Chapter 6

Automated live cell imaging of adaptive stress responses for assessment of drug-induced liver injury (DILI) liabilities.

This chapter is based on a manuscript in preparation:

Steven Wink*‡, Steven W. Hiemstra*‡, Suzanna Huppelschoten‡*, Janna E. Klip*, Bob van de Water*§

‡Authors contributed equally

*Division of Toxicology, Leiden Academic Centre for Drug Research, Leiden University, Leiden, The Netherlands

1. Abstract

Drug-induced liver injury remains a major concern in the clinic and during drug development. There is an urgent need for improved prediction of DILI liabilities. Many DILI compounds activate cellular adaptive stress response pathways. We have evaluated the application of three BAC-GFP HepG2 reporter cell lines representing oxidative stress (Srxn1-GFP), endoplasmic reticulum stress (CHOP-GFP) and p53-related signalling (p21-GFP) for DILI assessment. More than 170 DILI compounds and reference control compounds (0, 1, 5, 10, 50 and 100 Cmax concentrations) were screened for reporter activation using automated high throughput high content live cell confocal imaging. Quantitative data analysis at the single level revealed activation of the Srxn1-GFP > CHOP-GFP > p21-GFP, with some compounds preferably activating individual reporters. Hierarchical clustering of time course dynamics of all individual reporter responses for all compounds and concentrations, allowed the refinement of primary mode-of-action. Combined integration of concentration responses for both reporters and cell death features resulted in a clustering of 53% of compounds with a most-DILI concern label. The most-DILI concern compounds activated both Srxn1-GFP and the CHOP-GFP. A strong association between HepG2 BAC-GFP reporter activation and primary human hepatocyte mRNA transcript induction was observed. We anticipate that the integration of imaging-based high throughput assays for adaptive stress pathway activation will contribute to DILI assessment and likely chemical safety assessment in general.

2. Introduction

Drug-induced liver injury (DILI) is an important problem during drug development, in the clinic and during post-marketing [309]. Various chemical, genetic and life style factors contribute to DILI making it a multi-facetted challenge to predict DILI. This includes compound-specific mode-of-action causing oxidative stress through mitochondrial perturbation, cholestasis due to bile acid transporter inhibition, or steatohepatitis due to altered fat metabolism. Different polymorphisms in cytochrome P450 enzymes or HLA molecules affect drug metabolism and immune responses [310]. While hepatitis or diabetes may affect the susceptibility to DILI onset [311]. The multitude of mechanism contributing to DILI forces a more mechanistic approach in the pre-clinical prediction of DILI liability.

Gene expression analysis has contributed significantly to improve our understanding of DILI [91, 240, 241, 312]. This has led to the identification of various signalling pathways that are activated during DILI and possibly predictive for chemical-induced liver injury. Some of these mechanistic insights have been integrated in high throughput assays, including phospholipidosis [313], cholestasis [314], cytokine induced synergistic apoptosis [38], mitochondrial damage [315], and oxidative stress [84]. However, so far a high throughput approach that more directly integrates the transcriptome-based mode-of-action is missing.

To integrate the mode-of-action in high throughput approaches we propose a more practical approach which focuses on a general biological theme evolved in evolution: the cellular adaptive stress response pathways [45]. Several key adaptive stress response pathways that are essential in the maintenance of cellular homeostasis are the oxidative stress response, the endoplasmic reticulum stress response (ER-stress) / unfolded protein response (UPR) and the DNA damage response (DDR). The TG-GATES transcriptomics dataset indicates that these pathways are often

affected during liver toxicity (Wink *et al.,* manuscript submitted), while mechanistic studies indicate the contribution of these pathways in the pathophysiology of chemical-induced liver failure [238, 316, 317]. Oxidative stress is caused by the oxidation/reduction reactions or alkylation from reactive intermediates or indirect oxidative stress induction by e.g. disruption of endogenous mitochondrial function [318]. Overall, this leads to the modification of KEAP1 followed by the stabilization of the transcription factor Nrf2 which activates the oxidative stress defense gene network including the expression of NQO1 and Srxn1 [217]. Protein alkylation, disruption of protein trafficking or calcium homeostasis in the ER disrupt normal protein folding in the ER and leads to activation of the unfolded protein response [319]. This involves the activation of the kinases IRE1α and PERK and the proteolytic activation of the transcription factor ATF6; the transcription factors Xbp1 and ATF4 are activated downstream of both kinases [320]. The UPR transcriptional response involves the expression of the chaperone BiP that rescues unfolded proteins in the ER lumen as well as the expression of CHOP/DDIT3, a transcription factor that promotes the expression of pro-apoptotic genes [321]. DNA damage is caused by genotoxic compounds, which are usually electrophiles that directly interact with DNA and form covalent bonds. This involves the activation of p53 followed by activation of downstream target genes including p21 and Btg2. We have previously integrated all the downstream target genes for these three main adaptive stress response pathways in bacterial artificial chromosome (BAC) GFP-based HepG2 reporter cell lines. We showed that these BAC reporters are sensitive and selective for mode-of-action evaluation and can be applied in high throughput – high content live cell imaging to capture the dynamic adaptive stress response activation at the single cell level (Wink *et. al,* manuscript in preparation).

Here we assessed the application of three BAC GFP HepG2 reporter cell lines that represent three major adaptive stress response pathways to predict DILI liability: i.e. Srxn1 (KEAP1/Nrf2), CHOP/DDIT3 (UPR) and p21 (DNA damage response). These three BAC reporters were exposed to more than 170 chemicals at different concentrations covering human exposure levels, and the dynamic stress pathway activation as well as onset of cytotoxicity was followed for 24 hr using automated live cell confocal imaging. All quantitative dynamic GFP reporter data as well as cytotoxicity measurements were integrated to assess DILI liability.

3. Materials and methods

3.1. Cell culture

Human hepatoma HepG2 cells were acquired from ATCC (clone HB8065). HepG2 Srxn1, DDIT3 (CHOP) and CDKN1A (p21) BAC GFP reporter were generated according to [48] and have been carefully characterized previously (Wink *et al*, manuscript submitted). HepG2 BAC GFP reporters were maintained and exposed to drugs in DMEM high glucose supplemented with 10% (v/v) FBS, 25U/mL penicillin and 25µg/mL streptomycin. The cell lines were used between passage 5 and 20. For live cell imaging, the cells were seeded in Greiner black µ-clear 384 wells plates, at 8,000 cells per well.

3.2. Reagents

All reference compound chemicals were acquired from Sigma-Aldrich and freshly dissolved in DMSO; except for metformin, fluphenazine, buthionine sulfoxamine, bromoethlyamine (all PBS), acetaminophen and phenobarbital (all DMEM). DILI compounds were a kind gift from the Dr. Weida Tong, NCTR-FDA [264]. All compounds were maintained as 500-fold stock such that final treatments did not exceed 0.2 % v/v DMSO.

3.3. Microscopy

Accumulation of BAC GFP-fusion levels, priopodium iodide (PI) and Annexin-V-Alexa633 (AV) staining was monitored using a Nikon TiE2000 confocal laser microscope (lasers: 647nm, 540nm, 488nm and 408nm), equipped with an automated stage and perfect focus system and at 37 degrees C and humidified atmosphere and 5% CO_2/air mixture. Prior to imaging at 20x magnification and either 1X, or 2X zoom, HepG2 cells were loaded for 45 minutes with 100 ng/mL $Hoechst_{33342}$ to visualize the nuclei, upon which the Hoechst-containing medium was washed away to avoid Hoechst phototoxicity [224] and replaced with medium containing PI and AV to monitor cell death. Each imaged 384-well plate containing one reporter cell line all the compounds used in the screen at one certain concentration (1, 5, 10, 50 or 100 Cmax); for each concentration 2-3 replicates were imaged per reporter cell line.

3.4. Reporter response quantification

Quantitative image analysis was performed with CellProfiler version 2.1.1 [49] with an in house developed module implementing the watershed masked algorithm for segmentation [193]. The watershed separates an image in regions with single cells followed by pixel classification for each region as fore- or background and this method performs well detecting the $Hoechst_{33342}$ stained nuclei of the closely packed HepG2 cells. The binary mask containing the segmented nuclei was fed to the *identify-primary-objects module*, *overlap-based-tracking* module and *intensity-nuclei-size-shape-measurement* modules of CellProfler. For the cytosol location of the Srxn1-GFP reporter the nuclear objects were used as seeds for the *identify-secondary-objects* module set to a propagation method with the MCT algorithm on adaptive (window size approximately 20 pixels) segmentation. Segmentation results were stored as png files for quality control purposes and CellProfiler pipelines were stored for reproducibility. Image analysis results were stored on the local machine as HDF5 files. Data analysis, quality control and graphics was performed using the in house developed R package H5CellProfiler (Wink *et al*., 2015, manuscript in preparation).

For each reporter the mean intensity and integrated intensity levels of the GFP signal were measured on the single cell level. The mean intensity is less sensitive to size variations of the objects and the integrated intensity is less sensitive for faulty background segmentations within the object. In addition, the nuclear $Hoechst_{33342}$ intensity was measured to observe variations in DNA content (e.g. due to apoptosis or necrosis), cell migration speed (which can be affected by compound specific effects), nuclear size (which can vary e.g. due to mitosis inhibition, flattening of cells or changes in the internal osmotic pressure), PI and AV staining per single cell (for cell death detection) and finally the overall cell count. All features were measured with an average one hour intervals; with the exception of the PI and AV which was only measured after completion of the 24 hour live imaging session.

3.5. Data analysis

The features of interest were extracted from the HDF5 files and further analyzed using the graphical user interface of the H5CellProfiler package. The mean of the features for each compound, concentration, cell line and replicate combination was calculated. In addition for each plate the mean and standard deviation of the DMSO treated single cell population was calculated to determine background control values: the 2X mean, 3X mean and the mean + 3 standard error values for each plate. Thereafter for each treatment the fraction of cells above these control-values was determined. To account for PI and AV background staining and noise the segmented PI and AV segmentations were masked by a 2 pixel dilated nuclei. The area of these nuclei and the PI and AV objects were divided to obtain the cell death stain to cell area ratio. These ratios were filtered to be at least 10% of the cell size and following this procedure each cells was either flagged as alive or dead in the final time point of the 24 live imaging session. In this manner the fraction of dead cells could be accurately determined. All resultant summarized data was stored as tab delimited text files and further processed for normalization and graphical presentation using R.

Due to automated confocal imaging over a one year period, the time course data required intensity variation plate normalization as well as modeling of the time course-data. The mean and integrated intensity features and the nuclei size and $Hoechst_{33342}$ intensity features were first transformed to fold change with respect to the plate- specific DMSO controls at time point 1 and the cell count and cell speed features were transformed to fold change with respect to the plate-specific DMSO controls average over time (diagram 1). Afterwards these value were scaled between 0 and 1 over the entire dataset with the formula $(x\text{-}x_{min\_screen})/(x_{max\_screen} - x_{min\_screen})$ for the purpose of proper heatmap display. Prior validation of negative and positive control responses preceded this scaling procedure. After the normalization steps the response specific features (integrated and mean intensity and the positive fraction features) were fit separately per replicate with the b-splines method with a degree of freedom of 10 and 3^{rd} degree polynomials using the base-r lm function and bs function of the splines package. This allowed resampling the data with equidistant time points for replicate statistics and higher density time point sampling (200 points) for smooth heatmap display. All b-spline fits were stored for verification purposes.

3.6. Dose response data transformation

The maximum value of the fold change and normalized, scaled and b-spline modeled mean and integrated intensity features and the b-spline modeled positive GFP fraction features per

compound and concentration combination were selected for the dose response curves. The time course cell numbers, cell speed, $Hoechst_{33342}$ intensity and nuclear area features were regressed to the line a*x* + b to determine the slope and mean values.

The fraction of dead cells did not require normalization as these features are inherently a plate normalized measure nor did they require transformation as only the final time point was measured. The cytotox features, cell count, $Hoechst_{33342}$ intensity and nuclear area also obtained from the final imaging round, were normalized per plate as fold change with respect to DMSO controls and scaled between 0 and 1 for the entire dataset.

