Toxicology 461 (2021) 152900

Contents lists available at ScienceDirect

Toxicology

journal homepage: www.elsevier.com/locate/toxicol

Reproducibility of adipogenic responses to metabolism disrupting chemicals in the 3T3-L1 pre-adipocyte model system: An interlaboratory study

Christopher D. Kassotis [a,b,*], Kate Hoffman [b], Johannes Völker [c], Yong Pu [d], Almudena Veiga-Lopez [d], Stephanie M. Kim [e], Jennifer J. Schlezinger [e], Patrizia Bovolin [f], Erika Cottone [f], Astrid Saraceni [f], Rosaria Scandiffio [f], Ella Atlas [g], Karen Leingartner [g], Stacey Krager [h], Shelley A. Tischkau [h], Sibylle Ermler [i], Juliette Legler [i,j], Vesna A. Chappell [k], Suzanne E. Fenton [k], Fahmi Mesmar [l], Maria Bondesson [l], Mariana F. Fernández [m], Heather M. Stapleton [b]

[a] Institute of Environmental Health Sciences and Department of Pharmacology, Wayne State University, Detroit, MI, 48236, USA
[b] Nicholas School of the Environment, Duke University, Durham, NC, 27708, USA
[c] Department of Biology, Norwegian University of Science and Technology (NTNU), Trondheim, 7491, Norway
[d] Department of Pathology, University of Illinois at Chicago, Chicago, IL, 60612, USA
[e] School of Public Health, Boston University, Boston, MA, 02118, USA
[f] Department of Life Sciences and Systems Biology, University of Torino, 10123, Torino, Italy
[g] Environmental Health Science and Research Bureau, Health Canada, 251 Sir Frederick Banting Drive, Ottawa, Canada
[h] Department of Pharmacology, Southern Illinois University School of Medicine, Springfield, IL, USA
[i] Department of Life Sciences, College of Health and Life Sciences, Brunel University London, Uxbridge, UB8 3PH, UK
[j] Institute for Risk Assessment Sciences, Department of Population Health Sciences, Faculty of Veterinary Medicine, Utrecht University, 3508 TD, Utrecht, Netherlands
[k] National Toxicology Program Laboratory, Division of the National Toxicology Program, National Institutes of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, USA
[l] Department of Intelligent Systems Engineering, Indiana University, Bloomington, IN, 47408, USA
[m] University of Granada, Center for Biomedical Research (CIBM), Granada, Spain

ARTICLE INFO

Handling Editor: Matthew Wright

Keywords:
Endocrine disrupting chemicals
Adipogenesis
Obesogen
Triglyceride accumulation
Metabolic disruption
Reproducibility
3T3-L1

ABSTRACT

The 3T3-L1 murine pre-adipocyte line is an established cell culture model for screening Metabolism Disrupting Chemicals (MDCs). Despite a need to accurately identify MDCs for further evaluation, relatively little research has been performed to comprehensively evaluate reproducibility across laboratories, assess factors that might contribute to varying degrees of differentiation between laboratories (media additives, plastics, cell source, etc.), or to standardize protocols. As such, the goals of this study were to assess interlaboratory variability of efficacy and potency outcomes for triglyceride accumulation and pre-adipocyte proliferation using the mouse 3T3-L1 pre-adipocyte cell assay to test chemicals. Ten laboratories from five different countries participated. Each laboratory evaluated one reference chemical (rosiglitazone) and three blinded test chemicals (tributyltin chloride, pyraclostrobin, and bisphenol A) using: 1) their Laboratory-specific 3T3-L1 Cells (LC) and their Laboratory-specific differentiation Protocol (LP), 2) Shared 3T3-L1 Cells (SC) with LP, 3) LC with a Shared differentiation Protocol (SP), and 4) SC with SP. Blinded test chemical responses were analyzed by the coordinating laboratory. The magnitude and range of bioactivities reported varied considerably across laboratories and test conditions, though the presence or absence of activity for each tested chemical was more consistent. Triglyceride accumulation activity determinations for rosiglitazone ranged from 90 to 100% across test conditions, but 30–70 % for pre-adipocyte proliferation; this was 40–80 % for triglyceride accumulation induced by pyraclostrobin, 80–100 % for tributyltin, and 80–100 % for bisphenol A. Consistency was much lower for pre-adipocyte proliferation, with 30–70 % active determinations for pyraclostrobin, 30–50 % for tributyltin, and 20–40 % for bisphenol A. Greater

* Corresponding author at: Wayne State University, Institute of Environmental Health Sciences and Department of Pharmacology, 2111 Integrative Biosciences Center, 6135 Woodward Avenue, Detroit, MI, 48202, USA.
E-mail address: christopher.kassotis@wayne.edu (C.D. Kassotis).

https://doi.org/10.1016/j.tox.2021.152900
Received 11 June 2021; Received in revised form 2 August 2021; Accepted 13 August 2021
Available online 17 August 2021

consistency was observed for the SC/SP assessment. As such, working to develop a standardized adipogenic differentiation protocol represents the best strategy for improving consistency of adipogenic responses using the 3T3-L1 model to reproducibly identify MDCs and increase confidence in reported outcomes.

1. Introduction

The 3T3-L1 murine pre-adipocyte cell line is an established model for *in vitro* screening of metabolism disrupting chemicals (MDCs) (Heindel et al., 2015, 2017). When exposed to adipogenic stimuli, pre-adipocytes will differentiate into mature adipocytes, undergo morphological changes, accumulate triglycerides, and eventually develop into a rounded white fat cell with a number of large lipid droplets often displacing the nucleus (Green and Meuth, 1974; Green and Kehinde, 1975). Adipocyte differentiation requires eventual activation of the peroxisome proliferator-activated receptor-gamma (PPARγ), often considered the "master regulator" of adipocyte differentiation (Rosen et al., 1999). Molecular pathways upstream that contribute to this activation are diverse and include modulation of thyroid receptor-beta (TRβ), glucocorticoid receptor (GR), estrogen receptor (ER), androgen receptor (AR), liver X receptor (LXR), retinoid X receptor (RXR), and others (Niemelä et al., 2008), including non-receptor mediated mechanisms (Bournat and Brown, 2010; Kassotis and Stapleton, 2019; Luz et al., 2018). These receptor pathways are highly conserved across vertebrate species (Fu et al., 2005; Zhao et al., 2015), suggesting that mechanisms of adipogenesis are highly translatable. Indeed, active MDCs identified using 3T3-L1 cells and other *in vitro* models (increased triglyceride accumulation/differentiation, pre-adipocyte proliferation, etc.) have been routinely shown to be active *in vivo*, such as bisphenol A and tributyltin chloride, among others (Angle et al., 2013; Chamorro-Garcia et al., 2013; Li et al., 2011; Masuno et al., 2005). Compounds that modulate these receptors belong to diverse chemical classes (Fang et al., 2015; Hamers et al., 2006; Orton et al., 2011), and many are frequently detected in indoor environments and in human tissues (Hoffman et al., 2015; Kitamura et al., 2005; Meerts et al., 2000; Shen et al., 2009; Stapleton et al., 2009, 2011; Takeuchi et al., 2005). A number of these chemicals have been associated with adiposity, obesity, type 2 diabetes, and other chronic metabolic health conditions in humans (Gore et al., 2015; Heindel et al., 2015, 2017; Ruiz et al., 2018). As such, there is a critical need to ensure robust and validated models that promote highly reproducible toxicological outcomes across laboratories.

Despite a need to accurately identify MDCs, relatively little research has been performed to comprehensively evaluate reproducibility across laboratories, and comprehensively assess factors that might contribute to varying degrees of differentiation among laboratories. Previous research has described diverse differentiation success and declining performance over time with various cell bank stocks of 3T3-L1 cells (Zebisch et al., 2012). Cell culture vessel size and proprietary tissue culture coatings have also been demonstrated to influence differentiation success of 3T3-L1 cells (Mehra et al., 2007), and various cell line suppliers provide disparate protocols and techniques for eliciting maximal differentiation success (American Type Culture Collection (ATCC) ATTC, 2011; Zenbio Inc, 2015). Other providers suggest an inability to differentiate their 3T3-L1 cell stocks, with timelines of 2–5 weeks and very limited differentiation (European Collection of Authenticated Cell Cultures (ECACC), 2020). Despite these notable gaps, reproducibility studies across toxicological studies are limited, and often conducted only within the establishment of guideline assays such as Organisation for Economic Co-operation and Development (OECD) test guidelines. Chemists, in contrast, have demonstrated robust success in improving methodology and laboratory-specific measurement reliability through participation in interlaboratory reproducibility programs (Boyer et al., 1985; Ikonomou et al., 2012; M.Weiss et al., 2013; Voet et al., 1999; Wong et al., 2010). For example, recent initiatives to improve measurements for novel brominated and organophosphate flame retardants have demonstrated high precision between technical replicates within laboratories but not as strong accuracy for measuring the provided values across laboratories (Melymuk et al., 2015, 2018). Importantly, analytical reproducibility studies benefit from the concrete nature of the outcome: chemicals can be included at specific, known concentrations, making the determination of the "correct" result more straightforward than possible in toxicological studies. A limited number of studies have attempted this with endocrine outcomes such as measurement of nuclear receptor activation (Hettwer et al., 2018; Mehinto et al., 2015; Zava et al., 1982), reporting variable consistency across trials and pathways, with variances often seemingly resulting from non-harmonized protocols.

We previously published an assessment of some disparities in adipogenic cell culture systems under various conditions (Kassotis et al., 2017b). Specifically, both cell line (3T3-L1 vs. OP9) and source (ATCC vs. Zenbio 3T3-L1) had a significant impact on the responses to various chemicals. Ligands for LXR, RXR, GR, and TR promoted disparate responses between cell sources (Kassotis et al., 2017b), in some cases apparently mediated through gene expression differences. We also noted significant differences based on the cell culture plastic utilized, with different 96-well plates contributing to <50 % reduction in maximal fold induction differences and appreciably altering chemical potencies (Kassotis et al., 2017b). Cytotoxicity and proliferative response differences were also observed among plates, in some cases negatively impacting the ability to even detect chemicals acting via pre-adipocyte proliferation (Kassotis et al., 2017b). Significant differences were also noted in both triglyceride accumulation and DNA content among different differentiation induction times (7, 10, 14 days) (Kassotis et al., 2017b), and it is still unclear what differences might stem from the wide degree of heterogeneity in media additives across varying differentiation protocols. For example, while many researchers do not include dexamethasone in the differentiation cocktail (Boucher et al., 2015; Kassotis et al., 2017a, 2018; Kassotis et al., 2019; Sargis et al., 2010), others use a 1000-fold range of concentrations (Li et al., 2011; Masuno et al., 2002; Temkin et al., 2016; Zebisch et al., 2012). Differing protocols and cell culture supplies utilized may contribute to a lack of reproducibility and bias in measuring adipogenic potency and efficacy of chemicals between laboratories. Importantly, while reported previously, these factors have never been assessed in a systematic manner across laboratories.

