Archives of Toxicology (2023) 97:1163–1175
https://doi.org/10.1007/s00204-023-03461-z

GENOTOXICITY AND CARCINOGENICITY

High-throughput micronucleus assay using three-dimensional HepaRG spheroids for in vitro genotoxicity testing

Ji-Eun Seo[1] · Xilin Li[1] · Yuan Le[1] · Nan Mei[1] · Tong Zhou[2] · Xiaoqing Guo[1]

Received: 29 December 2022 / Accepted: 16 February 2023 / Published online: 27 February 2023

Abstract

The in vitro micronucleus (MN) assay is a component of most test batteries used in assessing potential genotoxicity. Our previous study adapted metabolically competent HepaRG cells to the high-throughput (HT) flow-cytometry-based MN assay for genotoxicity assessment (Guo et al. in J Toxicol Environ Health A 83:702–717, 2020b, https://doi.org/10.1080/15287394.2020.1822972). We also demonstrated that, compared to HepaRG cells grown as two-dimensional (2D) cultures, 3D HepaRG spheroids have increased metabolic capacity and improved sensitivity in detecting DNA damage induced by genotoxicants using the comet assay (Seo et al. in ALTEX 39:583–604, 2022, https://doi.org/10.14573/altex.22011212022). In the present study, we have compared the performance of the HT flow-cytometry-based MN assay in HepaRG spheroids and 2D HepaRG cells by testing 34 compounds, including 19 genotoxicants or carcinogens and 15 compounds that show different genotoxic responses in vitro and in vivo. 2D HepaRG cells and spheroids were exposed to the test compounds for 24 h, followed by an additional 3- or 6-day incubation with human epidermal growth factor to stimulate cell division. The results demonstrated that HepaRG spheroids showed generally higher sensitivity in detecting several indirect-acting genotoxicants (require metabolic activation) compared to 2D cultures, with 7,12-dimethylbenzanthracene and *N*-nitrosodimethylamine inducing higher % MN formation along with having significantly lower benchmark dose values for MN induction in 3D spheroids. These data suggest that 3D HepaRG spheroids can be adapted to the HT flow-cytometry-based MN assay for genotoxicity testing. Our findings also indicate that integration of the MN and comet assays improved the sensitivity for detecting genotoxicants that require metabolic activation. These results suggest that HepaRG spheroids may contribute to New Approach Methodologies for genotoxicity assessment.

Keywords HepaRG · 3D spheroids · Micronucleus assay · Genotoxicity · Benchmark dose

Introduction

Batteries of genetic toxicity tests are employed to evaluate the potential of regulated substances for inducing cancer and other diseases (Cimino 2006). The in vitro micronucleus (MN) assay is commonly included in the test batteries, including those recommended in guidances from the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH), the Organization for Economic Co-operation and Development (OECD), and the International Cooperation on Harmonization of Technical Requirements for Registration of Veterinary Medicinal Products (VICH) (FDA 2000; ICH 2013; OECD 2015; VICH 2014). The MN assay detects acentric chromosome fragments or whole chromosomes that are not incorporated into daughter nuclei during anaphase of cell division (OECD 2016) and thus serves to evaluate the clastogenic and aneugenic potential of test substances. Extensive data indicate that the MN assay is robust and effective in measuring genotoxicity in a number of rodent and human cell lines (i.e., L5178Y, V79, CHO, and TK6), while data also suggest that p53 status, genetic stability, DNA repair capacity, metabolic activation and species differences play roles in the responses measured

✉ Ji-Eun Seo
Jieun.Seo@fda.hhs.gov

✉ Xiaoqing Guo
Xiaoqing.Guo@fda.hhs.gov

[1] Division of Genetic and Molecular Toxicology, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, AR 72079, USA

[2] Center for Veterinary Medicine, U.S. Food and Drug Administration, Rockville, MD 20855, USA

by the assay (OECD 2016). In order to improve the value of MN findings, recommendations have been made to use TP53-competent, human-derived cell lines for the in vitro MN and chromosomal aberration tests (Pfuhler et al. 2011). The human hepatoma HepaRG cell line addresses both these recommendations.

Previously, we compared MN induction by 12 genotoxic carcinogens in HepaRG and HepG2 cells grown as two-dimensional (2D) or attached cell cultures and demonstrated that HepaRG cells were more sensitive than HepG2 cells in detecting genotoxicants that require metabolic activation (indirect acting) (Guo et al. 2020b). 2D HepaRG cells, however, showed similar or slightly less sensitivity than HepG2 cells for detecting MN formation induced by seven direct-acting genotoxicants. Recently, we optimized a 3D HepaRG spheroid model which maintains a stable phenotype and high levels of albumin secretion and cytochrome P450 (CYP) enzyme activity for at least 30 days, indicating exceptional liver functionality and metabolic capacity (Seo et al. 2022). These 3D cultures also displayed mitotic figures and were positive for a proliferation marker (Ki67) by histological and immunofluorescent staining, suggesting their potential use in genotoxicity assays that require cell proliferation, such as the MN assay and mutation assays.

The liver-spheroid-based MN assays are only at a very early stage of development. In contrast, the MN assay conducted in the 3D reconstructed human skin tissue models is considered sufficiently validated to serve as a 'tier 2' assay for dermally exposed compounds that are positive in standard in vitro genotoxicity assays (Pfuhler et al. 2020, 2021). Spheroids of the human hepatoblastoma HepG2 cell line constructed by a hanging-drop method were used for the cytokinesis-block MN (CBMN) assay (Shah et al. 2018). Indirect-acting carcinogens, i.e., aflatoxin B1 (AFB1), benzo[a]pyrene (B[a]P), and 2-amino-1-methyl-6-phenylimidazo(4,5-b) pyridine (PhIP), induced much higher MN frequencies at lower concentrations in 3D HepG2 spheroids compared to 2D HepG2 cells, while the direct-acting genotoxicant methyl methanesulfonate (MMS) induced similar levels of MN formation in both 2D and 3D HepG2 models (Conway et al. 2020; Shah et al. 2018). This study, however, also concluded that 3D HepaRG spheroids were unsuitable for conducting the MN assay due to a low induction rate of binucleated cells at the 24-h and 30-h timepoints (Conway et al. 2020). To overcome this issue, Rose et al. dissociated the HepaRG spheroids with trypsin following the treatment and cultured the cells in medium supplemented with 50 ng/ml recombinant human epidermal growth factor (rhEGF) for additional 72 h (Rose et al. 2022). With this additional incubation, 24-h or 14-day exposures to three indirect-acting carcinogens, AFB1, B[a]P, and cyclophosphamide (CPA), all induced positive responses in the CBMN assay.

High-throughput (HT) genotoxicity assays have the advantage of efficiently testing the growing numbers of compounds to which humans are exposed in occupational and environmental exposures, as well as in their food, drugs, and consumer products throughout their daily lives (Chen et al. 2022). The inventory of the U.S. Toxic Substances Control Act (TSCA) contains 86,631 chemicals, of which 42,039 are active and the number of compounds is increasing over time (EPA 2022). Although they have shown potential benefits in terms of sensitivity and accuracy (Barranger and Le Hegarat 2022; Conway et al. 2020), compared to assays conducted with 2D cell cultures, 3D genotoxicity assays generally are considered difficult to perform and have relatively low throughput (Pfuhler et al. 2020), making their use in screening large numbers of test substances problematic.

In order to address the throughput issue, we have conducted a 3D MN assay using 384-well ultra-low attachment (ULA) plates for HepaRG spheroid formation. Such plates have been reported to facilitate the development of homogeneous spheroids in a HT manner (Ivanov et al. 2014). In addition, we have adapted the flow-cytometry-based MN assay to the spheroids in order to automate MN scoring. We previously tested a total of 34 compounds in HepaRG spheroids and 2D cultures using the CometChip assay (Seo et al. 2022); these same 34 test substances were evaluated in this current study using 2D and 3D HepaRG MN assays. Positive responses in these assays were quantified using benchmark dose (BMD) potency ranking (Wills et al. 2016). The BMDs and their upper and lower bounds of the 90% confidence intervals (CIs; BMDU and BMDL, respectively) were used for quantitative comparison between the MN concentration–response data generated from 2D and 3D HepaRG models. In addition, data combined from the MN and comet assays were used to evaluate the genotoxic potency of positive compounds and compare the sensitivity of 2D and 3D HepaRG models for detecting genotoxicants or carcinogens.

Materials and methods

Test chemicals

Thirty-four test chemicals, including 8 direct-acting and 11 indirect-acting genotoxicants or carcinogens as well as 15 compounds that show different genotoxic responses in vitro and in vivo, were used for the study (Table 1). Except for 2-amino-3-methyl-3H-imidazo[4,5-f]quinoline (IQ) and PhIP, which were obtained from Toronto Chemical Research (Toronto, Canada), the test substances were purchased from Sigma-Aldrich (St. Louis, MO).