3.7. Data representation

All HCI data representations were generated or modified with Illustrator CS6, Fiji, ggplot2 [226], the aheatmap function of the NMF package [227]. For some of the data clustering the equidistant sample time profile features from the b-spline model were used to calculate a distance matrix for each feature separately using Euclidean distance. A mean distances matrix was calculated and subjected to clustering with the ward.D method of the hclust function. The columns representing the different features were not subjected to clustering. The same method was applied to the row-clustering, only now the dose response vectors were used for the distance matrixes.

3.8. Gene expression analysis

CEL files were downloaded from the Open TG-GATEs database: “Toxicogenomics Project and Toxicogenomics Informatics Project under CC Attribution-Share Alike 2.1 Japan” http://dbarchive.biosciencedbc.jp/en/open-tggates/desc.html. Probe annotation was performed using the hthgu133pluspmhsentrezg.db package version 17.1.0 and Probe mapping was performed with hthgu133pluspmhsentrezgcdf downloaded from NuGO (http://nmg-r.bioinformatics.nl/NuGO_R.html). Probe-wise background correction (Robust Multi-Array Average expression measure), between-array normalization within each treatment group (quantile normalization) and probe set summaries (median polish algorithm) were calculated with the rma function of the Affy package (Affy package, version 1.38.1) (Irizarry, R.A. et al.). The normalized data were statistically analyzed for differential gene expression using a linear model with coefficients for each experimental group within a treatment group[258, 259]. A contrast analysis was applied to compare each exposure with the corresponding vehicle control. For hypothesis testing the empirical bayes statistics for differential expression was used followed by an implementation of the multiple testing correction of [254]using the LIMMA package [259].

4. Results

4.1. High content adaptive stress response screen with DILI compounds.

To assess the application of adaptive stress pathway activation measurements for the assessment of DILI liability a large high content live cell screen was performed on three BAC reporter cell lines (Srxn1, CHOP and p21) that were described previously (Wink *et. al*, manuscript submitted), with 176 compounds covering mostly DILI related compounds and various reference control compounds. DILI compounds were classified for most-DILI-concern (58), less-DILI-concern (51) or no-DILI-concern (37) (see Table 1). The control reference compounds included negative controls (i.e. DMSO and medium) and positive controls (i.e. alkylating agents, mitochondrial toxicants, inducers of the UPR and DNA damaging agents). All reporters were exposed to five concentrations (1, 5, 10 , 50 and 100 Cmax) followed by live cell imaging and automated multi-parametric image analysis (Fig. 1 and 2).

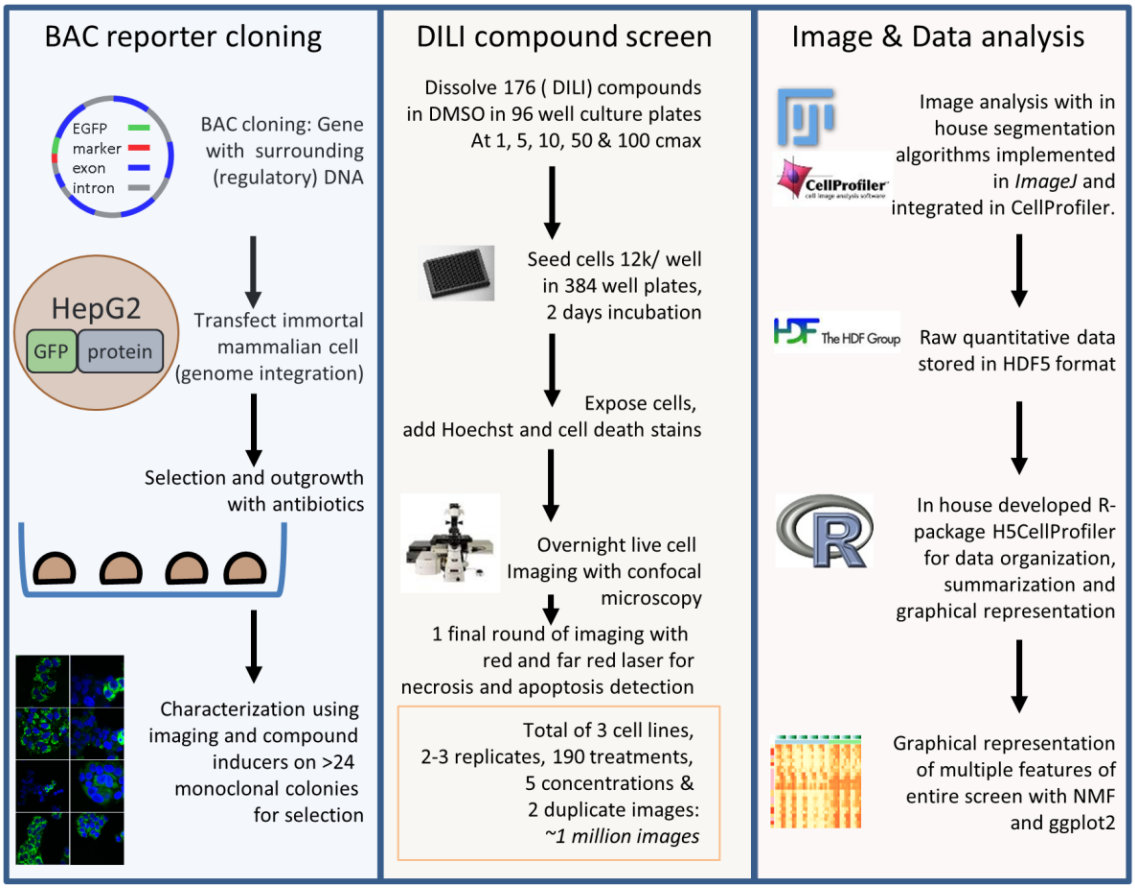


Figure 1: **BAC cloning, BAC reporter DILI screen and analysis pipeline.** Left panel) BAC cloning technology is used to maintain endogenously regulated reporter protein levels and regulation. Monoclonal reporter selection from a high number of clones to ensure endogenous response to positive control stimuli and suitability of reporter for imaging. Middle panel) High content live cell screen of 176 compound at 1, 5, 10, 50 and 100 cmax at 2 or 3 replicates. Right panel) Image and data analysis is performed with CellProfiler/Fiji and R, respectively. Some in-house tools were developed in CellProfiler and R to assist in the quality and analysis of the large data output.

4.2. Single cell analysis allows fine tuning of sensitivity versus dynamic range.

For all images single cell analysis was performed to extract a diverse set of quantitative data, including GFP reporter activity, cell migration speed, cell number and cytotoxicity (see Figure 2 for overview of the dataset). The GFP reporter single cell data was used to derive quantitative fitted

data for five different determinants of reporter activity: mean intensity, mean integrated intensity, fraction of cells with GFP intensity 2-fold (GFP_pos.2m) or 3-fold (GFP_pos.3m) above average of control, population mean plus three times the standard deviation (GFP_pos.m3d). Systematic evaluation of these descriptors for the lowest and strongest responding compound for each individual reporter, allowed fine tuning of the sensitivity versus the dynamic range (Figure 3). Based on the mean GFP intensity over the single cell population hydroxurea would not have been defined as a positive in the Srxn1-GFP reporter cell line, because only in a small proportion of cells that contain a higher level of Srxn1-GFP the signal was detected; the GFP_pos.2m was a more sensitive descriptor that also allowed evaluation of the time course dynamics. Similar observations were made for nitrofurantoin and clozapine for the GFP-CHOP and GFP-p21 reporters, respectively. However for strong inducers of oxidative stress (diethylmaleate; DEM), UPR (thapsigargine) and DNA damage (etoposide), GFP mean intensity already allowed the sensitive detection of the reporter responses, while GFP_pos.2m caused an early saturation, thereby lowering the dynamic range.

4.3. Temporal order of adaptive stress response activation as indication of primary mode-of-action.

For evaluation of the reporter activation for the entire compound screen GFP_pos.2m was selected as the most sensitive initial readout. The b-spline fits of the GFP_pos.2m time courses were used to calculate the mean of the replicates for Srxn1-GFP, CHOP-GFP and p21-GFP reporter responses for all compounds (Figure 4). Some compounds showed a response in all three reporters, while others only showed a response in one cell line only. Often a specific order in the reporters can be seen in time. For brefeldin A (BFA), CCCP, FCCP, doxycycline (DX), oligomycin A and B (OMA, OMB), thapsigargin (THG) , tunicamycin (TUN) and zimelide (ZMI) CHOP-GFP activation was followed by the Srxn1-GFP activation. Yet in other cases Srxn1-GFP did precede CHOP-GFP activation (azathioprine (AZA), benzbromarone (BB), bromfenac (BFC), ethacrynic acid (ETA) and sodium-arsenite (SA)). Simultaneous activation of two reporters was also observed. Interestingly, for some compounds the stress response activation was transient (iodoacetamide; IAA) implicating reversal of the stress response, possibly due to toxicokinetics and/or true cellular adaptation. Few compounds showed direct fluorescence increase starting from the first time point (e.g. tetracyclin (TET), doxorubicin (DOX) and dantrolene (DAN)), suspecting compound autofluorescence. Several compounds showed activation at a lower C-max while a response was absent at higher concentrations indicating that an adaptive stress response preceded cell death. This was verified with the additional markers captured during the live cell imaging: cell number, cell migration speed, nuclear area and DNA content. Although most compounds did not affect these markers, several compounds did, typically in a concentration-dependent manner. Cell migration speed was the most often affected parameter as exemplified by the actin cytoskeleton disrupting agent cytochalasin (Supplemental Figure S1 and S2).

Table 1: **Test compound set.** Alphabetically sorted list of test compounds screened in this study including their c-max values, abbreviations, DILI-concern label and metabolic potential.

