Archives of Toxicology (2018) 92:1295–1310
https://doi.org/10.1007/s00204-017-2126-3

ORGAN TOXICITY AND MECHANISMS

Prediction of metabolism-induced hepatotoxicity on three-dimensional hepatic cell culture and enzyme microarrays

Kyeong-Nam Yu[1] · Sashi Nadanaciva[2] · Payal Rana[2] · Dong Woo Lee[3] · Bosung Ku[4] · Alexander D. Roth[1] · Jonathan S. Dordick[5] · Yvonne Will[2] · Moo-Yeal Lee[1]

Received: 27 September 2017 / Accepted: 15 November 2017 / Published online: 22 November 2017

Abstract

Human liver contains various oxidative and conjugative enzymes that can convert nontoxic parent compounds to toxic metabolites or, conversely, toxic parent compounds to nontoxic metabolites. Unlike primary hepatocytes, which contain myriad drug-metabolizing enzymes (DMEs), but are difficult to culture and maintain physiological levels of DMEs, immortalized hepatic cell lines used in predictive toxicity assays are easy to culture, but lack the ability to metabolize compounds. To address this limitation and predict metabolism-induced hepatotoxicity in high-throughput, we developed an advanced miniaturized three-dimensional (3D) cell culture array (DataChip 2.0) and an advanced metabolizing enzyme microarray (MetaChip 2.0). The DataChip is a functionalized micropillar chip that supports the Hep3B human hepatoma cell line in a 3D microarray format. The MetaChip is a microwell chip containing immobilized DMEs found in the human liver. As a proof of concept for generating compound metabolites in situ on the chip and rapidly assessing their toxicity, 22 model compounds were dispensed into the MetaChip and sandwiched with the DataChip. The IC_{50} values obtained from the chip platform were correlated with rat LD_{50} values, human C_{max} values, and drug-induced liver injury categories to predict adverse drug reactions in vivo. As a result, the platform had 100% sensitivity, 86% specificity, and 93% overall predictivity at optimum cutoffs of IC_{50} and C_{max} values. Therefore, the DataChip/MetaChip platform could be used as a high-throughput, early stage, microscale alternative to conventional in vitro multi-well plate platforms and provide a rapid and inexpensive assessment of metabolism-induced toxicity at early phases of drug development.

Keywords Metabolism-induced hepatotoxicity · Three-dimensional (3D) cell culture array · Metabolizing enzyme microarray · DataChip/MetaChip · High-throughput toxicity screening

Introduction

Modern drug discovery is a multidisciplinary enterprise consisting of disease-based target identification and validation, and high-throughput screening of chemical and natural product libraries (Kennedy et al. 2008). This is followed by the careful optimization of selected lead compounds, in vitro and in vivo pharmacokinetics, toxicology and bioavailability testing, and finally, preclinical and clinical studies (Lee and Dordick 2006; Hughes et al. 2011). These new drug candidates can be eliminated at different stages of drug development for many reasons, resulting in only 1 of 5000 lead candidates that pass the discovery process ever reaching the market. The total capitalized development cost per drug now approaches $2.6 billion, of which a large portion is attributable to drug candidate failures (DiMasi and Grabowski 2012).

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00204-017-2126-3) contains supplementary material, which is available to authorized users.

✉ Moo-Yeal Lee
m.lee68@csuohio.edu

[1] Department of Chemical and Biomedical Engineering, Cleveland State University, 455 Fenn Hall (FH), 1960 East 24th Street, Cleveland, OH 44115-2214, USA

[2] Compound Safety Prediction, Pfizer Inc., Groton, CT 06340, USA

[3] Department of Biomedical Engineering, Konyang University, Daejeon, Republic of Korea

[4] Central R & D Center, Medical & Bio Device (MBD) Co., Ltd, Suwon, Republic of Korea

[5] Department of Chemical and Biological Engineering, and Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

A wide range of emerging in vitro technologies have begun to impact the assessment of chemical and drug candidate toxicity in high-throughput with an aim to reduce the need for animal models (Shukla et al. 2010). While many of these technologies are still nascent, a roadmap toward ultimate validation and industry adoption is becoming clearer. One area where such a roadmap is critically needed is in the confluence of chemical toxicity and human metabolism. Recent advances in genomics and proteomics coupled with sequencing of the human genome have dramatically increased the number of drug targets and their lead compounds (Schadt et al. 2009). Combinatorial and diversity-oriented synthesis programs along with increased access to natural products and their structural scaffolds have provided vast numbers of compounds to screen for identifying lead candidates. Conventional models to elucidate drug toxicity and human metabolism in vitro include isolated liver slices (Westra et al. 2016), primary hepatocytes (Hewitt et al. 2007), transformed hepatoma cell lines (Watanabe et al. 2003), immortalized liver cells expressing P450s (Gustafsson et al. 2014), as well as human liver microsomes (HLMs) and isolated recombinant cytochrome P450 (CYP450) isoforms (Brandon et al. 2003; Hariparsad et al. 2006). Hepatocytes and liver slices most closely resemble the in vivo system, and cryopreserved primary hepatocytes have been extensively used for in vitro drug testing, and are considered as the gold standard for drug screening (Soldatow et al. 2013). Primary hepatocytes provide a complete set of drug-metabolizing enzymes (DMEs) and pathways, and therefore, offer an appropriate system to test for toxicity, metabolite production, and drug stability and partitioning (Bale et al. 2016). Nevertheless, primary hepatocytes are expensive and difficult to obtain in large quantity for high-throughput toxicity screening (Soldatow et al. 2013). More problematic, these cells rapidly lose liver specific functions when maintained under standard in vitro cell culture conditions and often variably express CYP450s and other metabolizing enzymes over time (Gómez-Lechón et al. 2004). In addition, primary hepatocytes are not easy to pool due to varying expression levels of DMEs, which often result in significant experimental variability. Many modifications to conventional culture methods have been developed to foster retention of hepatocyte function. However, the current biotransformation functions of the liver are likely difficult to be mimicked at desired in vivo levels (Sivaraman et al. 2005; Hewitt et al. 2007; Huch et al. 2015).

While these problems represent a major gap in the development of high-throughput in vitro techniques for concordance between in vitro assays and in vivo responses, and consequently, there is a significant opportunity for new technologies to fill this gap. Complicating the difficulty of in vitro systems to mimic human metabolism and toxicity, predicting human responses in drug testing and disease research in vivo with animal tests have poor outcomes. For example, animal models have failed to reproduce human liver toxicity of troglitazone (Rezulin™), which was withdrawn from the market because of CYP450-medicated hepatotoxicity in humans (Reddy et al. 2005; Masubuchi et al. 2006). Mibefradil was also withdrawn due to hepatotoxicity, cardiovascular toxicity, and drug interactions via CYP450 isoforms (Bui et al. 2008). Moreover, drug candidates that fail in clinical trials due to toxicity concerns, by definition, were not flagged by animal models as being potentially toxic. Thus, there remains a gap in our ability to identify toxic drug candidates before clinical testing.

To address these needs, we have developed the Data Analysis Toxicology Assay Chip (DataChip) and the Metabolizing Enzyme Toxicology Assay Chip (MetaChip) technologies that link metabolism and cell-based screening (Lee et al. 2005, 2008). However, the earlier version of the DataChip/MetaChip (i.e., the DataChip 1.0 and MetaChip 1.0) used chemically functionalized microscope glass slides, which required direct contact between cell spots and metabolizing enzyme spots through a liquid layer on sandwiched glass slides to transfer compounds and their metabolites to target cells. It had several technical limitations such as the difficulty of accurately aligning cell/enzyme spots on the glass slides and limited incubation times (typically 6 h). To overcome these technical issues, a plastic micropillar/microwell chip platform (Lee et al. 2013, 2014a, b) was developed and used in the DataChip/MetaChip platform. By inserting a micropillar chip (DataChip) into a microwell chip (MetaChip), the new DataChip/MetaChip platform 2.0 eliminates the spot alignment issue and provides sufficient growth media with compounds in the microwell for cell culture (typically 1–3 days). In addition, the composition of DME solutions on the MetaChip was changed from individual DMEs to mixtures of DMEs to better mimic DME conditions in the liver. In this study, we report in vitro toxicity data obtained for a set of 22 compounds and their in situ-generated metabolites using the high-throughput DataChip/MetaChip platform 2.0 and correlate the in vitro IC_{50} values from the chip with rat LD_{50} values as well as human C_{max} values to better predict drug-induced liver injury (DILI) in humans.

Materials and methods

Preparation of the micropillar and microwell chip

A micropillar chip made of poly (styrene-co-maleic anhydride) (PS-MA) contains 532 micropillars (0.75 mm pillar diameter and 1.5 mm pillar-to-pillar distance). In addition, a microwell chip made of co-polymer of polystyrene and polybutadiene has a complementary set of 532 microwells

(1.2 mm microwell diameter and 1.5 mm well-to-well distance). PS-MA provides a reactive functionality to covalently attach poly-L-lysine (PLL) and, ultimately, alginate spots by ionic interactions. Plastic molding was performed by the SODIC PLUSTECH injection molder in Samsung Electro-Mechanics Company (SEMCO, Suwon, South Korea).

Human liver cell culture and preparation of cell suspension for spotting

Hep3B human hepatoma cell line [American Type Culture Collection (ATCC), Manassas, VA, USA] at passage numbers between 15 and 23 was grown in RPMI 1640 (Mediatech, Manassas, MA, USA) supplemented with 10% fetal bovine serum (FBS, Sigma-Aldrich, St. Louis, MO, USA) and 1% Penicillin–Streptomycin (P/S, ThermoFisher Scientific, Waltham, MA, USA) in T-75 cell culture flasks in a humidified 5% CO_2 incubator (ThermoFisher Scientific) at 37 °C. Suspensions of Hep3B cells were prepared by trypsinizing a confluent layer of the cells from the culture flask with 0.6 mL of 0.05% trypsin-0.53 mM EDTA (ThermoFisher Scientific), and re-suspending the cells in 7 mL of 10% FBS-supplemented RPMI. After centrifugation at 300×*g* for 4 min, the supernatant was removed and the cell pellets were re-suspended with 10% FBS-supplemented RPMI to a final concentration of 6×10^6 cells/mL.

2D cell viability assessment

For toxicity testing on Hep3B monolayers in 96-well plates (i.e., 2D culture), 5.0×10^3 cells were seeded with 200 µL RPMI media in each well and incubated in the CO_2 incubator. Following overnight pre-incubation, the cells were treated with test compounds at varying concentrations for 72 h. After incubation, the cells were incubated with 50 µL of 2.5 mg/mL MTT solution in PBS for 3 h at 37 °C. Purple-colored MTT-formazan crystals generated in metabolically active cells were measured by completely removing the MTT solution and adding 150 mL of DMSO. After shaking for 30 min at 150 rpm, absorbance was measured at 590 nm using a microtiter plate reader (Synergy H1, BioTek instruments, VT, USA).