As such, the objectives of this study were to evaluate the interlaboratory variability in the response of 3T3-L1 cells to the exposures of several chemical compounds. Given previous reports of inconsistencies in responses using this model, we sought to comprehensively evaluate the underlying factors and how they might influence differences in efficacy (magnitude of effects) and/or potency (concentration of effects) for both triglyceride accumulation (marker of differentiation success) and pre-adipocyte proliferation (marker of cell number). To accomplish this, we assessed three blinded test chemicals (bisphenol A, BPA; tributyltin chloride, TBT; pyraclostrobin) between ten laboratories in the United States, Canada, Italy, Norway, and the United Kingdom. These data should provide comprehensive insight into the most important factors that influence the assay's responses and to inform strategies to increase interlaboratory reproducibility.

2. Materials and methods

2.1. Chemicals

Chemicals for use in bioassays were purchased as follows:

rosiglitazone (Sigma cat # R2408, ≥ 98 %), pyraclostrobin (Sigma cat # 33696, 99.9 %), tributyltin chloride (TBT; Aldrich cat # T50202, 96 %), and bisphenol A (BPA; Sigma cat # 239658, >99 %). Stock solutions were prepared in 100 % cell-culture grade DMSO (Sigma cat # D2650) and stored at − 20 °C between uses. Laboratories were recruited via a scientific conference discussion and coordination via a metabolism disruption research listserv. Rosiglitazone (1 mM) was provided as a labeled amber glass vial, and laboratories were instructed to dilute the solution 1000-fold and then perform four 10-fold dilutions (0.1 nM – 1 μM in contact with cells). Pyraclostrobin, TBT, and BPA (10 mM) were provided as blinded chemicals (Chemicals A, B, and C, respectively) in amber glass vials, and laboratories were instructed to perform the same 1000-fold dilution and four 10-fold dilution scheme (1 nM – 10 μM).

2.2. Sample shipments

Packages were shipped to participating laboratories in insulated Styrofoam shipping boxes with between one and five kilograms of dry ice, depending on distance of shipment. All packages were shipped priority overnight, generally resulting in next day domestic delivery; however, international shipments often took as long as ten days to clear customs and be delivered by local couriers. Additional dry ice was added by Fedex as necessary to ensure packages arrived frozen, and this was confirmed by receiving laboratories upon receipt. Given the long length of delivery for certain participating international laboratories, a cryo-shipper was utilized (Chart MVE BL-7) to allow for improved assurance on the frozen cell stock.

2.3. Testing, differentiation, and evaluation protocol

The following materials were provided to each laboratory (Fig. 1). One vial of murine 3T3-L1 cells (Zenbio cat# SP-L1-F, lot# 3T3062104; Research Triangle Park, NC), one vial NucBlue® Live ReadyProbes® Reagent (Thermo cat # R37605), one vial of laboratory prepared Nile Red reagent (40 μg/mL Nile Red in acetone; Sigma 72485−100MG), one sleeve containing six black clear-bottom 96-well tissue culture plates (Greiner cat # 655090; four for assays and two replacements), one vial of DMSO for use as a solvent control, one vial of rosiglitazone for use as a positive control, and three vials containing the blinded test chemicals. A shared differentiation protocol, differentiation data sheet, and test instructions were emailed to each participating laboratory. Differentiation data sheets included spaces to include raw fluorescence values from Nile Red and NucBlue outputs, a plate map to denote placement of all test chemicals and replicates, and spaces to specify differentiation details, media constituents, and other assay details. The shared differentiation protocol was adapted based on a previously published protocol (Zebisch et al., 2012) and has been described in detail previously (Kassotis et al., 2017a, b). A short protocol was provided to detail the desired test conditions. Laboratories were requested to prepare media according to the shared differentiation protocol and according to their own protocols and then test blinded test chemicals under four sets of conditions: 1) using their laboratory-specific differentiation protocol, media, and supplies along with laboratory-specific stock of 3T3-L1 cells (LC/LP), 2) using their laboratory-specific differentiation protocol, media, and supplies along with the provided stock of 3T3-L1 cells (SC/LP), 3) using the provided differential protocol, media prepared according to the shared protocol, and using the included cell culture plates along with laboratory-specific stock of 3T3-L1 cells (LC/SP, and 4) using the provided differential protocol, media prepared according to the shared protocol, and using the included cell culture plates along with the provided stock of 3T3-L1 cells (SC/SP, Fig. 1).

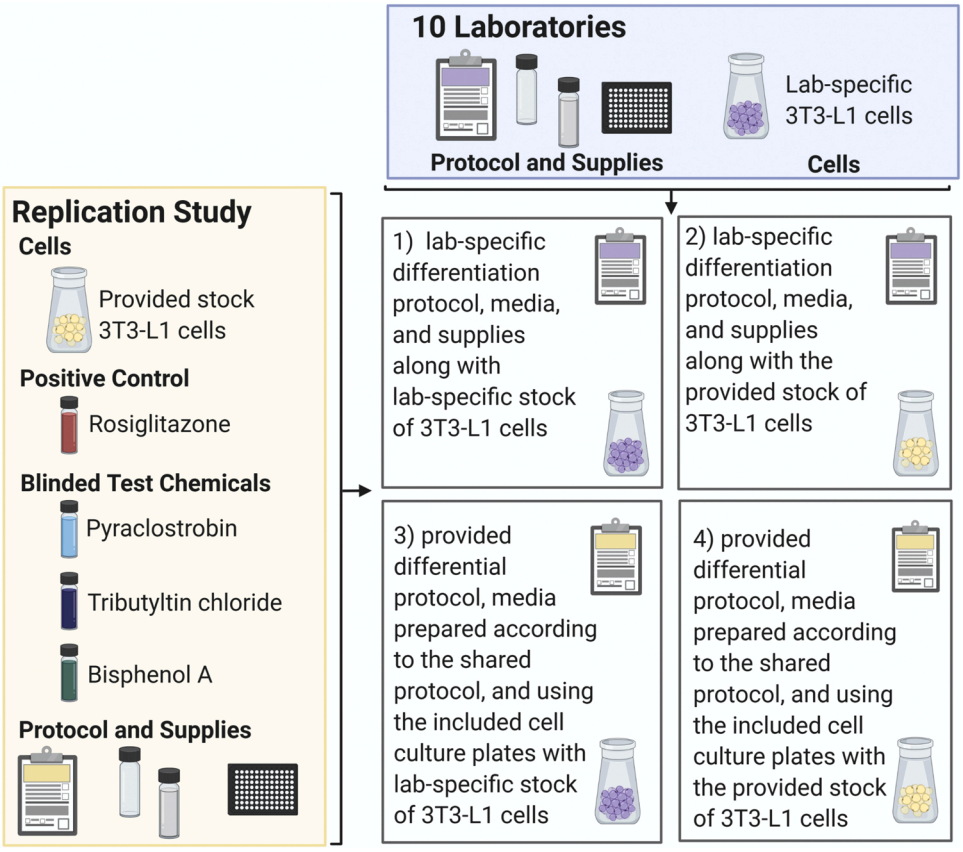

Fig. 1. Study Design and Test Conditions. Schematic of the study design and test conditions. Ten laboratories were recruited from the United States and other countries. The following materials were provided to each laboratory: one vial of 3T3-L1 cells, one vial NucBlue® Live ReadyProbes® Reagent, one vial of laboratory prepared Nile Red reagent, black clear-bottom 96-well tissue culture plates, DMSO, rosiglitazone, and three vials containing the blinded test chemicals. A shared differentiation protocol, differentiation data sheet, and test instructions were emailed to each participating laboratory. A short protocol was provided to detail the desired test conditions. Laboratories were requested to prepare media according to the shared differentiation protocol and according to their own protocols and then test blinded test chemicals under four sets of conditions: 1) using their laboratory-specific differentiation protocol, media, and supplies along with laboratory-specific stock of 3T3-L1 cells, 2) using their laboratory-specific differentiation protocol, media, and supplies along with the provided stock of 3T3-L1 cells, 3) using the provided differential protocol, media prepared according to the shared protocol, and using the included cell culture plates along with laboratory-specific stock of 3T3-L1 cells, and 4) using the provided differential protocol, media prepared according to the shared protocol, and using the included cell culture plates along with the provided stock of 3T3-L1 cells (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

2.4. 3T3-L1 cell care and differentiation assays

Zenbio 3T3-L1 cells were provided to all laboratories and were used for SC test conditions. Laboratory-specific cell sources of 3T3-L1 cells (LC) varied depending on the laboratory and are detailed within Table 1. The shared protocol called for cells to be maintained in pre-adipocyte media (Dulbecco's Modified Eagle Medium – High Glucose; DMEM-HG; Gibco cat# 11995, supplemented with 10 % bovine calf serum and 1% penicillin and streptomycin; Gibco cat# 15140). These cells were seeded in pre-adipocyte media into 96-well tissue culture plates (Greiner cat # 655090) at approximately 30,000 cells/well and grown to confluency; after confluency, cells were allowed 48 h to undergo growth arrest and initiate clonal expansion. Media was then replaced with controls, and/or blinded test chemicals using a DMSO vehicle (at 0.1 %) in differentiation media (DMEM-HG with 10 % fetal bovine serum, 1% penicillin/streptomycin, 1.0 μg/mL human insulin, and 0.5 mM 3-isobutyl-1-methylxanthine, IBMX). After 48 h of differentiation induction, media was replaced with fresh dilutions of test chemicals in adipocyte maintenance media (differentiation media without IBMX), and this media was refreshed every 2–3 days until assay, ten days after induction. Laboratory-specific protocols varied among laboratories, with variations to this general protocol specified within Table 1. 3T3-L1 cells were utilized between passages 8 (at shipment) and 12 (provided Zenbio cells) or as noted in Table 1 for laboratory-specific cells, and were maintained in a sub-confluent state until differentiation.

2.5. 3T3-L1 triglyceride accumulation, cell proliferation, and cell viability measurements

Fluorescence endpoint measurements were measured using a plate reader for the shared protocol tests and using standard laboratory practice for laboratory-specific protocol tests. For shared protocol experiments, media was removed from plates, and cells were rinsed with Dulbecco's phosphate-buffered saline (DPBS; Gibco cat # 14040) before replacing with 200 μL of a live-cell dye mixture (19 mL DPBS, 1 drop/mL NucBlue® Live ReadyProbes® Reagent (cell proliferation/cytotoxicity measure of DNA content; Thermo cat # R37605) and 500 μL Nile Red (intracellular lipid measure of triglyceride accumulation; 40 μg/mL in acetone; Sigma 72485−100MG) per plate). Plates were protected from light and incubated at room temperature for approximately forty minutes; fluorescence was then measured on plate readers with excitation 485 nm/emission 572 nm (previously demonstrated to be ideal wavelengths for intracellular neutral lipids (Greenspan et al., 1985)) and/or 485/535 (more accessible, generally used) for Nile Red and 360/460 for NucBlue®. Measurement wavelengths varied for laboratory-specific protocols, with most laboratories also providing their data from plate readers using these wavelengths.