compound	cmax (µM)	abbreviation	DILI classification	metabolism
acarbose	0.15	ACA	Most-DILI-Concern	NM
acetaminophen	100	APAP	Most-DILI-Concern	YES
adefovir	0.09	ADV	No-DILI-Concern	NM
aflatoxin B1	1.00	AFB1	control	NA
allopurinol	13.8	ALLO	Most-DILI-Concern	NM
allyl alcohol	2.00	AA	control	NA
altretamine	3.76	ALM	Less-DILI-Concern	YES
amiodarone	0.81	AMIO	Most-DILI-Concern	NM
amoxicillin	22.3	AX	Less-DILI-Concern	NM
antimycin A	0.20	AA	control	NA
aspirin	1650	ASA	No-DILI-Concern	NM
azathioprine	0.34	AZA	Most-DILI-Concern	YES
azidothymidine	3.97	AZT	No-DILI-Concern	NA
benzbromarone	4.34	BB	Most-DILI-Concern	YES CYP2C9
betaine	940	BET	No-DILI-Concern	NA
brefeldin A	1.00	BFA	control	NA
bicalutamide	1.97	BAT	Most-DILI-Concern	YES CYP3A4
bisphenol A	0.80	BPA	No-DILI-Concern	NA
bosentan	3.43	BOS	Most-DILI-Concern	YES
bromfenac	18.0	BFC	Most-DILI-Concern	NA
buspirone	0.02	BUS	Less-DILI-Concern	YES CYP3A4
busulfan	0.28	BU	Most-DILI-Concern	YES
buthionine sulfoxamine	0.25	BSO	control	NA
butylated hydroxyanisole	1.00	BHA	control	NA
captopril	8.88	CPL	Less-DILI-Concern	YES
carbamazepine	50.8	CBZ	Most-DILI-Concern	NM
carmustine	60.0	BCNU	Most-DILI-Concern	YES
CCCP	1.00	CCCP	control	NA
CCl4	6.50	CCl4	control	NA
CDDO-me	0.01	CDDO	control	NA
chloramphenicol	46.4	CAMP	No-DILI-concern	NA
chlormezanone	10.6	CMZ	Most-DILI-Concern	NA
chloroquine diphosphate	0.00	CHQ	No-DILI-concern	MINIMAL
chlorpromazine	0.94	CPZ	Less-DILI-Concern	YES
chlorpropamide	130	CHL	Less-DILI-Concern	NM
cimetidine	9.00	CMT	Less-DILI-Concern	YES
ciprofloxacin	6.58	CIPX	Most-DILI-Concern	NM
cisplatin	2.00	CDDP	Less-DILI-Concern	NA
clofibrate	25.0	CLO	Less-DILI-Concern	YES
clotrimazole	0.09	CTZ	No-DILI-Concern	NM
clozapine	0.98	CLZ	Most-DILI-Concern	YES
colchicine	0.02	CLC	Less-DILI-Concern	YES
coumarin	0.80	COH	control	NA
cromolyn	0.02	CRO	No-DILI-Concern	NA
cycloheximide	50.0	CHX	Most-DILI-Concern?	NA
cyclophosphamide	375	CTX	Less-DILI-Concern	YES
cyclosporin A	0.20	CSA	Less-DILI-Concern	YES
cytochalasin D	1.00	CCD	control	NA
dacarbazine	20.6	DTIC	Most-DILI-Concern	NM
danazol	0.11	DNZ	Most-DILI-Concern	NM
dantrolene	7.90	DAN	Most-DILI-Concern	YES
deferoxamine	17.8	DFX	No-DILI-Concern	NM
diethyl maleate	10.0	DEM	control	NA
dexamethasone	0.22	DXS	Less-DILI-Concern	NM
dextromethorphan HBr	0.01	DXM	No-DILI-Concern	NA
diclofenac	4.20	DCF	Most-DILI-Concern	YES
didanosine	9.83	DDL	Most-DILI-Concern	NM
digoxin	0.00	DIG	No-DILI-Concern	NA
diltiazem	96.0	DTZ	Most-DILI-Concern	YES
disulfiram	5.40	DIS	Most-DILI-Concern	YES
DMNQ	0.04	DMNQ	control	NA
DMOG	7.50	DMOG	control	NA
doxorubicin	1.18	DOX	Less-DILI-Concern	YES
doxycycline	11.3	DX	Less-DILI-Concern	NM
edrophonium	60.2	EDR	No-DILI-Concern	NA
enalapril	0.40	ENP	Less-DILI-Concern	YES
entacapone	3.93	ECP	No-DILI-Concern	YES
epinephrine	0.00	EPI	No-DILI-Concern	NA
erythromycin	11.0	ERYC	Most-DILI-Concern	NM
ethacrynic acid	33.0	ETA	No-DILI-Concern	YES
ethambutol	24.5	EMB	Less-DILI-Concern	NM
ethionine	10.0	ETH	Less-DILI-Concern	NA
etodolac	68.5	ELAC	Most-DILI-Concern	YES
etoposide	4.04	ETO	Less-DILI-Concern	YES
famotidine	0.31	FAM	Less-DILI-Concern	YES
FCCP	1.00	FCCP	Less-DILI-Concern	NA
fenofibrate	4.10	FF	Less-DILI-Concern	NM
fenoprofen	58.2	FPF	Most-DILI-Concern	NA
fialuridine	1.00	FIAU	Most-DILI-Concern	NA
fluoxetine	0.05	FLX	Less-DILI-Concern	YES
flurbiprofen	57.3	FBP	Most-DILI-Concern	NA
folic acid	0.04	FAM	No-DILI-Concern	NA
furosemide	3.29	FUR	Less-DILI-Concern	YES
galactosamide	100	GALN	control	NA
ganciclovir	4.62	GCV	No-DILI-Concern	MINIMAL
glimepiride	1.12	GLP	Less-DILI-Concern	NM
griseofulvin	4.54	GF	Most-DILI-Concern	YES
haloperidol	0.01	HDL	Less-DILI-Concern	YES
hydrazine	14.7	HYD	control	NA
hydroxyurea	348	HU	Most-DILI-Concern	YES
hydroxyzine	0.27	HYZ	No-DILI-Concern	NM
ibuprofen	70.1	IBU	Less-DILI-Concern	YES
imipramine	0.29	IM	Less-DILI-Concern	YES
indomethacin	5.59	IMN	Most-DILI-Concern	YES
interleukin 6	0.40	IL6	control	NA
iodoacetamide	1.00	IAA	control	NA
isoniazid	45.4	INH	Most-DILI-Concern	YES
isoproterenol	0.01	IPR	No-DILI-Concern	NA
kanamycin	60.1	KM	No-DILI-Concern	NA
ketoconazole	6.59	KTZ	Most-DILI-Concern	NM
ketorolac	3.53	KTL	Less-DILI-Concern	NM
labetalol	2.68	LABE	Most-DILI-Concern	YES
maprotiline	0.18	MPT	No-DILI-Concern	NA
mebendazole	0.13	MBZ	Less-DILI-Concern	NM
meclizine	0.03	MCZ	No-DILI-Concern	NM
menadione	2.50	MEN	control	NA
mercaptopurine	0.48	6MP	Most-DILI-Concern	YES
metformin	116	MF	No-DILI-Concern	YES
methimazole	2.62	MTZ	Most-DILI-Concern	YES
methotrexate	0.77	MXT	Most-DILI-Concern	NM
methyldopa	18.9	MD	Most-DILI-Concern	YES
metoprolol	0.56	MTPL	Less-DILI-Concern	YES
metronidazole	237	MTDZ	No-DILI-Concern	NM
mexiletine	3.83	MXT	Most-DILI-Concern	YES
mitomycin C	7.18	MIT-C	No-DILI-Concern	NM
moxisylyte	0.16	MOX	Most-DILI-Concern	NA
naproxen	0.20	NPX	Less-DILI-Concern	YES
nefazodone	0.86	NFZ	Most-DILI-Concern	YES
neomycin	0.44	NEO	No-DILI-Concern	NM
nifedipine	0.43	NFP	Less-DILI-Concern	YES
nimesulide	21.1	NMS	Most-DILI-Concern	NA
nitrofurantoin	6.00	NTF	Most-DILI-Concern	YES
nizatidine	4.00	NIZ	Less-DILI-Concern	YES
nocodazole	1.00	NOC	control	NA
ochratoxin A	0.05	OTA	No-DILI-Concern	NA
ofloxacin	9.96	OFX	Most-DILI-Concern	NM
oligomycin A	0.50	OMA	control	NA
oligomycin B	1.00	OMB	control	NA
omeprazole	4.70	OMZ	Less-DILI-Concern	YES
oxytetracycline	3.26	OTC	No-DILI-Concern?	NA
paroxetine	0.06	PXT	Less-DILI-Concern	YES
penicillamine	27.5	DP	Less-DILI-Concern	NM
perhexiline	2.16	PHX	Most-DILI-Concern	NA
phenacetin	0.62	POD	Less-DILI-Concern	NA
phenobarbital	1.46	PBT	Less-DILI-Concern	NM
phenytoin	21.7	PT	Most-DILI-Concern	YES
pioglitazone	2.95	PGZ	Less-DILI-Concern	NM
prednisolone	0.68	PRD	Less-DILI-Concern	NM
primaquine	0.61	PQ	No-DILI-Concern	NM
primidone	4.67	PRI	No-DILI-Concern	YES
procyclidine	0.40	PCD	No-DILI-Concern	NA
propranolol	0.20	PPL	Less-DILI-Concern	YES
propylthiouracil	9.10	PTU	Most-DILI-Concern	YES
ranitidine	1.79	RNT	Less-DILI-Concern	YES
ribavirin	2.61	RBV	No-DILI-Concern	MINIMAL
rifampicin	15.0	RFP	Less-DILI-Concern	YES
rotenone	0.40	ROT	control	NA
simvastatin	0.08	SVN	Less-DILI-Concern	YES
sodium-arsenite	10.0	SA	control	NA
staurosporin	0.01	STAU	control	NA
succinylcholine	138	SUCCS	No-DILI-Concern	NA
sulfamethoxazole	217	SMZ	No-DILI-Concern?	YES
sulindac	32.0	SUL	Most-DILI-Concern	YES
tacrine	0.08	TC	Most-DILI-Concern	YES
tacrolimus	0.04	TAC	No-DILI-Concern	YES
tamoxifen	0.16	TMX	Most-DILI-Concern	YES
terbinafine	4.00	TRB	Most-DILI-Concern	NM
tetracycline	21.0	TET	Less-DILI-Concern	NM
thapsigargin	0.10	THG	control	NA
thioridazine	0.55	TDZ	Less-DILI-Concern	YES
ticlopidine	8.07	TPD	Most-DILI-Concern	YES
TNF-a	0.50	TNF	control	NA
TO901317	1.25	TO9	control	NA
tolbutamide	233	TOLB	Less-DILI-Concern	NM
tolcapone	22.0	TC	Most-DILI-Concern	YES
trazodone	5.06	TZ	Most-DILI-Concern	YES
troglitazone	6.39	TRG	Most-DILI-Concern	YES
tunicamycin	1.00	TUN	control	NA
valacyclovir	17.4	VACV	No-DILI-Concern	NM
valproic acid	243	VA	Most-DILI-Concern	YES
verapamil	0.16	VRP	Less-DILI-Concern	NM
warfarin	4.86	WAR	Less-DILI-Concern	NM
WY14643	1.25	WY14	control	NA
ximelagatran	16.9	XML	Most-DILI-Concern	NA
zafirlukast	1.21	ZFL	Most-DILI-Concern	YES
zimelidine	60.0	ZMI	Less-DILI-Concern	NA

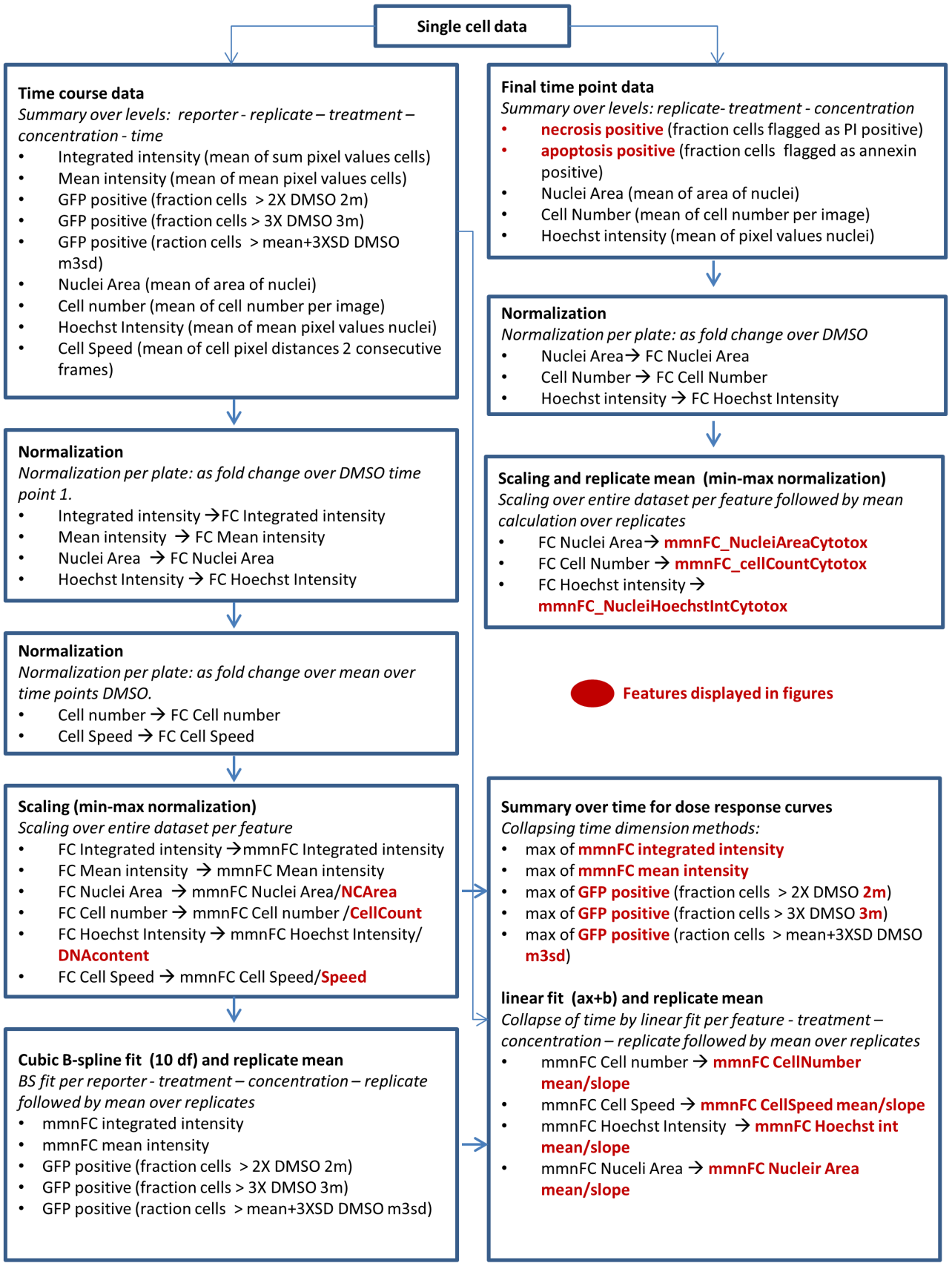


Figure 2: **Data analysis workflow.** The features in red are displayed in the figures of the results section.