Preparation of a miniaturized 3D cell culture array (DataChip) on a micropillar chip

To attach cell spots on the micropillar chip, a mixture of poly-L-lysine (PLL, Sigma-Aldrich) and $BaCl_2$ (Sigma-Aldrich) was prepared by mixing an equal volume of 0.01% (w/v) PLL and 100 mM $BaCl_2$. The DataChip was prepared by spotting 60 nL/micropillar of the PLL/$BaCl_2$ mixture onto each of the 532 micropillars using a microarray spotter (S + MicroArrayer, Samsung ElectroMechanics, Co. (SEMCO)) and allowed to dry for 24 h. This was followed by printing 60 nL/micropillar of Hep3B cells suspended in 0.75% (w/v) alginate on top of the dried PLL/$BaCl_2$ spots. While printing Hep3B cells, the micropillar chip was placed on a chilling deck at 4 °C to retard evaporation of water in the spots. The suspension of Hep3B cells in low-viscosity alginate (Sigma-Aldrich) was prepared by mixing 500 µL of the Hep3B cell suspension in 10% FBS-supplemented RPMI (6×10^6 cells/mL), 250 µL of 3% alginate in distilled water, and 250 µL of RPMI so that the final concentration of the cells and alginate were 3×10^6 cells/mL and 0.75%, respectively. After nearly instantaneous gelation, each Hep3B cell spot was immersed in 800 nL of RPMI growth medium in the complementary microwell by sandwiching the micropillar chip with the cells (DataChip) and the microwell chip containing growth media together ("stamping"). The stamped chips placed in a gas-permeable incubation chamber for 30 min to remove excess $BaCl_2$ were then separated and the DataChip was re-stamped onto the microwell chip containing fresh growth media. Finally, the stamped chips were incubated in the CO_2 incubator at 37 °C for 18 h prior to toxicity assessment.

Preparation of a miniaturized enzyme array (MetaChip) on a microwell chip

The MetaChip, a complementary array of encapsulated metabolizing enzymes that was designed to emulate the metabolic reactions in the human liver, was prepared on a microwell chip made of a co-polymer of polystyrene and polybutadiene. Fresh metabolizing enzyme solutions in Matrigel were prepared in a 96-well plate on ice (Table 1) and 120 nL of metabolizing enzyme mixtures in Matrigel were printed on the microwell chip laid on a chilling deck at 4 °C. The MetaChip was transversely divided into four regions (I–IV in Fig. 1c). Specifically, regions I–IV contained no enzyme as a test compound only control, a mixture of human CYP450 isoforms (P450 Mix), a mixture of P450 Mix and human Phase II metabolizing enzymes (All Mix), and human liver microsomes (HLM). Immediately after enzyme printing, the MetaChip was placed in a Petri dish (4 MetaChips per 150 mm-diameter Petri dish) and stored in a − 80 °C freezer until use.

Stamping the DataChip onto the MetaChip with test compounds

Metabolism-induced toxicity assays were performed by printing compounds into the MetaChip and then stamping the DataChip onto the MetaChip. The compounds selected were acetaminophen (as a positive control), benzbromarone, fenoterol, flutamide, diclofenac, labetalol, imipramine,

Table 1 Composition of enzyme mixtures for preparing the MetaChip

Regions	Enzymes	Co-substrates	Matrix
I	120 µL Baculosome®[a]	120 µL RPMI	80 µL Matrigel
II	120 µL P450 Mix[b]	120 µL Phase I cofactor[e]	80 µL Matrigel
III	120 µL All Mix[c]	120 µL All cofactor[f]	80 µL Matrigel
IV	120 µL HLM[d]	120 µL All cofactor	80 µL Matrigel

Phase II mixture contained 12% UGT1A1 (5 mg/mL), 12% T1A3 (5 mg/mL), 12% UGT1A4 (5 mg/mL), 12% UGT1A9 (5 mg/mL), 12% UGT2B4 (5 mg/mL), 12% UGT2B7 (5 mg/mL) from BD Gentest, 4% SULT1A1 (250 µg/50 µL), 4% SULT1A3 (250 µg/50 µL), 4% SULT1B1 (250 µg/50 µL) from Cypex, 6% GST (5 mg/200 µL) from Sigma, 5% NAT1 (2.5 mg/mL), and 5% NAT2 (2.5 mg/mL) from BD Gentest

[a]Baculosome® negative control was purchased from Invitrogen and used without dilution

[b]P450 Mix contained 52% 3A4 (1 µM), 20% 2D6 (1 µM), 8% 2C9 (1 µM), 5% 2E1 (1 µM), 4% 1A2 (1 µM), 4% 3A5 (1 µM), 3% 2C8 (1 µM), 3% 2C19 (1 µM), and 1% 2B6 (1 µM), all from BD Gentest

[c]All Mix contained 50% P450 Mix and 50% Phase II enzyme mixture

[d]Human liver microsome (HLM) were purchased from BD biosciences and used without dilution

[e]Phase I cofactor solution (containing 24.8 mM NADP+, 52.8 mM glucose-6-phosphate, and 8 U/mL glucose-6-phosphate dehydrogenase) was prepared by mixing 4 mL NADP regenerating system solution A and 1 mL solution B from BD biosciences

[f]All cofactor solution contained 50% Phase I cofactor and 50% Phase II cofactor. Phase II cofactor solution contained 50% UDP-GA (2 mL BD solution A + 3 mL BD solution B = 10 mM in 50 mM Tris–HCl buffer, pH 7.5), 20% GSH (100 mM in 20 mM PBS buffer, pH 8), 20% PAPS (25 mM in 20 mM PBS buffer, pH 8), and 10% acetyl CoA (25 mM in 10 mM PBS buffer, pH 8)

phentolamine, risperidone, oxybendazole, sulindac, propranolol, promazine, trazodone, buspirone, carbidopa, bosentan, chlorpropamide, phenazopyridine, estradiol, mefenamic acid, and fluoxetine, all of which were water soluble at the highest dosages to avoid issues with precipitation over time and adsorption on the chip surfaces. Briefly, compound stock solutions were prepared by dissolving compounds in DMSO. Typically, higher than 100 mM of compound stock solutions were required to maintain final DMSO content less than 0.5%. Approximately 40 µL of test compound solutions were prepared in 200-fold higher concentrations than the desired final concentration (5 dosages plus 1 control) by serially diluting compound stock solutions in DMSO in a 384-well plate. As a control, 100% DMSO without compound was used. After that, 300 µL of diluted test compound solutions in a round-bottom 96-well plate was prepared by mixing 1.5 µL of the diluted compounds in DMSO with 298.5 µL of RPMI (typically 0–1000 µM of final concentrations). Frozen MetaChips were removed from the freezer and immediately placed on the cold slide deck at 4 °C, and then 720 nL of the test compound solutions in RPMI were printed into each well of the MetaChip using the microarray spotter. Six different compounds were printed in sections 1–6 of the MetaChip, each region containing a 12 × 6 mini-array. The stamped chips were placed, with the DataChip on top, in the gas-permeable chamber with 20 mL of sterile distilled water to prevent water evaporation during incubation, and then incubated for 24 h in the CO_2 incubator at 37 °C for cytotoxicity assays. After 24 h incubation with compounds, the MetaChip was discarded, the DataChip was stamped onto the pre-warmed microwell chip with 800 nL/well of fresh RPMI, and then the stamped chips were incubated in the gas-permeable chamber in the CO_2 incubator at 37 °C for additional 48 h.

Cell staining, chip scanning, and data analysis

At the end of the 48 h culture period post-MetaChip stamping, the DataChip was washed twice by immersing the micropillars with cell spots in a deep-well staining plate containing 5 mL of 140 mM NaCl with 20 mM $CaCl_2$ for 5 min each. $CaCl_2$ was supplemented to prevent degradation of alginate spots by excess phosphate. A staining dye solution was prepared by adding 1.0 µL of calcein AM (4 mM stock from ThermoFisher Scientific) and 4.0 µL of ethidium homodimer-1 (2 mM stock from ThermoFisher Scientific) in 8 mL of 140 mM NaCl supplemented with 20 mM $CaCl_2$. To stain the cell spots, 2 mL of the dye solution was dispensed on a shallow-well staining plate and then the DataChip was placed on the top of the shallow-well staining plate, and incubated in the dark for 45 min at room temperature. The DataChip was then washed twice by immersing micropillars with cell spots in the deep-well staining plate containing 5 mL of 140 mM NaCl with 20 mM $CaCl_2$ for 15 min each to remove excess dye in the alginate spots. After drying the DataChip in dark for at least 2 h, the location of each cell spot where compounds added was detected by imaging the entire DataChip using a blue laser (488 nm) and a standard blue filter for green dye (PMT gain: 180 and power: 10) and a blue laser and a 645AF75/594 filter for red dye (PMT gain: 200 and power: 10) in a GenePix® Professional 4200A scanner (MDS Analytical Technologies). Due to the scanning

height difference of the micropillar chip from standard glass slides, the DataChip was scanned at focus position 120. Data files were saved as single images for analysis. The green fluorescence intensity was quantified from the scanned images using the S + Chip Analysis (SEMCO) program by extracting fluorescent intensity from each cell spot and plotting the percentage of live cells against the concentration of the compound tested. We used a background subtraction of dead cells (cells immersed in 70% methanol for 1 h), which was negligible compared to the total fluorescence. The percentage of live cells was calculated using the following equation:

$$\%\ \text{Live cells} = \frac{F_{\text{Reaction}}}{F_{\text{Max}}} \times 100$$

where F_{Reaction} is the green fluorescence intensity of the reaction spot and F_{Max} is the green fluorescence intensity of untreated viable cells. To produce a conventional sigmoidal dose–response curve, with response values normalized to span the range from 0 to 100% plotted against the logarithm of test concentration, the green fluorescence intensities of all cell spots were normalized with the fluorescence intensity of 100% live cell spot (i.e., cell spots contacted with no compound) and the test compound concentration was converted to their respective logarithms. The sigmoidal dose–response curves and IC_{50} values (concentration of the compound where 50% of cell growth inhibited) were obtained using the following equation:

$$Y = \text{Bottom} + \left(\frac{\text{Top} - \text{Bottom}}{1 + 10^{(\log \text{IC}_{50} - X) \times H}} \right)$$

where IC_{50} is the midpoint of the curve, H is the hill slope, X is the logarithm of test concentration, and Y is the response (% live cells), starting at Bottom and going to Top with a sigmoid shape.