All laboratories provided raw data, and as such, normalizations and activity calculations and determinations were made in a uniform manner for all data. For triglyceride accumulation data, percent activities were calculated relative to the maximal rosiglitazone-induced fold induction over intra-assay differentiated vehicle control (0.1 % dimethylsulfoxide, DMSO) responses, after correcting for background fluorescence. Rosiglitazone was utilized as the positive control herein (provided as an unblinded chemical stock to all laboratories) due to selective, robust, and potent activation of PPARγ (Lehmann et al., 1995; Seimandi et al., 2005; Spiegelman, 1998) and in order to provide ease of comparisons across laboratories and experiments. DNA content was calculated as percent change from differentiated vehicle control (0.1 % DMSO) responses for each chemical at each concentration and was then used to normalize total triglyceride values to obtain triglyceride content per unit DNA (proxy for triglyceride accumulation per cell). Four technical (replicates within each assay plate) and two biological replicates (separate cell passages/assays) were requested for every test chemical and concentration herein.

2.6. Statistical analysis

Data for adipogenic activities (triglyceride accumulation and pre-adipocyte proliferation) are presented as means ± standard error of the mean (SEM) from replicates. First, the median of four technical replicates (within plates) was determined and then medians of two or three (depending on laboratory) biological replicates (separate experimental plates) were averaged to provide a final value for each chemical for each laboratory under each test condition. Efficacy values were defined as the percent maximal activity relative to the rosiglitazone-induced maximal response for each laboratory and test condition. Relative potency values (effective concentration, EC_{20}; concentration of each chemical that exhibits 20 % of assay maximal activity, respectively) values were estimated from raw fluorescence data, setting the axis to 20 % of response and estimating the concentration at which the response curve passes this activation value. Values were extrapolated as necessary for efficacy and potency values for samples approaching the cut-off; potency values were not extrapolated when there was no apparent activity (samples not approaching 20 % activity), as potencies cannot be calculated for inactive chemicals/samples. Sensitivity was defined as the lowest concentration that exhibited a significant effect for each chemical above its own solvent control under each set of test conditions for each laboratory. To ease comparisons among laboratories, a uniform limit of quantification was set between laboratories as follows: a biological activation threshold approach was utilized, where the variation in the differentiated solvent (0.1 % DMSO) control was calculated as the average differentiated solvent value plus three times the standard deviation of the differentiated solvent control response. All reproducibility metrics were based on the raw, unadjusted data to assess variance within experimental groups and across laboratories; they were calculated by subtracting the average experimental group response from each individual laboratory response and dividing by the standard deviation. Values further from zero in either direction represent greater variation from the average experimental group response. Responses were analyzed using a one-way ANOVA and Dunnett's post-hoc test. Differences between treatment and control groups were considered statistically significant at $p < 0.05$.

3. Results

Ten participating laboratories were asked to test three blinded test chemicals and one standard positive control chemical (rosiglitazone) under a defined concentration range and four defined sets of conditions (Fig. 1). Laboratories were asked to test using a shared differentiation protocol to assess the potential differences contributed by differentiation protocol variations and using a shared source of 3T3-L1 cells to assess potential differences contributed by variations in cell sources. Activity determinations were assessed as presence of significant activity above the biological activation threshold as described in the methods, and were based on triglyceride accumulation (standard marker for extent of differentiation) and pre-adipocyte proliferation (increase in DNA content relative to differentiated solvent controls).

3.1. Rosiglitazone responses across laboratories and test conditions

Rosiglitazone was tested under four sets of conditions within each laboratory (Fig. 2). Given that triglyceride accumulation efficacies for rosiglitazone were normalized to the maximal intra-assay rosiglitazone-induced response, maximal efficacies (percent activation) could not be compared and fold induction responses relative to the differentiated solvent control were used instead. Maximal triglyceride accumulation fold inductions for rosiglitazone did not vary considerably across the four test conditions (2.5–3.2-fold), with the highest fold induction responses observed in the shared protocol groups (Table S1, Fig. 2). More variation was observed among laboratories relative to among test conditions, with fold inductions ranging from 1.3–6.0 (agnostic of test

Table 1
Descriptive Statistics for Laboratory-Specific Differentiation Protocols.

	Laboratory cells/laboratory protocol														
Participating Laboratory	Cell Source	Cell Lot #	Passage	Test Chem Length (d)	Diff. Length (d)	Cocktail Length (d)	TC Plate	TC plate #	Base Media	IBMX	Insulin	Dexamethasone	Biotin	Ca Pantothenate	FBS
Lab 1	ATCC	Batch 6	u + 3	10	10	2	Falcon	353219	DMEM-LG	0.5 mM	0.6 mg/mL	250 nM	–	–	Wisent #090150, Lot#115680
Lab 2	ATCC	700009858	u + 3, 9	10	10	2	Greiner BioOne	82050−748	DMEM-HG	0.5 mM	1 mg/mL	250 nM	–	–	Euroclone, ECS0180 L, lot # EUS0040912
Lab 3	ATCC	63343749	u + 3	10	10	2	Corning	07−200-565	DMEM-HG	0.5 mM	1 mg/mL	250 nM	–	–	Sigma F6178, Lot #013K8411
Lab 4	ATCC	MBX clone[1]	u + 5	10	10	2	Greiner BioOne	82050−748	DMEM/F12	0.5 mM	1 mg/mL	–	33 mM	17 mM	Fisher Scientific (Gibco) 10437−028
Lab 5	ATCC	2268173	u + 9	14	14	2	Greiner BioOne	82050−748	DMEM-HG	0.5 mM	1 mg/mL	20 nM	–	–	Sigma #F2442, lot 1982183
Lab 6	ATCC	–	u + 3, 8	10	10	2	Greiner BioOne	82050−748	DMEM-HG	0.5 mM	10 mg/mL	1 mM	–	–	GE Healthcare life sciences, SH30910.03HI
Lab 7	ATCC	63891946	u + 3	10	10	2	Greiner CellStar	655090	DMEM-HG	0.5 mM	10 mg/mL	–	–	–	GIBCO; Ref: 10270−106; Lot: 41G1780K
Lab 8	Green[2]	–	11	8	8	2	Falcon	353219	DMEM-HG	125 mM	2.5 ug/mL	1.25 mM	2 mg/L	1 mg/L	Gemini #100−106
Lab 9	Zenbio	3T3L1062104	10	10	10	2	Greiner CellStar	M0562−32EA	DMEM-HG	0.5 mM	1 mg/mL	–	–	–	Sigma F9665
Lab 10	ATCC	–	u + 4	11	11	2	USA Scientific	5665−5087	DMEM-HG	0.5 mM	10 mg/mL	1 mM	–	–	Hyclone SH30541.03
Shared	Zenbio	3T3L1062104	8	10	10	2	Greiner BioOne	82050−748	DMEM-HG	0.5 mM	1 mg/mL	–	–	–	

Descriptive statistics for laboratory-specific differentiation protocols provided by each participating laboratory. These cells and differentiation details were utilized for the LC and LP conditions only. The "Shared" cell source was utilized for all SC conditions, and the shared protocol (Supplemental information) for all SP conditions. All laboratories completed tests using each of the four defined test conditions.
u = unknown passage number.
Differentiation length specifies the total duration of adipocyte differentiation and adipocyte maintenance media treatment, whereas cocktail length specifies just the duration of the differentiation induction media treatment.

[1] 3T3-L1 MBX clone utilized for these experiments: designed to ensure more complete adipocyte differentiation and insulin sensitivity; unknown lot #.

[2] Green H (isolating laboratory) source as gift from Philip Pekala (same apparent source as commercial Zenbio cells); no reported lot #.