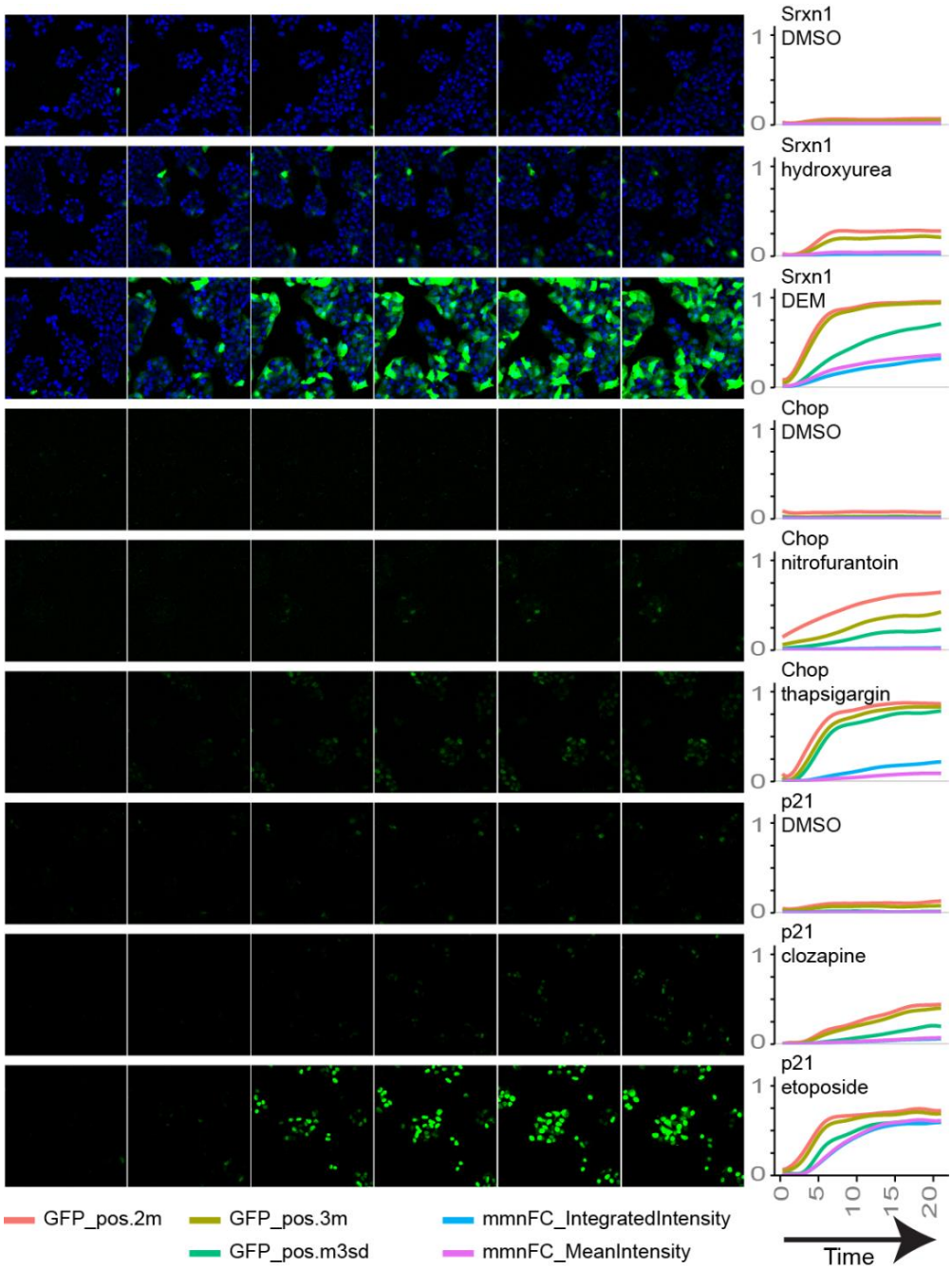


Figure 3: **Dynamics of GFP-Srxn1, GFP-Chop and GFP-p21 reporter activation.** Left panel) Time lapse stills of the reporters exemplifying the importance of single cell analysis which allows fine tuning sensitivity versus dynamic range of BAC-reporter readout. Right panel) quantification of GFP signal of the 3 reporters from a control, a weak reporter-activating compound and a strong reporter-activating compound. The single cell population means of integrated or mean intensity values show no effect for the weak-activating compounds. By counting and determining the fraction of GFP positive cells above a plate-specific threshold; GFP_pos.2m (twice the DMSO population mean), GFP_pos.3m (thrice the DMSO population mean) & GFP_pos.m3sd (population mean plus three times the standard deviation) the sensitivity is increased, at the cost of dynamic range.

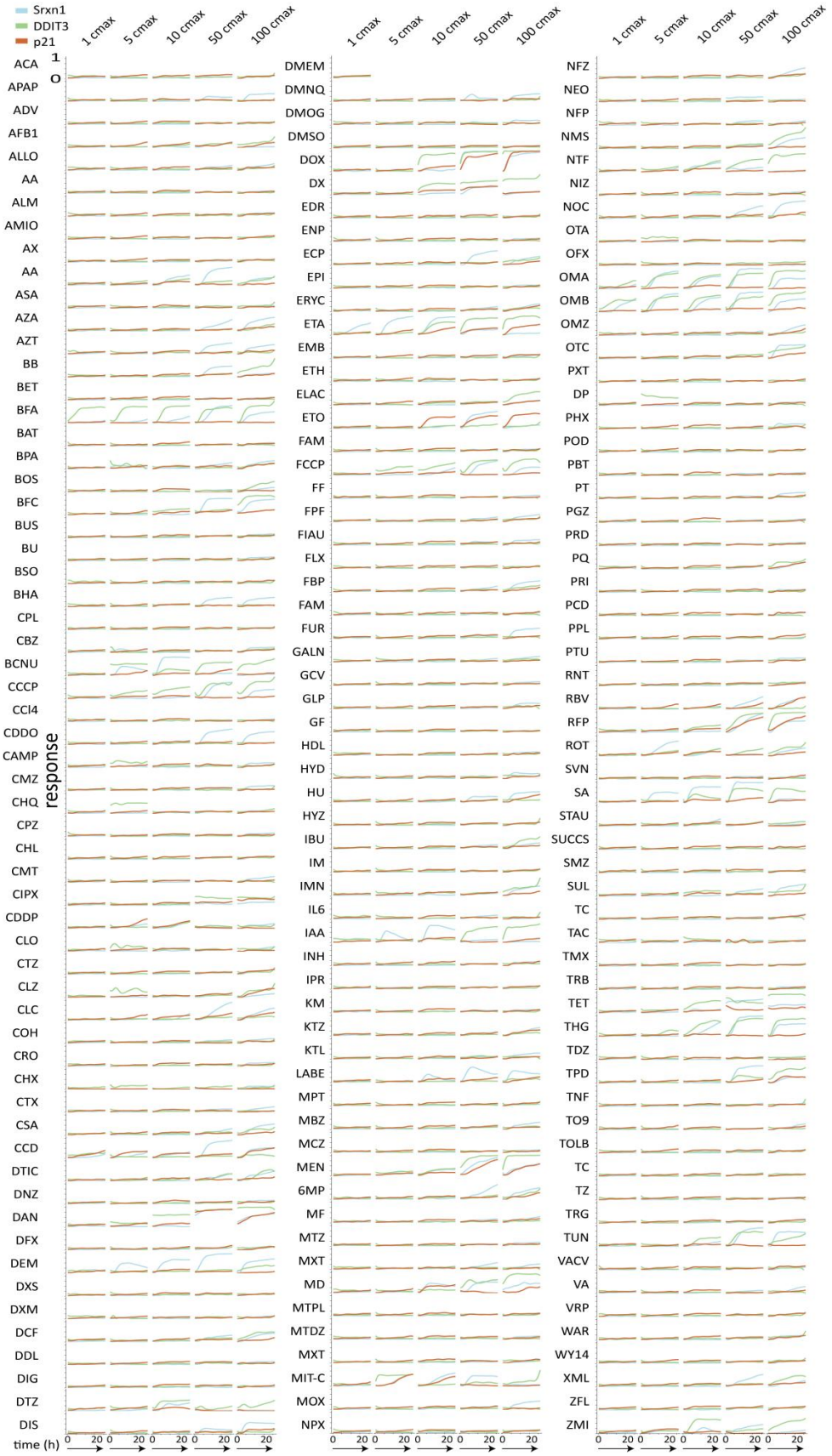
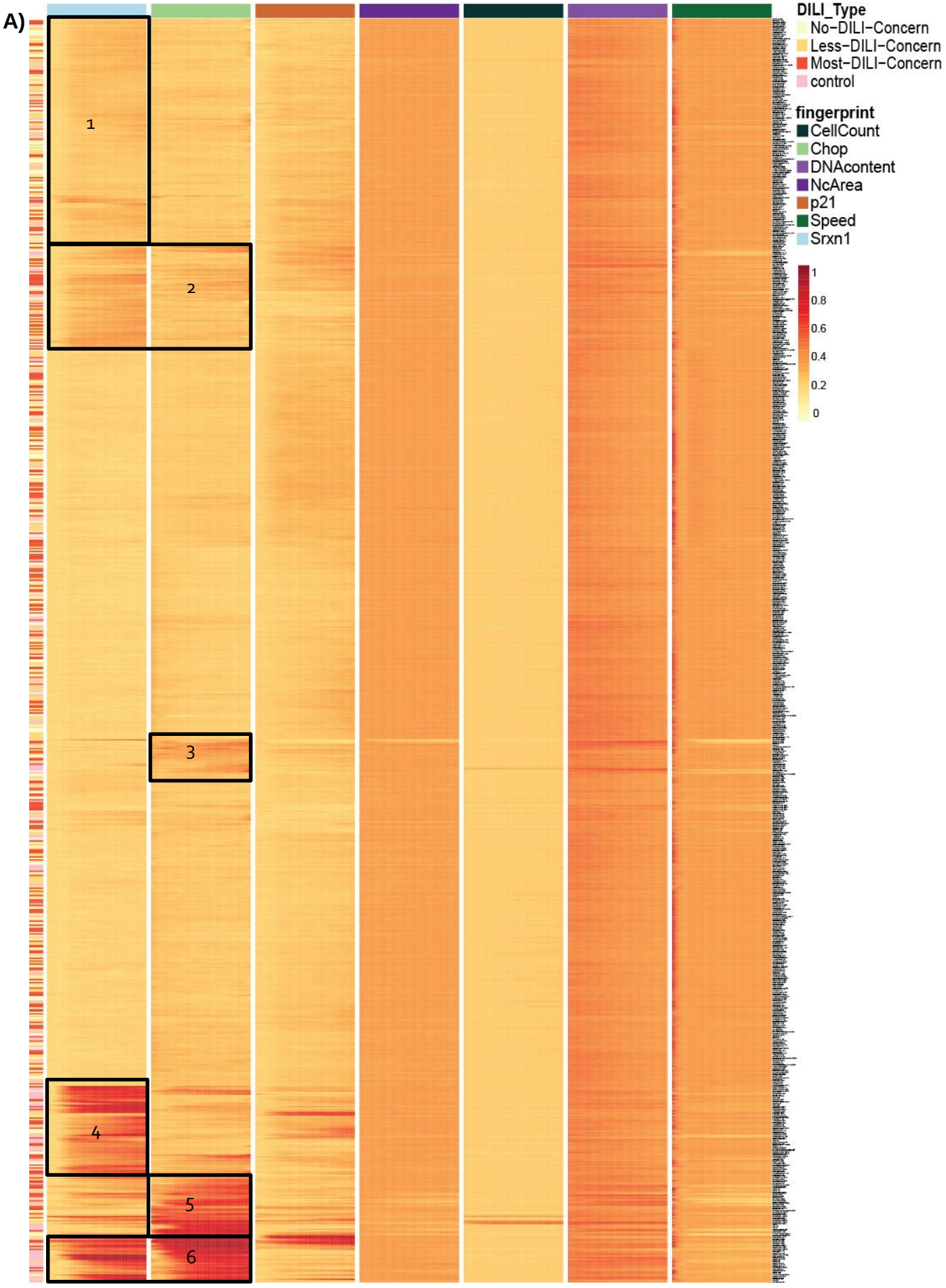


Figure 4: **Concentration time course responses of all compounds for GFP-Srxn1, GFP-Chop and GFP-p21 reporter activation.** For all compounds the reporter activity for the individual reporters was fitted and scaled as described in material and methods section. Shown are the responses for individual concentration (C-max 1, 5, 10, 50 and 100) for the different reporter (GFP-Srxn1: blue; GFP-Chop/DDIT3: green; GFP-p21; red).