Toxicity prediction with sensitivity and specificity analysis

To assess the predictivity of metabolism-induced compound toxicity, sensitivity and specificity were calculated using IC_{50} values from the DataChip/MetaChip platform, human C_{max} values, and rat LD_{50} values determined by oral administration. Briefly, test compounds that exhibited an IC_{50} value less than or equal to an arbitrary IC_{50} cutoff at a given cell/enzyme condition were categorized as toxic. Similarly, test compounds that exhibited a human C_{max} value or a rat LD_{50} value less than or equal to an arbitrary C_{max} or LD_{50} cutoff were categorized as toxic. Based on the results of IC_{50}, C_{max}, and LD_{50} evaluation, the test compounds were classified into four categories: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). For example, when arbitrary cutoffs of LD_{50} of 300 mg/kg and IC_{50} of 250 μM are used, TP, FP, TN, and FN are determined as follows:

- True positive (TP): $LD_{50} \leq 300$ mg/kg (toxic) and $IC_{50} \leq 250$ μM (toxic)
- False positive (FP): $LD_{50} > 300$ mg/kg (nontoxic) and $IC_{50} \leq 250$ μM (toxic)
- True negative (TN): $LD_{50} > 300$ mg/kg (nontoxic) and $IC_{50} > 250$ μM (nontoxic)
- False negative (FN): $LD_{50} \leq 300$ mg/kg (toxic) and $IC_{50} > 250$ μM (nontoxic)

The predictive performance of the DataChip/MetaChip technology from test compounds was assessed by calculating sensitivity and selectivity as follows:

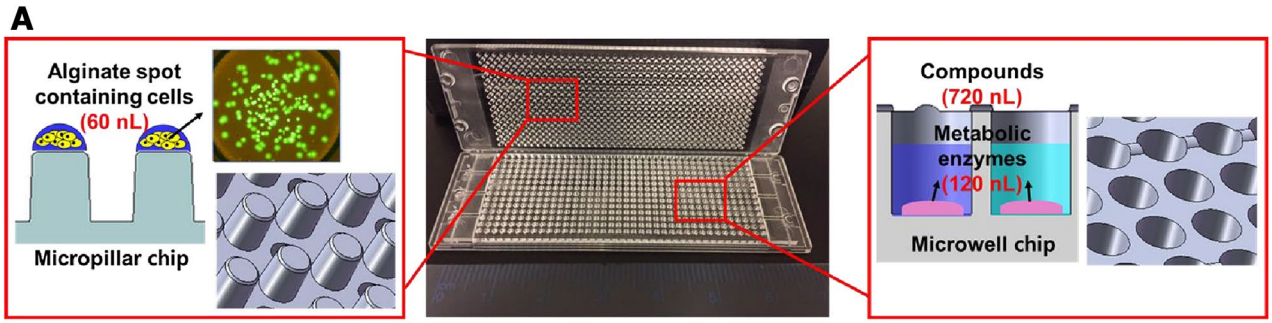
- Sensitivity (%) = [Number of in vitro toxic test compounds (TP)]/[Number of in vivo toxic test compounds (TP + FN)] × 100
- Specificity (%) = [Number of in vitro nontoxic test compounds (TN)]/[Number of in vivo nontoxic test compounds (TN + FP)] × 100
- Overall predictivity (%) = [sensitivity + specificity]/2

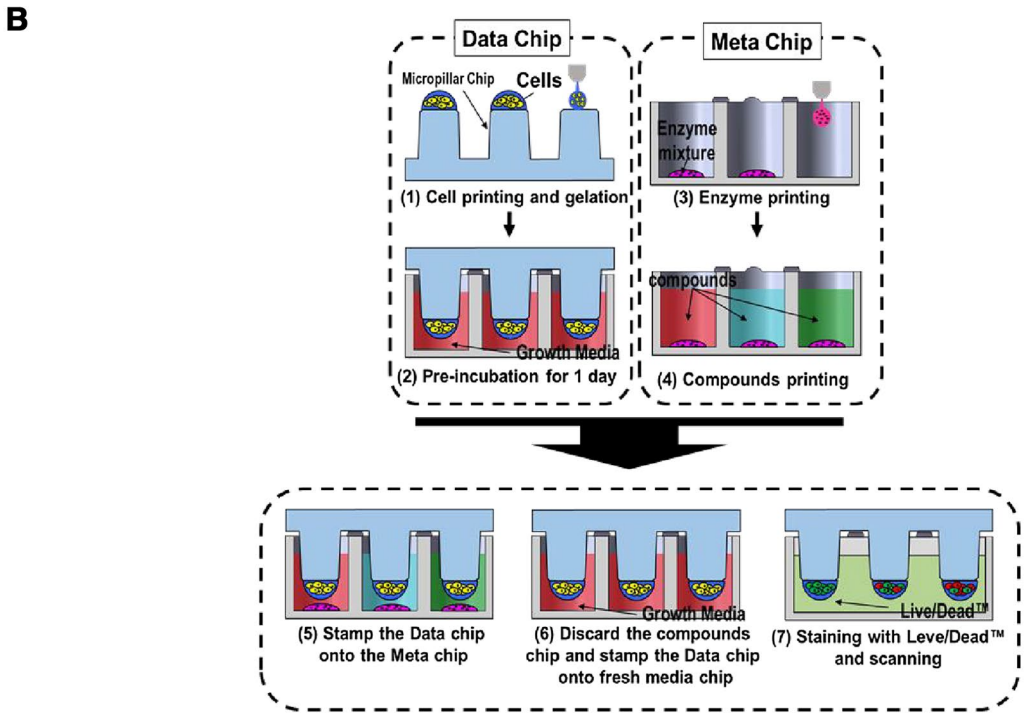
The arbitrary LD_{50} cutoffs were determined based on OECD categories for testing in vivo compound toxicity (OECD 2002). Since identifying optimum cutoffs for obtaining high predictivity are paramount importance, we either varied both in vivo and in vitro cutoffs simultaneously (variable cutoffs) or varied in vitro cutoffs at a fixed in vivo cutoff (fixed cutoffs). For both variable and fixed cutoffs, the acceptance limit for both sensitivity and specificity was set for greater than 50%.

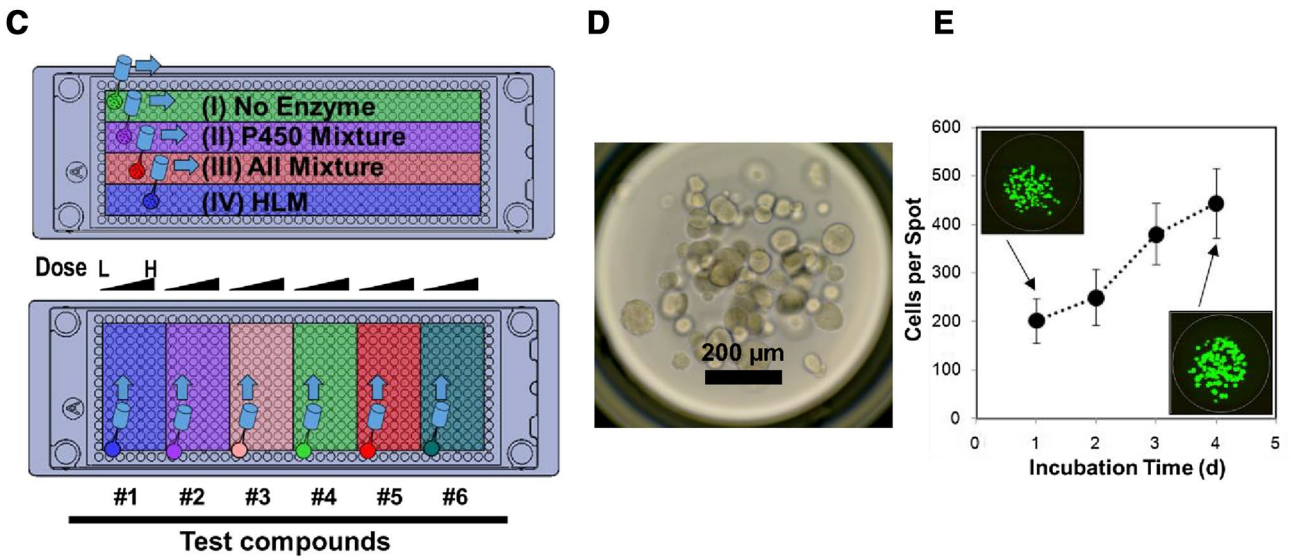
Results

Miniaturized 3D cell culture on the DataChip

The highly versatile DataChip/MetaChip platform is based on micropillar/microwell structures made by plastic injection molding, which is a robust and flexible system for mammalian cell culture, enzymatic reactions, and compound screening (Fig. 1a, b). For 3D cell culture, small Hep3B cell spots were printed on the micropillar chip and strongly attached through robust surface chemistry, as in our previous studies (Lee et al. 2008). The maleic anhydride group in PS-MA was used to covalently attach PLL with amine groups, which led to the negatively charged alginate attaching to positively charged PLL by ionic interactions. $BaCl_2$ was used for the gelation of the alginate matrix on the micropillar chip. After incubating the stamped DataChip onto the microwell chip with RPMI for 3 days, a unique 3D morphology of Hep3B


A
Alginate spot containing cells
(60 nL)
Micropillar chip
Compounds
(720 nL)
Metabolic enzymes
(120 nL)
Microwell chip


B
Data Chip
Micropillar Chip
Cells
(1) Cell printing and gelation
Growth Media
(2) Pre-incubation for 1 day
Meta Chip
Enzyme mixture
(3) Enzyme printing
compounds
(4) Compounds printing
(5) Stamp the Data chip onto the Meta chip
Growth Media
(6) Discard the compounds chip and stamp the Data chip onto fresh media chip
Live/Dead™
(7) Staining with Leve/Dead™ and scanning


C
(I) No Enzyme
(II) P450 Mixture
(III) All Mixture
(IV) HLM
Dose L H
#1 #2 #3 #4 #5 #6
Test compounds
D
200 µm
E
Cells per Spot
600
500
400
300
200
100
0
0 1 2 3 4 5
Incubation Time (d)

◂**Fig. 1** Schematics and photographs of the micropillar/microwell chip with cells and enzymes printed. **a** The structure of the micropillar and microwell chip. The inset shows the scheme of each chip, including cells and drug-metabolizing enzymes (DMEs) printed. **b** Experimental procedures to prepare the DataChip/MetaChip for metabolism-induced toxicity assays. After cell printing on the micropillar chip, the DataChip was sandwiched with the microwell chip containing growth media for 3D cell culture. For the MetaChip, DME mixtures and compounds were printed into the microwell chip sequentially. This was followed by the DataChip sandwiched with the MetaChip and incubated for cytotoxicity assays. **c** The layout of DMEs and compounds printed in the microwell chip to prepare the MetaChip and test metabolism-induced toxicity. Regions I–IV contained no enzyme as a test compound only control, a mixture of human cytochrome P450 isoforms (P450 Mix), a mixture of P450 Mix and human Phase II metabolic enzymes (All Mix), and human liver microsomes (HLM), respectively. Regions 1–6 contained six different compounds in triplicate microwells. From left to right, the concentration of each compound was increased (6 concentrations per compound in each region). **d** Microscopic picture of Hep3B cell growth on the micropillar chip after 3 days. **e** The growth of Hep3B cells encapsulated in alginate spots on the DataChip over time