Table 1 Comparison of cytotoxicity and micronuclei (MN) responses in 2D and 3D HepaRG cultures

Compound		Highest conc (μM)[a]		Relative survival (%)[b]; [%ATP][c]		MN (%)[d]		Fold increase[e]		Outcome[f]	
		2D	3D	2D	3D	2D	3D	2D	3D	2D	3D
Genotoxicants/carcinogens											
Direct-acting genotoxicants/carcinogens	4-NQO	5	10	93.0	53.0 [65.4]	2.0	2.0	2.7*	2.1*	+	+
	$CdCl_2$	10	20	66.5	87.0 [51.7]	0.7	0.9	0.8	1.0	−	−
	Cisplatin	40	12.5	57.0	60.7 [63.8]	11.6	3.5	11.6*	4.0*	+++	+
	Colchicine	2.6	1	52.0	49.5 [63.0]	5.1	2.2	6.5*	2.2*	++	+
	ENU	2048	3200	55.5	60.0 [56.3]	2.0	2.1	1.9*	2.4*	±	+
	Etoposide	25	25	65.7	51.6 [60.5]	13.9	9.8	19.8*	12.1*	+++	+++
	HQ	160	300	99.2	78.2 [53.8]	1.4	1.3	1.8*	1.4	±	−
	MMS	500	500	53.7	55.0 [51.3]	6.7	4.3	10.4*	4.6*	+++	+
Indirect-acting genotoxicants/carcinogens	2,4-DAT	10,000	5000	70.0	62.1 [76.7]	0.4	1.0	0.9	1.2	−	−
	2-AAF	1000	1000	69.4	53.1 [41.9]	1.3	1.7	1.6	2.1*	−	+
	Acrylamide	5000	5000	74.6	47.3 [53.3]	0.6	1.4	0.8	1.7	−	−
	AFB1	2	1	45.9	50.2 [62.1]	4.9	3.6	6.2*	4.4*	++	+
	B[a]P	50	100	46.8	58.8 [41.9]	2.6	3.1	3.6*	3.5*	+	+
	CPA	10,000	5000	65.6	57.8 [59.0]	5.0	4.9	6.7*	5.4*	++	++
	DMBA	500	125	64.2	45.9 [49.5]	2.6	5.4	3.6*	5.9*	+	++
	NDMA	10,000	5000	67.8	44.1 [50.8]	2.2	5.0	3.0*	5.3*	+	++
	IQ	400	500	59.9	77.1 [83.2]	1.1	1.1	1.5	1.4	−	−
	PhIP	1000	1000	75.8	49.4 [35.5]	1.8	1.3	2.7*	1.6*	+	±
	Styrene	10,000	10,000	82.1	87.5 [86.5]	1.2	0.9	1.5	1.1	−	−
Compounds that show different genotoxic responses in vitro and in vivo											
In vitro (+) but in vivo (−), and Ames (+)	3-MCPD	10,000	10,000	86.9	76.3 [67.6]	1.3	1.6	1.2	1.8	−	−
	DFPBA	1000	1000	58.8	83.3 [51.5]	1.1	1.2	1.0	1.2	−	−
	EDAC	200	300	73.7	66.5 [65.4]	1.5	1.5	1.5	1.5	−	−
	HOPO	800	750	63.3	77.3 [61.4]	1.8	1.3	1.7	1.4	−	−
	PBA	5000	5000	56.5	68.3 [44.2]	1.2	1.2	1.1	1.3	−	−
In vitro (+) but in vivo (−), and Ames (−)	4-Nitrophenol	400	500	60.0	62.6 [46.9]	1.0	1.9	1.0	1.7	−	−
	Ethyl acrylate	5000	7500	71.7	60.5 [59.5]	1.1	2.0	1.1	1.9	−	−
	Phthalic anhydride	10,000	10,000	82.7	72.9 [82.6]	1.4	1.3	1.4	1.2	−	−
	Sodium xylenesulfonate	10,000	10,000	87.7	77.1 [77.2]	1.2	1.1	1.2	1.0	−	−
	TBHQ	320	375	67.7	71.0 [44.0]	1.8	1.6	2.0	1.5	−	−
In vitro (−) but in vivo (+)	1,4-Dioxane	10,000	10,000	94.5	65.4 [89.8]	1.2	0.8	1.2	0.7	−	−
	Dicyclanil	2000	3000	50.9	51.6 [47.8]	1.6	1.5	1.7	1.6	−	−
	DMTP	1000	1000	48.2	73.2 [67.2]	1.2	1.2	1.1	1.1	−	−
	Estragole	5000	10,000	64.7	97.3 [82.0]	1.4	1.0	1.4	1.1	−	−
	LMG	1000	1000	81.7	65.1 [55.0]	1.4	1.6	1.5	1.6	−	−

4-NQO 4-nitroquinoline 1-oxide, *$CdCl_2$* cadmium chloride, *ENU* *N*-ethyl-*N*-nitrosourea, *HQ* hydroquinone, *MMS* methyl methanesulfonate, *2,4-DAT* 2,4-diaminotoluene, *2-AAF* 2-acetylaminofluorene, *AFB1* aflatoxin B1, *B[a]P* benzo[a]pyrene, *CPA* cyclophosphamide, *DMBA* 7,12-dimethylbenzanthracene, *NDMA* *N*-nitrosodimethylamine, *IQ* 2-amino-3-methylimidazo[4,5-f]quinoline, *PhIP* 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine, *3-MCPD* 3-chloro-1,2-propanediol, *DFPBA* 3,5-difluorophenylboronic acid, *EDAC* 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride, *HOPO* 2-pyridinol 1-oxide, *TBHQ* tertiary-butylhydroquinone, *DMTP* dimethyl terephthalate, and *LMG* leucomalachite green

[a]The highest concentration tested in the MN assay

[b]Relative survival is used as an indicator for cytotoxicity induced by compounds, relative to the vehicle control

[c]Relative cell viability was determined using one 5 K spheroid by the ATP assay

[d]Values represent the percentage of MN relative to intact nuclei

[e]The fold increases of chemical-induced MN over the vehicle control

[f]−, the ratio < 1.5-fold ($p \geq 0.05$ vs. vehicle control); ±, $1.5 \leq$ ratio < 2; +, $2 \leq$ ratio < 5; ++, $5 \leq$ ratio < 10, and +++, the ratio ≥ 10 ($p < 0.05$)

*$p < 0.05$ vs. vehicle control as evaluated by one-way ANOVA followed by Dunnett's test

HepaRG cell culture and spheroid formation

The HepaRG cell line was purchased from Biopredic International (Saint Grégoire, France). HepaRG cells were cultured and 3D HepaRG spheroids were formed as described in our previous study (Seo et al. 2022). Briefly, undifferentiated HepaRG cells at passages 14–20 were cultured in growth medium for 14 days and then differentiated in differentiation medium for another 14 days according to the supplier's protocol with minor modifications (Fig. 1A). The growth and differentiation media were prepared by adding growth and differentiation additives (Lonza; Walkersville, MD), respectively, to William's E Medium (Thermo Fisher; Waltham, MA) supplemented with 2 mM GlutaMax (Thermo Fisher) and 100 µg/ml primocin (InvivoGen; San Diego, CA). For 2D cultures, fully differentiated HepaRG cells were dissociated with TrypLE Express (Thermo Fisher) and plated into a 96-well flat bottom plate (Corning Inc.; Corning, NY) at a density of 5×10^4 cells/well in the differentiation medium (Fig. 1B). For 3D cultures, dissociated HepaRG cells were resuspended in fresh medium and passed through a 40 µm nylon mesh cell strainer (Corning) to remove cell aggregates. Then the cells were seeded into a 384-well ultra-low attachment (ULA) plate (Corning) at a density of 5×10^3 cells/well to form spheroids, each containing approximately 5000 (5 K) cells (Fig. 1C). The plates were incubated at 37 °C in a humidified atmosphere with 5% CO_2 in air. The medium was refreshed every 2–3 days using a VIAFLO 96/384 Electronic Pipette (INTEGRA Biosciences; Hudson, NH).

ATP assay

The ATP assay was used for monitoring cell proliferation and chemical-induced cytotoxicity following treatment. The cellular ATP levels in 2D and 3D HepaRG models were determined using the CellTiter-Glo Luminescent Cell Viability Assay kit (Promega; Madison, WI) and CellTiter-Glo 3D Cell Viability Assay kit (Promega), respectively. The ATP reagent was added into each well at a ratio of 1:10 and incubated for 10 min at room temperature. The luminescence from the luciferase reaction was recorded with a Cytation 5 Cell Imaging Multi-Mode Reader (BioTek; Winooski, VT). The relative cell viability was calculated by comparing the signal intensity of the treated cells to those of the vehicle controls.

Albumin secretion

Albumin secretion was measured in 2D HepaRG cells at Day 4 and in 3D HepaRG spheroids at Day 11, with and without the addition of 100 ng/ml hEGF (Sigma-Aldrich). One hundred microliters of culture medium were collected 24 h after the medium was refreshed and the secreted albumin was quantified using the Human Serum Albumin DuoSet Enzyme-linked Immunosorbent Assay (ELISA) kit (R&D Systems; Minneapolis, MN), following the manufacturer's instructions. In addition, 2D cells and 3D spheroids were lysed with RIPA buffer (Thermo Fisher) and the protein concentrations were measured using the Pierce BCA Protein Assay kit (Thermo Fisher). The albumin levels were normalized to protein concentration and presented as ng/mg protein.