4.4. DILI compounds show specific reporter activation with distinct time dynamics and magnitude.

As a next step we performed hierarchical clustering of all the dynamic information of the 176 concentration-response datasets representing the reporter activities from all BAC-GFP reporter cell lines as well as the viability markers. For this the distance between all time-vectors for each feature were computed separately followed by calculating the mean distances over all features (Figure 5). These mean-distances were used for the ward-clustering. The BAC-GFP reporter responses contributed most to the clustering; as expected the viability measures did only contribute to a minor extent. The clustering distinguished 6 clearly distinct groups: 1) mild Srxn1 activation only; 2) mild Srxn1 and CHOP activation, but no p21 activation; 3) mild CHOP activation only; 4) predominant strong Srxn1 activation; 5) predominantly strong CHOP activation; and 6) strong Srxn1 and CHOP activation. Several compounds also induce p21 strongly (arrows). Since these data include all different concentrations of 1 up to 100 times C-max, we would not expect a defined clustering of all DILI compound exposures, since the lower concentrations were expected not to give a strong response. Regardless, in cluster 4-6 we could see strong responses of 13 of in total 51 (25%) less-DILI concern and 26 of in total 58 (45%) most-DILI concern compounds. 8 of total 37 (22%) non-DILI-concern compounds were in this group, of which ethacrynic acid, mitomycin C, bisphenol A and entecapone are known to be directly cytotoxic to cells and would be defined as reference control compounds. Of relevance, we observed that group 4, i.e. predominant fast and strong activation of the Srxn1-GFP reporter, contained reference control compounds that directly affect the KEAP1/Nrf2 pathway (CDDO-Me and DEM) and other compounds, including DILI compounds, that based on their structure and (in)direct reactivity are likely to affect cellular thiol residues and thus may affect the KEAP1/Nrf2 pathway as well (e.g. carmustine are disulfiram). This illustrates the direct and specific assessment of mode-of-action. Additionally, we observed that several compounds that directly affect mitochondrial function through different mechanisms (oligomycin A and B, FCCP and CCCP) clustered together and strongly affected both the Srxn1-GFP and CHOP-GFP. The p21-GFP was only mildly affected in this screen: only (in)direct DNA damaging agents (etoposide, doxorubicin, mitomycin C, menadione).


A)
1
2
3
4
5
6
DILI_Type
No-DILI-Concern
Less-DILI-Concern
Most-DILI-Concern
control
fingerprint
CellCount
Chop
DNAcontent
NcArea
p21
Speed
Srxn1
1
0.8
0.6
0.4
0.2
0

B)


DILI_Type
No-DILI-Concern
Less-DILI-Concern
Most-DILI-Concern
control
fingerprint
Srxn1
Chop
p21
1
0.8
0.6
0.4
0.2
0
coumarin 0.8
isoniazid 45
succinylcholine 138
carmustine 600
iodoacetamide 10
sodium arsenite 100
oxytetracycline 326
DEM 1000
tunicamycin 100
CDDO-me 1
CDDO-me 0.5
DEM 100
disulfiram 540
labetalol 134
ethacrynic acid 165
rotenone 2
DEM 500
ticlopidine 404
antimycin A 10
bromfenac 898
nocodazole 100
etoposide 202
etoposide 40
etoposide 404
ethacrynic acid 33
mercaptopurine 24
colchicine 1.6
nefazodone 86
azathioprine 17
ximelagatran 845
nocodazole 50
omeprazole 470
hydroxyurea 34843
ribavirin 131
mitomycin C 72
ribavirin 261
colchicine 0.8
cytochalasin D 50
cytochalasin D 100
methyldopa 189
iodoacetamide 5
DEM 50
sodium arsenite 50
moxisylyte 16
acetaminophen 10000
DMNQ 4.1
furosemide 329
butylated hydroxyanisole 100
butylated hydroxyanisole 50
azidothymidine 198
azidothymidine 397
indomethacin 559
dacarbazine 2064
nitrofurantoin 60
antimycin A 2
menadione 25
bisphenol A 80
entacapone 393
benzbromarone 217
entacapone 197
labetalol 268
mitomycin C 359
azathioprine 34
sulindac 3199
mercaptopurine 48
antimycin A 20
cyclosporin A 20
flurbiprofen 5731
ibuprofen 7010
ximelagatran 1690
bosentan 342
tunicamycin 10
benzbromarone 434
nitrofurantoin 300
etodolac 6850
FCCP 10
CCCP 10
tetracycline 210
methyldopa 947
diltiazem 960
diclofenac 420
rotenone 40
zimelidine 6000
carmustine 3000
iodoacetamide 50
sodium arsenite 1000
zimelidine 600
ciprofloxacin 329
ciprofloxacin 658
dantrolene 79
chloroquine diphosphate 0.01
penicillamine 138
carmustine 6000
doxycycline 113
tetracycline 1049
carmustine 300
mitomycin C 36
tetracycline 2096
doxycycline 1126
doxycycline 563
BFA 1
BFA 10
BFA 5
iodoacetamide 100
nitrofurantoin 600
doxorubicin 12
ethacrynic acid 1649
ethacrynic acid 3299
doxorubicin 118
doxorubicin 59
menadione 250
dantrolene 395
dantrolene 790
rifampicin 1500
menadione 125
bromfenac 1796
oligomycin B 100
methyldopa 1894
thapsigargin 10
BFA 100
oligomycin A 50
sodium arsenite 500
thapsigargin 5
oligomycin A 25
oligomycin B 50
CCCP 50
FCCP 50
oligomycin A 5
oligomycin B 10
oligomycin B 1
nimesulide 2108
thapsigargin 1
ticlopidine 807
CCCP 100
FCCP 100
ethacrynic acid 330
oligomycin A 2.5
oligomycin B 5
rifampicin 750
BFA 50
tunicamycin 50

Figure 5: **Hierarchial clustering of concentration time course adaptive stress responses of 173 compound treatments.** (A) Time course heatmap of the three response features of the 3 BAC_GFP adaptive stress response reporters [Srxn1 (blue), Chop (light green) and p21 (brown)] and the four viability markers [nuclear area (dark purple), cell count (black), DNA content (light purple) and cell count (dark green)]. Each of the 7 columns represents a time course of 24 hours. Each line is a separate treatment [compound/concentration]. The red intensity-level represents the magnitude of the feature. All treatments are annotated with the vertical bar on the left as reference control compound (pink), no-DILI (white), less-DILI (light orange) and most-DILI (dark orange). (B) A detail of the clusters 4, 5 and 6 is shown.

4.5. Dose response curves and multiple feature clustering increases mechanistic insight of compound induced cytotoxicity.

As a next step the time profiles of all relevant features were transformed (see also Figure 2). Several features were independent of time (such as cell death fraction at the end of the live imaging session); others required transformation to the maximum in time (to include early response profiles); and for yet other features such as cell speed or cell count the averages and the slopes were determined to get an indication of these features with respect to treatments (with the average) and in time (with the slope). These data transformations allowed visualization of dose response relationships for all compounds and the three BAC-GFP reporters (Figure 6). Half of the compounds induce SRXN1-GFP, about one-third induce CHOP-GFP; much fewer compounds activate p21-GFP. Most responses follow a typical increase concentration level dynamic. Yet, several compounds lead to a decrease in response at higher concentrations; this is likely caused by cytotoxicity and cell death caused by the compounds at the higher concentrations (e.g. ethacrynic acid, nitrofurantoin and iodoacetamide; see also cluster 1 in Figure 7). Interestingly, several compounds at lower concentrations initiate oxidative stress and only at higher concentration the 'secondary' UPR is activated (e.g. dilitiazem, ethacrynic acid, iodoacetamide and simvastatin). Reversely, several compounds activate the UPR CHOP-GFP reporter at lower C-max values followed by 'secondary' SRXN1-GFP induction (e.g. brefeldin A, bromfenac, CCP, FCCP, rifampicin, thapsigargin and tunicamycin).

4.6. Clustering of compound reporter fingerprints allows mode-of-action discrimination.

Next we used the concentration response data from all reporters as well as the biological relevant reporter-independent features for multiple feature-based unsupervised hierarchical clustering of all the DILI compounds together with reference control compounds with known mode-of-action. For the reporter responses we included: 2X mean of DMSO, 3X mean of DMSO, mean + 3X standard deviation of DMSO, mean and integrated intensity; for the reporter-independent responses we included information on: nuclear area, Hoechst intensity, cell speed, number of cells and the PI (necrotic) and AV (apoptosis) positive fraction of cells (Figure 7A and B). Clustering showed two groups with strong reporter activation (see group 1 and 2 in Figure 7A); of importance: cytotoxicity responses are in particular observed in cluster 1.

Group 1 contains 12 DILI compounds (6 less-DILI & 6 most-DILI), 2 non-DILI compounds (ethacrynic acid and mitomycin C) and 10 control compounds. Ethacrynic acid is labeled as non-DILI although this compound is well known to deplete glutathione [322] and rather reflects a reference control compound. Also mitomycin is an alkylating agent causing genotoxicity [323]. The compounds in group 1 showed overt toxicity at higher C-max values and high Srxn1 and CHOP