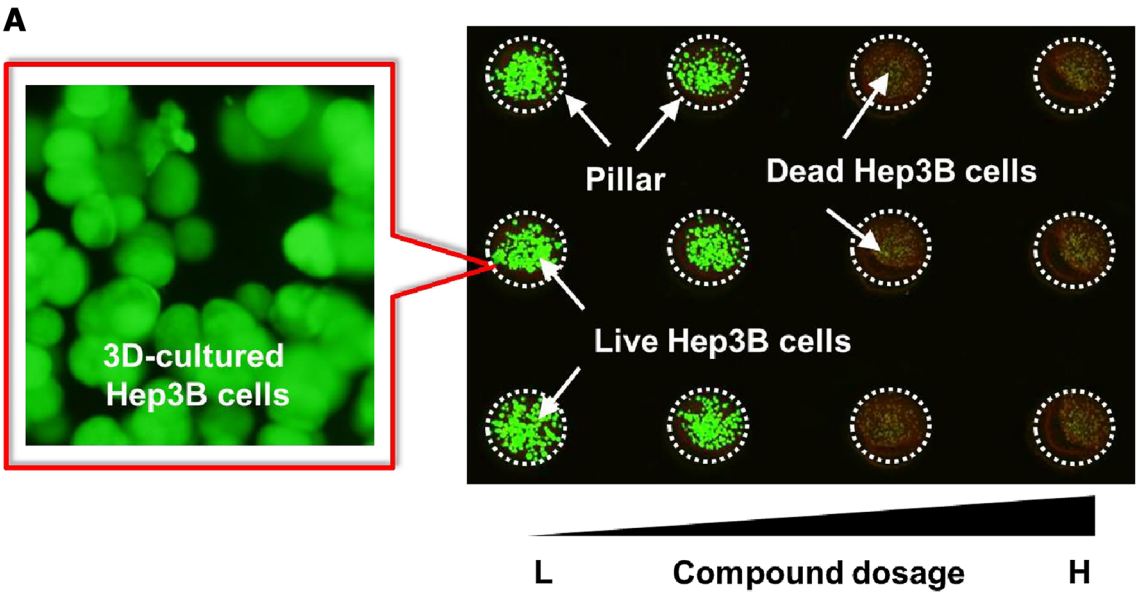
cells in 60 nL alginate spots was observed on the micropillar chip. The Hep3B cells in the spots were stained with calcein AM and ethidium homodimer for assessing live and dead cells and determining cell viability. Based on the calculation of the green fluorescence intensity from stained cell spots on the DataChip, the population of Hep3B cells on each micropillar was very uniform with a 14% coefficient of variability. To determine 3D cell growth quantitatively on the chip, changes in the green fluorescence of Hep3B cells at 3 million cells/mL seeding density (i.e., 180 Hep3B cells per 60 nL spot) were monitored over time. As evidenced by increase in green fluorescence over time, Hep3B cells in alginate spots grew linearly, forming unique 3D spheroids (Fig. 1d, e). The doubling time of Hep3B cells on the chip calculated from green fluorescence intensities measured was approximately 60 h. 3D-cultured Hep3B cells on the chip at the high seeding density grew approximately half as fast as in the 2D counterpart (32 h doubling time), presumably due to the nature of 3D cell culture and limited space available for cell growth within small alginate spots.

Metabolism-induced toxicity assessment in combination with the DataChip and the MetaChip

To mimic human metabolism in high-throughput screening, the DataChip containing 3D-cultured hepatic cells was coupled with the MetaChip containing DMEs and model compounds to generate their metabolites in situ on the chip and assess metabolism-induced toxicity of the compounds in Hep3B cell spheroids. The Hep3B cells within 12×6 mini-arrays were exposed to six different dosages of a compound and four different DME conditions, including no DME control, P450 Mix, All Mix, and HLM (Table 1). Thus, a single Data chip combined with a single MetaChip had the capability to generate 24 dose response curves for 6 compounds and their metabolites from DMEs (Fig. 1c). To study metabolism-induced toxicity with model compounds, IC_{50} values were determined for a parent test compound and its enzyme-generated metabolites against Hep3B cells by staining the DataChip with a Live/Dead® cell viability kit. As demonstrated in Fig. 2a, cell death occurred when Hep3B cells were exposed to high concentrations of acetaminophen in the presence of P450 Mix. The dotted circles represent the boundary of the micropillars onto which Hep3B cell spots were encapsulated. Green dots indicate live cells, whereas dark-red dots represent dead cells. The activity of metabolizing enzymes on the frozen MetaChip was stable for at least 6 months.

The fundamental question we wanted to address in this study was whether or not we could predict in vivo adverse drug responses on the chip platform. Acetaminophen (an analgesic and antipyretic drug) that is known to be hepatotoxic by CYP450 catalysis, and is a major cause of liver failure, was selected as a key model compound and included on each chip to monitor chip-to-chip and day-to-day variability (Supplementary Table 1). As expected, acetaminophen demonstrated metabolism-induced toxicity on the chip, as evidenced by the toxic response of Hep3B cells when they were exposed to P450 Mix and All Mix. This result indicates that the DataChip/MetaChip platform could predict hepatotoxicity caused by active metabolites of acetaminophen (most likely *N*-acetyl-*p*-benzoquinone imine) (Andersson et al. 2011) produced by human liver CYP450 isoforms. The degree of cytotoxicity in the All Mix was less than that of P450 Mix, presumably due to Phase II DMEs included in All Mix, which could reduce toxicity of toxic metabolites generated in situ on the chip through conjugation reactions. Similar results were obtained from Hep3B cells exposed to acetaminophen and HLM (Supplementary Table 1). The degree of cytotoxicity in the HLM was less than that of All Mix, presumably due to larger amounts of Phase II DMEs included in HLM. HLM purchased from BD Biosciences contained approximately 5–10 times large amount of UDP-glucuronosyltransferase (UGT) isoforms compared to All Mix, but did not contain sulfotransferase (SULT), glutathione S-transferase (GST), and *N*-acetyltransferase (NAT) isoforms. Thus, All Mix is a better mimic of human liver.

To further validate the concept and calculate predictivity of in vivo hepatotoxicity, the 22 compounds were printed on the MetaChip in triplicate and tested under the four DME conditions. As a result, several compounds showed augmented toxicity or were detoxified by the DMEs. For example, carbidopa, which is used to manage the symptoms of Parkinson's disease, was activated in P450 Mix and All Mix on the chip presumably due to formation of toxic metabolites (Fig. 2b). Eighteen compounds were found to be toxic against Hep3B cells on the chip, out of which nine



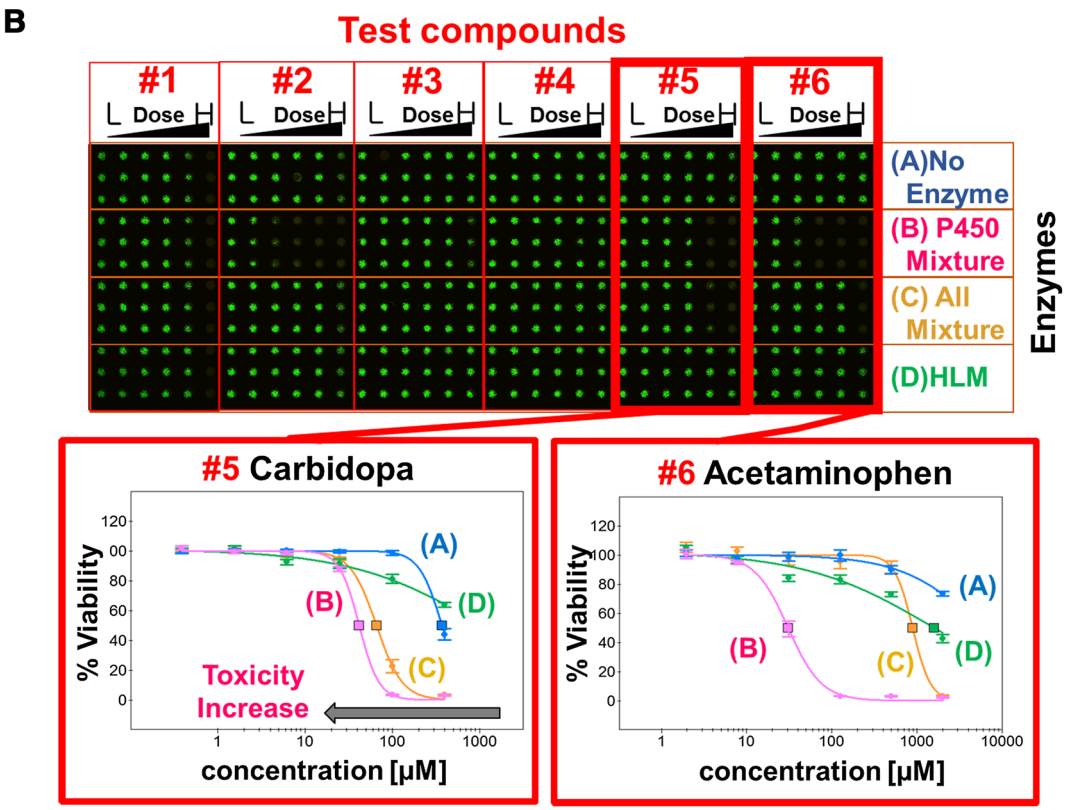


Fig. 2 The scanned images of Hep3B cells stained after compound treatment. **a** Images of Hep3B cells on the DataChip after exposure to the MetaChip containing P450 Mix and different concentrations of troglitazone (3–250 μM). Live cells are stained in green and dead cells are stained in red. **b** The scanned image of the DataChip with Hep3B cells after exposure to the MetaChip containing metabolic variance (no enzyme, P450 Mix, All Mix, and HLM) and compound variance (6 compounds at 6 different concentrations per compound). Representative dose response curves shown were obtained from carbidopa (Compound 5) and acetaminophen (Compound 6). Hep3B cell spots in the 24 distinct regions (each region containing a 3×6 mini-array, triplicates with 6 varied concentrations) were exposed to various combinations of compounds and DMEs. (Color figure online)

compounds showed statistically significant, augmented toxicity in the P450 Mix, indicating that toxic metabolites could be generated on the chip by CYP450 isoforms. Two compounds (carbidopa and oxybendazole) showed augmented toxicity in All Mix compared to their parent compounds. In addition, in situ, on-chip metabolism of flutamide, sulindac, and mefenamic acid led to statistically less toxicity in the All Mix vs. the respective parent compounds (Table 2). The cytotoxicity profiles under varying DME conditions on the chip were well correlated with representative toxic metabolites of the compounds generated.

Prediction of hepatotoxicity in vivo by comparing rat LD_{50} and IC_{50} values from the chip

A common way to predict in vivo toxicity using in vitro data is to compare LD_{50} values with IC_{50} values at arbitrary cutoffs and determine the number of compounds that can be classified into TP, FP, TN, and FN. Thus, we calculated sensitivity and specificity to assess in vivo metabolism-induced hepatotoxicity by comparing rat oral LD_{50} values with IC_{50} values from the DataChip/MetaChip. We initially tested a range of cutoffs to identify an optimum cutoff that can provide high sensitivity and specificity from the chip (Supplementary Fig. 1), thus providing high predictivity of in vivo hepatotoxicity from the IC_{50} values. The acceptance level of sensitivity and specificity was set for greater than 50%.

