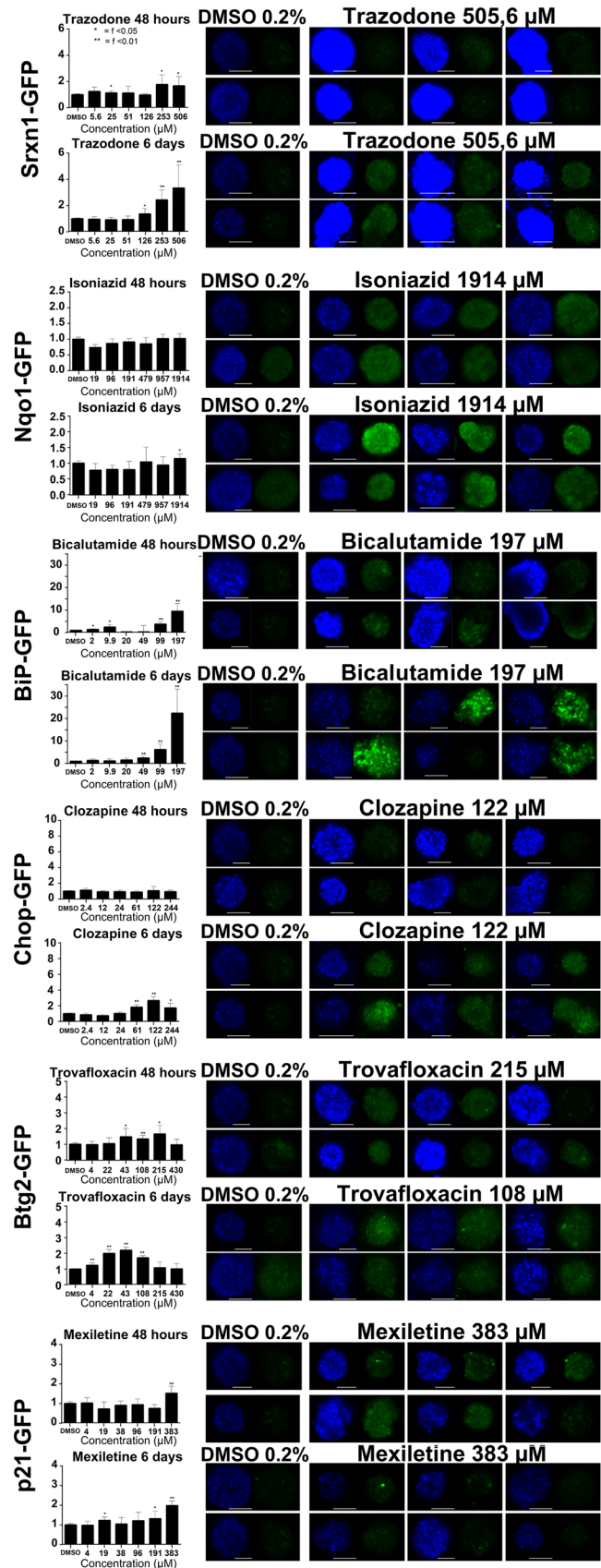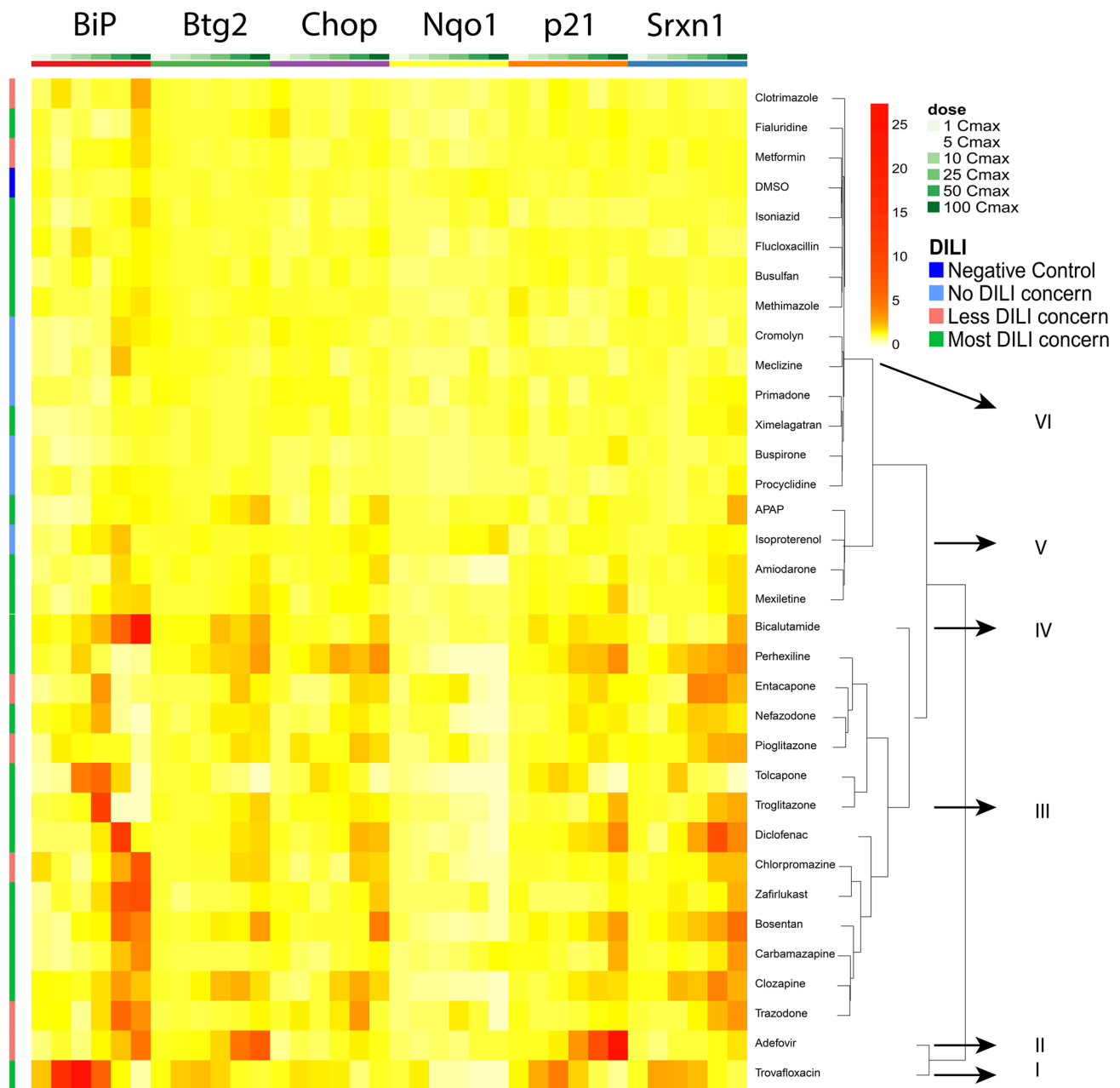
Fig. 6 Hierarchical clustering of reporter activation of different reporters by compounds with different DILI-concern. Heatmap was based on GFP fold-change responses depicted in Supplemental Fig. 4; red intensity is indicative for higher FC response. Ward.d hierarchical clustering was performed with all GFP responses of all 6 day repeated exposure tested. Left color bar indicates type of compound: purple for negative control DMSO, light blue for no-DILI-concern drugs, red for less-DILI-concern drugs, and green for most-DILI-concern drugs. The six reporters are depicted in alphabetical order, with concentration responses from 1, 5, 10, 25, 50, and 100 C_{max} (from left to right). Arrows indicate the branch where the cluster diverges from the main tree (color figure online)