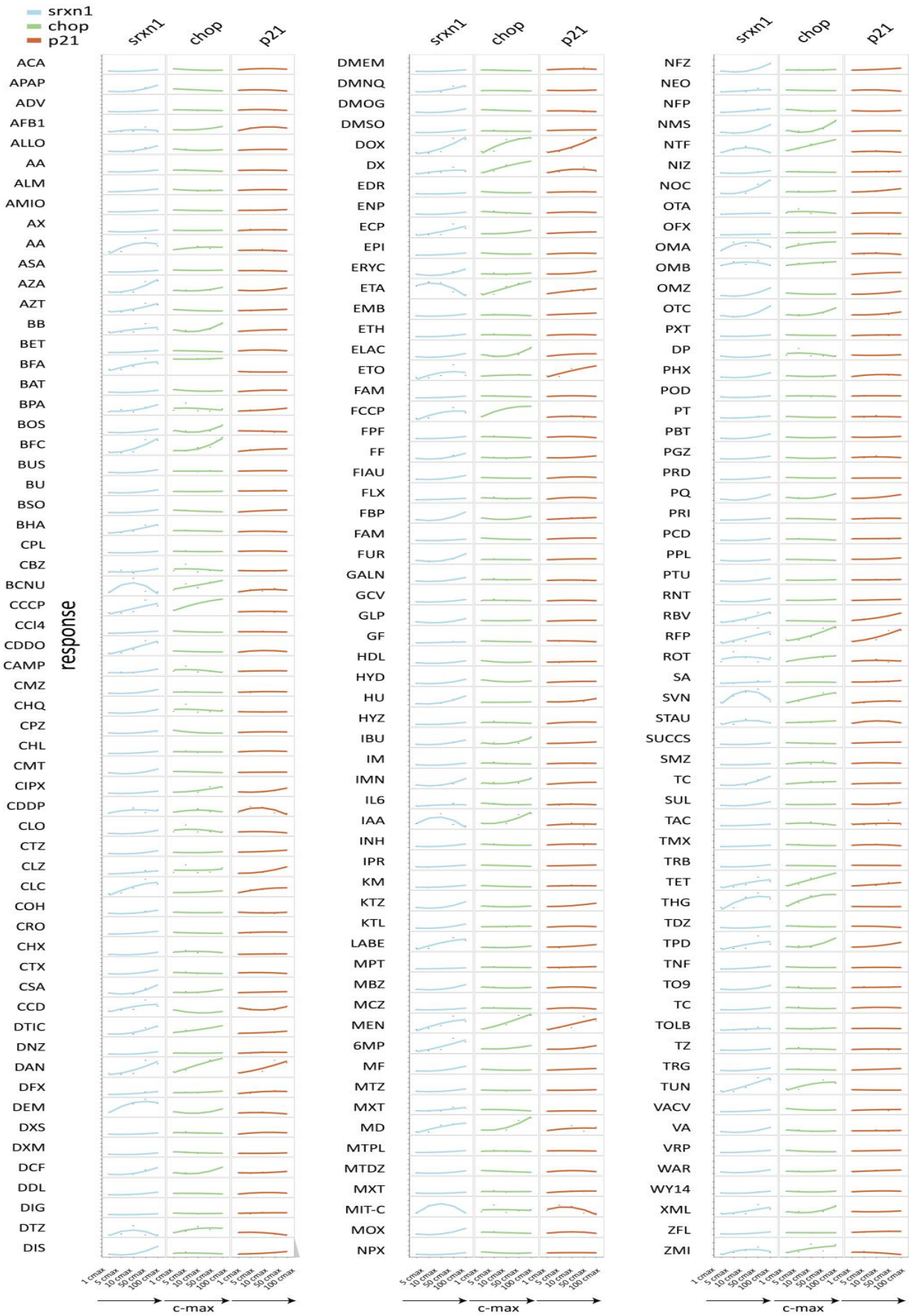
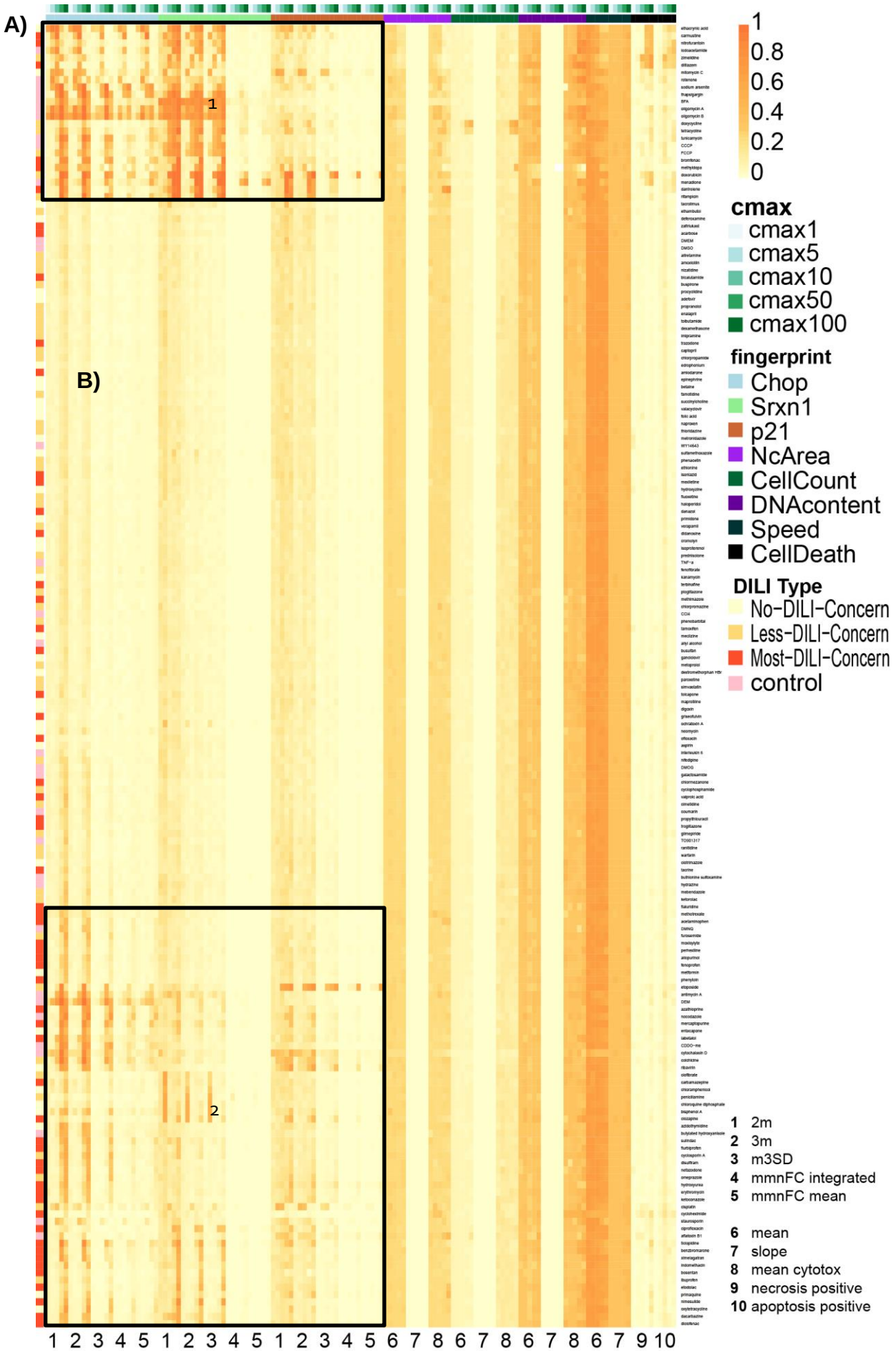


Figure 6: **Dose response curves for all individual compounds and reporters.** Dose response is based on max-in-time of fraction of cells above 2X mean DMSO level for the adaptive stress response proteins Srxn1 (blue), Chop (green) and p21 (red). The compounds are ordered alphabetically, compound abbreviations can be found in Table 1.


A)
B)
1
2
1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 6 7 8 6 7 8 6 7 8 6 7 9 10
1
0.8
0.6
0.4
0.2
0
cmax
cmax1
cmax5
cmax10
cmax50
cmax100
fingerprint
Chop
Srxn1
p21
NcArea
CellCount
DNAcontent
Speed
CellDeath
DILI Type
No-DILI-Concern
Less-DILI-Concern
Most-DILI-Concern
control
1 2m
2 3m
3 m3SD
4 mmnFC integrated
5 mmnFC mean
6 mean
7 slope
8 mean cytotox
9 necrosis positive
10 apoptosis positive

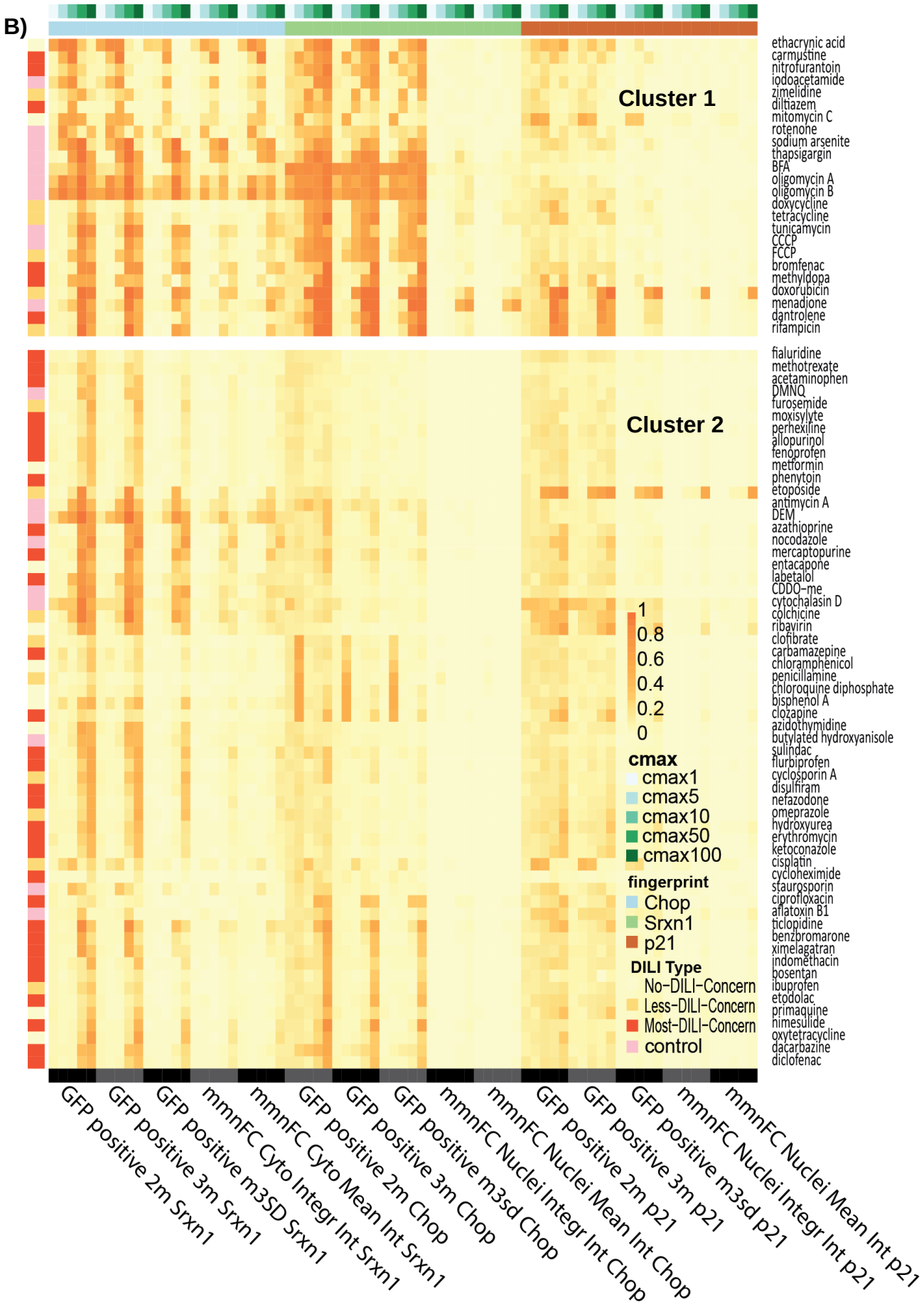
responses at lower C-max values. Strikingly both structure-related (e.g. oligomycine A and B, CCCP and FCCP) as well as mode-of-cation-related compounds (brefeldin A, tunicamycin and thapsigargin) clustered together in this group (Figure 7B). Since the reference control compounds in this group are quite reactive and cytotoxic (e.g. iodoacetamide, ethacrynic acid, carmustine, sodium arsenite, rotenone, oligomycin, CCCP, and menadione), we anticipate that the DILI compounds that fall within this cluster are likely to be equally reactive and/or disrupt vital metabolic processes in the cell leading to oxidative stress and UPR activation.

Within group 2 little cytotoxicity was observed within the entire concentration range, indicating that the adaptive stress responses well precede cytotoxic dose ranges. The group 2 compounds are highly enriched with most-DILI concern compounds (31 (53%) of all most-DILI compounds). Several key groups can be identified: i) compounds that predominately induce the SRXN1-GFP reporter, ii) compounds that predominately induce CHOP-GFP reporter activity; and iii) compounds that activate both pathways at closely similar C-max values.