As a result, we were able to obtain a good in vivo LD_{50}-in vitro IC_{50} correlation at cutoffs of LD_{50} of 300 mg/kg and IC_{50} of 250–450 µM depending on DME conditions tested. Among all DME conditions tested, including no enzyme control in 2D and 3D-cultured Hep3B cells, P450 Mix in 3D, All Mix in 3D and HLM in 3D (Fig. 3), All Mix in 3D appeared to be a better predictor of rat in vivo acute toxicity with 60% sensitivity and 60% specificity (60% overall predictivity). This outcome is likely due to

Table 2 Summary of IC_{50} values obtained from 3D Hep3B cells on the DataChip/MetaChip and 2D Hep3B cell monolayers in 96-well plates

Compounds	DILI category	2D cultured Hep3B cells in 96-well plates	3D cultured Hep3B cells on the Data/MetaChip					1000 × Human C_{max} (µM)	Tox prediction (All Mix vs. DILI)
			Highest dosage used	Enzyme condition used					
		IC_{50} (µM)	(µM)	No Enzyme	P450 Mix	All Mix	HLM		
Acetaminophen	P1	300 ± 0	1200	1200 ± 0	68 ± 30	1010 ± 160	1200 ± 0	138,900	TP
Benzbromarone	P1	58 ± 22	1200	260 ± 80	170 ± 50	330 ± 70	370 ± 20	4300	TP
Fenoterol	P2	300 ± 0	1200	1200 ± 0	290 ± 140	1200 ± 0	1200 ± 0	0.6	FN
Flutamide	P2	45 ± 18	1200	190 ± 10	200 ± 20	270 ± 10	200 ± 10	360	TP
Diclofenac	P2	120 ± 51	1200	780 ± 190	520 ± 100	790 ± 140	860 ± 100	7990	TP
Labetalol	P2	150 ± 34	1200	370 ± 40	200 ± 10	360 ± 70	310 ± 50	2680	TP
Imipramine	P2	5 ± 2	1200	1200 ± 0	1200 ± 0	1200 ± 0	1200 ± 0	87	FN
Phentolamine	P2	48 ± 9	1200	310 ± 50	300 ± 40	380 ± 30	360 ± 30	85	TP
Risperidone	P2	270 ± 60	100	100 ± 0	84 ± 7	100 ± 0	100 ± 0	80	FN
Oxybendazole	O1	24 ± 18	100	71 ± 20	37 ± 10	41 ± 10	88 ± 20	30	TP
Sulindac	P2	300 ± 0	1200	460 ± 90	170 ± 70	1100 ± 160	840 ± 160	31,900	TP
Propranolol	N1	56 ± 11	1200	470 ± 40	3 ± 0.3	530 ± 70	500 ± 50	200	TN
Promazine	N3	21 ± 3	1200	1200 ± 0	1200 ± 0	1200 ± 0	1200 ± 0	490	TN
Trazodone	P2	110 ± 26	400	400 ± 0	250 ± 60	400 ± 0	400 ± 0	5050	FN
Buspirone	N1	170 ± 74	400	400 ± 0	400 ± 0	400 ± 0	400 ± 0	4.9	TN
Carbidopa	N1	46 ± 28	400	370 ± 30	41 ± 3	65 ± 10	400 ± 0	660	FP
Bosentan	P1	55 ± 20	1200	180 ± 30	120 ± 30	230 ± 50	250 ± 40	7430	TP
Chlorpropamide	P2	300 ± 0	1200	1200 ± 0	980 ± 370	1200 ± 0	1200 ± 0	122,800	FN
Phenazopyridine	P2	55 ± 13	800	800 ± 0	350 ± 120	800 ± 0	800 ± 0	124	FN
Estradiol	N3	53 ± 20	1200	1200 ± 0	1200 ± 0	1200 ± 0	1200 ± 0	0.1	TN
Mefenamic acid	P2	260 ± 74	1200	450 ± 20	320 ± 40	770 ± 120	900 ± 140	26,900	TP
Fluoxetine	N3	13 ± 0.3	1200	98 ± 10	89 ± 30	100 ± 10	130 ± 30	97	TN

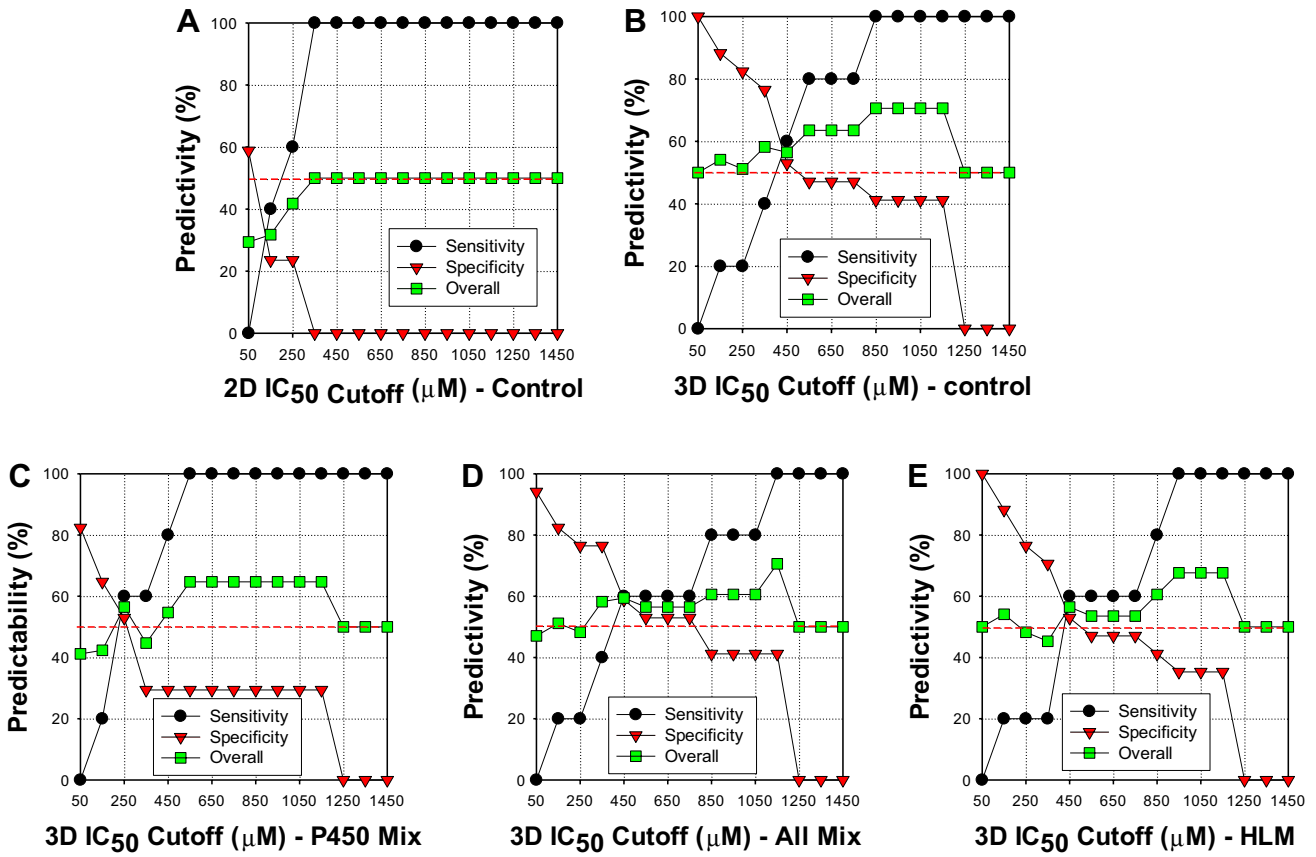


Fig. 3 Calculation of sensitivity and specificity using LD_{50} values at 300 mg/kg cutoff and IC_{50} values at variable cutoffs: LD_{50} compared with **a** IC_{50} from 2D-cultured Hep3B cells in the 96-well plate, **b** IC_{50} from 3D-cultured Hep3B cells without enzymes on the chip, **c** IC_{50} from 3D Hep3B cells with P450 Mix, **d** IC_{50} from 3D Hep3B cells with All Mix, and **e** IC_{50} from 3D Hep3B cells with HLM. To calculate sensitivity, specificity, and overall predictivity, LD_{50} and IC_{50} values were compared at different cutoffs, and then TP, FP, TN, and FN were determined. The red dashed line represents 50% acceptance limit for sensitivity, specificity, and overall predictivity, and all of which have to be above the line to be acceptable. (Color figure online)

the presence of both Phase I and II DMEs in the All Mix, and hence, being more representative of the in vivo situation. Thus, our approach could be applied to predict human acute toxic potential of drug candidates in the liver. Interestingly, the standard in vitro toxicity assessment in 2D without DMEs gave noticeably relatively poor predictivity. In addition, no enzyme control in 3D-cultured Hep3B cells on the chip at cutoffs of LD_{50} of 300 mg/kg and IC_{50} of 450 μM could produce 60% sensitivity and 53% specificity, which indicate that 3D cell culture is superior to 2D cell culture in terms of predicting in vivo hepatotoxicity (Fig. 3a, b). Overall, comparing in vivo LD_{50} values with in vitro IC_{50} values from the chip led to as high as 60% predictivity, which suggests that our DataChip/MetaChip platform could predict in vivo rat hepatotoxicity. Of course, these were mixed species correlations involving human DMEs and a transformed cell line in comparison with rat LD_{50} literature. To achieve greater human predictivity, human in vivo information is needed.

Prediction of hepatotoxicity in vivo by comparing human C_{max} and IC_{50} values from the chip

We proceeded to evaluate a human pharmacokinetic endpoint to address the discrepancy between animals and humans in terms of toxicity evaluation. To this end, we used in vivo human C_{max} values, which represent the maximum allowable concentration of a drug in serum. C_{max} values can be considered an indirect indicator of drug toxicity, as the concentration above C_{max} could cause harmful side effects in the body (Jang et al. 2001). Since C_{max} values are much lower than IC_{50} values, in general, we compared 10-, 100-, and 1000-fold human C_{max} values to IC_{50} values determined using the chip platform to calculate specificity and sensitivity. Overall, use of the 1000-fold human C_{max} values resulted in higher predictivity compared to use of 10- and 100-fold counterparts (Supplementary Fig. 2). Interestingly, the 1000-fold human C_{max} and LD_{50} comparison generated only 50% sensitivity and 65% specificity at 150 variable cutoffs

(Fig. 4a). Not surprisingly, this result indicates that there is a poor correlation between rat in vivo data and human in vivo data. To better understand this outcome and identify optimum cutoffs for in vivo animal and human correlations, we calculated sensitivity and specificity in detail at fixed cutoffs. The highest predictivity (67% sensitivity and 68% specificity) was obtained at 150 µM cutoff for 1000-fold human C_{max} and 200 mg/kg cutoff for rat LD_{50}, which is still lower than the predictivity obtained from the chip platform (Supplementary Fig. 3). As opposed to relatively poor animal predictivity, and the poor correlation of in vitro 2D results (Fig. 4a, b), our chip data outperformed in terms of toxicity prediction under control, P450 Mix, All Mix, and HLM conditions (Fig. 4c–f). In particular, the 1000-fold human C_{max} and All Mix IC_{50} comparison generated remarkable 100% sensitivity and 86% specificity at 50 variable cutoffs (93% overall predictivity). These results indicate that combining 3D hepatic cell culture with drug metabolism on the chip platform could provide better predictivity of hepatotoxicity in vivo as compared to animal and in vitro 2D counterparts. Overall, maximum predictivity achieved at optimum cutoffs by comparing LD_{50} and IC_{50} values with 10, 100, and 1000-fold human C_{max} values is summarized in Fig. 5. As indicated in Fig. 5, the highest sensitivity and specificity was obtained from All Mix compared with 1000-fold human C_{max} values. This outcome implies once again that All Mix containing both Phase I and II DMEs could be a better indicator for predicting hepatotoxicity in vivo. The All Mix was better than the P450 Mix, thereby showing the importance of a full complement of DMEs in predicting hepatotoxicity in vivo. Indeed, the P450 Mix was worse than the no enzyme control.