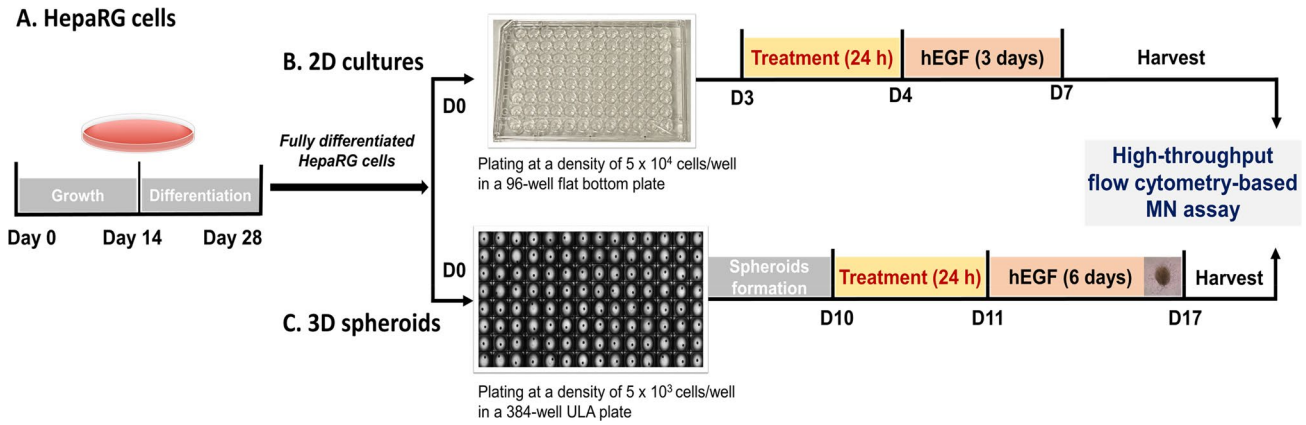

Fig. 1 High-throughput MN assay study design. HepaRG cells were cultured in growth medium for 14 days and then differentiated in differentiation medium for another 14 days (**A**). Fully differentiated cells were plated into a 96-well flat bottom plate or a 384-well ultra-low attachment plate at a density of 5×10^4 or 5×10^3 cells/well for forming 2D monolayer cultures (**B**) or 3D spheroids (**C**), respectively. Following a 24-h treatment, 2D and 3D HepaRG models were cultured for additional periods of time in order to go through 1.5 to 2 cell population doublings prior to conducting the high-throughput flow-cytometry-based MN assay

Gene expression of Phase I and Phase II enzymes

After 100 ng/ml hEGF treatment, 2D cultured cells and 3D HepaRG spheroids were harvested for comparison of basal gene expression levels of 13 Phase I and 5 Phase II enzymes using previously described methods (Seo et al. 2022). Briefly, total RNA from 2D and 3D HepaRG samples was isolated using the RNeasy Mini kit (Qiagen; Valencia, CA). RNA concentration and quality were measured with a NanoDrop 8000 spectrophotometer (Thermo Fisher) and 1 µg RNA was transcribed to cDNA with a High-Capacity cDNA Reverse Transcription kit (Applied Biosystems; Foster City, CA). Quantitative real-time PCR (qPCR) was performed using a ViiA 7 Real-Time PCR system (Applied Biosystems) with the FastStart Universal Probe Master (Rox) (Roche Applied Science; Indianapolis, IN) and 18 TaqMan probes (Supplementary Table 1). The amplification was carried out in triplicate 20 µl reactions using the following conditions: 2 min at 50 °C, 10 min at 95 °C, 40 cycles of 15 s at 95 °C and 1 min at 60 °C. The endogenous control gene, glyceraldehyde 3-phosphate dehydrogenase (GAPDH), was used as a reference gene to normalize levels of gene expression. The cycle threshold (Ct) values for all tested genes are shown in Supplementary Table 1 and Ct values greater than 35 were considered non-detectable. The expression value of each gene in 2D and 3D HepaRG models was defined by the equation: $E = 2^{-(\text{Ct of test gene} - \text{Ct of reference gene})} \times 10{,}000$. The expression value represents the relative mRNA expression abundance of a gene, arbitrarily assuming an expression level of GAPDH at 10,000 copies.

Chemical treatment

Approximately 7–8 spheroids (each containing 5 K cells) were transferred from ULA plates into each well of a 96-well round-bottom plate (TPP, Switzerland). 2D HepaRG cells at Day 3 and 3D spheroids at Day 10 were exposed to various concentrations of 34 test chemicals (Table 1) in a total volume of 100 µl for 24 h at 37 °C. All stock solutions were dissolved in DMSO and stored at −20 °C, except for cadmium chloride ($CdCl_2$) and *N*-nitrosodimethylamine (NDMA), which were dissolved in deionized water, and cisplatin, which was dissolved in 0.9% NaCl. Four chemicals, cisplatin, *N*-ethyl-*N*-nitrosourea (ENU), hydroquinone (HQ), and tertiary-butylhydroquinone (TBHQ), were freshly prepared before each experiment. Working solutions were prepared by serial dilution in differentiation medium. The final concentration of DMSO did not exceed 1%. Following a 24-h treatment, 100 µl of treatment media was removed from the 96-well plates, taking caution not to disturb the spheroids, and replaced with 200 µl of fresh differentiation medium supplemented with 100 ng/ml hEGF to stimulate cell proliferation. The 2D and 3D plates were then incubated at 37 °C for 3 and 6 additional days, respectively, prior to conducting the MN assay. Both the 2D and 3D HepaRG treatments were repeated independently at least three times for each chemical.

HT flow-cytometry-based MN assay

HT flow-cytometry-based MN analysis was performed using the In Vitro MicroFlow kit (Litron Laboratories; Rochester, NY) as previously described (Guo et al. 2020b). Briefly, after the hEGF treatments, 2D cultured cells and 3D spheroids were stained under visible light for 30 min on ice with 50 µl ethidium monoazide bromide (EMA) solution to label apoptotic/necrotic cells. After removing supernatants, 100 µl each of Lysis Solutions I and II containing SYTOX Green were added sequentially into each well to label the chromatin. MN events were scored using a FACSCanto II flow cytometer (BD Biosciences; San Jose, CA) equipped with an HT sampler and using a stopping gate set to record 10,000 intact nuclei. The percentage of MN, calculated as the ratio of MN events to the total number of nucleated events, was used for genotoxicity evaluation. Cytotoxicity was evaluated by relative survival which was calculated using the ratio of intact nucleated events in treated cells relative to those of vehicle controls at a specified time point. Relative survival includes a measure of cell growth during the 24-h treatment and the stimulation period with hEGF.

Benchmark dose (BMD) analysis

Positive MN data were quantified using BMD analysis by the PROAST web-based software (version 70.1). The concentration–response data were analyzed using both the exponential and Hill models, as recommended for the assessment of continuous data by the European Food Safety Authority (EFSA), and based on superior fit, the exponential model was selected for calculating the BMDs. A critical effect size (CES) of 0.5, corresponding to 50% increases in %MN over vehicle control (BMD_{50}), was used to compare the MN concentration-responses generated from 2D and 3D HepaRG cultures. The upper (BMDU) and lower bounds (BMDL) of the BMD_{50} were calculated simultaneously for each data set. Responses having non-overlapping BMDUs and BMDLs were considered to have different relative potencies.

Statistical analysis

Data are expressed as the mean ± standard deviation (SD) or the standard error of the mean (SEM) from at least three independent experiments. The statistical significance of albumin secretion was evaluated by one-way analysis of variance (ANOVA) followed by the Holm-Sidak test; a two-tailed Student's t-test to compare mRNA gene expression

data for Phase I and Phase II enzymes measured in 2D and 3D cultures. One-way ANOVA followed by Dunnett's post hoc test was used for evaluating the ATP levels and %MN comparisons between different concentrations of the test compounds and the vehicle control. Analyses were performed using SigmaPlot 13.0 (Systat Software; San Jose, CA, USA). Differences were considered significant for all tests when $p < 0.05$.

Results

Characterization of hEGF-stimulated 2D HepaRG cells and 3D spheroids

According to OECD Test Guideline 487 for the in vitro mammalian cell MN test, MN should be scored in cells that have gone through mitosis during or after exposure to the test chemical (OECD 2016). The proliferative capacity of 2D HepaRG cells and 3D spheroids was monitored to ensure that negative control (untreated), differentiated HepaRG models underwent approximately 1.5–2 cell population doublings. Cell proliferation was based on comparing the ATP content after extended incubation in differentiation medium supplemented with 100 ng/ml hEGF at Day 4 for 2D cells and at Day 11 for 3D spheroids after seeding. These were the timings used in the present study for cultures exposed to the test substances for 24 h. In 2D HepaRG cells, ATP levels increased by 1.6-, 1.7-, and 1.9-fold after 1-, 2-, and 3-day hEGF stimulation, respectively; while ATP levels of 3D spheroids increased by 1.4-, 1.7-, 2.1-fold after 4-, 6-, and 8-day hEGF stimulation, respectively (Fig. 2A). Based on these observations, cell harvesting time for conducting the MN assay occurred following 3- and 6-day incubations