less-DILI-concern compounds are in activated clusters I–V. Cluster VI also contained DMSO control as well as 5 out of 6 no-DILI-concern compounds; also 6 out of 20 most-DILI-concern compounds. Thus, hierarchical clustering resulted in an overrepresentation of most-DILI-concern compounds in activated clusters, suggesting 70% sensitivity and 83% specificity.

Finally, we compared the response in our 3D HepG2 GFP reporter DILI screen with our recently reported 2D HepG2 GFP reporter DILI screen data (Wink et al. 2018). Benchmark concentrations (BMC) were calculated, defined as the concentration where the GFP fold change exceeded the 1.5 threshold (Ritz et al. 2015). Nqo1-GFP was left out of this analysis due to a lack of activation. For only one

no-DILI-concern compound, BMC values could be determined for Srxn1-GFP and Btg2-GFP. Compounds with a relatively low BMC in relation to the human plasma C_{max} value were typically most-DILI-concern compounds (trovafloxacin, clozapine, perhexiline, and bicalutamide). In addition, the dotted regression line shows a lowering trend in BiP-GFP, Btg2-GFP, and Srxn1-GFP for most-DILI-concern drugs compared to less- and no-DILI-concern drugs. For four drugs that were not detected in 2D monolayer, bicalutamide, mexiletine, trazodone, and zafirlukast, the

