MEMORANDUM

To: Eric Hooker, Matthew Cho, and Kristina Hatlelid, CPSC

From: Samantha Snow, Joei Robertson, and Elizabeth Martin, ICF
Lynne Haber, Mark Bradley, University of Cincinnati

Date: December 6, 2024

Re: Task 1C: For chemicals without sufficient studies to support derivation of toxicity reference values (see Task 1B), identify alternative data types and methods and review alternative method toxicity and potency data for two subclasses
(CPSC Contract Number 61320622A0005, CPSC Order Number 61320623F2030)

---

1. Overview of Memorandum

Under CO-9, Subtask 1C seeks to identify approaches to derive toxicity reference values (TRVs) from new approaches/alternative methods (NAMs) data. To find these methods, a targeted literature search was conducted followed by title/abstract and full text screening to identify relevant references.

Analysis of transcriptomic data to derive a point of departure (POD) was selected as the method to calculate TRVs for two subclasses of organohalogen flame retardants (OFRs). Datasets for whole transcriptomic data following exposure to chemicals in the polyhalogenated organophosphate (PHOP) and polyhalogenated bisphenol aliphatics and functionalized (PHBAF) subclasses were identified, moved through a workflow to derive a POD, and used to calculate TRVs. These TRVs were then compared to those identified in Subtask 1A and derived in Subtask 1B. This memo describes the methods used in the derivation of TRVs from NAMs data and a comparison of these TRVs to those derived from apical endpoints.

2. Literature Search for Identifying Approaches to Derive TRVs from NAMs Data

ICF and UC conducted a search for relevant methods papers for this subtask based on three data streams:

1) Literature identified via searches of scientific literature databases (i.e., PubMed and Web of Science) with no date restrictions;
2) Literature identified via a search of the grey literature, or publications produced outside of traditional scientific publishing avenues, including reports from government agencies and authoritative bodies not generally available in databases of published literature; and
3) Publications and reports identified by subject matter experts.

Under Subtask 2A, we conducted a literature search to identify approaches for deriving TRVs using NAMs data. Following deduplication and prioritization (Section 3.1), the papers identified were screened in the title/abstract step (Section 3.2). The search strings and strategies for this subtask are presented in Appendix A. Refer to the Subtask 2A report for the complete literature search methods.

2.1. Literature Search Strategies

The following sections describe literature search strategies used for databases and additional sources. The literature search strategies included searches within core literature databases (i.e., PubMed and Web of Science) as well as relevant domestic and international non-periodical "grey" literature, such as technical reports, monographs, and symposium proceedings prepared by select committees or bodies (e.g., National Academy of Sciences).

2.2. Seed Studies

Subject matter experts identified 33 references to use as positive seed studies to help verify the accuracy and efficiency of our literature search (Attachment 1; Figure 1). These relevant references were identified from reference compilations identified by the report authors as part of other projects or from general monitoring of the literature. The goal was to identify key references related to the use of NAMs in risk assessment, with particular focus on dose-response, read across, and addressing data-poor chemicals. These studies were used as seed studies for the database searches described in Section 2.3 and in the ensemble supervised clustering described in Section 3.1. Of the 33 seed studies, 7 studies were captured in our database search (Section 2.3) and 1 study was captured in our grey literature search (Section 2.4). Although the low capture rate for the seed studies in the initial literature search may raise concerns about the adequacy of the initial literature searches, the initial searches were heavily supplemented by expert review of the grey literature and tree searching, as described below. In addition, missing individual studies are of less concern for this task than for an assessment, since the goal was to capture the general approaches that are in the literature.

2.3. Database Searches

The literature searches for peer-reviewed publications were conducted on December 4, 2023 in Pubmed and Web of Science with no limit on the date of publication. Removing duplicates across the searches resulted in a total of 4,672 unique references (Attachment 1; Figure 1). These references were moved to a topic extraction step (Section 3.1), before going forward into title/abstract screening using Litstream®.

2.4. Identification of Grey Literature

There were 11 agency websites identified by ICF and UC, and confirmed by CPSC, for review of relevant grey literature (Appendix A). Grey literature was pulled from these relevant websites using an ICF-developed Google web scraping tool. This tool searched the Google results page for a search that had been limited to a specific domain and pulled links and metadata associated with the search results into an Excel file. Each search was limited to pull in only the top 10 hits per source (for 11 sources) per search string (for 3 topic specific search strings), meaning a maximum of 330 grey literature search results would be pulled by the web scraper.