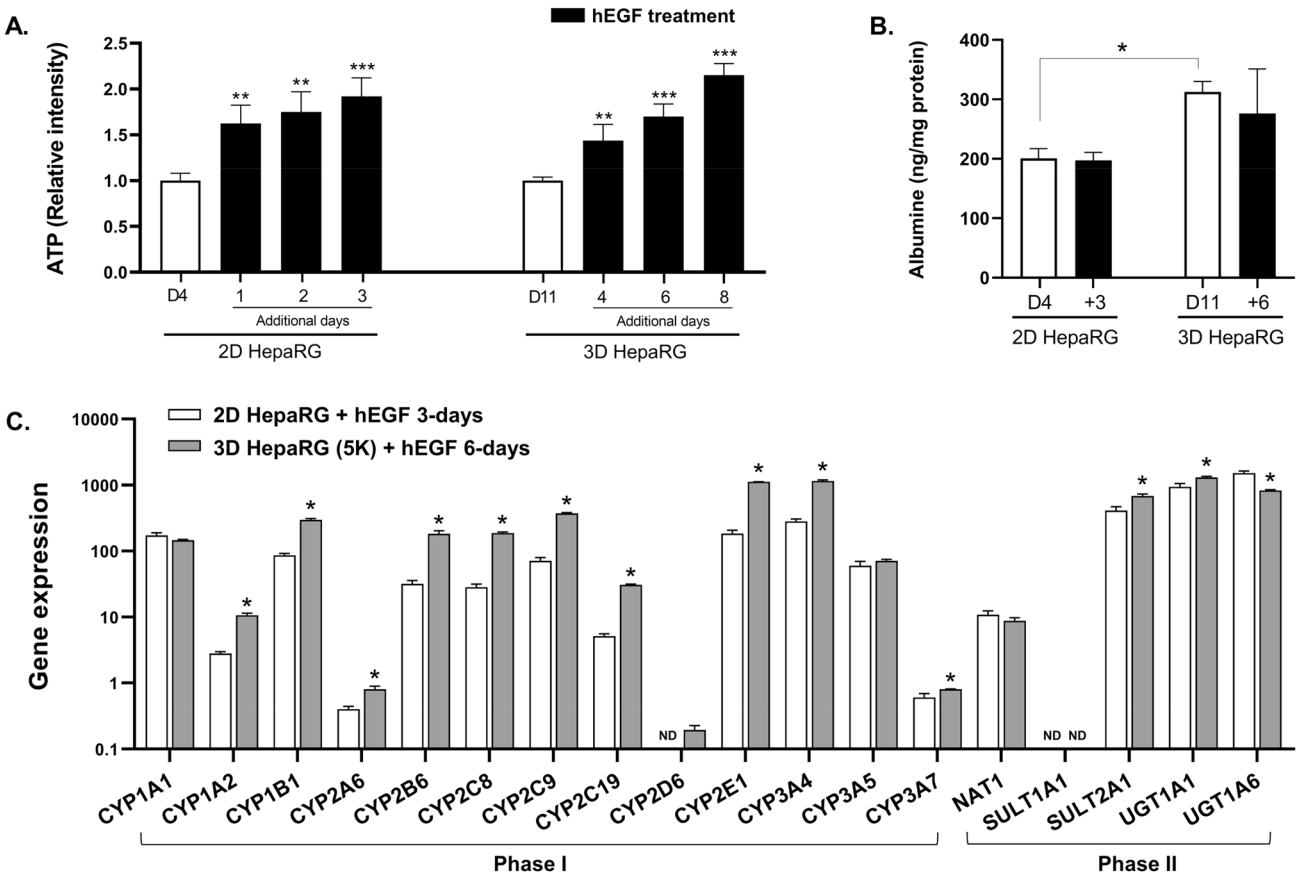

Fig. 2 Confirmation of the proliferative and metabolic capacity of 2D and 3D HepaRG models. (**A**) Cell proliferation was monitored by measuring the relative ATP levels after additional incubation with human epidermal growth factor (hEGF, 100 ng/ml). $^{**}p < 0.01$, $^{***}p < 0.001$ by one-way ANOVA followed by Dunnett's test. (**B**) Albumin secretion was measured by ELISA and expressed as ng/mg protein. $^{*}p < 0.05$ by one-way ANOVA followed by Holm-Sidak test. (**C**) Gene expression of Phase I and Phase II enzymes in 2D HepaRG cells and 3D HepaRG spheroids (each containing 5 K cells) with additional hEGF incubation was measured by quantitative real-time PCR (qPCR). qPCR data was normalized to GAPDH expression. $^{*}p < 0.05$ by two-tailed Student's t-test. The data are presented as the mean ± SD ($n \geq 3$). *CYP* cytochrome P450, *NAT N*-acetyltransferase, *SULT* sulfotransferase, *UGT* UDP-glucuronosyltransferase, *GAPDH* glyceraldehyde 3-phosphate dehydrogenase, *ND* non-detectable

with hEGF-stimulation in 2D cultured cells and 3D spheroids, respectively.

Albumin secretion in both 2D and 3D HepaRG cultures was examined within 24 h after differentiation medium was refreshed. 3D HepaRG spheroids at Day 11 showed significantly higher albumin secretion than 2D HepaRG cells at Day 4 (313 ng/mg protein in 3D vs. 201 ng/mg protein in 2D); and albumin levels remained stable during the hEGF stimulation periods in both 2D and 3D HepaRG cultures (Fig. 2B).

Gene expression of Phase I and II enzymes was determined using qPCR in 2D and 3D HepaRG hEGF-stimulated cultures (Fig. 2C). Compared to hEGF-stimulated 2D HepaRG cells, hEGF-stimulated 3D spheroids exhibited significantly higher gene expression values for 10 Phase I enzymes, including CYP1A2 (3.8-fold), 1B1 (3.4-fold), 2A6 (1.9-fold), 2B6 (5.7-fold), 2C8 (6.6-fold), 2C9 (5.3-fold), 2C19 (6.0-fold), 2E1 (6.1-fold), 3A4 (4.1-fold), and 3A7 (1.5-fold), and 2 Phase II enzymes, SULT2A1 (1.7-fold) and UGT1A1 (1.4-fold). CYP2D6 was not detected in 2D cultures but the expression level showed noticeable increases in 3D spheroids. No significant changes were found between 2D and 3D HepaRG models for CYP1A1, CYP3A5, or NAT1, while UGT1A6 had lower expression in hEGF-stimulated 3D spheroids than in hEGF-stimulated 2D HepaRG cells. SULT1A1 gene expression was not detected in the 2D or 3D models in the present study.

Comparison of cytotoxicity of the 34 compounds in 2D HepaRG cells and 3D spheroids

A wide range of concentrations was used to test the cytotoxicity of 34 compounds including 8 direct-acting and 11 indirect-acting genotoxicants or carcinogens as well as 15 compounds that show different genotoxic responses in vitro and in vivo (Table 1). Relative survival (% of control) served as an indicator of cytotoxicity and was assessed simultaneously with flow cytometric MN scoring in 2D HepaRG cells and 3D spheroids (Supplementary Figs. 1 and 2). In addition, the cytotoxicity of the 34 test compounds was determined concurrently by the ATP assay in 3D spheroids after a 6-day hEGF stimulation (Supplementary Figs. 3 and 4). When no cytotoxicity was observed, 10 mM was used as the maximum concentration as recommended by OECD guidance for genetic toxicology testing (OECD 2015). In addition, concentrations that resulted in < 40% relative survival were excluded to avoid excessive cytotoxicity for MN evaluation (OECD 2016).

The highest test concentration and corresponding cytotoxicity (% relative survival and %ATP) for each compound are shown in Table 1. The top test concentrations were lower in 3D HepaRG spheroids than in 2D cultures for 2 direct-acting genotoxicants (cisplatin and colchicine), 5 indirect-acting genotoxicants (2,4-DAT, AFB1, CPA, DMBA, and NDMA), and one compound that showed different genotoxic responses in vitro and in vivo (HOPO). $CdCl_2$, ENU, IQ, and estragole showed higher cytotoxicity in 2D cells than in 3D spheroids.

Comparison of MN induction of the 34 compounds in 2D HepaRG cells and 3D spheroids

Seven out of the eight direct-acting genotoxicants or carcinogens induced significant MN formation in 2D HepaRG cells, while six out of eight compounds were positive in 3D HepaRG spheroids (Fig. 3A; Table 1). Two compounds, 4-NQO and ENU, produced 1.9–2.7-fold increases in %MN over vehicle control in both 2D and 3D HepaRG models. Four compounds, cisplatin, colchicine, etoposide, and MMS, induced much higher %MN frequencies in 2D HepaRG cells (11.6-, 6.5-, 19.8-, and 10.4-fold, respectively) than in 3D spheroids (4.0-, 2.2-, 12.1-, and 4.6-fold, respectively) at the highest test concentration. $CdCl_2$ was negative for MN induction in both 2D and 3D HepaRG models, with lower cytotoxicity observed in 3D spheroids. HQ induced a marginal positive response (1.8-fold) in 2D HepaRG cells but was negative in 3D spheroids.

For the 11 indirect-acting genotoxicants or carcinogens, six and seven of them showed positive MN responses in 2D and 3D HepaRG cultures, respectively (Fig. 3B; Table 1). Under the experimental conditions that we employed, four compounds, 2,4-DAT, acrylamide, IQ, and styrene, produced negative MN responses in both 2D and 3D HepaRG models, while 2-AAF was positive only in 3D HepaRG spheroids (2.1-fold increase over the vehicle control). Six compounds, AFB1, B[a]P, CPA, DMBA, NDMA, and PhIP, induced significant MN production in both the 2D and 3D HepaRG models. At the highest concentrations, the relative MN frequency fold-induction in 2D and 3D cultures for AFB1, B[a]P, and CPA was 6.2 and 4.4, 3.6 and 3.5, and 6.7 and 5.4, respectively. PhIP showed moderate increases in %MN (2.7- and 1.6-fold) over the vehicle controls in 2D and 3D cultures, respectively. Two compounds, DMBA and NDMA, induced relatively higher MN production in 3D spheroids than in 2D HepaRG cells (5.9- and 5.3-fold vs. 3.6- and 3.0-fold increase over the controls, respectively).