4.7. Multiple adaptive stress response pathway activation and higher C-max values indicate higher risk for DILI.

We further investigated the relationship between concentration and reporter activity. We therefore defined for each compound the actual concentration at which a compound caused at least 25% or 50% induction of GFP-positive cells for individual reporters. Most compounds passed the criteria for the Srxn1-GFP reporter; in particular reference control compounds caused activation of this reporter at sub-µM range, whereas most DILI compounds (either less or most concern) caused the activation in the range of 10-100 µM; also some non-DILI compounds fell in this range (Figure 8A). Similar observations were made for CHOP-GFP and p21-GFP reporters. Further enrichment of most-DILI compounds occurred for minimally 2 reporter responses with both >25% GFP positive cells, with fewer non-DILI compounds remaining (Figure 8B). Further increasing the stringency (50% GFP-positive cells) overall decreased the number of compounds, and in particular remained Srxn1-GFP and CHOP-GFP, responses (Figure 8C). Interestingly, dual activation under these stringent conditions further reduced the number of compounds, as well as

Figure 7: **Hierarchical clustering of compounds based on concentration responses for all reporter and cell viability features.** A) Clustering results of all features were transformed to single values as described in material and methods. Fingerprint-sets Srxn1 (blue), Chop (light green), p21 (brown), nuclei area (light purple), cell count (dark green), DNA content(dark purple), cell speed (grey) and cell death (black) consist of different features. For reporters: 1) **2m** (fraction of cells above 2 times the mean), 2) **3m** (fraction of cells above 3 times the mean), 3) **m3SD** (3 times the mean or mean plus 3 times standard deviation DMSO population), 4) **mmnFC integrated** (normalized fold change with respect to DMSO population integrated), 5) **mmnFC mean** (normalized fold change with respect to DMSO population integrated). For non-reporter features: 6) **mean** (mean of linear fits of time course data), 7) **slope** (slope of linear fits of time course data(, 8) **mean cytotox** (population mean of PI and annexin), 9) **necrosis positive** (fraction of cells stained for PI), 10) **apoptosis positive** (fraction of cells stained annexin-V positive). Each feature shows the concentration range (shades of blue-green [marked on top]) from left to right: 1, 5, 10, 50 and 100 C-max. On the left the compound DILI annotation is indicated as before. B) Zoom of group 1 and group 2 for the different reporter features. B) Zoom of boxes 1 and 2 (displayed on next page)


B)
Cluster 1
Cluster 2
ethacrynic acid
carmustine
nitrofurantoin
iodoacetamide
zimelidine
diltiazem
mitomycin C
rotenone
sodium arsenite
thapsigargin
BFA
oligomycin A
oligomycin B
doxycycline
tetracycline
tunicamycin
CCCP
FCCP
bromfenac
methyldopa
doxorubicin
menadione
dantrolene
rifampicin
fialuridine
methotrexate
acetaminophen
DMNQ
furosemide
moxisylyte
perhexiline
allopurinol
fenoprofen
metformin
phenytoin
etoposide
antimycin A
DEM
azathioprine
nocodazole
mercaptopurine
entacapone
labetalol
CDDO-me
cytochalasin D
colchicine
ribavirin
clofibrate
carbamazepine
chloramphenicol
penicillamine
chloroquine diphosphate
bisphenol A
clozapine
azidothymidine
butylated hydroxyanisole
sulindac
flurbiprofen
cyclosporin A
disulfiram
nefazodone
omeprazole
hydroxyurea
erythromycin
ketoconazole
cisplatin
cycloheximide
staurosporin
ciprofloxacin
aflatoxin B1
ticlopidine
benzbromarone
ximelagatran
indomethacin
bosentan
ibuprofen
etodolac
primaquine
nimesulide
oxytetracycline
dacarbazine
diclofenac
1
0.8
0.6
0.4
0.2
0
cmax
cmax1
cmax5
cmax10
cmax50
cmax100
fingerprint
Chop
Srxn1
p21
DILI Type
No-DILI-Concern
Less-DILI-Concern
Most-DILI-Concern
control
GFP positive 2m Srxn1
GFP positive 3m Srxn1
GFP positive m3SD Srxn1
mmnFC Cyto Integr Int Srxn1
mmnFC Cyto Mean Int Srxn1
GFP positive 2m Chop
GFP positive 3m Chop
GFP positive m3sd Chop
mmnFC Nuclei Integr Int Chop
mmnFC Nuclei Mean Int Chop
GFP positive 2m p21
GFP positive 3m p21
GFP positive m3sd p21
mmnFC Nuclei Integr Int p21
mmnFC Nuclei Mean Int p21

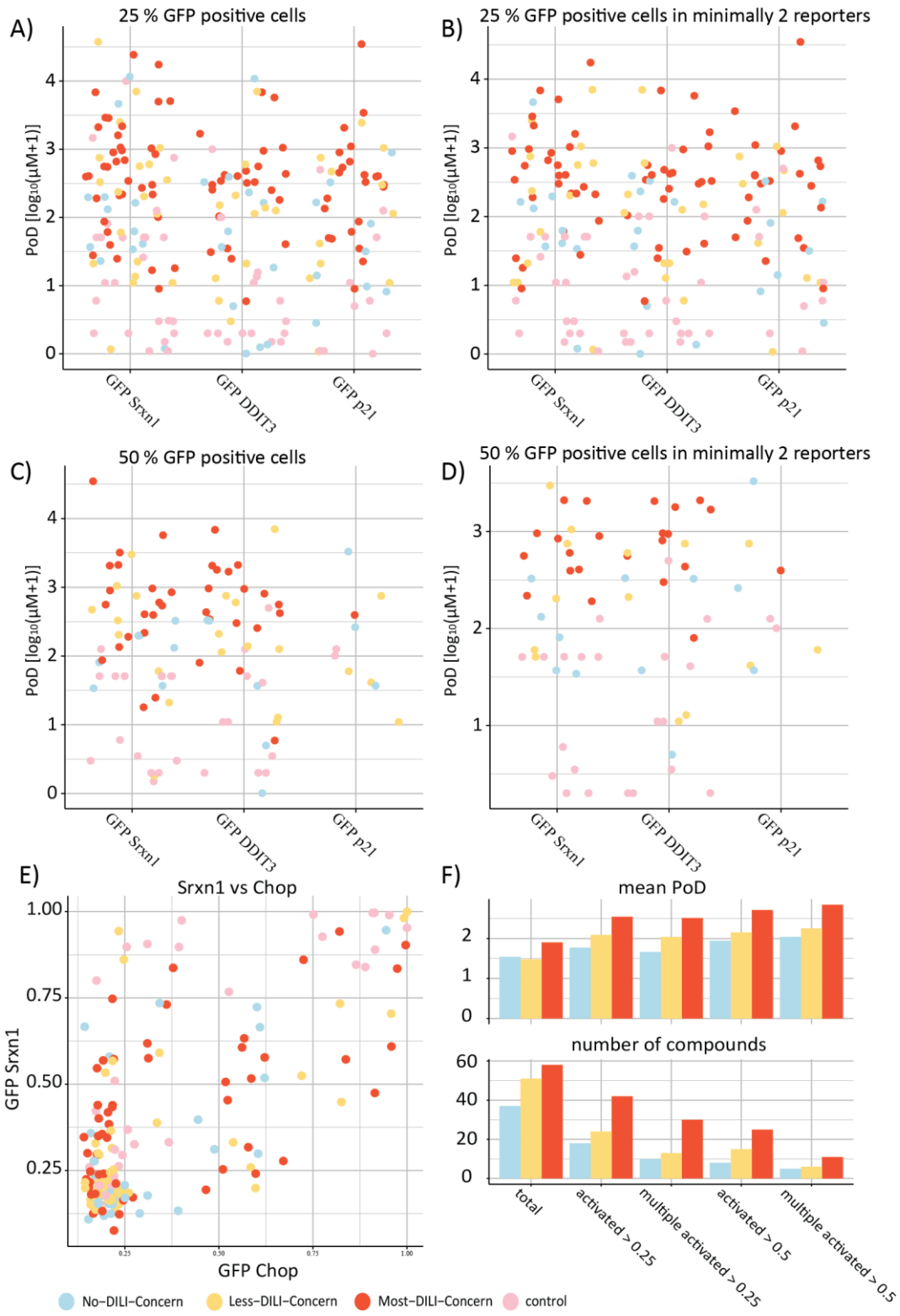


Figure 8: **Relationship between reporter response and DILI labeling.** A) Shown is the lowest concentration where at least 25% cells is "GFP positive" as defined by 2X above population mean of DMSO; only compounds are included for which this is true. B) Lowest concentration where at least 25% of cells is GFP positive in at least two different reporters. C & D) Same as panel A and B, respectively, but now with at least 50% of cells GFP positive. F) summary of panels A-E: top shows the mean Point of Departure (PoD) as the mean of the lowest concentrations of all treatments that activates the reporters. Bottom shows the total number of compounds per DILI category that are activated under these conditions.

further increased the overall concentration at which such a double response was observed (Figure 8D and F). When plotting for all compounds the Srxn1-GFP and CHOP-GFP reporter activities, both most informative, we observed that strong activation of Srxn1 and CHOP was primarily observed for less and most DILI concern compounds (Figure 8E); only one non-DILI compound belonged to this group which was actually ethycrynic acid, a strong soft electrophilic compound directly affecting cellular thiols.

4.8. Concordance between HepG2 BAC-GFP reporters and primary human hepatocytes (PHH).

To assess the concordance between our reporters and PHH we compared our BAC-GFP reporter data (maximal % GFP positive cells observed) and the fold transcript data in PHH based on the TG-GATEs dataset (highest concentration at 24 hr). In general the response of PHH transcript levels is highly comparable to the HepG2 BAC reporters, in particular for the Srxn1-GFP and CHOP-GFP reporters (Figure 9A). The concordance increased when strong responses were observed in the reporter cells. Interestingly BAC reporters seemed more sensitive than transcript analysis in PHH (Figure 9B). In addition treatments that activate both the HepG2 BAC reporters as well as increase transcript levels are highly enriched in most-DILI concern compounds.

5. Discussion

Here we investigated the integrated application of a panel of three adaptive stress response BAC reporters in high content high throughput screening as a method for DILI liability assessment. We focused on adaptive stress pathway signalling as these defense programs are a universal theme in all life forms, and respond before the onset of overt toxicity. We monitored three downstream target genes for oxidative stress (Srnx1), ER-stress/UPR (CHOP) and DNA damage (p21); these are selective targets for these pathways. Using CellProfiler and R-package H5CellProfiler we quantified all features and performed detailed analysis on the individual contribution to DILI classification. We demonstrate how advanced analysis of live-single cell data can provide key information on the concentration-time course reporter responses that can be applied for DILI liability evaluation. We showed that integration of such mode-of-action assessment using different reporter strongly improves the classification of less- and most-DILI concern compounds.

The current data demonstrate the strength in the application of time-resolved live single cell reporter data. To date toxicity screening efforts using high content imaging have mostly focused on single time point fluorescent dyes or anti-bodies [241] with several real-time based toxicity screening efforts [324]. However with the use of our reporter cell lines biological signalling can be visualized with a high time resolution to more accurately pin-point the primary mode-of-action. The use of dyes and anti-bodies brings additional noise to already very noisy systems as fixation and anti-body binding are likely additional sources of variability; this is not an issue using our reporter models. Time course signalling data also greatly benefits computational modeling efforts as these require detailed time and dose response dynamics, this is only feasible using live cell imaging data.