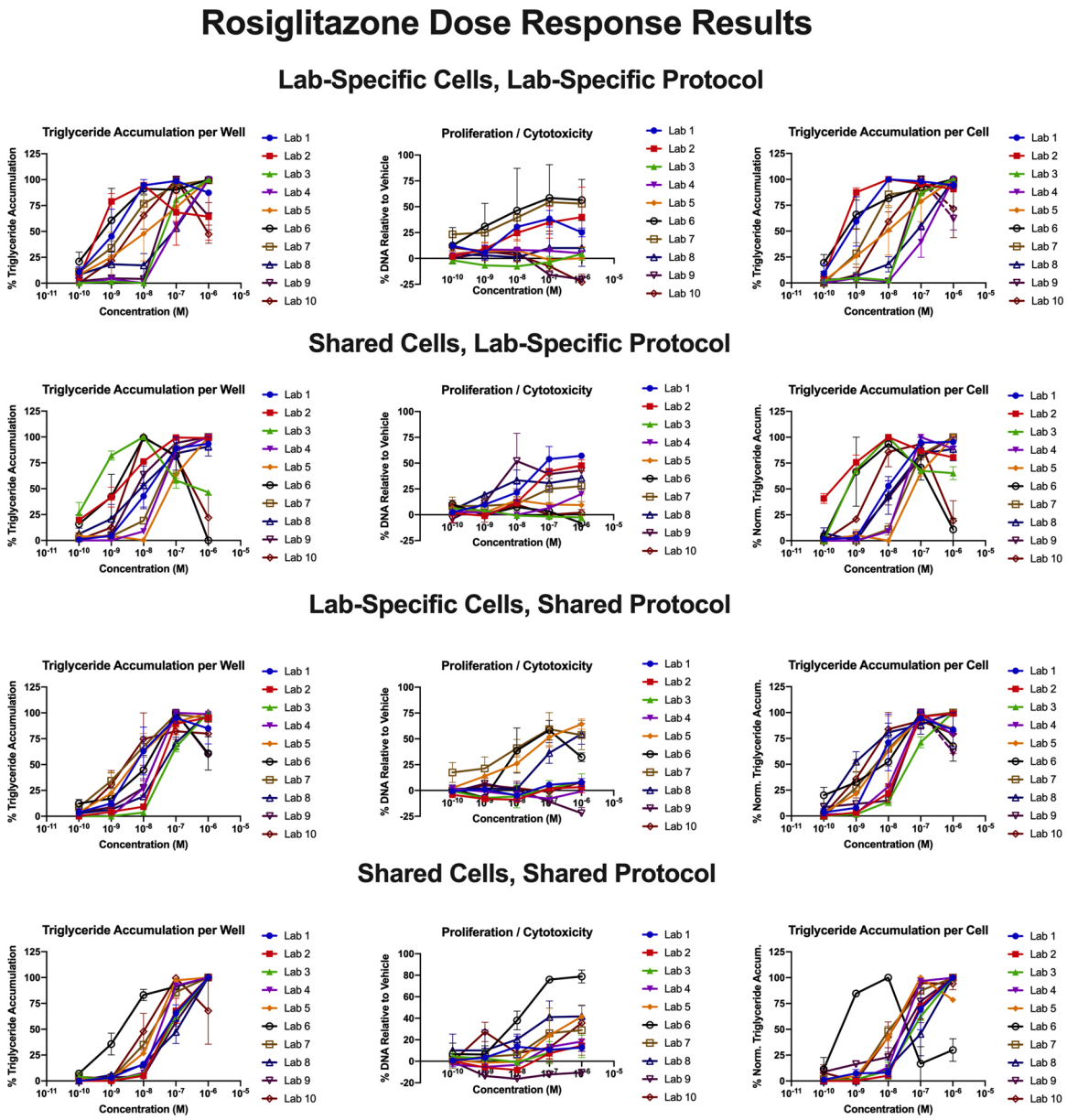

Fig. 2. Rosiglitazone Responses Across Laboratories and Conditions. Comparison of dose responses for rosiglitazone (provided, not blinded) across the ten participating laboratories. Responses are provided as raw triglyceride accumulation per well of tissue culture plate, normalized to maximal rosiglitazone-induced response using those test conditions (left column); cell proliferation and/or cytotoxicity as per Hoechst DNA dye (middle column); and normalized triglyceride accumulation per cell, normalized to DNA content of that treatment (right column). Laboratories were asked to test equivalent concentrations using their laboratory stock of 3T3-L1 cells and their laboratory differentiation protocol (top row), using the provided stock of 3T3-L1 cells and their laboratory protocol (second row), using their laboratory stock of 3T3-L1 cells and the provided protocol (third row), and using the provided 3T3-L1 cells and provided protocol (bottom row).

conditions; Table S2) and Z-scores exceeding ± 1.0. Lower variances were again observed in the shared protocol groups. With some exceptions, most laboratories reported the highest fold induction responses using the shared cell/shared protocol (SC/SP) test conditions. Dose responses were varied when laboratories utilized laboratory-specific cells and laboratory-specific protocols (LC/LP), with more than two orders of magnitude difference in potencies, three orders of magnitude difference in lowest observed effect level (LOEL; lowest tested concentration with significant increase above baseline), two orders of magnitude in maximal effective concentration, and a wide range of fold inductions relative to differentiated solvent controls (0.1 % DMSO; Fig. 2, Table S2). Similar variances were observed when all laboratories used the SC, with no apparent improvement in responses across laboratories, though variances improved slightly in the SP conditions (e.g. maximal concentrations demonstrated much greater consistency). Mean potencies (EC_{20}) for triglyceride accumulation were very similar, with values ranging from 0.009 μM for SC/SP, 0.004 for LC/SP, 0.006 for SC/LP, and 0.008 for LC/LP, though these trends were not evidence for pre-adipocyte proliferation (mean potencies of 18.45 μM for SC/SP, 146.72 for LC/SP, 0.17 for SC/LP, and 6.00 for LC/LP). For most laboratories, fold induction increased when using the shared protocol (with or without shared cells). Mean SC/SP metrics for rosiglitazone performance (triglyceride accumulation) were 0.04 ± 0.01 μM for LOEL, 0.72 ± 0.14 μM for maximal response concentration, 3.2 ± 0.4 for fold induction of response, and 0.02 ± 0.01 μM for potency (Table S2).

Pre-adipocyte proliferation means/deviations did not appreciably change across the test groups, though a greater proportion of laboratories reported significant adipogenic activity when using SC (with the most consistency using SC/SP; Fig. 1, Table S1, Table S2). Only one laboratory reported inactivity for pre-adipocyte proliferation using the

SC/SP, while 30–60 % of laboratories reported inactivity when assessed under the other test conditions. Efficacy ranged from < limit of detection (LOD) to ~80 % increased DNA content relative to the differentiated solvent control, and the magnitude of proliferation seemed to be highest in the SC/SP group (Table S1, Table S2). Average reproducibility metrics for the SC/SP conditions included LOEL (mean: 0.26 ± 0.13 μM), maximal efficacy (mean: 32.1 % ± 6.7 %), concentration at which maximal response was induced (mean: 0.9 ± 0.1 μM), and potency (18.45 ± 15.63 μM; Table S2). Maximal response concentration was nearly unanimous using the SC/SP, but large variances and potency shifts were observed in the other test groups (Fig. 1, Table S2). Some low-level cytotoxicity was observed for certain laboratories under certain test conditions, only at 10 mM, and all using the LC test conditions.

3.2. Pyraclostrobin responses across laboratories and conditions

Pyraclostrobin is a fungicide that has been reported previously to exhibit adipogenic effects (Kassotis et al., 2017a; Luz et al., 2018), though reportedly not through PPARγ activation (Luz et al., 2018). Maximal triglyceride accumulation efficacies for pyraclostrobin varied considerably across test conditions (<LOD – 400 %), with generally lower variances observed in the SP groups (Table S3, Table S4, Figure S1). Under each set of test conditions, there was at least one laboratory below or near the LOD for triglyceride accumulation. There appeared to be two groupings among laboratories for maximal efficacy

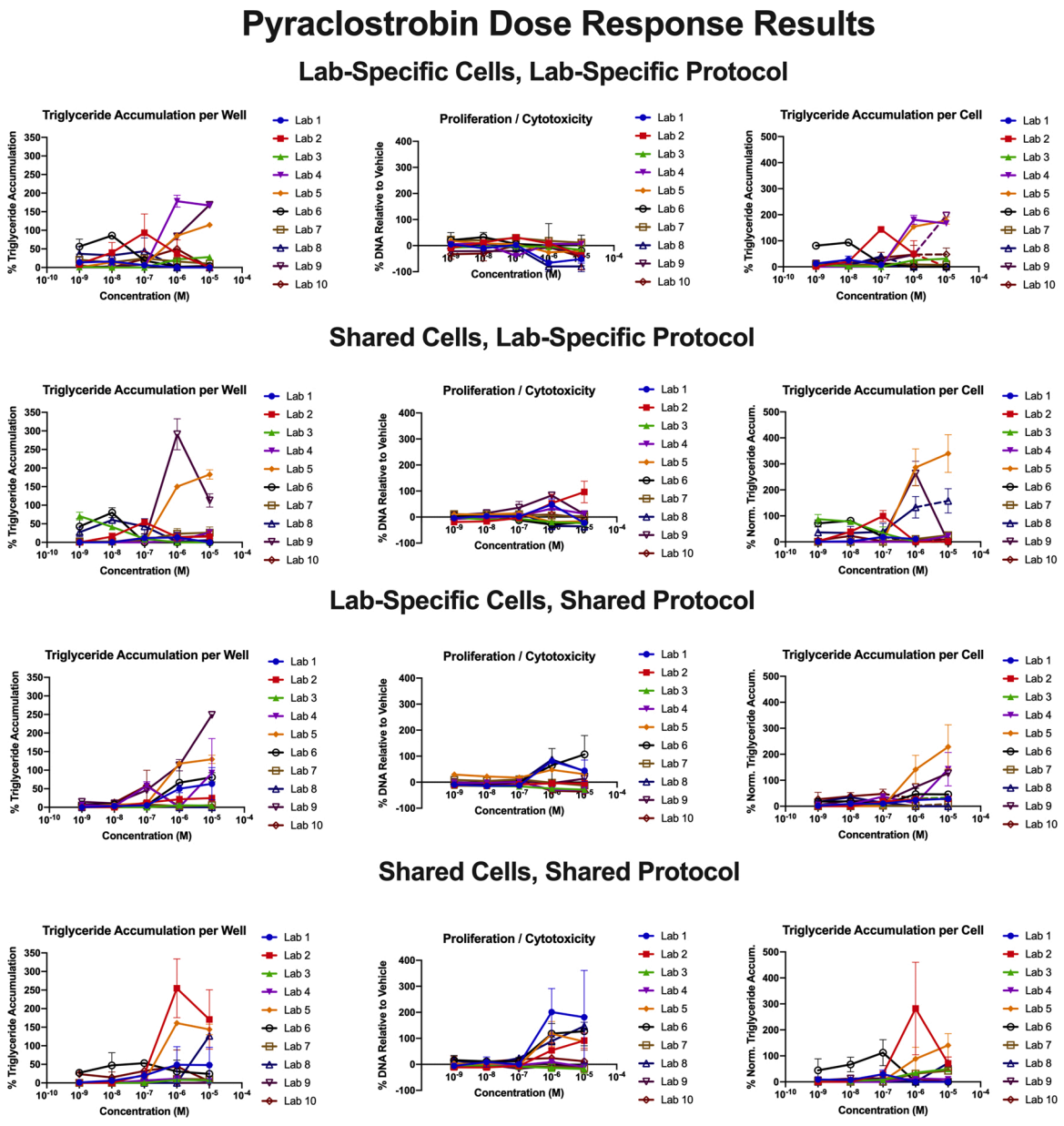

Fig. 3. Pyraclostrobin Responses Across Labs and Conditions. Comparison of dose responses for pyraclostrobin across the ten participating laboratories. Mean responses ± standard error of the mean (SEM) are provided as raw triglyceride accumulation per well of tissue culture plate, normalized to maximal rosiglitazone-induced response using those test conditions (left column); cell proliferation and/or cytotoxicity as per Hoechst DNA dye (middle column); and normalized triglyceride accumulation per cell, normalized to DNA content of that treatment (right column). Laboratories were asked to test equivalent concentrations using their lab stock of 3T3-L1 cells and their laboratory differentiation protocol (top row), using the provided stock of 3T3-L1 cells and their laboratory protocol (second row), using their lab stock of 3T3-L1 cells and the provided protocol (third row), and using the provided 3T3-L1 cells and provided protocol (bottom row). Dashed lines represent concentrations at which cytotoxicity was observed (significant decreased DNA content at that concentration).

results (low/no activity responders and high activity responders; Fig. 3). Three laboratories were consistently low responders, one was consistently a high responder, four laboratories were high responders in three of the four test conditions (test condition varied by laboratory), and two laboratories were split in responses across test conditions. The LOEL and maximal response concentrations appeared to be more consistent when using the SP, albeit with slightly lower potencies (particularly for the SC/SP condition; Fig. 3, Table S4). Almost all laboratories identified this chemical as active for inducing triglyceride accumulation with the exception of one each using LC/SP and SC/SP (blue boxes denote active determinations, whereas red denote inactive; Fig. 3, Table S3, Table S4).