concentration response curves for Srxn1-GFP, Chop-GFP, and p21-GFP reporters were compared between our current 3D spheroid screen and our previous 2D live reporter screen (Wink et al. 2018). In 2D, only few drugs exceeded the 25% GFP positive cell threshold; in 3D spheroids, almost all four compounds reached the 1.5-fold change threshold (Fig. 7b). We also compared the 2D and 3D BMC values for the Srxn1-GFP, Chop-GFP, and p21-GFP reporters, and defined the compounds for which a BMC could be derived in either 2D and/or 3D in any of the three reporters (Supplemental
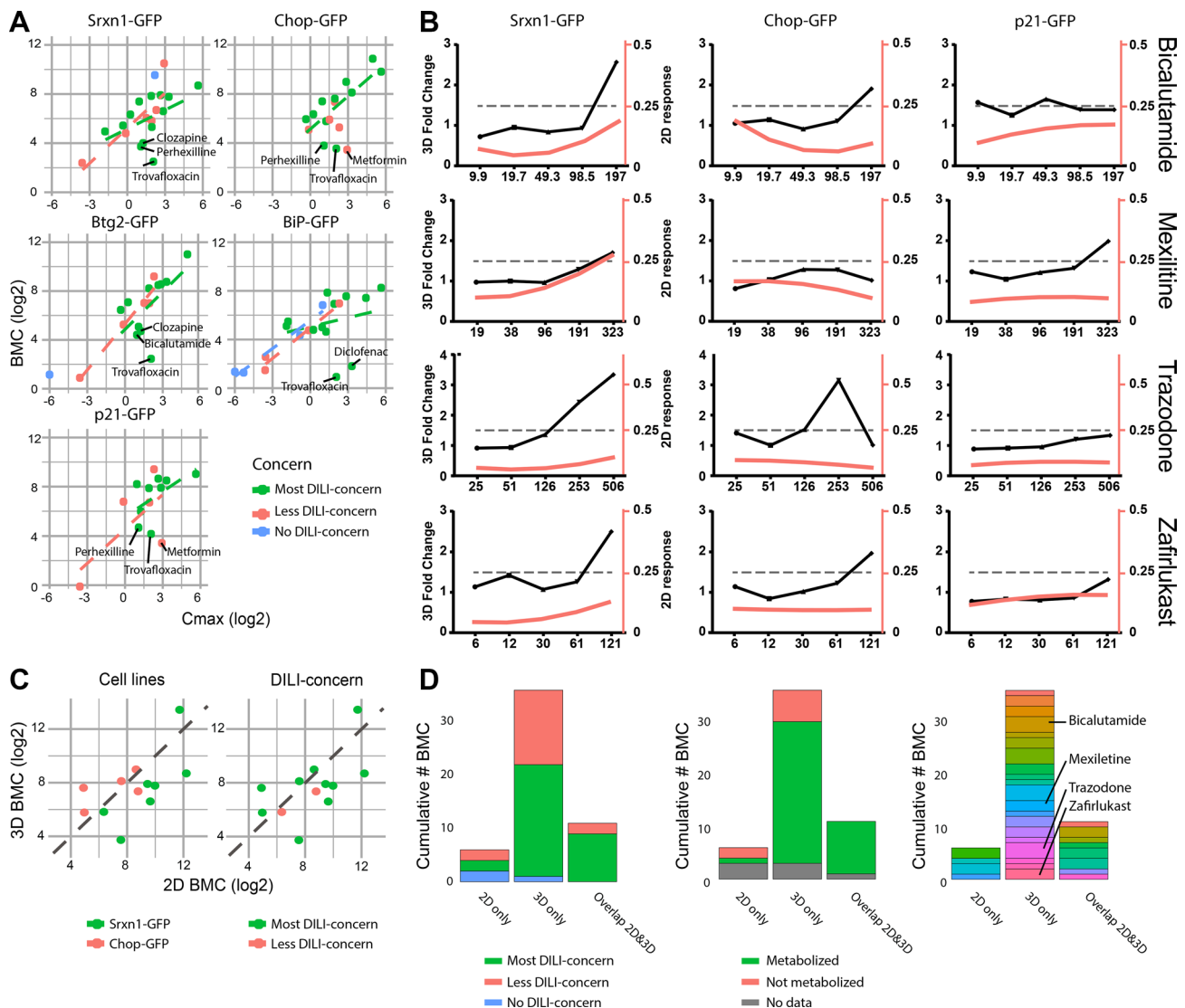

Fig. 7 Comparison of 2D monolayer and 3D spheroid GPF reporter responses by DILI compounds. **a** 3D 6 day repeated treatment BMC values (*y*-axis) depicted with corresponding C_{max} values (*x*-axis). Dotted lines indicate regression separated for the different DILI-concern classes. Most-DILI-concern in green, less-DILI-concern in red, and no-DILI-concern in blue. **b** 2D versus 3D dose-range comparison for bicalutamide, mexiletine, trazodone, and zafirlukast. The black line represents the 3D concentration curve; the red line depicts the 2D response as percentage GFP positive cells as most sensitive marker