Overall, 2D HepaRG cells showed a slightly higher sensitivity (88%) than 3D HepaRG spheroids (75%) in detecting 8 direct-acting genotoxicants or carcinogens, whereas 3D HepaRG models exhibited a slightly higher sensitivity (64%) than 2D HepaRG models (55%) in detecting 11 indirect-acting genotoxicants or carcinogens (Table 2). Both cell models produced consistently negative calls for the 15 compounds that show different genotoxic responses in vitro and in vivo (Fig. 4; Table 1).

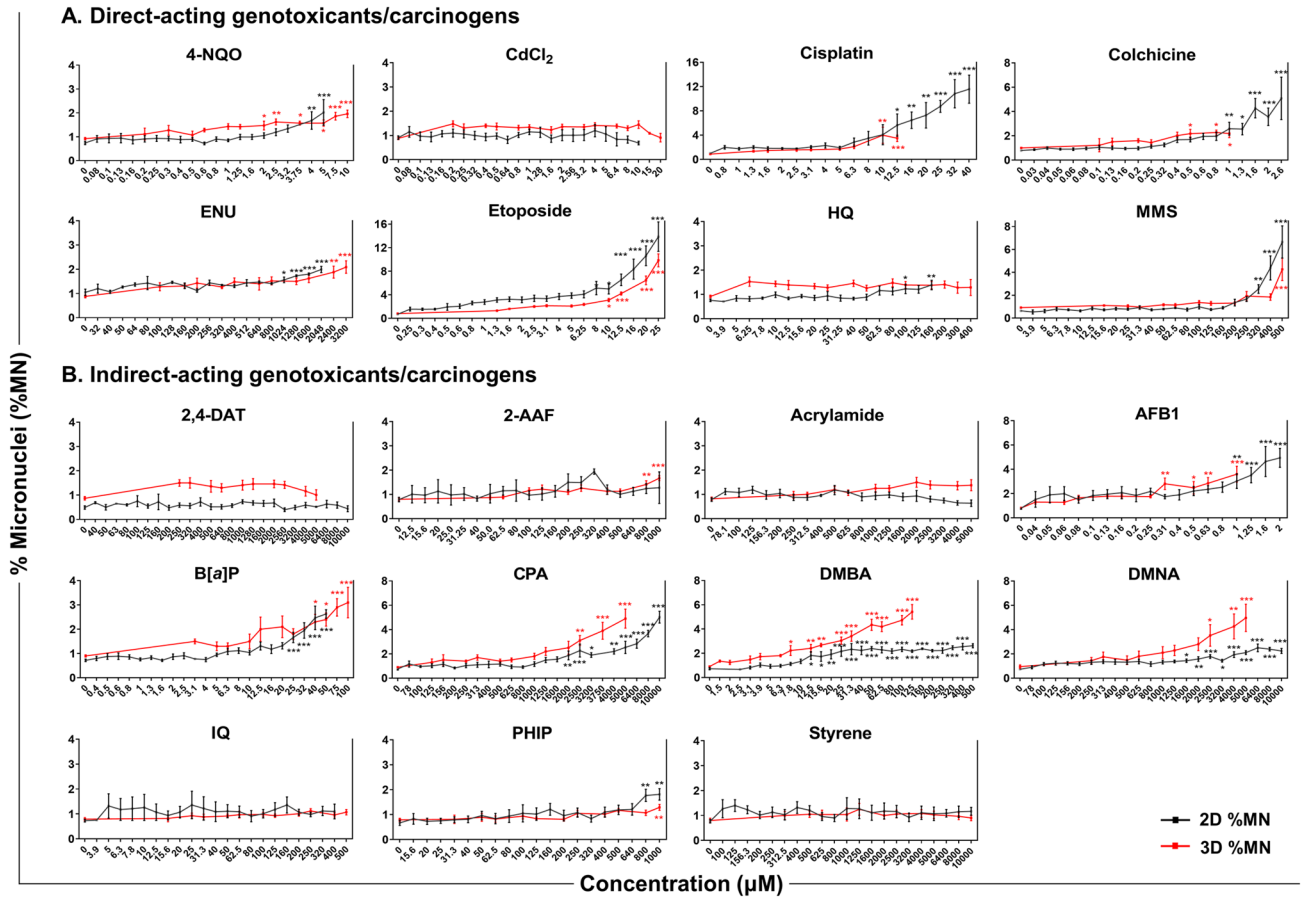

Fig. 3 Comparison of MN induction by 19 genotoxicants or carcinogens in 2D and 3D HepaRG cultures. 2D and 3D HepaRG cultures were exposed to 8 direct-acting (**A**) and 11 indirect-acting genotoxicants/carcinogens (**B**) for 24 h. MN frequency is presented as the percentage of MN relative to intact nuclei (% MN). Part of %MN data in 2D HepaRG cells were obtained from our previous study (Guo et al. 2020b). The red and black lines represent the results of 3D and 2D HepaRG cultures, respectively. The data are expressed as the mean ± SEM ($n \geq 3$). Significant difference was determined by one-way ANOVA followed by Dunnett's test (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$ *vs.* vehicle control). See Table 1 for abbreviations of the compounds tested (color figure online)

Table 2 Sensitivity of 2D and 3D HepaRG cultures for detecting genotoxicants or carcinogens using MN and comet data

Compound	MN assay		Comet assay[a]		Overall[b]	
	2D	3D	2D	3D	2D	3D
8 direct-acting	88% (7/8)	75% (6/8)	63% (5/8)	75% (6/8)	88% (7/8)	88% (7/8)
11 indirect-acting	55% (6/11)	64% (7/11)	45% (5/11)	73% (8/11)	64% (7/11)	91% (10/11)
19 genotoxicants or carcinogens	68% (13/19)	68% (13/19)	53% (10/19)	74% (14/19)	74% (14/19)	89% (17/19)

[a]Comet data were obtained from our previous study (Seo et al. 2022)

[b]Sensitivity for detecting the in vitro genotoxicity of compounds

Quantification of MN induction potency using benchmark dose (BMD) analysis

Quantitative analysis was performed on the 12 genotoxicants or carcinogens producing positive responses for MN induction in both 2D and 3D HepaRG models. BMD_{50} values calculated from the exponential model were used for comparing the MN induction potency of the compounds in 2D and 3D HepaRG models (Fig. 5 and Supplementary Table 2). The potency of the 12 compounds ranked similarly in 2D and 3D HepaRG cell models, with a few exceptions. Specifically, colchicine and AFB1 were the most potent compounds, followed by etoposide, cisplatin, and 4-NQO; ENU was the least potent in both HepaRG models. For the

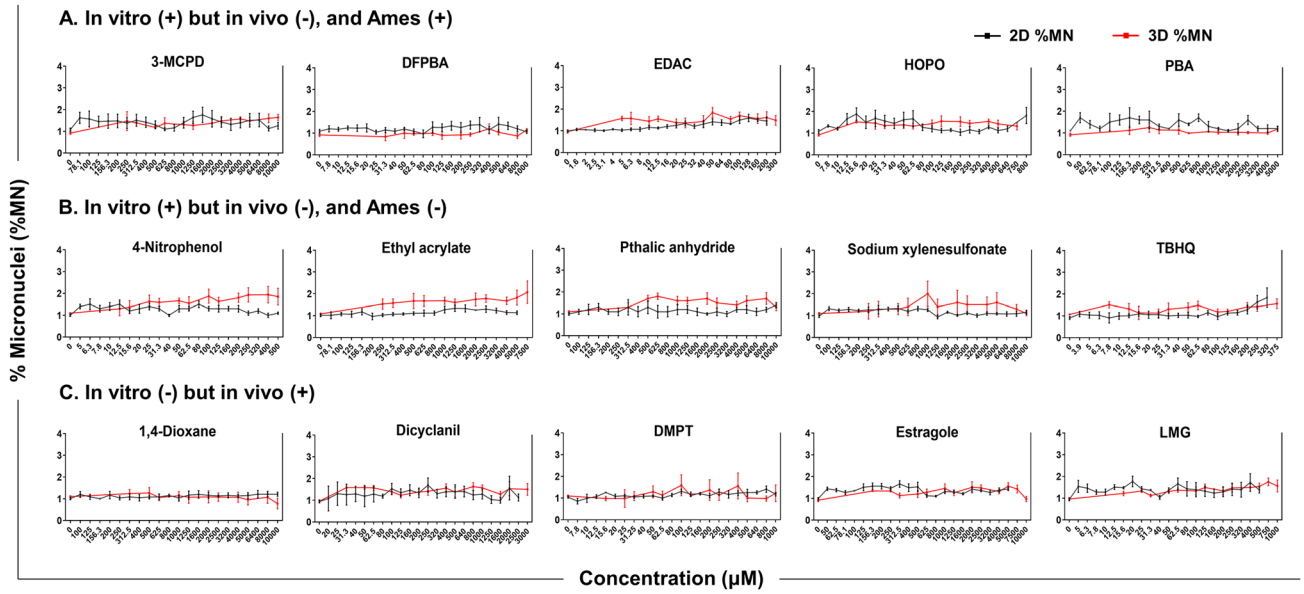

Fig. 4 Comparison of MN induction by 15 compounds in 2D and 3D HepaRG cultures. 2D and 3D HepaRG cultures were exposed to 15 compounds that show different genotoxic responses in vitro and in vivo including five in vitro (+) but in vivo (−) and Ames (+) compounds (**A**), five in vitro (+) but in vivo (−) and Ames (−) compounds (**B**), and five in vitro (−) but in vivo (+) compounds for 24 h. MN frequency is presented as the percentage of MN relative to intact nuclei (% MN). The red and black lines represent the results of 3D and 2D HepaRG cultures, respectively. The data are expressed as the mean ± SEM ($n \geq 3$). See Table 1 for abbreviations of the compounds tested (color figure online)