For pre-adipocyte proliferation, maximal efficacies also varied widely (<LOD – 201 %; Fig. S1, Fig. 3, Table S4). A greater proportion of laboratories reported significant proliferative activity using the SC: 30 % using LC/LP, 40 % using LC/SP, 70 % using SC/LP, and 60 % using SC/SP. In most cases, greater efficacies for proliferation were observed when in SP groups. Greater consistency but lower sensitivity/potency were observed for the SC/SP group. At the level of activity determination, there was an almost even split in responses (Fig. 3D). Only five laboratories reported pyraclostrobin as active for proliferation in three or more four conditions (generally reported as inactive using LC/LP). Significant toxicity was observed for some laboratories under certain test conditions. No toxicity was observed for any laboratory in the SC/SP condition, but 40 % of laboratories reported significant toxicity in both the SC/LP (one at 1 and 10 mM and three at 10 mM only) and LC/SP (three at 1 and 10 mM and one at 10 mM only) conditions, and 50 %

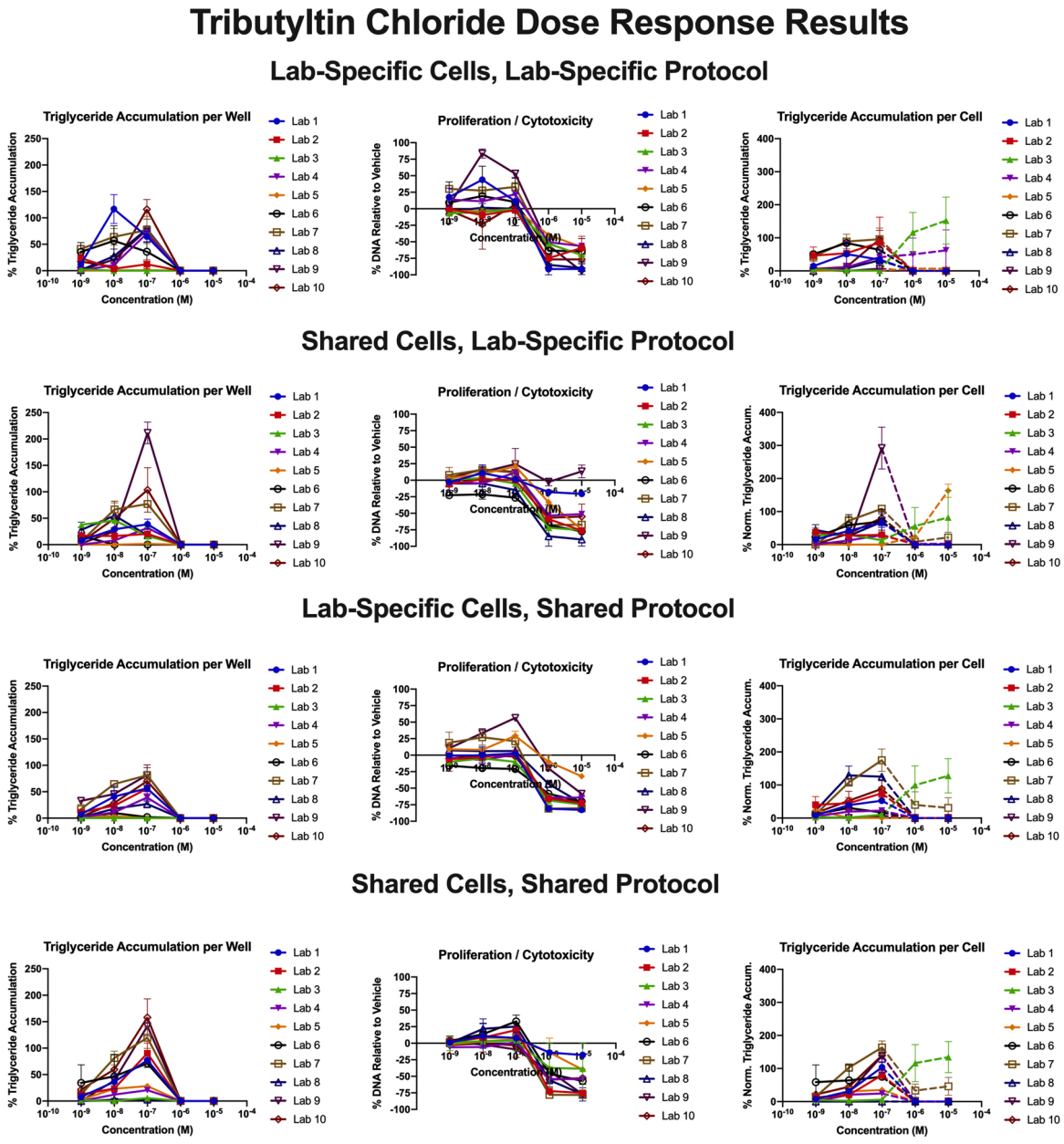

Fig. 4. Tributyltin Chloride Responses Across Labs and Conditions. Comparison of dose responses for tributyltin chloride across the ten participating laboratories. Mean responses ± standard error of the mean (SEM) are provided as raw triglyceride accumulation per well of tissue culture plate, normalized to maximal rosiglitazone-induced response using those test conditions (left column); cell proliferation and/or cytotoxicity as per Hoechst DNA dye (middle column); and normalized triglyceride accumulation per cell, normalized to DNA content of that treatment (right column). Laboratories were asked to test equivalent concentrations using their lab stock of 3T3-L1 cells and their laboratory differentiation protocol (top row), using the provided stock of 3T3-L1 cells and their laboratory protocol (second row), using their lab stock of 3T3-L1 cells and the provided protocol (third row), and using the provided 3T3-L1 cells and provided protocol (bottom row). Dashed lines represent concentrations at which cytotoxicity was observed (significant decreased DNA content at that concentration).

reported significant toxicity using LC/LP (two at 0.1, 1, and 10 mM; two at 1 and 10 mM; and one at 10 mM only).

3.3. Tributyltin chloride (TBT) responses across laboratories and conditions

TBT is a biocide that has been reported previously to exhibit robust adipogenic effects *in vitro* (Grun et al., 2006; Li et al., 2011; Pereira-Fernandes et al., 2013) and *in vivo* (Chamorro-Garcia et al., 2013; Penza et al., 2011). Maximal triglyceride accumulation efficacies for TBT varied considerably across test conditions (<LOD – 429 %), with lower variances in the LC/LP and SC/SP groups (Fig. 4, Fig. S2, Table S5, Table S6). Relative to pyraclostrobin, wide variances were observed even within laboratories, and two response groupings (high and low consistency) were observed. Five laboratories ranged widely in triglyceride accumulation responses across test conditions, and five laboratories were more consistent (moderate to high triglyceride accumulation). No differences were observed in LOELs, though maximum response concentrations and potencies appeared more consistent in the SC/SP group, albeit with slightly lower potencies (Fig. 4, Table S6). Almost all laboratories identified this chemical as active for inducing triglyceride accumulation, with the exception of two laboratories using the LC and one using the SC/SP (Fig. 4D).

For pre-adipocyte proliferation, maximal efficacies had low consistency (<LOD – 84 %; Fig. S2, Fig. 4, Table S6), but the highest consistencies for efficacy, LOEL, and potencies were observed in the SC/SP group. At the level of activity determination, 50–70 % of laboratories reported TBT as inactive for proliferation across all test conditions

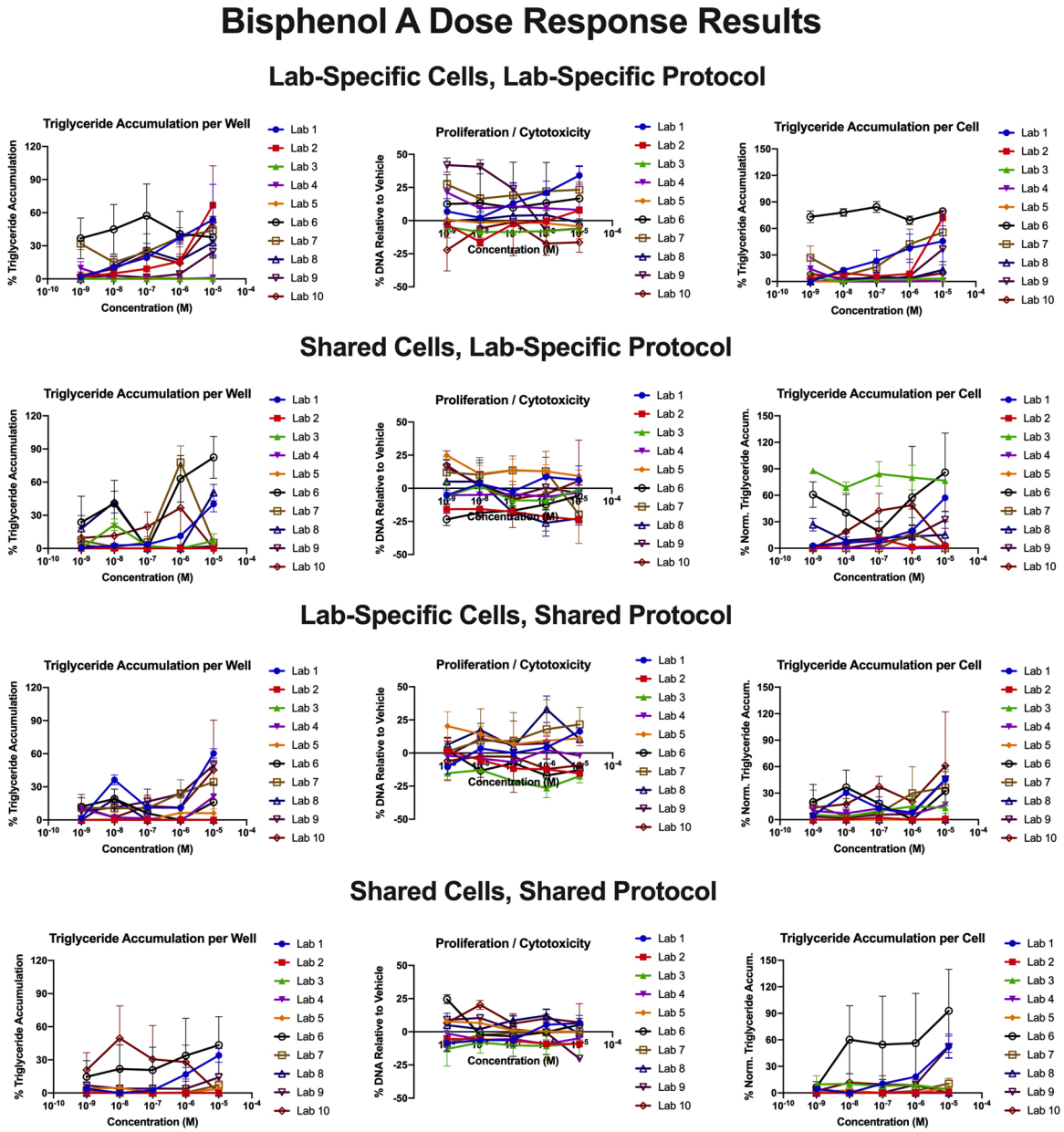

Fig. 5. Bisphenol A Responses Across Labs and Conditions. Comparison of dose responses for bisphenol A across the ten participating laboratories. Mean responses ± standard error of the mean (SEM) are provided as raw triglyceride accumulation per well of tissue culture plate, normalized to maximal rosiglitazone-induced response using those test conditions (left column); cell proliferation and/or cytotoxicity as per Hoechst DNA dye (middle column); and normalized triglyceride accumulation per cell, normalized to DNA content of that treatment (right column). Laboratories were asked to test equivalent concentrations using their lab stock of 3T3-L1 cells and their laboratory differentiation protocol (top row), using the provided stock of 3T3-L1 cells and their laboratory protocol (second row), using their lab stock of 3T3-L1 cells and the provided protocol (third row), and using the provided 3T3-L1 cells and provided protocol (bottom row).