A total of 259 references/websites were returned from the Google web scraping (Attachment 1; Figure 1). These items underwent a first pass screening by a subject matter expert to assign a priority rating (i.e., not relevant, maybe relevant, include) and to identify the topic or key take-away of the references identified as maybe relevant or include. If the initial web scraping tool identified a general website, that website was reviewed to identify any useful/relevant references, and those references were flagged for downloading if they were not included in the list from the web scraping tool. A second review was conducted of the literature marked as maybe relevant or include in order to identify the most recent relevant references. Key topics were those that addressed methods for read across, class-based assessments, cumulative risk/combined exposure, and use of NAMs in risk assessment. Please note that not all literature was reviewed in detail.

2.5. Expert Identified Literature

Key agency guidance documents that the authors were aware of but had been missed by the web scraping tool were downloaded and categorized as expert identified literature. In addition, subject matter experts identified additional peer-reviewed or grey literature sources during their synthesis writing (e.g., as a reference cited in a

paper that was captured in our literature search) that resulted in them pulling in expert identified literature. A total of 35 expert identified literature sources were included in this report (Figure 1; Attachment 1).

3. Prioritization and Literature Screening

3.1. Prioritization of Literature for Screening

The peer-reviewed literature identified via database searches (n = 4,672) underwent a prioritization step prior to title/abstract screening. To prioritize these 4,672 studies, three unique approaches were applied: (1) keyword analysis, (2) unsupervised clustering, and (3) supervised clustering. Keyword analysis enabled us to identify specific keywords that were directly relevant to identifying approaches that derive TRVs from NAMs data within the title or abstract of a paper. The unsupervised clustering method groups papers that have similar keywords in both the title and abstract. In contrast to keyword analysis, which only reports the incidence of pre-specified words, unsupervised clustering analysis relies on a computational algorithm to group papers without any *a priori* information. Supervised clustering uses expert identified seeds and multiple clustering algorithms to categorize papers by giving them an ensemble score to indicate likelihood of relevancy. Keyword analysis was selected as the best approach to optimize inclusion of likely relevant literature and resulted in 1,489 papers to be screened at the title/abstract level. See Subtask 2A report Appendix B for complete details on screening prioritization.

3.2. Title/Abstract Screening

A total of 1,489 studies were identified using keyword analysis and uploaded into ICF's Litstream® software. A screening form was developed using inclusion/exclusion criteria to guide decisions about the relevance of the literature. Screeners reviewed the title and abstract to determine whether the reference should be included, excluded, or marked unclear. If a reference was marked as unclear, it underwent a QC step by a senior screener to make the final decision. For a reference to be included for this subtask, it needed to describe an approach to derive TRVs using NAMs data.