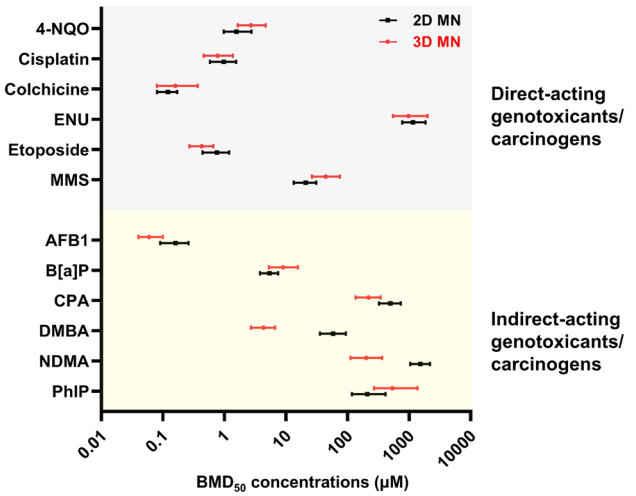

Fig. 5 Comparison of benchmark dose (BMD) values with their 90% confidence intervals (BMDU and BMDL) between 2D and 3D HepaRG cultures. BMD_{50} values with their BMDUs and BMDLs were calculated on the high-throughput positive MN data using exponential models of PROAST. The bars represent the uncertainty of BMD_{50} estimates (the range between BMDUs and BMDLs). Black square, 2D HepaRG cells; Red circle, 3D HepaRG spheroids. See Table 1 for abbreviations of the compounds tested (color figure online)

other 6 compounds, the potency ranking was B[a]P > MMS > DMBA > PhIP > CPA > NDMA in 2D HepaRG cells, while the ranking was DMBA > B[a]P > MMS > NDMA ≈ CPA > PhIP in 3D HepaRG spheroids (Supplementary Fig. 5).

For all 6 direct-acting compounds positive in both 2D and 3D cultures, the 90% CIs for the BMD_{50} values calculated from the MN responses in 2D and 3D HepaRG cultures overlapped each other (100%). Conducting a similar potency analysis on the 6 indirect-acting compounds, DMBA and NDMA produced significantly lower BMD_{50} values in 3D spheroids than those in 2D HepaRG models, whereas BMD_{50} CIs calculated from the 2D and 3D HepaRG MN dose response data for four compounds (67%), AFB1, B[a]P, CPA, and PhIP, overlapped.

Performance of 2D and 3D HepaRG models in the MN and comet assays

The sensitivity of 2D and 3D HepaRG cultures in detecting the genotoxicity of 19 genotoxicants or carcinogens after a 24-h treatment were compared using both MN data in the present study and CometChip data generated from our previous study (Seo et al. 2022). 3D and 2D cultures revealed the same sensitivity (68%) in the MN assay while comet assays conducted with 3D cultures showed a higher sensitivity than comet assays in 2D cultures (74 vs. 53%; Table 2). When the MN and comet assays were combined, 3D HepaRG spheroids showed the same sensitivity as 2D cultures (both 88%) in detecting the 8 direct-acting

compounds but had increased sensitivity than 2D cultures (91 vs. 64%) in detecting the 11 indirect-acting compounds. Collectively, 3D spheroids demonstrated an overall higher sensitivity than 2D cultures (89 vs. 74%) in detecting the 19 genotoxicants or carcinogens.

The in vitro genotoxic potencies of the 9 compounds positive in both the MN and comet assays were further compared between 2 and 3D HepaRG cultures using the BMD_{50} values and their upper and lower 90% CIs (Fig. 6). Among the five positive direct-acting compounds, cisplatin, 4-NQO, etoposide, and ENU showed similar or slightly higher genotoxic potency in 3D spheroids than in 2D cultures, while MMS had slightly increased genotoxic potency in 2D cultures compared to 3D spheroids for both the MN and comet assays. Three out of the four positive indirect-acting compounds, CPA, DMBA, and NDMA, demonstrated much greater genotoxic potency in 3D spheroids compared to 2D cultures for both the MN induction and DNA damage responses.

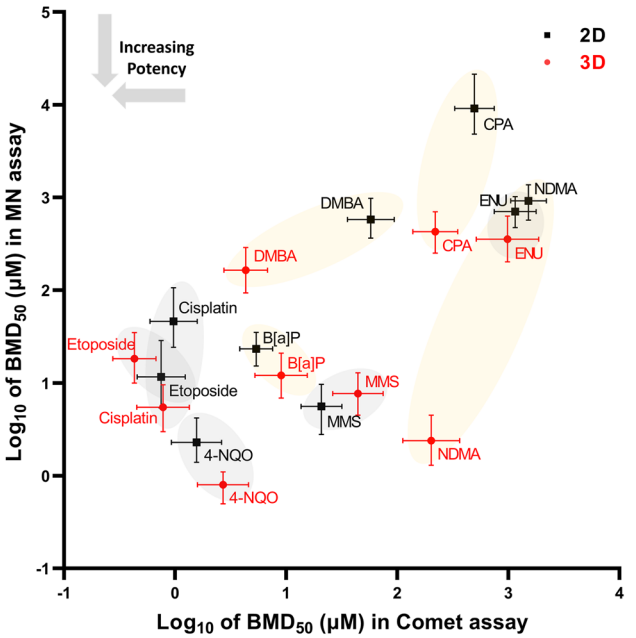

Fig. 6 Comparison of in vitro genotoxic potency between 2D and 3D HepaRG cultures. BMD_{50}-derived genotoxic potency on the nine positive compounds in both MN and comet assays was compared between 2D and 3D HepaRG cultures. BMD_{50} values of comet assays in 2D and 3D HepaRG models were obtained from our previous study (Seo et al. 2022). BMD_{50} values of MN and comet assays are displayed on log_{10} scales and the bars represent the uncertainty of BMD_{50} estimates in both the x-axis (DNA damage) and y-axis (MN induction). Black square, 2D HepaRG cells; Red circle, 3D HepaRG spheroids. See Table 1 for abbreviations of the compounds tested (color figure online)

Discussion

The scientific community has increasingly become aware of two issues concerning current approaches to in vitro genotoxicity testing that limit their value in human risk assessment: (1) traditional 2D monolayer cultures of mainly rodent cell lines are not fully representative of natural physiological or systemic conditions; and (2) animals are not perfectly predictive of human toxicity (Pridgeon et al. 2018). In vitro 3D cell culture models and microphysiological systems are considered more biologically relevant than 2D cell cultures and hold promise for bridging the gap between in vitro and in vivo toxicity testing. For instance, these advanced models have displayed improved cellular signal transduction and tissue-like structures that may result in better evaluation of the toxic properties of regulated substances (Cirit and Stokes 2018; Guo et al. 2020a; Vernon et al. 2022). Based upon our recent success in adapting 3D HepaRG spheroids to the HT CometChip assay for detecting DNA strand breaks (Seo et al. 2022), the present study used 3D HepaRG spheroids for conducting the MN assay, which evaluates two important mechanisms of genotoxicity, clastogenicity and aneugenicity (Fenech 2008).

As the MN assay requires actively proliferating cells for MN formation, the present study first confirmed that replicating cells could be generated from 3D HepaRG spheroids. As anticipated from the work of Rose et al (2022), a 3-day hEGF stimulation increased the metabolic activity in 2D HepaRG cultures by 1.9-fold as quantified by cellular ATP content (Fig. 2A); the extent of this increase was comparable to a previous report of a 2.3-fold increase in cell population measured by cell counting (Josse et al. 2012). As a variety of cell lines, such as CACO-2, DLD-1, HT-29, and SW-480 cells, showed a slower proliferation rate in 3D cultures relative to the same cells cultured in 2D (Luca et al. 2013), we extended the hEGF stimulation period for 3D HepaRG cultures to as long as 8 days (Fig. 2A). A 6-day incubation with hEGF produced a 1.7-fold increase in ATP content, which was anticipated to be the shortest period for the spheroids to go through 1.5–2 cell population doublings. Thus, 6 days of stimulation was chosen as a reasonable sampling time for conducting the MN assay. At the end of the hEGF stimulation period, 3D spheroids also maintained high levels of metabolic capacity compared to 2D cultures as demonstrated by significantly increased gene expression of 11 major CYP enzymes (Fig. 2C). This observation indicates that the 3D spheroids remained stable, and metabolically active, throughout the experiment.