(Wink et al. 2016). The dotted line is the BMC threshold in 2D and 3D. **c** 2D (*x*-axis) plotted against the 3D BMC (*y*-axis) for the overlapping compounds in the different cell lines. The compounds were color coded for distribution for the cell lines (left) and the DILI-concern (right). **d** Histogram of accumulative number of compounds with available BMC determined in 2D only, 3D only, or overlap-2D&3D. The compounds were colored for DILI-concern (left), metabolism information (middle), and the treatments (right) (color figure online)

Table 8). This resulted in a group of compounds for which a BMC was derived in "2D only", "3D only", or "overlap 2D&3D". For the "overlap 2D3D" group, typically, a lower BMC was derived in the 3D spheroid condition, which was particular in the Srxn1-GFP reporter and largely reflecting severe-DILI-concern compounds (Fig. 7c). No overlap was observed between 2D and 3D for the p21-GFP reporter; only one less-DILI-concern drug showed an overlap in 2D and 3D and none of no-DILI-concern was detected under both conditions. Finally, we compared the overall positive calls for both 2D and 3D reporter screen results based on BMC levels; for this, we addressed three different labeling conditions: FDA DILI-concern label, involvement of drug metabolism, and DILI compound (Fig. 7d). The majority of BMCs were derived under 3D spheroid reporters conditions, and reflected in particular less-DILI-concern compounds, that involve liver drug metabolism.

Discussion

Here, we have systematically evaluated the application of HepG2 BAC-GFP cellular stress response reporters as 3D differentiated liver-like spheroid systems for the assessment of DILI liability. Our results indicate that various cellular stress response reporters have a differentiated phenotype in 3D matrigel culture, show enhanced metabolic capacity, and are amenable for confocal imaging-based reporter activation assessment in a high-throughput manner. The systematic analysis of the reporter activation of a subset of DILI compounds indicates that improved identification of DILI liability in 3D spheroids, compared to 2D monolayer. The BiP-GFP and Srxn1-GFP reporter cell lines seem most promising for implementation in 3D spheroid-based DILI assessment.

To our knowledge, this is the first fluorescent protein-based reporter system to follow and quantify the activation of endogenous stress response protein expression in spheroid systems. While primary human hepatocyte spheroids have been applied for DILI assessment (Proctor et al. 2017), the biomarkers available to cost-effectively quantify endogenous biomarkers for mode-of-action at a single spheroid level in a high-throughput setting have not been possible so far. We combined our GFP reporter platform and the enhanced liver-like properties of HepG2 3D spheroid cultures. We reported that our 3D HepG2 spheroid model has enhanced differentiated liver-like properties which is related to increase in transcriptional activity of HNF4α, CEBP, and STAT1, which are involved in liver differentiation, in association with decreased activity of E2F4 and EST1, that are involved in liver cell proliferation. HNF4α promotes hepatocyte cell polarity and differentiation (Parviz et al. 2003), and drives the activation of promoter regions of G6PC and

PCK1. The two latter genes are considered key markers of differentiated hepatocytes (Hall et al. 1995; Boustead et al. 2003; Parviz et al. 2003). Of relevance, we found that G6PC and PCK1 were top two up-regulated genes at day 28 in 3D HepG2 spheroids with ~ 235- and 132-fold increase in gene expression compared to 2D HepG2 cultures. C/EBPα also plays a major role in differentiation and 'hepatogenesis' (Yuan et al. 2015). STAT1 was shown to negatively regulate cell proliferation of hepatocellular carcinoma and in differentiation of metanephric mesenchymal progenitor cells during kidney development (Wang et al. 2010; Chen et al. 2013). These data support the more liver-like phenotype of the 3D HepG2 system. While the 3D HepG2 spheroids showed close association with other more differentiated liver test systems, our data indicate that the cells reflect a fetal hepatocyte status. We demonstrated that GFP HepG2 reporter cell lines also differentiate in a comparable manner as the HepG2 wild-type cells, as evidenced by various differentiation markers including formation of MRP2-positive bile canaliculi-like structures, albumin protein expression, and enhanced phase 2 metabolizing enzyme activity. We established a high-throughput confocal microscopy method to automatically acquire and analyze GFP reporter activity in 3D spheroids. Well-established positive reference compounds showed robust up-regulation 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 to 100%. By establishing this concordance, we make note of the downsides of the TG-GATES data set, since this data set contains only two replicates and different TG GATES compounds with unusual concentration responses. Grinberg et al. (2014) listed SV3 genes which do show convincing results. Four of the six genes that were compared with our HepG2 data were part of this list. Therefore, we conclude a high concordance with primary human hepatocytes.