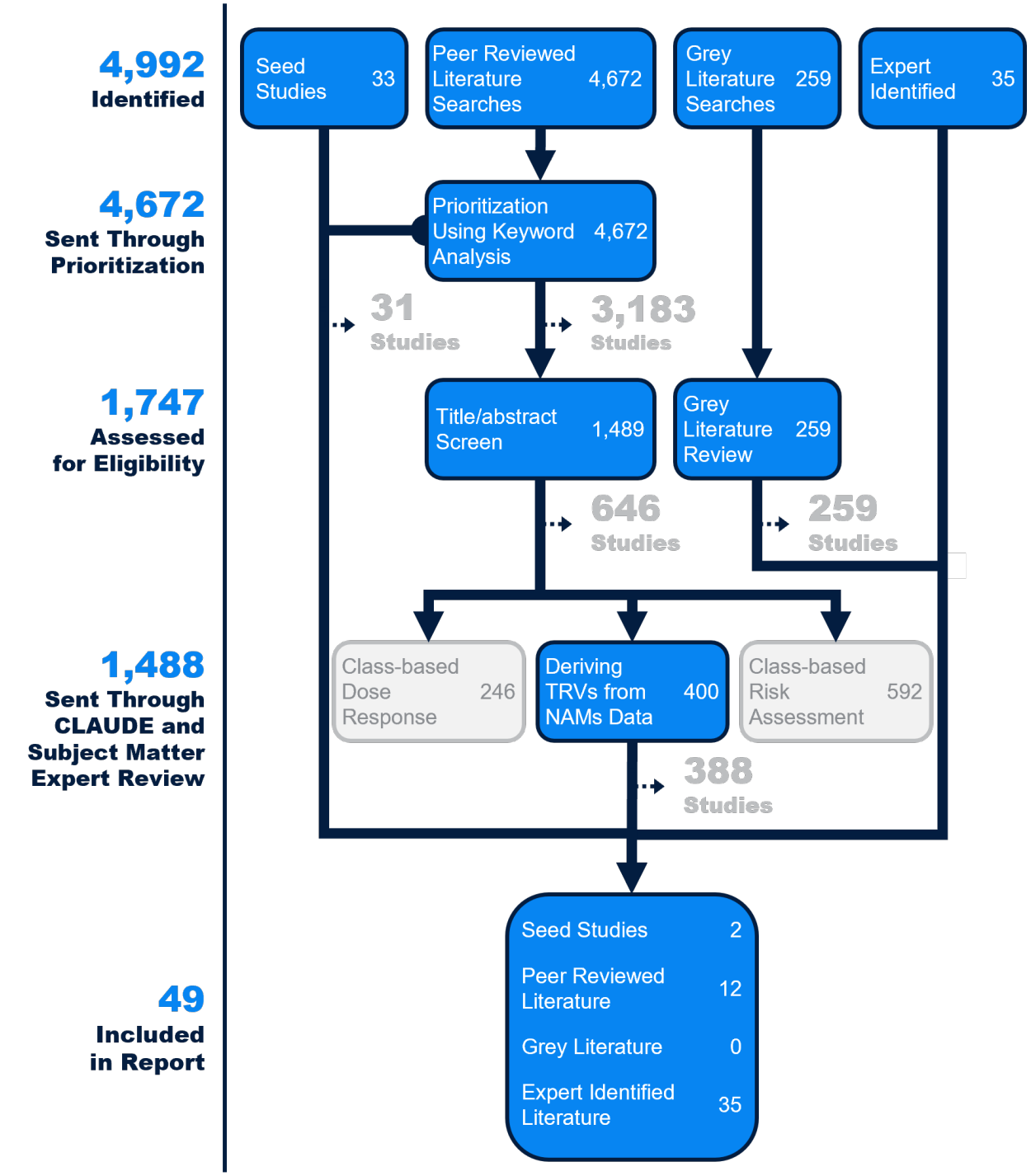
Each screener completed a pilot of 30 references, which were then QC'd. Feedback was provided if there were any discrepancies between the primary and senior screeners. After successfully completing the pilot, screeners continued working until all 1,489 references were reviewed. Each reference was only reviewed by one screener following the pilot. At the completion of title/abstract screening, there were 400 references tagged to approaches to deriving a TRV from NAMs data.

3.3. Full Text Screening

From title/abstract screening, 400 peer-reviewed studies were marked as 'include' for the topic of approaches to deriving a TRV from NAMs data. Of those, 363 had PDFs that were retrievable. To prioritize full text PDFs, a generative artificial intelligence (GenAI) workflow utilizing the large language model Claude 3.0 was deployed. This is neither an unintended nor prohibited use of Claude 3.0. Specifically, in the pipeline, PDFs were converted to text by Azure Document Intelligence. 12 PDFs were too large for conversion and were removed from subsequent analysis leaving 351 summarized studies. The 12 large PDFs were reviewed by subject matter experts to prevent loss of information. Next, a set of 10 prompts were tested against 10 expert curated PDFs (Appendix Table B-1). Additionally, parameters used for this prioritization included "Acting as a toxicologist" and Claude 3.0 (Appendix Table B-2). Specifically, these prompts were then graded for accuracy (correct, mostly correct, incorrect, or hallucinating[1]) and were determined to be accurately describing the topic queried. These 10 prompts were then queried against each paper individually. Each paper with a "yes" response was then analyzed by subject matter experts for inclusion in the specific report section for which it was identified to be relevant.

[1] Hallucination = When a large language model makes up an answer that is completely unrelated to the source material.

Figure 1. Literature Search Results



4. Methods to Derive TRVs from NAMs Data

There are many types of NAMs data that could be used to derive a TRV. The Subtask 2A report thoroughly discussed one of these approaches, derivation of a TRV from transcriptomic data. Below, we highlight other data types (i.e., computational methodologies, *in vitro* assays, non-mammalian *in vivo* assays) that are less well described in the Subtask 2A report. Importantly, the data from these NAMs are generally described in terms of throughput and complexity of the system. Computational models are the highest throughput models as millions of chemicals can be screened on relatively short time scales. Next, *in vitro* methods can be used to analyze thousands of chemicals and represent an increased level of biological complexity. Lastly, non-mammalian *in vivo* studies can be used to screen hundreds to thousands of chemicals and represent exposure in a whole animal, an even more complex system type.