The present study also adapted the HT flow-cytometry-based MN assay for measuring MN in 3D HepaRG

spheroids. This modification produced a significant improvement in assay throughput compared to previous studies in 3D cultures using microscope observations in the CBMN assay (Conway et al. 2020; Rose et al. 2022; Shah et al. 2018). Unlike a previous study in which treated spheroids were dissociated into single cells and replated in monolayer culture for a 72-h recovery period during which the cells were stimulated to divide (Rose et al. 2022), the 3D spheroids in our study remained intact in the same wells of 96-well plates in which they were treated, which facilitated throughput. In addition, the use of an HT flow-cytometry-based MN assay meant that there was no need to harvest cells by trypsinization, as the spheroids were lysed directly for the MN analysis. In addition, this strategy provided a cytotoxicity endpoint whose results were remarkably comparable with cytotoxicity assessed by the relative values for treated and control spheroids measured in the ATP assay (Table 1, Supplementary Figs. 1−4).

The present study assessed the performance of 3D cultures in the MN assay using the same set of compounds used for our HepaRG spheroid CometChip study (Seo et al. 2022), i.e., 8 direct-acting and 11 indirect-acting genotoxicants or carcinogens and 15 compounds that show different genotoxic responses in vitro and in vivo (Table 1). The data accumulated in our studies indicate that the 2D and 3D HepaRG assays were equally sensitive for detecting MN induction by the 8 direct-acting compounds. Most of the compounds (4-NQO, cisplatin, colchicine, etoposide, HQ, and MMS) induced slightly lower %MN frequencies in the 3D assay than in 2D cultures. However, the BMD_{50} CIs calculated from the 2D and 3D model dose response data for each compound overlapped (Fig. 5), suggesting that the two cell models generated comparable genotoxicity responses.

Among the 11 indirect-acting genotoxicants or carcinogens, four compounds (AFB1, CPA, DMBA, and NDMA) produced greater levels of cytotoxicity and/or genotoxicity in HepaRG spheroids (Table 1; Fig. 3B). This finding is likely due to the increased bioactivation of the compounds by the spheroid cultures into their genotoxic metabolites by CYP450 enzymes, mainly via CYP1A2, CYP3A4, CYP2B6, CYP1B1, and CYP2E1 (Fig. 2C). Despite increased CYP expression levels, HepaRG spheroids did not display increased sensitivity for all indirect-acting compounds, e.g., 2,4-DAT, acrylamide, IQ, and styrene were negative in both 3D spheroids and 2D cells (Table 1). 2,4-DAT and IQ became genotoxic after bioactivation by CYP1A2 in combination with *N*-acetyl transferase (NAT) (Chevereau et al. 2017), and both UDP glucuronosyltransferase (UGT) and glutathione transferase (GST) were involved in the detoxification of their *N*-hydroxylated metabolites. In addition, CYP2E1-mediated metabolites of acrylamide and styrene are detoxified by these Phase II enzymes (Dasari et al. 2018; Delos 2021; Pacifici et al. 1987; Turesky and Le Marchand 2011). It warrants further investigation if these conjugating enzymes as well as other factors such as nuclear receptors and transporters account for the negative responses of these compounds in HepaRG cells, treated both as spheroids and 2D cultures. 2-AAF was positive in 3D spheroids but negative in 2D cells in the present study. However, 2-AAF was reported to induce MN formation at concentrations ≥ 250 µM in the 2D HepaRG CBMN assay (Le Hegarat et al. 2014). Our results showed slight increases in %MN at concentrations of 250 and 320 µM in 2D cultures, but these responses were not statistically significant. This discrepancy might be explained by the fact that the CBMN assay detects MN along with nucleoplasmic bridges and nuclear buds, and is thus considered more sensitive than the MN assay as conducted by us (Fenech 2000, 2007).

Genotoxicity can be induced via various mechanisms, e.g., DNA strand breaks, other forms of chromosomal damage, and gene mutations. A battery of assays detecting different mechanisms of genotoxicity is recommended by international regulatory bodies for genotoxicity testing (ICH 2013; OECD 2015). Integration of two or three of the most commonly used in vitro genotoxicity tests (the Ames test, mouse lymphoma assay, and MN or chromosomal aberration assay) resulted in > 90% sensitivity for detecting over 700 rodent carcinogens (Kirkland et al. 2005). When the MN and comet assays were combined in the present study, 3D and 2D HepaRG models showed the same sensitivity (88%) for detecting the 8 direct-acting compounds, but 3D spheroids had a higher sensitivity than 2D cultures (91 vs. 64%) for detecting the 11 indirect-acting compounds (Table 2). Moreover, the potency analysis for MN induction using BMD_{50} values and the upper and lower bounds of their 90% CIs indicated that three out of the four indirect-acting compounds, CPA, DMBA, and NDMA, had much greater genotoxic potencies in 3D spheroids compared to 2D cultures, while the five direct-acting compounds showed similar genotoxic potency between 3D and 2D cultures (Fig. 6). These observations are consistent with the hypothesis that any advantage of the HepaRG spheroid model in detecting genotoxicants lies mainly with compounds that require metabolic activation and it is likely due to higher concentrations of reactive metabolites generated by the improved metabolic capacity of HepaRG spheroids.

With the increasing international trend to develop robust New Approach Methodologies (NAMs) for reducing and eventually eliminating animal toxicity testing (Kavlock et al. 2018), data from several in vitro NAMs that measure genetic toxicity (including the MN assay) have been evaluated for use in in-vitro-to-in-vivo (IVIVE) extrapolations for deriving administered equivalent doses (AEDs) from in vitro concentration–response data (Beal et al. 2022). The results indicate that the IVIVE-derived AEDs of the majority of compounds had lower values

relative to in vivo point of departure metrics (PODs) from animal studies. In agreement with other studies showing that 3D HepG2 and HepaRG spheroids generally had lower BMDs and/or the lowest observed genotoxic effect concentrations than their 2D counterparts (Barranger and Le Hegarat 2022; Conway et al. 2020), the BMD_{50} values generated from the 3D HepaRG MN and comet data from our study (Fig. 5) also were lower or similar to comparable values derived from 2D cells (Seo et al. 2022). Although there is concern about the smaller number of cells in 3D cultures that were exposed to the same concentrations of chemicals as 2D cultures (2–5 K vs. 50–70 K cells per well in 96-well plates) affecting the results (Barranger and Le Hegarat 2022), under the treatment conditions in the present study, all six direct-acting compounds had overlapping BMD_{50} 90% CIs, while a portion of indirect-acting compounds produced significantly lower BMD_{50} values for data generated with 3D spheroids vs. 2D cells. As indicated above, this observation is consistent with the relative metabolic capacity of the cell models for activating genotoxicants. We conjecture that AEDs derived from HepaRG spheroids, especially for compounds that require metabolic activation, will be more conservative than values derived from 2D monolayer cultures, which serves a protective role in preserving human health.

In summary, the present study demonstrated the feasibility of conducting the MN assay with HepaRG spheroids. With levels of human-relevant metabolism comparable to those found in primary human hepatocytes (PHHs) (Tascher et al. 2019), 3D spheroids demonstrated advantages over 2D cells in detecting indirect-acting genotoxicants or carcinogens using the MN endpoint for evaluating genotoxicity. We conclude that the HepaRG spheroid model is a promising NAM that may serve as a surrogate to PHHs for genotoxicity assessment based upon the following considerations. First, genotoxicant BMD values derived from the comet assay using HepaRG spheroids are much closer to those in PHHs than are BMDs generated in 2D cultures (Seo et al. 2020, 2022). Second, unlike PHHs, which lack proliferative capacity and quickly lose their identity and function in vitro (Rose et al. 2021), HepaRG cells can be stimulated to divide, and the model remains structurally and functionally stable for at least 30 days, facilitating their application for extended treatments and in experiments that require cell division, i.e., the MN assay and mutation assays. Third, the adaptation of spheroids in HT genotoxicity assays significantly increased assay efficiency and expanded the utilization of human hepatic cells in the field of genotoxicity. Further investigation is needed to evaluate whether HepaRG spheroids can be used confidently as NAMs for human-relevant genotoxicity predictions and for setting health-based guidance values for human exposures.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00204-023-03461-z.

Acknowledgements This study was supported by the U.S. Food and Drug Administration (FDA), National Center for Toxicological Research. We greatly appreciate Drs. Robert Heflich, Mugimane Manjanatha, and Lei Guo for their critical review of this article.

Data availability The data that support the findings of this study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Disclaimer This manuscript reflects the views of the authors and does not necessarily reflect those of the U.S. Food and Drug Administration. Any mention of commercial products is for clarification only and is not intended as approval, endorsement, or recommendation.