Prediction of hepatotoxicity in vivo by comparing drug-induced liver injury (DILI) index and IC_{50} values from the chip

The test compounds have been previously categorized according to their ability to cause drug-induced liver injury (DILI) in humans (Xu et al. 2008). For example, our test

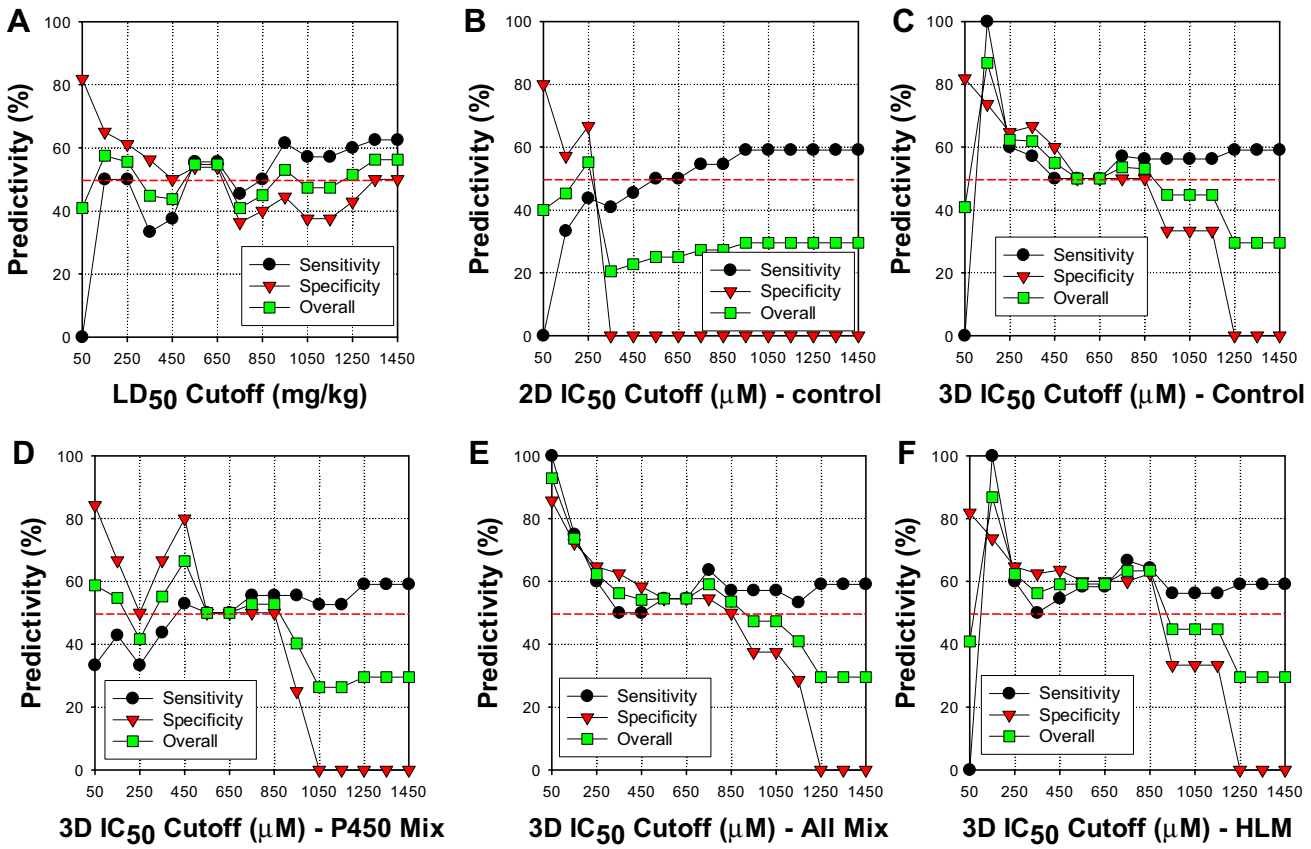


Fig. 4 Calculation of sensitivity and specificity using 1000-fold human C_{max} values at variable cutoffs and LD_{50} and IC_{50} values at variable cutoffs: 1000-Fold C_{max} compared with **a** rat LD_{50}, **b** IC_{50} from 2D-cultured Hep3B cells, **c** IC_{50} from 3D-cultured Hep3B cells without enzymes, **d** IC_{50} from 3D Hep3B cells with P450 Mix, **e** IC_{50} from 3D Hep3B cells with All Mix, and **f** IC_{50} from 3D Hep3B cells with HLM. To calculate sensitivity, specificity, and overall predictivity, 1000-fold human C_{max} values were compared with LD_{50} and IC_{50} values at different cutoffs, and then TP, FP, TN, and FN were determined

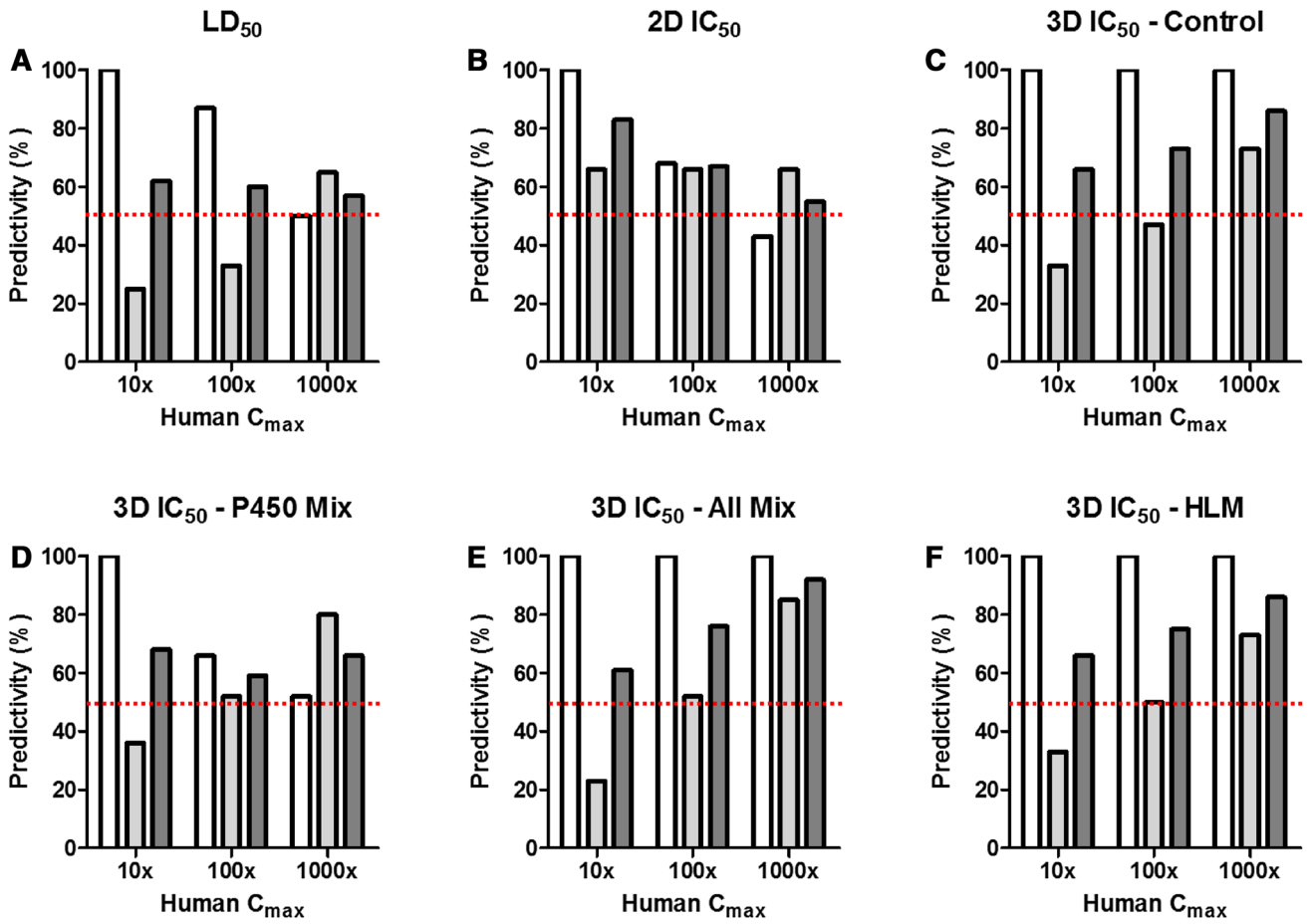


Fig. 5 Maximum predictivity achieved using 10, 100, and 1000-fold human C_{max} with LD_{50} and IC_{50} values: C_{max} compared with **a** rat LD_{50}, **b** IC_{50} from 2D-cultured Hep3B cells (2D IC_{50}—Control), **c** IC_{50} from 3D-cultured Hep3B cells without enzymes on the chip (3D IC_{50}—Control), **d** IC_{50} from 3D Hep3B cells with P450 Mix (3D IC_{50}—P450 Mix), **e** IC_{50} from 3D Hep3B cells with All Mix (3D IC_{50}—All Mix), and **f** IC_{50} from 3D Hep3B cells with HLM (3D IC_{50}—HLM). Color coding of bars indicate as follows: white—sensitivity, light gray—specificity, and black—overall predictivity

compounds fell into one of seven DILI categories: (a) P1 if it is associated with DILI in either animals or humans in a dose-dependent manner, (b) P2 if it is associated with idiosyncratic DILI, (c) O1 if it is hepatotoxic in animals, but untested in humans, (d) O2 if it causes elevated liver enzymes in humans, but generally safe, (e) N3 if it causes sporadic cases of DILI, but generally safe, (f) N2 if it is unknown to cause DILI but known to cause other organ injury, and (g) N1 if it is not known to cause DILI (Table 3). In general, compounds in P1, P2, and O1 categories are considered as hepatotoxic, and a compound O2, N1, N2, and N3 categories is considered as minimally or none hepatotoxic.

To predict DILI from the in vitro chip data, 1000-fold C_{max} values were used as a threshold to differentiate compounds causing DILI from compounds not causing DILI (non-DILI). We hypothesized that compounds potentially causing DILI would be due to toxicity from the parent compounds or their metabolites. In addition, the highest predictivity would be obtained from All Mix. Thus, we defined a DILI-causing compound if it has an IC_{50} value from All Mix < 1000-fold C_{max}. On the other hand, a compound that has an IC_{50} value from All Mix ≥ 1000-fold C_{max} was considered as not causative of DILI. Finally, TP, FP, TN, and FN were determined by comparing DILI/non-DILI outcomes from the chip with a compound's DILI category. For example, if a DILI-causing compound from the chip is in one of the P1, P2, and O1 categories, then it is a TP compound. Similarly, if a compound that does not cause DILI from the chip is in one of the P1, P2, and O1 categories, then it is a FN compound. The sensitivity of the DataChip/MetaChip platform was defined as the ability of the chip platform to predict the P1, P2, and O1 compounds as hepatotoxic [i.e., TP/(TP + FN)]. The specificity of the chip platform was defined as the ability to predict O2, N1, N2, and N3 compounds as nontoxic for DILI [i.e., TN/(TN + FP)].