We have successfully applied a limited set of features for DILI liability assessment. Yet,

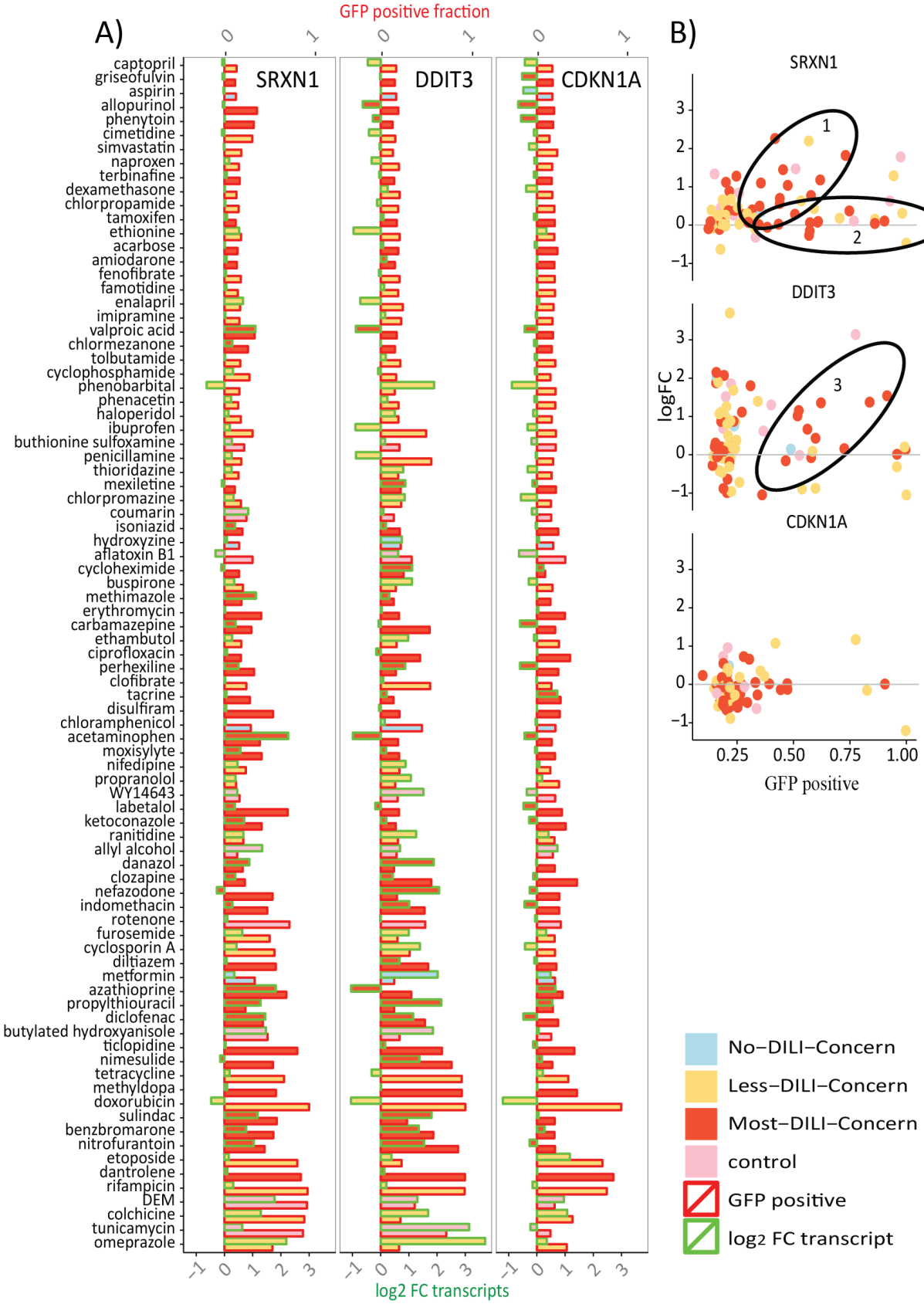


Figure 9: **Concordance between BAC-GFP HepG2 reporter activation and primary human hepatocyte transcript levels.** A) Compounds are ordered from lowest average over Srxn1, Chop and p21 to the highest (top). BAC GFP intensity levels are red-lined, transcript fold changes are green-lined. Bar colors correspond to No-DILI-concern (blue), Less-DILI-concern (light orange), Most-DILI-Concern(dark orange) and control compounds (pink). Panel B) GFP positive cells (x-axis) vs transcript levels (y-axis).

another approach to high content imaging is the use of large amounts (up to 700) of measurements/features [325, 326] and use supervised clustering, dimension reduction techniques and/or feature selection to obtain a classifier to predict toxicity or specific compound effects. However several challenges exist with such an approach. It is often difficult to define a biological meaning over certain results beyond co-clustering of compounds of which the toxicity/mechanism is presumed to be known (so called guilt by association). A high-dependency of most of these features exist and this magnifies certain phenotypes over others likely diluting more subtle and possibly important phenotypic changes. We focused on a smaller set of biological interpretable features we believe to be of interest in DILI. This small set of biological interpretable features is a 'fingerprint' as to the cellular state and in the current study includes the activation of: the oxidative stress response, i.e. Srxn1-GFP activation; the ER-stress response (UPR), i.e. CHOP activation; the DNA-damage response, i.e. p21 activation; nuclear area; number of cells; Hoechst intensity; cell speed and cell death (necrosis and apoptosis). This strategy captures together several adaptive stress responses as well as general cellular features. We acknowledgde the limitation of these measures, since DILI phenotypes that involve innate and adaptive immune responses, cholestasis or steatosis are likely not covered in our reporters. This possibly explains why the overall sensitivity of the combined assays is not yet sufficient to correctly identify all less- and most-DILI concern compounds. Future integration of other GFP-BAC reporters that would cover these DILI toxicity programs will still be required. At that stage supervised clustering using the pre-defined DILI labels followed by feature selection to establish a most-DILI-concern fingerprint should improve our current approach. Since several DILI compounds induce stress responses with distinct time and reporter activation dynamics, it would be relevant to include the order of reporter activation, the time of the maximum response, the magnitude of the response and the C-max value associated to these dynamics in such supervised clustering efforts. This will be our future research focus. Regardless, our current strategy already allows a near perfect fingerprinting, since oligomycin A and oligomycin B as well as CCCP and FCCP do cluster together. Additional reporter will further improve the compound mode-of-actiom fingerprinting.

Our data indicate a striking overlap between HepG2 BAC-GFP responses and transcript levels in primary human hepatocyte. Although HepG2 is metabolically incompetent compared to primary hepatocytes, several compounds that involve biotransformation-dependent toxicity do show a Srxn1-GFP oxidative stress response (e.g. acetaminophen and sulindac). Possibly the low metabolic capacity of the HepG2 model is offset to some extent by the increased sensitivity as it is not toxicity that is measured but an adaptive stress responses. Future studies will establish whether our HepG2-based reporter platform would further benefit from improved metabolic competence.

The anti-oxidant response is critical in the protection against oxidative stress and inhibition of this pathways propagates DILI for various compounds [327]. In our study the Srxn1-GFP reporter was activated by about half of the DILI compounds, which fits with the observation that oxidative stress in a major contributor to DILI [328]. Our test compounds may affect this pathway through direct interaction with KEAP1, thus unleashing Nrf2, or through modulation of the mitochondrial respiratory chain. Indeed various reference control compounds that either affect KEAP1 (CDDO, DEM, ethacrynic acid) or the mitochondria (rotenone, CCCP, oligomycin A) do induce Srxn1-GFP activation, supporting both direct and indirect mechanisms for Srxn1-GFP induction. Most notably oxidative stress following UPR activation was also seen for several ER-

stress inducing compounds, including thapsigargin, tunicamycin and brefeldin A. Reversely, oligomycine and CCCP strongly activated the CHOP-GFP activation. Altogether this underscores the connectivity and complexity of adaptive stress response regulation.

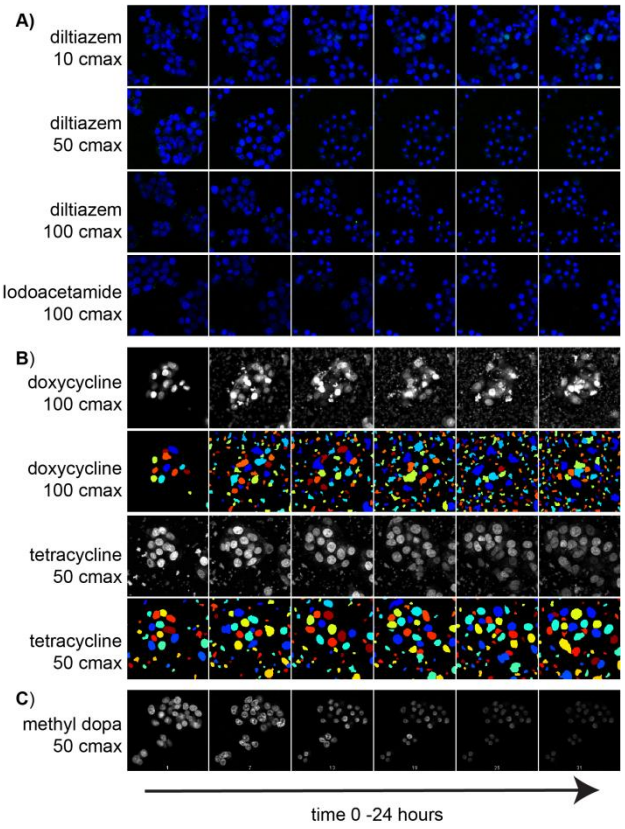
Only several DILI compounds mildly activated the p21 reporter. This is fortunate and expected since direct DNA damaging compounds are terminated early in drug development. Although p21-GFP is activated by DNA damaging agents in our study (e.g. doxorubicin and etoposide), p21 cannot be considered as a sole DNA damage reporter, since its expression is also regulated through other stress responses. Regardless, in our setting, the p21-GFP reporter contributed to the evaluation of DILI responses, be it to a limited level than Srxn1-GFP and CHOP-GFP.

In conclusion, we have shown that BAC-GFP reporter cell lines are a sensitive tool to provide detailed mechanistic information regarding the adaptive stress response activation in a broad compound screening setting using high-content live single cell imaging. Such detailed insights in the perturbations of signalling pathways after chemical exposure provides key information for predictive purposes. We anticipate that our BAC-GFP reporter platform will contribute to the early pre-clinical screening for DILI liabilities and possibly also other chemical safety assessment paradigms.

Supplemental figures

compound[µM]	nuclei area	cell count	DNA content	migration speed
aflatoxin B1 [50]	normal	normal	low	normal
aflatoxin B1 [100]	normal	high	high	normal
benzbromarone [440]	normal	normal	normal	low
carmustine [3000]	low	normal	normal	low
carmustine [6000]	normal	normal	normal	very low
cisplatin [20]	normal	normal	low	normal
cisplatin [100]	normal	normal	normal	bit low
cisplatin [200]	normal	normal	normal	bit low
cyclohexamide [50]	normal	normal	normal	bit low
cyclohexamide [250]	normal	normal	normal	bit low
cyclohexamide [500]	normal	normal	normal	bit low
cyclohexamide [2500]	normal	normal	normal	bit low
cyclohexamide [50]	normal	normal	normal	normal
cyclohexamide [5000]	normal	normal	high	low
cytochalasin D [5]	normal	normal	normal	low
cytochalasin D [10]	normal	normal	normal	low
cytochalasin D [50]	normal	normal	normal	very low
cytochalasin D [100]	normal	normal	normal	very low
digoxin [0.28]	normal	normal	normal	bit low
diltiazem [480]			high	
diltiazem [960]	normal	normal	high	low
diltiazem [4800]	low	low	high	very low
diltiazem [9600]	low	low	high	very low
doxycycline [112]	normal	high	normal	normal
doxycycline [560]	normal	high	normal	normal
doxycycline [1125]	normal	high	high	normal
ethacrynic acid [1650]	low	normal	normal	very low
ethacrynic acid [3300]	low	normal	normal	very low
iodoacetmide [50]	low	normal	high	very low
iodoacetmide [100]	low	normal	normal	very low
methyldopa [950]	low	normal	low	very low
methyldopa [1900]	normal	normal	low	normal
nitrofurantoin [300]	normal	normal	normal	low
nitrofurantoin [600]	low	normal	normal	very low
perhexiline [215]	normal	normal	normal	low
rotenone [4]	normal	normal	normal	low
rotenone [20]	normal	normal	high	very low
rotenone [40]	normal	normal	high	very low
staurosporin [0.5]	normal	normal	normal	low
staurosporin [1]	normal	normal	normal	very low
tetracycline [1050]	normal	high	normal	normal
tetracycline [2100]	normal	high	normal	normal
tunicamycin [100]	normal	normal	normal	bit low
zimelidine [600]	normal	normal	high	low
zimelidine [3000]	low	normal	normal	very low
zimelidine [6000]	low	normal	normal	very low

Supplemental Fig. S1



Supplemental Fig. S2

Supplemental Fig. S1: **Cell viability markers-quality control features.** Compound-concentration conditions that were flagged for at least 2 features are displayed. These are flagged as either normal (colored green) or very low (>20% increase), bit low (<30%), low (<50%) and high (>50%) (high-lighted in orange) as determined by visual inspection of figure 7.

Supplemental Fig. S2: **Example images of quality control flagged treatments. A)** Diltiazem at higher c-max values induces cell death; the cells stop moving, the nuclei decrease in number, shrink and the Hoechst intensity increases (top 3 rows). Nuclei shrinkage with migration stop of iodoacetamide at 100-cmax (bottom row). B) Unusual large number of cells of doxycycline and tetracycline due to auto fluorescence. C) low DNA content/ low Hoechst intensity is observed for methyldopa.