References

Barranger A, Le Hegarat L (2022) Towards better prediction of xenobiotic genotoxicity: CometChip technology coupled with a 3D model of HepaRG human liver cells. Arch Toxicol 96(7):2087–2095. https://doi.org/10.1007/s00204-022-03292-4

Beal MA, Audebert M, Barton-Maclaren T et al (2022) Quantitative in vitro to in vivo extrapolation of genotoxicity data provides protective estimates of in vivo dose. Environ Mol Mutagen. https://doi.org/10.1002/em.22521

Chen R, Lin YT, Fornace AJ Jr, Li HH (2022) A high-throughput and highly automated genotoxicity screening assay. Altex 39(1):71–81. https://doi.org/10.14573/altex.2102121

Chevereau M, Glatt H, Zalko D, Cravedi JP, Audebert M (2017) Role of human sulfotransferase 1A1 and *N*-acetyltransferase 2 in the metabolic activation of 16 heterocyclic amines and related heterocyclics to genotoxicants in recombinant V79 cells. Arch Toxicol 91(9):3175–3184. https://doi.org/10.1007/s00204-017-1935-8

Cimino MC (2006) Comparative overview of current international strategies and guidelines for genetic toxicology testing for regulatory purposes. Environ Mol Mutagen 47(5):362–390. https://doi.org/10.1002/em.20216

Cirit M, Stokes CL (2018) Maximizing the impact of microphysiological systems with in vitro-in vivo translation. Lab Chip 18(13):1831–1837. https://doi.org/10.1039/c8lc00039e

Conway GE, Shah UK, Llewellyn S et al (2020) Adaptation of the in vitro micronucleus assay for genotoxicity testing using 3D liver models supporting longer-term exposure durations. Mutagenesis 35(4):319–330. https://doi.org/10.1093/mutage/geaa018

Dasari S, Ganjayi MS, Meriga B (2018) Glutathione *S*-transferase is a good biomarker in acrylamide induced neurotoxicity and genotoxicity. Interdiscipl Toxicol 11(2):115–121. https://doi.org/10.2478/intox-2018-0007

Delos M (2021) Cell culture models as an in vitro alternative to study the absorption and biotransformation of drugs and mycotoxins in humans and animals. Ghent University, Ghent, Belgium. https://libstore.ugent.be/fulltxt/RUG01/003/010/626/RUG01-003010626_2021_0001_AC.pdf. Accessed 02/22/2023

EPA (2022) TSCA Chemical Substance Inventory. https://www.epa.gov/tsca-inventory. Accessed 11/28/2022

FDA (2000) Redbook 2000: IV.C.1. Short-term tests for genetic toxicity. http://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/IngredientsAdditivesGRASPackaging/ucm078321htm. Accessed 12/27/2022

Fenech M (2000) The in vitro micronucleus technique. Mutat Res 455(1–2):81–95. https://doi.org/10.1016/S0027-5107(00)00065-8

Fenech M (2007) Cytokinesis-block micronucleus cytome assay. Nat Protoc 2(5):1084–1104. https://doi.org/10.1038/nprot.2007.77

Fenech M (2008) The micronucleus assay determination of chromosomal level DNA damage. Methods Mol Biol 410:185–216. https://doi.org/10.1007/978-1-59745-548-0_12

Guo X, Seo JE, Li X, Mei N (2020a) Genetic toxicity assessment using liver cell models: past, present, and future. J Toxicol Environ Health Part B 23(1):27–50. https://doi.org/10.1080/10937404.2019.1692744

Guo X, Seo JE, Petibone D et al (2020b) Performance of HepaRG and HepG2 cells in the high-throughput micronucleus assay for in vitro genotoxicity assessment. J Toxicol Environ Health A 83(21–22):702–717. https://doi.org/10.1080/15287394.2020.1822972

ICH (2013) ICH guidance S2(R1) on genotoxicity testing and data interpretation for pharmaceuticals intened for human use. https://www.ema.europa.eu/en/ich-s2-r1-genotoxicity-testing-data-interpretation-pharmaceuticals-intended-human-use. Accessed 11/28/2022

Ivanov DP, Parker TL, Walker DA et al (2014) Multiplexing spheroid volume, resazurin and acid phosphatase viability assays for high-throughput screening of tumour spheroids and stem cell neurospheres. PLoS ONE 9(8):e103817. https://doi.org/10.1371/journal.pone.0103817

Josse R, Rogue A, Lorge E, Guillouzo A (2012) An adaptation of the human HepaRG cells to the in vitro micronucleus assay. Mutagenesis 27(3):295–304. https://doi.org/10.1093/mutage/ger076

Kavlock RJ, Bahadori T, Barton-Maclaren TS, Gwinn MR, Rasenberg M, Thomas RS (2018) Accelerating the pace of chemical risk assessment. Chem Res Toxicol 31(5):287–290. https://doi.org/10.1021/acs.chemrestox.7b00339

Kirkland D, Aardema M, Henderson L, Muller L (2005) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. Mutat Res 584(1–2):1–256. https://doi.org/10.1016/j.mrgentox.2005.02.004

Le Hegarat L, Mourot A, Huet S et al (2014) Performance of comet and micronucleus assays in metabolic competent HepaRG cells to predict in vivo genotoxicity. Toxicol Sci 138(2):300–309. https://doi.org/10.1093/toxsci/kfu004

Luca AC, Mersch S, Deenen R et al (2013) Impact of the 3D microenvironment on phenotype, gene expression, and EGFR inhibition of colorectal cancer cell lines. PLoS ONE 8(3):e59689. https://doi.org/10.1371/journal.pone.0059689

OECD (2015) Guidance document on revisions to OECD genetic toxicology test guidelines. OECD Workgroup of National Coordinators for Test 42 Guidelines (WNT) https://www.oecd.org/chemicalsafety/testing/Genetic%20Toxicology%20Guidance%20Document%20Aug%2031%202015.pdf. Accessed 11/28/2022

OECD (2016) In vitro mammalian cell micronucleus test, OECD Guidelines for the Testing of Chemicals, No. 487. Organisation for Economic Co-operation and Development (Paris, France) https://doi.org/10.1787/9789264264861-en. Assessed 11/28/2022

Pacifici GM, Warholm M, Guthenberg C, Mannervik B, Rane A (1987) Detoxification of styrene oxide by human-liver glutathione transferase. Hum Toxicol 6(6):483–489. https://doi.org/10.1177/096032718700600606

Pfuhler S, Fellows M, van Benthem J et al (2011) In vitro genotoxicity test approaches with better predictivity: summary of an IWGT workshop. Mutat Res 723(2):101–107. https://doi.org/10.1016/j.mrgentox.2011.03.013

Pfuhler S, van Benthem J, Curren R et al (2020) Use of in vitro 3D tissue models in genotoxicity testing: Strategic fit, validation status and way forward. Report of the working group from the 7(th) International Workshop on Genotoxicity Testing (IWGT). Mutat Res 850–851:503135. https://doi.org/10.1016/j.mrgentox.2020.503135

Pfuhler S, Downs TR, Hewitt NJ et al (2021) Validation of the 3D reconstructed human skin micronucleus (RSMN) assay: an animal-free alternative for following-up positive results from standard in vitro genotoxicity assays. Mutagenesis 36(1):1–17. https://doi.org/10.1093/mutage/geaa035

Pridgeon CS, Schlott C, Wong MW et al (2018) Innovative organotypic in vitro models for safety assessment: aligning with regulatory requirements and understanding models of the heart, skin, and liver as paradigms. Arch Toxicol 92(2):557–569. https://doi.org/10.1007/s00204-018-2152-9

Rose S, Ezan F, Cuvellier M et al (2021) Generation of proliferating human adult hepatocytes using optimized 3D culture conditions. Sci Rep 11(1):515. https://doi.org/10.1038/s41598-020-80019-4

Rose S, Cuvellier M, Ezan F et al (2022) DMSO-free highly differentiated HepaRG spheroids for chronic toxicity, liver functions and genotoxicity studies. Arch Toxicol 96(1):243–258. https://doi.org/10.1007/s00204-021-03178-x

Seo JE, Wu Q, Bryant M et al (2020) Performance of high-throughput CometChip assay using primary human hepatocytes: a comparison of DNA damage responses with in vitro human hepatoma cell lines. Arch Toxicol 94(6):2207–2224. https://doi.org/10.1007/s00204-020-02736-z

Seo JE, He X, Muskhelishvili L et al (2022) Evaluation of an in vitro three-dimensional HepaRG spheroid model for genotoxicity testing using the high-throughput CometChip platform. Altex 39(4):583–604. https://doi.org/10.14573/altex.2201121

Shah UK, Mallia JO, Singh N, Chapman KE, Doak SH, Jenkins GJS (2018) A three-dimensional in vitro HepG2 cells liver spheroid model for genotoxicity studies. Mutat Res 825:51–58. https://doi.org/10.1016/j.mrgentox.2017.12.005

Tascher G, Burban A, Camus S et al (2019) In-depth proteome analysis highlights HepaRG cells as a versatile cell system surrogate for primary human hepatocytes. Cells 8(2):192. https://doi.org/10.3390/cells8020192

Turesky RJ, Le Marchand L (2011) Metabolism and biomarkers of heterocyclic aromatic amines in molecular epidemiology studies: lessons learned from aromatic amines. Chem Res Toxicol 24(8):1169–1214. https://doi.org/10.1021/tx200135s

Vernon AR, Pemberton RM, Morse HR (2022) A novel in vitro 3D model of the human bone marrow to bridge the gap between in vitro and in vivo genotoxicity testing. Mutagenesis 37(2):112–129. https://doi.org/10.1093/mutage/geac009

VICH (2014) Studies to evaluate the safety of residues of veterinary drugs in human food: genotoxicity testing, VICH GL23(R). International Cooperation on Harmonization of Technical Requirements for Registration of Veterinary Medicinal Products (VICH) https://vichsec.org/en/guidelines/pharmaceuticals/pharma-safety/toxicology.html. Accessed 11/28/2022

Wills JW, Johnson GE, Doak SH, Soeteman-Hernandez LG, Slob W, White PA (2016) Empirical analysis of BMD metrics in genetic toxicology part I: in vitro analyses to provide robust potency rankings and support MOA determinations. Mutagenesis 31(3):255–263. https://doi.org/10.1093/mutage/gev085