Our current 3D spheroid reporter platform seems to outperform the same reporter cells when used in a 2D monolayer configuration in the context of identification of DILI liabilities. In our current 3D screen, we purposely included severe-DILI-concern compounds that were previously observed as negative in our 2D reporter setting (Wink et al. 2018), including bicalutamide, trazodone, and zafirlukast. These compounds caused a significant concentration-dependent activation of different reporters in the 3D spheroid configuration. Comparison of BMCs between 3D spheroids and 2D monolayer configurations further supported the superiority of the 3D setting. Of relevance, drugs that are metabolized in the liver did represent a large portion of the compounds where BMC definition was only possible in 3D spheroid reporter. This suggests a role for the more

differentiated phenotype, included drug metabolizing capacity, of the 3D spheroids, although we cannot exclude that the repeated dosing treatment regimen also contributes to this difference. Yet, repeated dosing in 2D monolayer is difficult due to the relatively high proliferation rate of HepG2 cells. Given the fact that 3D HepG2 spheroid cultures are differentiated and lack proliferating cells, we anticipate that the 3D spheroids reflect a more physiological situation with respect to stress response activation. As such, we would propose a tiered testing approach where the reporters in 2D monolayer reflect a rapid and cost-effective tier 1 testing setting allowing dynamic imaging of stress reporter activation (Wink et al. 2018) and the 3D spheroid would mimic a more physiological relevant conditions, but are more time consuming and costly.

Several of the selected most-DILI-concern compounds were inactive in our GFP reporter system, including 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 (Kindmark et al. 2008; Daly et al. 2009; Cheung et al. 2015; Heidari et al. 2015; Metushi et al. 2016). We anticipate that there are clear rationales why we would fail to identify these most-DILI-concerns compound in our current reporters. Flucloxacillin, ximelagatran and methimazole are more specifically linked to single-nucleotide polymorphisms in the human leukocyte antigen (HLA) allele, typically associated with iDILI responses. Genome-wide association study (GWAS) has identified HLA-B*5701 single-nucleotide polymorphisms involved in flucloxacillin-mediated liver injury(Daly et al. 2009). Similarly, for ximelagatran, a strong association was found between elevated ALAT levels and single-nucleotide polymorphisms in DRB1*07 and DQA1*02 (Kindmark et al. 2008). Also, for methimazole, an increased risk was observed for HLA-B*38:02:01 carriers and onset of agranulocytosis (Cheung et al. 2015). Also for isoniazid, involvement of immune-mediated mechanisms of liver injury may be of importance (Metushi et al. 2016). The mechanism of fialuridine induced liver injury has been suggested to be mitochondrial related (Mckenzie et al. 1995) 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 component downstream of cytokine signaling (Fredriksson et al. 2011), although more complex adaptive immune response signaling would not allow integration in reporter systems.

Isoproterenol is the only no-DILI-concern compound which is clustered in an activated cluster. Isoproterenol is a β-adrenal-receptor agonist which has been labeled with no-DILI-concern. In rats, isoproterenol induces TNF-α and IL-6 elevation indicating induction of an innate immune response (Deng et al. 2015). Possibly, this is related to mild stress responses which are observed in our reporter assay; but this would require more detailed investigations.

The less-DILI-concern drugs represent a group of drugs which induces different classes of mild liver injury. To analyze these mild liver injuries, these classes of DILI are added as were previously established (Supplemental Table 6) (Wink et al. 2018). Two less-DILI-concern drugs (clotrimazole and metformin) are part of cluster VI, which did not show elevation of stress responses in are screen. It should be noted that both drugs show only very mild clinical indications of DILI. Interestingly, less-DILI-concern drugs with hepatotoxicity ranking "cholestasis steatohepatitis" (adefovir and chlorpromazine) have increased stress responses in our assay, suggesting that these genes are up-regulated upon cholestasis or steatosis. In addition, elevations of stress responses by individual compounds can be evaluated in more detail in favor of mechanistic insight. These observations can be anchored in already existing adverse outcome pathways (AOPs) to elucidate the complexity of DILI toxicity. For the AOP that describes bile salt export pump inhibition was developed (Vinken et al. 2013), bosentan and troglitazone used in our study are listed. Oxidative stress was marked as one of the key event parts of the AOP. We found increased levels of Srxn1-GFP caused by both compounds, indicative for oxidative stress induction.