Table 3 Drug-induced liver injury (DILI) categories sorted by hepatotoxicity levels

Group	Category	Description
Hepatotoxic	P1	Associated with drug-induced liver injury, type 1 (hepatotoxic in animals and/or humans in a dose-dependent manner)
	O1	Hepatotoxic in animals, untested in humans
	P2	Associated with drug-induced liver injury, type 2 (hepatotoxic in animals and/or humans in a dose-independent manner, generally regarded as idiosyncratic hepatotoxicity)
Minimally or not hepatotoxic	O2	Elevated liver enzymes in humans, but generally regarded as safe
	N3	Sporadic cases of liver injury in humans, but generally safe
	N2	Not known to cause liver injury, but known to cause other organ injury
	N1	Not known to cause liver injury

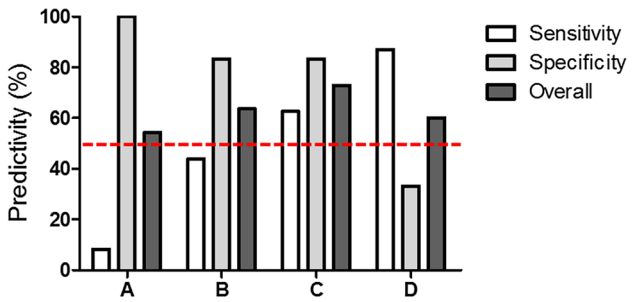


Fig. 6 Predictivity calculated by comparing IC_{50} values from 3D Hep3B cells-All Mix with human C_{max} values and then determining TP, FP, TN, and FN by further comparing with DILI categories: **a** 10-fold C_{max}, **b** 100-fold C_{max}, and **c** 1000-fold C_{max} compared with IC_{50} from 3D Hep3B cells-All Mix. **d** 1000-fold C_{max} compared with IC_{50} from 2D Hep3B cells. To calculate sensitivity, specificity, and overall predictivity, first, human C_{max} values were compared with IC_{50} values from 3D Hep3B cells-All Mix to determine temporary toxicity prediction of the compound. This toxicity prediction was further compared with DILI categories to determine TP, FP, TN, and FN. For example, when the prediction from human C_{max} and IC_{50} is toxic and the compound is classified as hepatoxic by DILI categories, then the prediction of the compound is true positive (TP). After determining TP, TN, FP, and FN for all compounds, sensitivity, specificity, and overall predictivity were calculated accordingly

Out of the 16 P1, P2, and O1 compounds tested, the DataChip/MetaChip platform could predict the DILI potential of ten compounds in the All Mix system. The sensitivity of the chip platform under the All Mix system was 63%. Out of the six N1 and N3 compounds tested, the DataChip/MetaChip could predict five compounds in All Mix that were not hepatotoxic. Thus, the specificity of the chip platform under the All Mix system was 83% (Table 2). Our overall predictivity from DILI index was 73%. The outcome was further compared with overall predictivity from 10-fold and 100-fold C_{max} values compared with All Mix IC_{50} values as well as 1000-fold C_{max} values compared with 2D counterpart IC_{50} values (Fig. 6). Interestingly, predictivity calculated with DILI categories and 1000-fold human C_{max} values in All Mix was much higher than other counterparts including 10-fold, 100-fold human C_{max} in All Mix, as well as 1000-fold human C_{max} with 2D Hep3B cells.

Discussion

Existing in vitro screening technologies for assessing drug metabolism and toxicology lack the ability to provide information on highly predictive metabolism-induced drug toxicity and the necessary throughput for early stage go/no-go decision for lead compounds, and therefore, do not address a critical bottleneck in the drug development process. The goal of this study was to understand whether the new DataChip/MetaChip platform could be used to screen metabolism-induced compound toxicity by correlating our in vitro chip data with known toxicity profiles of test compounds.

There may be several reasons for poor predictivity of adverse drug reactions (ADRs) in humans using current in vitro assays as well as in vivo animal models. First, current in vitro cell models, including 2D hepatoma cell monolayers and sandwiched hepatocytes, may not adequately represent at human liver tissue, thus lacking accurate biochemical and cellular responses in vivo. Several hepatic cell models lack key hepatic properties, including metabolism competence, drug transporters, and cell–cell interactions between hepatocytes and immune cells. Second, current in vitro assays may not provide proper biological circumstances necessary to predict toxicological reactions. For example, defensive pathways such as nuclear factor erythroid 2-related factor 2 (Nrf-2) and nuclear factor-kappa B (NF-kB) can affect the toxic response depending on their levels of activation (Osburn and Kensler 2008; Tak and Firestein 2001). Third, current in vitro assays may not give sufficient information on surrounding cell types, proteins affected by metabolism, and toxicological pathways. Fourth, the number of potentially hepatotoxic compounds identified from individual assays may not be sufficient to capture the broad array of mechanisms leading to in vivo toxicity. In the case of animal models, there are significant cross-species differences between

animals and humans (Shanks et al. 2009). Thus, predicting human toxicity with study outcomes from rats, mice, and rabbits is challenging in general.

To address these issues, there have been several in silico approaches developed in recent years to predict human toxicity directly from in vitro toxicity data. One of good examples of in silico approaches is the in vitro–in vivo correlation (IVIVC) model, which is provided by the U.S. Food and Drug Administration (FDA) with diverse formulations and guidelines for predicting in vitro and in vivo pharmaceutical correlations. It has been used as a surrogate to reduce bioavailability studies of new drugs (Emami 2006; Sakore and Chakraborty 2011). Several research groups reported IVIVC results using a correlation between bioavailability variables and in vitro data (Emara et al. 2000; Mahayni et al. 2000; Balan et al. 2001). Another pioneering approach is the in vitro–in vivo extrapolation (IVIVE) model, which refers to computational simulation to predict in vivo pharmacokinetics (PK) data such as C_{max} values from in vitro experimental data such as IC_{50} values (Yoon et al. 2015; Yoon and Clewell 2016). For example, Johnson et al. used IC_{50} values from 11 drugs to calculate the predictivity of in vivo clearance in neonates, infants, and children (Johnson et al. 2006). In addition, US Environmental Protection Agency (EPA) applied the IVIVE model to grapple with potential human toxicity of environmental toxicants such as risk assessment of ToxCast chemicals in early age children (Wetmore et al. 2014). Since both IVIVC and IVIVE models require more predictive in vitro data to better predict in vivo outcomes, there have several attempts made to incorporate metabolism competence in their in vitro assays and consider metabolic stability and metabolism-induced toxicity of drugs (Yoon et al. 2015; Yoon and Clewell 2016). The biotransformation of drugs can produce metabolites that have different toxicity profiles from their parent compounds (Combes et al. 2002). In particular, CYP450 reactions can generate metabolites that are more reactive and can induce toxicity through a variety of mechanisms (e.g., covalent binding to macromolecules or contributing to oxidative damage) (Costas 2008). Differences in individual responses to compounds are common among the human population and can be attributed to genetic variations that limit the expression or activity of certain DMEs (Astrid et al. 2007). Thus, determining which enzymes activate or deactivate a compound is essential to understand population variances in drug and drug candidate toxicity.

In recognition of these issues, we developed the new DataChip/MetaChip platform to incorporate 3D hepatic cell culture as well as metabolism competence with an array of DMEs, which in turn can decipher metabolism-induced compound toxicity. In our results, the DataChip/MetaChip identified 18 compounds whose reactions with Phase I and II DMEs resulted in IC_{50} values significantly lower than that of the parent compounds, indicating that these compounds were directly metabolized and activated by the Phase I and/or II DMEs on the chip. In addition, 11 compounds were detoxified by Phase II DMEs. For example, metabolism of sulindac resulted in increased IC_{50} in the All Mix relative to that with the No Enzyme control and P450 Mix, suggesting that this compound was transformed by Phase II DMEs. Similar results were obtained from compounds exposed to HLMs. By simply comparing human C_{max} values and All Mix IC_{50} values from the chip at different cutoffs, we achieved 100% sensitivity, 86% specificity, and 93% overall predictivity. This outcome implies that our DataChip/MetaChip platform could provide high predictivity of human hepatotoxicity as compared to animal and in vitro 2D counterparts. In addition, our in vitro chip data might be used in IVIVC and IVIVE models to provide more predictive information on metabolism competence. In summary, the chip technology could be used at early stages of drug development, not to predict the extent and nature of all possible in vivo toxic effects, but rather to estimate the risk of failure if a new lead compound is transformed into metabolites that can be toxic to cells and potentially to humans.

Conclusions

Drugs react and form metabolites in the body via various metabolic pathways. Metabolites formed from Phase I and II DME reactions can cause ADRs, which may not be detected easily in animal models due difference in genetic makeups between animals and humans. Predictivity of in vivo hepatotoxicity of 22 test compounds obtained from the DataChip/MetaChip containing 3D-cultured Hep3B cells and Phase I and II DMEs demonstrated that the chip platform could provide a better correlation with in vivo human C_{max} values compared to in vivo rat LD_{50} data. This result is presumably due to the in situ generation of compound metabolites on the MetaChip for more accurate assessment of metabolism-induced toxicity and the 3D culture environment on the DataChip that may better mimic the tissue architecture and enhance functionality of the hepatic cells. The DataChip/MetaChip platform could ultimately be tailored to accommodate an individual's DME inventory and be used in conjunction with various high-content imaging assays to provide specific mechanisms of metabolic toxicity profiles in different populations of individuals as a component of broader precision medicine. With more in vitro IC_{50} data from the chip platform for further validation, the DataChip/MetaChip platform could represent a promising, high-throughput microscale alternative to conventional in vitro multi-well plate platforms and may create new opportunities for rapid and inexpensive assessment of human toxicology at very early phases of drug development.

Acknowledgements We acknowledge support from the National Institute of Environmental Health Sciences (ES018022, ES012619, and ES025779) and National Science Foundation (IIP-0740592). This work was partly supported by Samsung Electro-Mechanics Co. (SEMCO), Ltd. The authors are grateful to Dr. Byeong-Cheon Koh (former Executive Vice President) and members of the cell chip research group in SEMCO for helpful suggestions and assistance with chip fabrication.

Compliance with ethical standards

Conflict of interest The research was partly supported by Samsung Electro-Mechanics Co. (SEMCO). Thus, the authors declare that there might be potential conflict of interest.