(Fig. 4D), with no laboratory reporting consistent activity across test conditions. Two laboratories identified TBT as inactive for proliferation for all conditions and two as inactive in three of four conditions. Clear cytotoxicity was observed for TBT at 1 and 10 mM, with high consistency across laboratories and conditions (Fig. S2).

3.4. Bisphenol A responses across laboratories and conditions

BPA is a synthetic chemical used often as a cross-linker in the synthesis of some plastics and can be found in some consumer products. (vom Saal et al., 2007; Welshons et al., 2006) BPA has been extensively described to disrupt metabolic health *in vitro* (Masuno et al., 2002; Sargis et al., 2010; Taxvig et al., 2012), *in vivo* (Angle et al., 2013; Somm et al., 2009; Vom Saal et al., 2012), and in epidemiological studies (Carwile and Michels, 2011; Rochester, 2013; Trasande et al., 2012). Maximal triglyceride accumulation efficacies for BPA varied across test conditions (<LOD – 93 %), with apparent lower variances in the SP groups (Table S7, Table S8, Fig. S3, Fig. 5). Relatively wide variances were observed even within laboratories, and three response groupings (inactive, low/moderate and moderate/high activity) were observed. Given lower reported activity, potency and LOEL comparisons were difficult to ascertain (Fig. 5, Table S8). Nearly all laboratories identified this chemical as active for triglyceride accumulation, though with considerable variation across test conditions and no apparent difference between groups (Fig. 5).

For pre-adipocyte proliferation, maximal efficacies were again less consistent but with a smaller dynamic range (<LOD – 42 %; Fig. S3, Fig. 5, Table S8) that hindered evaluation of several reproducibility metrics. Greater agreement was observed for inactivity of BPA on the proliferation metric (Fig. 5D). Three laboratories reported BPA as inactive for proliferation across test conditions, three laboratories as inactive in three of four test conditions, and the remaining laboratories reported higher rates of activity. Within test conditions, the greatest consistency was observed using the SC/LP and SC/SP conditions, with 80 % of laboratories reporting BPA as inactive for proliferation. No significant toxicity was reported for BPA at any test concentration by any laboratory and in any test condition.

4. Discussion

These results confirm that repeatability of adipogenic (pre-adipocyte proliferation) and lipogenic (triglyceride accumulation) responses utilizing 3T3-L1 cells are highly variable across laboratories, which can be problematic for reproducibility and data comparability. There have been previous reports of inconsistencies in adipogenic determinations for specific chemicals (Kassotis et al., 2017b), but this has not previously been evaluated in a blinded, systematic manner. While standard in most areas, reproducibility studies are relatively uncommon for toxicological outcomes, and as a consequence, reproducibility across laboratories is not well-appreciated. While bioactivities (efficacies/potencies) varied considerably across laboratories and test conditions, activity determinations (active/inactive) were more consistent, which suggests that most laboratories can accurately identify MDCs. Though importantly, even these determinations were less consistent when using LC/LP, suggesting that standardization may greatly improve reproducibility between laboratories (at least for triglyceride accumulation) and thus confidence in reported outcomes. The most consistent results across test conditions and chemicals tested were generally observed in the LC/SP and SC/SP groups. While potencies were lower for rosiglitazone-induced triglyceride accumulation in the SC/SP group, this pattern did not carry over to other test chemicals or the pre-adipocyte proliferation metric, and this group had the most potent responses for other test chemicals.

Pyraclostrobin was selected for screening given the previously reported non-traditional mechanism(s) of action (antagonism of TRb and/or mitochondrial dysfunction) and robust adipogenic response (Kassotis et al., 2017a; Luz et al., 2018). A high degree of consistency was observed for triglyceride accumulation activity determination between laboratories and test conditions for this chemical, though magnitude and potency of responses were much more variable. Efficacy did not seem to vary based on test conditions, but instead was more laboratory-specific, with laboratories reporting more consistent responses regardless of test condition. These responses appeared independent of cell source and protocol, which could suggest more overarching variables such as fetal bovine serum source/consistency that may impact results across these conditions. Maximal response concentrations and LOELs were more consistent when using the SP, albeit with slightly lower potencies. Pre-adipocyte proliferation was much more variable (50 % of laboratories and test conditions reported active), though twice the laboratories reported significant proliferative activity using the SP. In most cases, greater proliferative responses were observed when using the SP, for which laboratories were provided tissue culture plates; it has been previously reported that the tissue culture plates used could drastically impact the proliferative response (Kassotis et al., 2017b; Mehra et al., 2007). Overall, greater consistency but lower sensitivity/potency were observed for the SC/SP group. It is worth noting that pyraclostrobin exhibited much greater activity

TBT was selected due to well-reported adipogenic effects via activation of PPAR and RXR (Chamorro-Garcia et al., 2013; Grun et al., 2006; le Maire et al., 2009; Li et al., 2011; Penza et al., 2011; Pereira-Fernandes et al., 2013). We previously demonstrated consistent expression for PPAR isoforms and RXRα in ATCC and Zenbio-sourced 3T3-L1 cells (Kassotis et al., 2017b), suggesting this chemical might have greater consistency across laboratories and test conditions. However, while nearly all laboratories and test conditions successfully identified TBT as active for promoting significant triglyceride accumulation, the broad range of efficacies, potencies, and sensitivities reported here would suggest that these methods may not be sensitive enough to accurately characterize less active chemicals. This could be due to small differences in cytotoxicity impacting our broad concentration response curves. Pre-adipocyte proliferation is a less frequently examined and/or reported endpoint in 3T3-L1 cells, with explicit reporting only becoming more standard in the last several years. Accounting for varying cell densities across wells and replicates can be achieved by normalizing the triglyceride content with the DNA content. Indeed, when examining the total triglyceride content per well only (Table S5), the number of laboratories and conditions reporting TBT inactive would increase from five to nine. Thus, the normalization to DNA content is important to accurately defining adipogenic activity, particularly given that varying cell densities across plates are common and since mature adipocytes detach easily from the plate bottom during media changes and rinses. Pre-adipocyte proliferation as its own metric, however, was considerably less consistent. Approximately half of the laboratories and test conditions reported TBT as inactive for proliferation (17/40) and half as active, demonstrating no clear correct determination for this endpoint. Greater consistency was observed when using SC/SP, however, demonstrating that greater concordance in testing may be possible if protocols were standardized.

BPA was selected due to the non-canonical mechanism of action and the less reproducible outcomes reported previously *in vitro* (Kassotis et al., 2017b). BPA has been reported to increase adipogenic gene expression (aP2) through an estrogen receptor-mediated mechanism (Boucher et al., 2014), which we reported to have differential expression between 3T3-L1 cell sources (Kassotis et al., 2017b). As expected, lower consistency was observed for BPA-induced triglyceride accumulation (65 % of laboratories/conditions reported as active). Lower variance was observed for DNA content measurements, with only 30 % of labs and test conditions reporting BPA as active for this metric. While test condition did not appreciably influence activity determinations for triglyceride accumulation, improved consistency was observed for the pre-adipocyte proliferation metric using the LC/SP and SC/SP conditions, suggesting that protocol and not cell differences were the primary

factors in these disparities.

Overall, cell source appeared to be a significant factor in variation observed between laboratories and test conditions. Zenbio-sourced cells often had lower variation than ATCC-sourced cells for chemicals, which we reported previously (Kassotis et al., 2017b). It is perhaps unsurprising that 3T3-L1 cells sourced from different companies had high variation in responses between them, though it is notable that considerable variation was observed even when comparing cells obtained from the same provider (Table 1, Figs. 2–5); cells sourced from the Green Lab were unsurprisingly more similar to the Zenbio-sourced cells. It has been previously established that ATCC maintains a variety of 3T3-L1 cell lots, which are described in ATCC protocols to differentiate to different extents (Kassotis et al., 2017b). This appears to be a common problem, as noted above, as the European Collection of Authenticated Cell Culture (ECACC) currently reports that their 3T3-L1 cells will not differentiate. It is likely, based on the varied responses reported herein and the differentiation issues noted by the providers, that continuity of this cell line has not been properly controlled across sources. This may be contributing to a portion of the divergent responses observed across laboratories. As observed here, this contributes to considerable variance even for laboratories ordering presumably the same cell line from the same supplier. This has been reported previously for MCF-7 cells obtained by two laboratories that ordered the same lot of MCF-7 cells from ATCC (Kleensang et al., 2016). Phenotypic, gene expression, metabolomic, and hormone-responsiveness were all demonstrated to clearly vary between these sub-clones even of the same lot of cells and were eventually linked to genetic variability in a single frozen lot of ATCC MCF-7 cells (Kleensang et al., 2016). As these authors recommended, future research should investigate these 3T3-L1 cell sources and lots through Direct Compararative Genome Hybridization and/or deep sequencing approaches to determine shifts in frozen cell line stocks that may be contributing to disparate responses. This should be pursued to improve response consistency and to increase confidence and transparency of results.

For a number of chemicals and laboratories, the most reproducible and least variable outcomes (lower Z-score ranges; Supplemental Tables) based on the performance metrics examined here were observed in the SC/SP group. Often, we observed reduced variance in the LC/SP group as well, suggesting that while cell source should be considered as a contributory factor in improving reproducibility, improvements can be made more readily through taking steps to harmonize differentiation protocols across laboratories. Numerous factors varied widely across laboratories (Table 1). While the differentiation timeline was often consistent, a variety of cell culture plastics were used, concentrations and presence of media additives varied considerably. The source of additives such as fetal bovine serum have been reported to vary substantially and may have contributed to varying degrees of differentiation. While isobutylmethylxanthine (IBMX) concentrations were quite consistent, insulin concentrations varied >1000-fold, and seven laboratories used dexamethasone (10-fold variation in concentrations). Even 10-fold variations in insulin concentration have been described to promote robust impacts on adipogenesis via both triglyceride accumulation and pre-adipocyte proliferation (Green and Kehinde, 1975). Additionally, dexamethasone has been demonstrated to promote potent and efficacious effects on triglyceride accumulation (Kassotis et al., 2017b). Taken together, the concentrations of these additives are likely contributory factors, though no clear trends in media additive use could be associated with specific patterns of activity (Fig. S4). While the SP groups utilized consistent detection and staining protocols, these were not consistent in the LP groups and may have contributed to some of the protocol-specific variances. Despite these factors, particularly variation likely contributed by fetal bovine serum sourcing, we need to utilize models such as this to evaluate the tens of thousands of chemicals requiring toxicological characterization. Appreciating these factors, and understanding the inherent variability, is key to making proper determinations based on studies utilizing these and similar models.