The enhanced liver-like properties present in 3D spheroids compared to 2D monolayer culturing not only provides the advantage of enhanced metabolizing capacity but also enables repeated dose exposures. After approximately 7–14 days, 3D spheroids stop proliferating and start differentiating (Ramaiahgari et al. 2014). 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 2 years after administration. During this period, patients obtain at least a daily dose of the drug. By applying fresh drug every day for 6 or 11 days, we resemble closer the human drug administration situation, which seems pivotal to hepatotoxicity screening. This is recognized by much higher sensitivity values in both 6 and 11 day repeated exposure compared to 48 h single exposure (Supplemental Table 4). 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 (C_{max}). For in vitro repeated dose studies, it has been shown that drugs can accumulate to high

intracellular concentrations (Wilmes et al. 2013). For the same drug, postmortem studies showed 53-fold increase compared to C_{max} levels (Lensmeyer et al. 1991). 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) C_{max} 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, BMC, viability, and TC50 values. Such a classifier will likely increase the sensitivity and specificity determinants.

The application of 3D spheroids for primary human hepatocytes has recently been systematically evaluated both assessment of adverse drug response and DILI liability (Bell et al. 2016; Proctor et al. 2017) as well as more recently to replicate disease states (Hendriks et al. 2016). These models show a strong physiological relevance with respect to gene and protein expression compared to human liver (Bell et al. 2018). These models currently rely on (commercial) (pooled) cryopreserved primary human hepatocyte sources with limited batch volume, thus continues requiring batch-to-batch validation and quality control. Moreover, the assessment of mechanistic mode-of-action biomarkers has yet not been integrated and validated in these physiological systems. Our 3D HepG2 fluorescent reporter platform may, therefore, be a robust, cheaper, and sustainable alternative solution and contributor to provide mechanistic information for DILI liability assessment.

References

Bell CC, Hendriks DFG, Moro SML et al (2016) Characterization of primary human hepatocyte spheroids as a model system for drug-induced liver injury, liver function and disease. Sci Rep 6:1–13. https://doi.org/10.1038/srep25187

Bell CC, Dankers ACA, Lauschke VM et al (2018) Comparison of hepatic 2D sandwich cultures and 3D spheroids for long-term toxicity applications: a multicenter study. Toxicol Sci. https://doi.org/10.1093/toxsci/kfx289

Bieller A, Pasche B, Frank S et al (2001) Isolation and characterization of the human forkhead gene FOXQ1. DNA Cell Biol 20(9):555–561. https://doi.org/10.1089/104454901317094963

Boustead JN, Stadelmaier BT, Eeds AM et al (2003) Hepatocyte nuclear factor-4α mediates the stimulatory effect of peroxisome proliferator-activated receptor γ co-activator-1 α (PGC-1 α) in H4IIE cells. Biochem J 369:17–22. https://doi.org/10.1042/BJ20021382

Chen G, Wang H, Xie S et al (2013) STAT1 negatively regulates hepatocellular carcinoma cell proliferation. Oncol Rep 29:2303–2310. https://doi.org/10.3892/or.2013.2398

Cheung C-L, Sing C-W, Tang CSM et al (2015) HLA-B*38:02:01 predicts carbimazole/methimazole-induced agranulocytosis. Clin Pharmacol Ther 00:1–7. https://doi.org/10.1002/cpt.309

Cui Y, Paules RSS (2010) Use of transcriptomics in understanding mechanisms of drug-induced toxicity. Pharmacogenomics 11:573–585. https://doi.org/10.2217/pgs.10.37

Daly AK, Donaldson PT, Bhatnagar P et al (2009) HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. Nat Genet 41:816–819. https://doi.org/10.1038/ng.379

Dawson S, Stahl S, Paul N et al (2012) In vitro inhibition of the bile salt export pump correlates with risk of cholestatic drug-induced liver injury in humans. Drug Metab Dispos 40:130–138. https://doi.org/10.1124/dmd.111.040758

Deng X-Y, Chen J-J, Li H-Y et al (2015) Cardioprotective effects of timosaponin B II from Anemarrhenae asphodeloides Bge on isoproterenol-induced myocardial infarction in rats. Chem Biol Interact 240:22–28. https://doi.org/10.1016/j.cbi.2015.08.001

Falix FA, Aronson DC, Lamers WH, Gaemers IC (2012) Possible roles of DLK1 in the Notch pathway during development and disease. Biochim Biophys Acta 1822(6):988–995. https://doi.org/10.1016/j.bbadis.2012.02.003

Fredriksson L, Herpers B, Benedetti G et al (2011) Diclofenac inhibits tumor necrosis factor-α-induced nuclear factor-kβ activation causing synergistic hepatocyte apoptosis. Hepatology 53:2027–2041. https://doi.org/10.1002/hep.24314