References

Andersson DA, Gentry C, Alenmyr L, Killander D, Lewis SE, Andersson A, Bucher B, Galzi JL, Sterner O, Bevan S, Högestätt ED, Zygmunt PM (2011) TRPA1 mediates spinal antinociception induced by acetaminophen and the cannabinoid Δ(9)-tetrahydrocannabiorcol. Nat Commun 22(2):551. https://doi.org/10.1038/ncomms1559

Astrid S, Helmut S, Roland S (2007) Drug metabolism as catalyzed by human cytochrome P450 systems. In: Metal ions in life science. Volume 3: the ubiquitous roles of cytochrome P450 proteins. Wiley Online Library, England

Balan G, Timmins P, Greene DS, Marathe PH (2001) In vitro–in vivo correlation (IVIVC) models for metformin after administration of modified-release (MR) oral dosage forms to healthy human volunteers. J Pharm Sci 90:1176–1185. https://doi.org/10.1002/jps.1071

Bale SS, Moore L, Yarmush M, Jindal R (2016) Emerging in vitro liver technologies for drug metabolism and inter-organ interactions. Tissue Eng Part B Rev 22:383–394. https://doi.org/10.1089/ten.teb.2016.0031

Brandon EF, Raap CD, Meijerman I, Beijnen JH, Schellens JH (2003) An update on in vitro test methods in human hepatic drug biotransformation research: pros and cons. Toxicol Appl Pharmacol 189:233–246. https://doi.org/10.1016/S0041-008X(03)00128-5

Bui PH, Quesada A, Handforth A, Hankinson O (2008) The Mibefradil derivative NNC55-0396, a specific T-type calcium channel antagonist, exhibits less CYP3A4 inhibition than mibefradil. Drug Metab Dispos 36:1291–1299. https://doi.org/10.1124/dmd.107.020115

Combes R, Balls M, Bansil L, Barratt M, Bell D, Botham P, Broadhead C, Clothier R, George E, Fentem J, Jackson M, Indans I, Loizou G, Navaratnam V, Pentreath V, Phillips B, Stemplewski H, Stewart J (2002) An assessment of progress in the use of alternatives in toxicity testing since the publication of the report of the second FRAME Toxicity Committee (1991). Altern Lab Anim 30:365–406

Costas I (2008) Cytochromes P450: role in the metabolism and toxicity of drugs and other xenobiotics. Royal Society of Chemistry, Cambridge

DiMasi JA, Grabowski HG (2012) R&D costs and returns to new drug development: a review of the evidence. Oxford University Press, Oxford, pp 21–46

Emami J (2006) In vitro–in vivo correlation: from theory to applications. J Pharm Pharm Sci 9:169–189

Emara LH, El-Menshavi BS, Estefan MY (2000) In vitro-in vivo correlation and comparative bioavailability of vincamine in prolonged-release preparation. Drug Dev Ind Phar 26:243–251. https://doi.org/10.1081/DDC-100100352

Gómez-Lechón MJ, Donato MT, Castell JV, Jover R (2004) Human hepatocytes in primary culture: the choice to investigate drug metabolism in man. Curr Drug Metab 5:443–462. https://doi.org/10.2174/1389200043335414

Gustafsson F, Foster AJ, Sarda S, Bridgland-Taylor MH, Kenna JG (2014) A correlation between the in vitro drug toxicity of drugs to cell lines that express human P450s and their propensity to cause liver injury in humans. Toxicol Sci 137:189–211. https://doi.org/10.1093/toxsci/kft223

Hariparsad N, Sane RS, Strom SC, Desai PB (2006) In vitro methods in human drug biotransformation research: implications for cancer chemotherapy. Toxicol In Vitro 20:135–153. https://doi.org/10.1016/j.tiv.2005.06.049

Hewitt NJ, Lechón MJ, Houston JB, Hallifax D, Brown HS, Maurel P, Kenna JG, Gustavsson L, Lohmann C, Skonberg C, Guillouzo A, Tuschl G, Li AP, LeCluyse E, Groothuis GM, Hengstler JG (2007) Primary hepatocytes: current understanding of the regulation of metabolic enzymes and transporter proteins, and pharmaceutical practice for the use of hepatocytes in metabolism, enzyme induction, transporter, clearance, and hepatotoxicity studies. Drug Metab Rev 39:159–234. https://doi.org/10.1080/03602530601093489

Huch M, Gehart H, van Boxtel R, Hamer K, Blokzijl F, Verstegen MM, Ellis E, van Wenum M, Fuchs SA, de Ligt J, van de Wetering M, Sasaki N, Boers SJ, Kemperman H, de Jonge J, Ijzermans JN, Nieuwenhuis EE, Hoekstra R, Strom S, Vries RR, van der Laan LJ, Cuppen E, Clevers H (2015) Long-term culture of genome-stable bipotent stem cells from adult human liver. Cell 160:299–312. https://doi.org/10.1016/j.cell.2014.11.050

Hughes JP, Rees S, Kalindjian SB, Philpott KL (2011) Principles of early drug discovery. Br J Pharmacol 162:1239–1249. https://doi.org/10.1111/j.1476-5381.2010.01127.x

Jang GR, Harris RZ, Lau DT (2001) Pharmacokinetics and its role in small molecule drug discovery research. Med Res Rev 21:382–396. https://doi.org/10.1002/med.1015

Johnson TN, Rostami-Hodjegan A, Tucker GT (2006) Prediction of the clearance of eleven drugs and associated variability in neonates, infants and children. Clin Pharmacokinet 45:931–956. https://doi.org/10.2165/00003088-200645090-00005

Kennedy JP, Williams L, Bridges TM, Daniels RN, Weaver D, Lindsley CW (2008) Application of combinatorial chemistry science on modern drug discovery. J Comb Chem 10:345–354. https://doi.org/10.1021/cc700187t

Lee MY, Dordick JS (2006) High-throughput human metabolism and toxicity analysis. Curr Opin Biotechnol 17:619–627. https://doi.org/10.1016/j.copbio.2006.09.003

Lee MY, Park CB, Dordick JS, Clark DS (2005) Metabolizing enzyme toxicology assay chip (MetaChip) for high-throughput microscale toxicity analyses. Proc Natl Acad Sci USA 102:983–987. https://doi.org/10.1073/pnas.0406755102

Lee MY, Kumar RA, Sukumaran SM, Hogg MG, Clark DS, Dordick JS (2008) Three-dimensional cellular microarray for high-throughput toxicology assays. Proc Natl Acad Sci USA 105:59–63. https://doi.org/10.1073/pnas.0708756105

Lee DW, Yi SH, Jeomg SH, Ku B, Kim J, Lee MY (2013) Plastic Pillar inserts for three-dimensional(3D) cell cultures in 96-well plates. Sensors Actuators B 177:78–85. https://doi.org/10.1016/j.snb.2012.10.129

Lee DW, Choi YS, Seo YJ, Lee MY, Jeon SY, Ku B, Kim S, Yi SH, Nam DH (2014a) High-throughput screening (HTS) of anticancer drug efficacy on a micropillar/microwell chip platform. Anal Chem 86:535–542. https://doi.org/10.1021/ac402546b

Lee DW, Lee MY, Ku B, Yi SH, Ryu JH, Jeon R, Yang M (2014b) Application of the DataChip/MetaChip technology for the evaluation of ajoene toxicity in vitro. Arch Toxicol 88:283–290. https://doi.org/10.1007/s00204-013-1102-9

Mahayni H, Rekhi GS, Uppoor RS, Marroum P, Hussain AS, Augsburger LL, Eddington ND (2000) Evaluation of external predictability of an in vitro–in vivo correlation for an extended-release formulation containing metoprolol tartrate. J Pharm Sci 89:1354–1361. https://doi.org/10.1002/1520-6017(200010)89:10<1354::AID-JPS13>3.0.CO;2-P

Masubuchi Y, Kano S, Horie T (2006) Mitochondrial permeability transition as a potential determinant of hepatotoxicity of antidiabetic thiazolidinediones. Toxicology 222:233–239. https://doi.org/10.1016/j.tox.2006.02.017

OECD (2002) Guidelines for the testing of chemicals/OECD series on testing and assessment harmonised integrated classification system for human health and environmental hazards of chemical substances and mixtures. OECD

Osburn WO, Kensler TW (2008) Nrf2 signaling: an adaptive response pathway for protection against environmental toxic insults. Mutat Res 659:31–39. https://doi.org/10.1016/j.mrrev.2007.11.006

Reddy VB, Karanam BV, Gruber WL, Wallace MA, Vincent SH, Franklin RB, Baillie TA (2005) Mechanistic studies on the metabolic scission of thiazolidinedione derivatives to acyclic thiols. Chem Res Toxicol 18:880–888. https://doi.org/10.1021/tx0500373

Sakore S, Chakraborty B (2011) In vitro–in vivo correlation (IVIVC): a strategic tool in drug development. J Bioequiv Availab S3. https://doi.org/10.4172/jbb.S3-001

Schadt EE, Friend SH, Shaywitz DA (2009) A network view of disease and compound screening. Nat Rev Drug Discov 8:286–295. https://doi.org/10.1038/nrd2826

Shanks N, Greek R, Greek J (2009) Are animal models predictive for humans? Philos Ethics Humanit Med 15:4:2. https://doi.org/10.1186/1747-5341-4-2

Shukla SJ, Huang R, Austin CP, Xia M (2010) The future of toxicity testing: a focus on in vitro methods using a quantitative high-throughput screening platform. Drug Discov Today 15:997–1007. https://doi.org/10.1016/j.drudis.2010.07.007

Sivaraman A, Leach JK, Townsend S, Iida T, Hogan BJ, Stolz DB, Fry R, Samson LD, Tannenbaum SR, Griffith LG (2005) A microscale in vitro physiological model of the liver: predictive screens for drug metabolism and enzyme induction. Curr Drug Metab 6:569–591. https://doi.org/10.2174/138920005774832632

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

Tak PP, Firestein GS (2001) NF-kappaB: a key role in inflammatory diseases. J Clin Invest 107:7–11. https://doi.org/10.1172/JCI11830

Watanabe T, Shibata N, Westerman KA, Okitsu T, Allain JE, Sakaguchi M, Totsugawa T, Maruyama M, Matsumura T, Noguchi H, Yamamoto S, Hikida M, Ohmori A, Reth M, Weber A, Tanaka N, Leboulch P, Kobayashi N (2003) Establishment of immortalized human hepatic stellate scavenger cells to develop bioartificial livers. Transplantation 75(11):1873–1880. https://doi.org/10.1097/01.TP.0000064621.50907.A6

Westra IM, Mutsaers HA, Luangmonkong T, Hadi M, Oosterhuis D, de Jong KP, Groothuis GM, Olinga P (2016) Human precision-cut liver slices as a model to test antifibrotic drugs in the early onset of liver fibrosis. Toxicol In Vitro 35:77–85. https://doi.org/10.1016/j.tiv.2016.05.012

Wetmore BA, Allen B, Clewell HJ, Parker T, Wambaugh JF, Almond LM, Thomas RS (2014) Incorporating population variability and susceptible subpopulations into dosimetry for high-throughput toxicity testing. Toxicol Sci 142:210–224. https://doi.org/10.1093/toxsci/kfu169

Xu JJ, Henstock PV, Dunn MC, Smith AR, Chabot JR, de Graaf D (2008) Cellular imaging predictions of clinical drug-induced liver injury. Toxicol Sci 105:97–105. https://doi.org/10.1093/toxsci/kfn109

Yoon M, Clewell HJ 3rd (2016) Addressing early life sensitivity using physiologically based pharmacokinetic modeling and in vitro to in vivo extrapolation. Toxicol Res 32:15–20. https://doi.org/10.5487/TR.2016.32.1.015

Yoon M, Kedderis GL, Yan GZ, Clewell HJ 3rd (2015) Use of in vitro data in developing a physiologically based pharmacokinetic model: carbaryl as a case study. Toxicology 332:52–66. https://doi.org/10.1016/j.tox.2014.05.006