In summary, we report poor reproducibility (efficacies, potencies, and sensitivities) for several blinded test chemicals across laboratories, though this largely did not impact determination of chemical activity classification (e.g. categorized as “active” or “inactive”). While activity determinations for triglyceride accumulation were quite consistent, even this gross level of bioactivity measurement was not consistent for pre-adipocyte proliferation. These results suggest that toxicologic reproducibility assessments are warranted for other endpoints (for other assays and other metrics of differentiation success such as percent of differentiated cells, etc.) and suggests some avenues for improvements through harmonization of cell sourcing and differentiation protocols. It is also important to note that analytical reproducibility efforts have also reported high variance across laboratories (Melymuk et al., 2015, 2018); we do not believe this should reduce confidence in bioassay findings, but should be acknowledged to support accurate determinations of activity for unknown test chemicals when considering regulatory next steps. We report that the differentiation protocol and the source of the 3T3-L1 cells both contributed to variance in responses and even some dissimilar determinations of activity, suggesting that these factors may provide opportunity for reducing inter-laboratory variability. While we did not detect specific differences that we could associate with individual media additives, further research should evaluate this more explicitly. It should be noted that pyraclostrobin exhibited much greater activities than BPA, though has received as of yet very limited relative research attention.

While human mesenchymal stem cell and human pre-adipocyte cell lines are increasingly available commercially, there are wide reported variations based on sex, race and ethnicity, as well as physiological status of the donor. These models should see increased use, and direct comparisons to 3T3-L1 results, in future research; however, the decades of research on 3T3-L1 cells support continued use of this model at least until there are clearly superior models and/or clear translation of prior findings to human models.

Given the increasing need for accurate and reliable metabolic health assays and the difficulty of policing cell line providers, we suggest that future efforts focus on establishing best practices and consistency in differentiation protocols to improve translation across laboratories. There are also notable efforts ongoing through the Horizon 2020 program in the European Union to develop and validate new and diverse metabolism disrupting chemical assays *in silico*, *in vitro,* and *in vivo* (Audouze et al., 2020; Legler et al., 2020). Towards that end, we have provided the detailed shared differentiation protocol used in this study (Supplemental File 1), though further efforts towards harmonization should be undertaken with experts to weigh inclusion or exclusion of specific factors in the protocol and to consider the issues inherent with specific sources of 3T3-L1 cells.

Funding

Over-arching project supported by grants [R01 ES016099 to HMS; R00 ES030405 to CDK] from the National Institute of Environmental Health Sciences (NIEHS). Laboratory-specific funding includes local funding 2018-19 from the University of Turin to PB. SE and JL received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement GOLIATH No. 825489 and from Brunel University London. 1K22ES026208 and R01ES027863 to AV-L from NIEHS and NIEHS Z0ES102785 to SEF. MFF received funding from the Spanish Institute of Health Carlos III (grant FIS-PI16/01812).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.tox.2021.152900.

References

American Type Culture Collection (ATCC) ATTC, 2011. Chemically-induced Differentiation of ATCC CL-173 (3T3-L1) Using Single-component Commercially-available Reagents. Available: http://www.atcc.org//media/6124AF1E4C2A47CF904435117909AC25.ashx [accessed 17 May 2016.

Angle, B.M., Do RP, Ponzi D., Stahlhut, R.W., Drury, B.E., Nagel, S.C., et al., 2013. Metabolic disruption in male mice due to fetal exposure to low but not high doses of bisphenol A (BPA): evidence for effects on body weight, food intake, adipocytes, leptin, adiponectin, insulin and glucose regulation. Reprod. Toxicol. 42, 13.

Audouze, K., Sarigiannis, D., Alonso-Magdalena, P., Brochot, C., Casas, M., Vrijheid, M., et al., 2020. Integrative strategy of testing systems for identification of endocrine disruptors inducing metabolic disorders-an introduction to the Oberon project. Int. J. Mol. Sci. 21.

Boucher, J.G., Boudreau, A., Atlas, E., 2014. Bisphenol A induces differentiation of human preadipocytes in the absence of glucocorticoid and is inhibited by an estrogen-receptor antagonist. Nutr. Diabetes 4, e102.

Boucher, J.G., Boudreau, A., Ahmed, S., Atlas, E., 2015. In vitro effects of bisphenol A beta-d-glucuronide (BPA-G) on adipogenesis in human and murine preadipocytes. Environ. Health Perspect. 123, 1287–1293.

Bournat, J.C., Brown, C.W., 2010. Mitochondrial dysfunction in obesity. Curr. Opin. Endocrinol. Diabetes Obes. 17, 446–452.

Boyer, K.W., Horwitz, W., Albert, R., 1985. Interlaboratory variability in trace element analysis. Anal. Chem. 57, 454–459.

Carwile, J.L., Michels, K.B., 2011. Urinary bisphenol A and obesity: nhanes 2003-2006. Environ. Res. 111, 825–830.

Chamorro-Garcia, R., Sahu, M., Abbey, R.J., Laude, J., Pham, N., Blumberg, B., 2013. Transgenerational inheritance of increased fat depot size, stem cell reprogramming, and hepatic steatosis elicited by prenatal exposure to the obesogen tributyltin in mice. Environ. Health Perspect. 121, 359–366.

European Collection of Authenticated Cell Cultures (ECACC), 2020. ECACC General Cell Collection: 3T3-L1. Available: https://www.phe-culturecollections.org.uk/products/celllines/generalcell/detail.jsp?refId=86052701&collection=ecacc_gc#medDoc [accessed 4 February 2021].

Fang, M., Webster, T.F., Stapleton, H.M., 2015. Activation of human peroxisome proliferator-activated nuclear receptors (PPARgamma1) by semi-volatile compounds (SVOCs) and chemical mixtures in indoor dust. Environ. Sci. Technol. 49, 10057–10064.

Fu, M., Sun, T., Bookout, A.L., Downes, M., Yu, R.T., Evans, R.M., et al., 2005. A nuclear receptor atlas: 3T3-L1 adipogenesis. Mol. Endocrinol. 19, 2437–2450.

Gore, A.C., Chappell, V.A., Fenton, S.E., Flaws, J.A., Nadal, A., Prins, G.S., et al., 2015. Edc-2: the Endocrine Society's second scientific statement on endocrine-disrupting chemicals. Endocr. Rev. 36, E1–E150.

Green, H., Kehinde, O., 1975. An established preadipose cell line and its differentiation in culture. II. Factors affecting the adipose conversion. Cell 5, 19–27.

Green, H., Meuth, M., 1974. An established pre-adipose cell line and its differentiation in culture. Cell 3, 127–133.

Greenspan, P., Mayer, E.P., Fowler, S.D., 1985. Nile red: a selective fluorescent stain for intracellular lipid droplets. J. Cell Biol. 100, 965–973.

Grun, F., Watanabe, H., Zamanian, Z., Maeda, L., Arima, K., Cubacha, R., et al., 2006. Endocrine-disrupting organotin compounds are potent inducers of adipogenesis in vertebrates. Mol. Endocrinol. 20, 2141–2155.

Hamers, T., Kamstra, J.H., Sonneveld, E., Murk, A.J., Kester, M.H., Andersson, P.L., et al., 2006. In vitro profiling of the endocrine-disrupting potency of brominated flame retardants. Toxicol. Sci. 92, 157–173.

Heindel, J.J., vom Saal, F.S., Blumberg, B., Bovolin, P., Calamandrei, G., Ceresini, G., et al., 2015. Parma consensus statement on metabolic disruptors. Environ. Health 14, 54.

Heindel, J.J., Blumberg, B., Cave, M., Machtinger, R., Mantovani, A., Mendez, M.A., et al., 2017. Metabolism disrupting chemicals and metabolic disorders. Reprod. Toxicol. 68, 3–33.

Hettwer, K., Jahne, M., Frost, K., Giersberg, M., Kunze, G., Trimborn, M., et al., 2018. Validation of arxula yeast estrogen screen assay for detection of estrogenic activity in water samples: results of an international interlaboratory study. Sci. Total Environ. 621, 612–625.

Hoffman, K., Butt, C.M., Chen, A., Limkakeng, A.T., Stapleton, H.M., 2015. High exposure to organophosphate flame retardants in infants: associations with baby products. Environ. Sci. Technol. 49, 14554–14559.

Ikonomou, M.G., Kelly, B.C., Blair, J.D., Gobas, F.A., 2012. An interlaboratory comparison study for the determination of dialkyl phthalate esters in environmental and biological samples. Environ. Toxicol. Chem. 31, 1948–1956.

Kassotis, C.D., Stapleton, H.M., 2019. Endocrine-mediated mechanisms of metabolic disruption and new approaches to examine the public health threat. Front. Endocrinol. (Lausanne) 10, 39.

Kassotis, C.D., Hoffman, K., Stapleton, H.M., 2017a. Characterization of adipogenic activity of semi-volatile indoor contaminants and house dust. Environ Sci Technol In 51, 8735–8745.

Kassotis, C.D., Masse, L., Kim, S., Schlezinger, J.J., Webster, T.F., Stapleton, H.M., 2017b. Characterization of adipogenic chemicals in three different cell culture systems: implications for reproducibility based on cell source and handling. Sci. Rep. 7, 42104.

Kassotis, C.D., Kollitz, E.M., Ferguson, P.L., Stapleton, H.M., 2018. Nonionic ethoxylated surfactants induce adipogenesis in 3T3-L1 cells. Toxicol. Sci. 162, 124–136.

Kassotis, C.D., Kollitz, E.M., Hoffman, K., Sosa, J.A., Stapleton, H.M., 2019. Thyroid receptor antagonism as a contributory mechanism for adipogenesis induced by environmental mixtures in 3T3-L1 cells. Sci. Total Environ. 666, 431–444.

Kitamura, S., Kato, T., Iida, M., Jinno, N., Suzuki, T., Ohta, S., et al., 2005. Anti-thyroid hormonal activity of tetrabromobisphenol A, a flame retardant, and related compounds: affinity to the mammalian thyroid hormone receptor, and effect on tadpole metamorphosis. Life Sci. 76, 1589–1601.