Fry EA, Inoue K (2018) Aberrant expression of ETS1 and ETS2 proteins in cancer. Cancer rep rev 2:1–23

Garcia-alonso L, Ibrahim MM, Turei D, Saez-rodriguez J (2018) Benchmark and integration of resources for the estimation of human transcription factor activities. bioRxiv 1–27. https://doi.org/10.1101/337915

Grinberg et al (2014) Toxicogenomics directory of chemically exposed human hepatocytes. Arch Toxicol 88(12):2261–2287. https://doi.org/10.1007/s00204-014-1400-x

Guerquin M, Charvet B, Nourissat G et al (2013) Transcription factor EGR1 directs tendon differentiation and promotes tendon repair. J Clin Invest 123:3564–3576. https://doi.org/10.1172/JCI67521DS1

Gustafsson F, Foster AJ, Sarda S et al (2014) A correlation between the in vitro drug toxicity of drugs to cell lines that express human p450 s and their propensity to cause liver injury in humans. Toxicol Sci 137:189–211. https://doi.org/10.1093/toxsci/kft223

Guzman-lepe J, Cervantes-alvarez E, Collin A et al (2018) Liver-enriched transcription factor expression relates to chronic hepatic

failure in humans. Hepotol Commun 2:582–594. https://doi.org/10.1002/hep4.1172

Hall RK, Sladekt FM, Granner DK (1995) The orphan receptors COUP-TF and HNF-4 serve as accessory factors required for induction of phosphoenolpyruvate carboxykinase gene transcription by glucocorticoids. Proc Natl Acad Sci USA 92:412–416

Heidari R, Niknahad H, Jamshidzadeh A et al (2015) An overview on the proposed mechanisms of antithyroid drugs-induced liver injury. Adv Pharm Bull 5:1–11. https://doi.org/10.5681/apb.2015.001

Hendriks DFG, Puigvert LF, Messner S, Mortiz W (2016) Hepatic 3D spheroid models for the detection and study of compounds with cholestatic liability. Nat Publ Gr. https://doi.org/10.1038/srep35434

Hsu J, Sage J (2016) Novel functions for the transcription factor E2F4 in development and disease. Cell Cycle 15:3183–3190. https://doi.org/10.1080/15384101.2016.1234551

Jenkins RE, Meng X, Elliott VL et al (2009) Characterisation of flucloxacillin and 5-hydroxymethyl flucloxacillin haptenated HSA in vitro and in vivo. Proteom Clin Approach. https://doi.org/10.1002/prca.200800222

Jennings P, Limonciel A, Felice L, Leonard MO (2013) An overview of transcriptional regulation in response to toxicological insult. Arch Toxicol 87:49–72. https://doi.org/10.1007/s00204-012-0919-y

Jiang J, Wolters JE, van Breda SG et al (2015) Development of novel tools for the in vitro investigation of drug-induced liver injury. Expert Opin Drug Metab Toxicol 5255:1–15. https://doi.org/10.1517/17425255.2015.1065814

Keisu M, Andersson TB (2010) Drug-induced liver injury in humans: the case of ximelagatran. In: Uetrecht J (ed) Adverse drug reactions. Springer, Berlin Heidelberg, pp 407–418

Khetani SR, Kanchagar C, Ukairo O et al (2013) Use of micropatterned cocultures to detect compounds that cause drug-induced liver injury in humans. Toxicol Sci 132:107–117. https://doi.org/10.1093/toxsci/kfs326

Kindmark A, Jawaid A, Harbron CG et al (2008) Genome-wide pharmacogenetic investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. Pharmacogenom J 8:186–195. https://doi.org/10.1038/sj.tpj.6500458

LeCluyse EL, Witek RP, Andersen ME, Powers MJ (2012) Organotypic liver culture models: meeting current challenges in toxicity testing. Crit Rev Toxicol 42:501–548. https://doi.org/10.3109/10408444.2012.682115

Lensmeyer GL, Wiebe DA, Carlson IH, Subramanian R (1991) Concentrations of cyclosporin a and its metabolites in human tissues postmortem. J Anal Toxicol 15:110–115. https://doi.org/10.1093/jat/15.3.110

Limonciel A, Ates G et al (2018) Comparison of base-line and chemical-induced transcriptomic responses in HepaRG and RPTEC/TERT1 cells using TempO-Seq. Arch Toxicol 92(8):2517–2531. https://doi.org/10.1007/s00204-018-2256-2

Mckenzie R, Fried M, Sallie R et al (1995) Investigational nucleoside analogue for chronic hepatitis B. N Engl J Med 333:1099–1105

Metushi IG, Uetrecht J, Phillips E (2016) Mechanism of isoniazid-induced hepatotoxicity: then and now. Br J Clin Pharmacol. https://doi.org/10.1111/bcp.12885

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. Arch Toxicol 80:580–604. https://doi.org/10.1007/s00204-006-0091-3

Ostapowicz G, Fontana RJ, Schiødt FV et al (2002) Results of a prospective study of acute liver failure at 17 tertiary care centers in the United States. Ann Intern Med 137:947–954

Parviz F, Matullo C, Garrison WD et al (2003) Hepatocyte nuclear factor 4 α controls the development of a hepatic epithelium and liver morphogenesis. Nat Genet 34:292–296. https://doi.org/10.1038/ng1175

Poser I, Sarov M, Hutchins JR et al (2008) BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals. Nat Methods 5:409–415. https://doi.org/10.1038/nmeth.1199.nmeth.1199

Proctor WR, Foster AJ, Vogt J et al (2017) Utility of spherical human liver microtissues for prediction of clinical drug induced liver injury. Arch Toxicol 91:2849–2863. https://doi.org/10.1007/s00204-017-2002-1

Quackenbush J (2002) Microarray data normalization and transformation. Nat Genet 32(suppl):496–501. https://doi.org/10.1038/ng1032

Ramaiahgari SC, den Braver MW, Herpers B et al (2014) A 3D in vitro model of differentiated HepG2 cell spheroids with improved liver-like properties for repeated dose high-throughput toxicity studies. Arch Toxicol 88:1083–1095. https://doi.org/10.1007/s00204-014-1215-9

Regenthal R, Krueger M, Koeppel C, Preiss R (1999) Drug levels: therapeutic and toxic serum/plasma concentrations of common drugs. J Clin Monit Comput 15:529–544. https://doi.org/10.1023/A:1009935116877

Ritz C, Baty F, Streibig JC, Gerhard D (2015) Dose-response analysis using R. PLoS One. https://doi.org/10.1371/journal.pone.0146021

Soldatow VVY, Lecluyse EEL, Griffith LLG, Rusyn I (2013) In vitro models for liver toxicity testing. Toxicol Res (Camb) 2:23–39. https://doi.org/10.1039/C2TX20051A.In

Takebe T, Sekine K, Enomura M et al (2013) Vascularized and functional human liver from an iPSC-derived organ bud transplant. Nature 499:481–484. https://doi.org/10.1038/nature12271

Tanaka S, Uchida S, Inui N et al (2014) Simultaneous LC–MS/MS analysis of the plasma concentrations of a cocktail of 5 cytochrome P450 substrate drugs and their metabolites. Biol Pharm Bull 37:18–25. https://doi.org/10.1248/bpb.b13-00401

Vinken M, Landesmann B, Goumenou M et al (2013) Development of an adverse outcome pathway from drug-mediated bile salt export pump inhibition to cholestatic liver injury. Toxicol Sci 136:97–106. https://doi.org/10.1093/toxsci/kft177

Wang H, Yang Y, Sharma N et al (2010) STAT1 activation regulates proliferation and differentiation of renal progenitors. Cell Signal 22:1717–1726. https://doi.org/10.1016/j.cellsig.2010.06.012

Wilke R, Lin D, Roden D (2007) Identifying genetic risk factors for serious adverse drug reactions: current progress and challenges. Nat Rev Drug Discov 6:904–916. https://doi.org/10.1038/nrd2423

Wilmes A, Limonciel A, Aschauer L et al (2013) Application of integrated transcriptomic, proteomic and metabolomic profiling for the delineation of mechanisms of drug induced cell stress. J Proteomics 79:180–194. https://doi.org/10.1016/j.jprot.2012.11.022

Wink S, Hiemstra S, Herpers B, van de Water B (2016) High-content imaging-based BAC-GFP toxicity pathway reporters to assess chemical adversity liabilities. Arch Toxicol. https://doi.org/10.1007/s00204-016-1781-0

Wink S, Hiemstra SW, Huppelschoten S et al (2018) Dynamic imaging of adaptive stress response pathway activation for prediction of drug induced liver injury. Arch Toxicol. https://doi.org/10.1007/s00204-018-2178-z

Xu JJ, Henstock PV, Dunn MC et al (2008) Cellular imaging predictions of clinical drug-induced liver injury. Toxicol Sci 105:97–105. https://doi.org/10.1093/toxsci/kfn109

Yuan H, Wen B, Liu X et al (2015) CCAAT/enhancer-binding protein α is required for hepatic outgrowth via the p53 pathway in zebrafish. Sci Rep 5:1–8. https://doi.org/10.1038/srep15838