Kleensang, A., Vantangoli, M.M., Odwin-DaCosta, S., Andersen, M.E., Boekelheide, K., Bouhifd, M., et al., 2016. Genetic variability in a frozen batch of MCF-7 cells invisible in routine authentication affecting cell function. Sci. Rep. 6, 28994.

le Maire, A., Grimaldi, M., Roecklin, D., Dagnino, S., Vivat-Hannah, V., Balaguer, P., et al., 2009. Activation of RXR-PPAR heterodimers by organotin environmental endocrine disruptors. EMBO Rep. 10, 367–373.

Legler, J., Zalko, D., Jourdan, F., Jacobs, M., Fromenty, B., Balaguer, P., et al., 2020. The GOLIATH project: towards an internationally harmonised approach for testing metabolism disrupting compounds. Int. J. Mol. Sci. 21.

Lehmann, J.M., Moore, L.B., Smith-Oliver, T.A., Wilkison, W.O., Willson, T.M., Kliewer, S.A., 1995. An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor y (PPARy). J. Biol. Chem. 270, 12953–12956.

Li, X., Ycaza, J., Blumberg, B., 2011. The environmental obesogen tributyltin chloride acts via peroxisome proliferator activated receptor gamma to induce adipogenesis in murine 3T3-L1 preadipocytes. J. Steroid Biochem. Mol. Biol. 127, 9–15.

Luz, A.L., Kassotis, C.D., Stapleton, H.M., Meyer, J.N., 2018. The high-production volume fungicide pyraclostrobin induces triglyceride accumulation associated with mitochondrial dysfunction, and promotes adipocyte differentiation independent of ppargamma activation, in 3T3-L1 cells. Toxicology 393, 150–159.

M.Weiss, J., Ivd, Veen, Boer, J., SPJv, Leeuwen, Cofino, W., Crum, S., 2013. Analytical improvements shown over four interlaboratory studies of perfluoroalkyl substances in environmental and food samples. Trac Trends Anal. Chem. 43, 204–216.

Masuno, H., Kidani, T., Sekiya, K., Sakayama, K., Shiosaka, T., Yamamoto, T., et al., 2002. Bisphenol A in combination with insulin can accelerate the conversion of 3T3-L1 fibroblasts to adipocytes. J. Lipid Res. 43, 676–684.

Masuno, H., Iwanami, J., Kidani, T., Sakayama, K., Honda, K., 2005. Bisphenol A accelerates terminal differentiation of 3T3-L1 cells into adipocytes through the phosphatidylinositol 3-kinase pathway. Toxicol. Sci. 84, 319–327.

Meerts, I.A., van Zanden, J.J., Luijks, E.A., van Leeuwen-Bol, I., Marsh, G., Jakobsson, E., et al., 2000. Potent competitive interactions of some brominated flame retardants and related compounds with human transthyretin in vitro. Toxicol. Sci. 56, 95–104.

Mehinto, A.C., Jia, A., Snyder, S.A., Jayasinghe, B.S., Denslow, N.D., Crago, J., et al., 2015. Interlaboratory comparison of in vitro bioassays for screening of endocrine active chemicals in recycled water. Water Res. 83, 303–309.

Mehra, A., Macdonald, I., Pillay, T.S., 2007. Variability in 3T3-L1 adipocyte differentiation depending on cell culture dish. Anal. Biochem. 362, 281–283.

Melymuk, L., Goosey, E., Riddell, N., Diamond, M.L., 2015. Interlaboratory study of novel halogenated flame retardants: INTERFLAB. Anal. Bioanal. Chem. 407, 6759–6769.

Melymuk, L., Diamond, M.L., Riddell, N., Wan, Y., Vojta, S., Chittim, B., 2018. Challenges in the analysis of novel flame retardants in indoor dust: results of the INTERFLAB 2 interlaboratory evaluation. Environ. Sci. Technol. 52, 9295–9303.

Niemelä, S., Miettinen, S., Sarkanen, J.R., Ashammakhi, N., 2008. Adipose tissue and adipocyte differentiation: molecular and cellular aspects and tissue engineering applications. In: Topics in tissue engineering, Vol. 4, 1–26.

Orton, F., Rosivatz, E., Scholze, M., Kortenkamp, A., 2011. Widely used pesticides with previously unknown endocrine activity revealed as in vitro anti-androgens. Environ. Health Perspect. 119, 794–800.

Penza, M., Jeremic, M., Marrazzo, E., Maggi, A., Ciana, P., Rando, G., et al., 2011. The environmental chemical tributyltin chloride (TBT) shows both estrogenic and adipogenic activities in mice which might depend on the exposure dose. Toxicol. Appl. Pharmacol. 255, 65–75.

Pereira-Fernandes, A., Vanparys, C., Hectors, T.L., Vergauwen, L., Knapen, D., Jorens, P. G., et al., 2013. Unraveling the mode of action of an obesogen: mechanistic analysis of the model obesogen tributyltin in the 3T3-L1 cell line. Mol. Cell. Endocrinol. 370, 52–64.

Rochester, J.R., 2013. Bisphenol a and human health: a review of the literature. Reprod. Toxicol. 42, 132–155. C.

Rosen, E.D., Sarraf, P., Troy, A.E., Bradwin, G., Moore, K., Milstone, D.S., et al., 1999. PPAR gamma is required for the differentiation of adipose tissue in vivo and in vitro. Mol. Cell 4, 611–617.

Ruiz, D., Becerra, M., Jagai, J.S., Ard, K., Sargis, R.M., 2018. Disparities in environmental exposures to endocrine-disrupting chemicals and diabetes risk in vulnerable populations. Diabetes Care 41, 193–205.

Sargis, R.M., Johnson, D.N., Choudhury, R.A., Brady, M.J., 2010. Environmental endocrine disruptors promote adipogenesis in the 3T3-L1 cell line through glucocorticoid receptor activation. Obesity Silver Spring (Silver Spring) 18, 1283–1288.

Seimandi, M., Lemaire, G., Pillon, A., Perrin, A., Carlavan, I., Voegel, J.J., et al., 2005. Differential responses of PPARalpha, PPARdelta, and PPARgamma reporter cell lines to selective PPAR synthetic ligands. Anal. Biochem. 344, 8–15.

Shen, O., Du, G., Sun, H., Wu, W., Jiang, Y., Song, L., et al., 2009. Comparison of in vitro hormone activities of selected phthalates using reporter gene assays. Toxicol. Lett. 191, 9–14.

Somm, E., Schwitzgebel, V.M., Toulotte, A., Cederroth, C.R., Combescure, C., Nef, S., et al., 2009. Perinatal exposure to bisphenol A alters early adipogenesis in the rat. Environ. Health Perspect. 117, 1549–1555.
Spiegelman, B.M., 1998. Ppar-gamma: adipogenic regulator and thiazolidinedione receptor. Diabetes 47, 507–514.
Stapleton, H.M., Klosterhaus, S., Eagle, S., Fuh, J., Meeker, J.D., Blum, A., et al., 2009. Detection of organophosphate flame retardants in furniture foam and U.S. House dust. Environ. Sci. Technol. 43, 7490–7495.
Stapleton, H.M., Klosterhaus, S., Keller, A., Ferguson, P.L., van Bergen, S., Cooper, E., et al., 2011. Identification of flame retardants in polyurethane foam collected from baby products. Environ. Sci. Technol. 45, 5323–5331.
Takeuchi, S., Iida, M., Kobayashi, S., Jin, K., Matsuda, T., Kojima, H., 2005. Differential effects of phthalate esters on transcriptional activities via human estrogen receptors alpha and beta, and androgen receptor. Toxicology 210, 223–233.
Taxvig, C., Dreisig, K., Boberg, J., Nellemann, C., Schelde, A.B., Pedersen, D., et al., 2012. Differential effects of environmental chemicals and food contaminants on adipogenesis, biomarker release and PPARgamma activation. Mol. Cell. Endocrinol. 361, 106–115.
Temkin, A.M., Bowers, R.R., Magaletta, M.E., Holshouser, S., Maggi, A., Ciana, P., et al., 2016. Effects of crude oil/dispersant mixture and dispersant components on PPARgamma activity and : identification of dioctyl sodium sulfosuccinate (DOSS; CAS #577-11-7) as a probable obesogen. Environ. Health Perspect. 124, 112–119.
Trasande, L., Attina, T.M., Blustein, J., 2012. Association between urinary bisphenol A concentration and obesity prevalence in children and adolescents. JAMA 308, 9.
Voet, Hd, JAHv, Rhijn, HJvd, Wiel, 1999. Inter-laboratory, time, and fitness-for-purpose aspects of effective validation. Anal. Chim. Acta 391, 159–171.
vom Saal, F.S., Akingbemi, B.T., Belcher, S.M., Birnbaum, L.S., Crain, D.A., Eriksen, M., et al., 2007. Chapel Hill bisphenol A expert panel consensus statement: integration of mechanisms, effects in animals and potential to impact human health at current levels of exposure. Reprod. Toxicol. 24, 131–138.
Vom Saal, F.S., Nagel, S.C., Coe, B.L., Angle, B.M., Taylor, J.A., 2012. The estrogenic endocrine disrupting chemical bisphenol A (BPA) and obesity. Mol. Cell. Endocrinol. 354, 74–84.
Welshons, W.V., Nagel, S.C., vom Saal, F.S., 2006. Large effects from small exposures. III. Endocrine mechanisms mediating effects of bisphenol A at levels of human exposure. Endocrinology 147, S56–69.
Wong, J., Hao, C., Zhang, K., Yang, P., Banerjee, K., Hayward, D., et al., 2010. Development and interlaboratory validation of a QuEChERS-based liquid chromatography-tandem mass spectrometry method for multiresidue pesticide analysis. J. Agric. Food Chem. 58, 5897–5903.
Zava, D.T., Wyler-von Ballmoos, A., Goldhirsch, A., Roos, W., Takahashi, A., Eppenberger, U., et al., 1982. A quality control study to assess the inter-laboratory variability of routine estrogen and progesterone receptor assays. Eur. J. Cancer Clin. Oncol. 18, 713–721.
Zebisch, K., Voigt, V., Wabitsch, M., Brandsch, M., 2012. Protocol for effective differentiation of 3T3-L1 cells to adipocytes. Anal. Biochem. 425, 88–90.
Zenbio Inc, 2015. 3T3-L1 Cell Care Manual - Maintenance and Differentiation of Preadipocytes to Adipocytes. Available: http://www.zen-bio.com/pdf/ZBM0009.013T3L1CareprotocolRV08.08.pdf [accessed 30 September 2015.
Zhao, Y., Zhang, K., Giesy, J.P., Hu, J., 2015. Families of nuclear receptors in vertebrate models: characteristic and comparative toxicological perspective. Sci. Rep. 5, 8554.