4.1. Computational Methods

There are two well-described computational toxicology methods that are used to derive TRVs when neither *in vivo* nor *in vitro* experimental data are available: (quantitative) structure-activity relationship ((Q)SAR) and physiologically based pharmacokinetic (PBPK) modeling (Demchuk et al., 2008). Overall, (Q)SAR modeling is most useful for cross-chemical extrapolation for a given endpoint, while PBPK modeling is most useful for cross-species, cross-route, and cross-duration extrapolation for a toxicant (Demchuk et al., 2008). Once a model relevant to the chemical or exposure scenario of interest is available, it can serve as an alternative for time consuming experimental methods of TRV prediction. Some limitations of these methods include lack of generalizability, use of a small training dataset and therefore a lack of chemical diversity, and adequately verified input data.

4.1.1. (Q)SAR

(Q)SAR methods may qualitatively describe the potential for a chemical to have a certain activity (e.g., mutagenicity or cancer) in a SAR model, or quantitatively predict the value of a property (e.g., solubility or LD_{50}) or the likelihood of a specific effect in a QSAR model. (Q)SAR models describe theoretical or mathematical relationships between chemical structure of a toxicant and its biological activity. (Q)SAR methods may be used to identify a POD for the chemical of interest based upon data from a surrogate chemical. For example, the POD of a surrogate chemical with structural, metabolic, or toxicity similarity may be used for the chemical of interest as the most probable value (Wang et al., 2012). Usually, expert judgement is needed to evaluate the similarity between the chemical of interest and its surrogate, specifically related to whether or not toxicity characteristics are similar; at this time, computational prediction of hazard characteristics is not possible based solely on structural similarities. Similarly, read-across methods may also use (Q)SAR to extrapolate information from source substances with measured TRVs to a target substance that lacks TRVs.

QSAR models that are used for toxicological assessment are usually constructed through three steps:

i) building a dataset of chemicals with experimentally measured toxicities,

ii) building a mathematical or statistical relationship between molecular properties (e.g., molecular weight, lipophilicity, presence of certain molecular substructures) of those chemicals and their biological effects, and

iii) validation of the model by using datasets of chemicals that have not been used in construction of the model.

When the model is validated, it can be used to predict toxicity of new chemicals. As an example, (Q)SAR models were used by Union Carbide Corporation for risk assessment of byproducts that leached into the groundwater from Reich Farm hazardous waste sites. QSAR models predicted 9 of 15 chemicals to be potential carcinogens, 6 to be developmental toxicants, and 6 to be mutagenic (NTP, 2007). Recently, machine learning approaches have evolved that use large experimental datasets to build QSAR models. Compared to traditional QSAR models, deep neural networks, and traditional machine learning methods, multitask machine learning models offer better predictivity (Wu et al., 2021).

4.1.2. PBPK

PBPK models are mathematical models that describe the absorption, distribution, metabolism, and excretion (ADME) of a chemical in the body, using physiologically-based parameters and equations to describe chemical flows and transformation. Individual organs or groups of organs (e.g., rapidly perfused tissues) are described as compartments in the model (U.S. EPA, 2006). Model parameters are usually optimized using a combination of experimental data and curve fitting. PBPK models predict internal doses of chemicals in body tissues for a given external (administered) dose. Once the internal dose of a chemical is determined for a specific route of exposure, the exposure conditions resulting in the same internal dose from other exposure routes can be determined if the model appropriately reflects those routes. Similarly, the internal dose for an exposure duration (e.g., subchronic) may be extrapolated to another duration (e.g., chronic), although large extrapolations, such as from acute to chronic duration are generally not considered appropriate. PBPK models have many possible uses related to deriving TRVs. They can be used to identify the relevant dose metric, and thus the mode of action of toxicity. PBPK models can also be used to improve interspecies extrapolation, or to aid in applying the toxicity testing results from one route to another route.

4.2. *In Vitro* Models

While *in silico* methods are useful in screening large numbers of compounds, they provide limited amounts of information and rely on a limited amount of input data; thus, additional source(s) of information are often needed. *In vitro* methods provide an additional layer of challenges: rather than relying on chemical structure or extrapolation, data is generated based on the interaction of chemicals with a specific cell type or types of interest, or with specific targets. Relevant *in vitro* models include cytotoxicity assays, flow cytometry, micronucleus assays, DNA damage assays, hormone production assays, and -omics data (Black et al., 2022; Embry et al., 2010; Gilmour et al., 2020; Sharma et al., 2017; Soeteman-Hernandez et al., 2015). These types of studies use benchmark dose (BMD) analysis to establish a POD. Additional endpoints of interest have also been developed in response to these new data types. For example, the no observed transcriptional effect level (NOTEL) is a new metric that represents the level at which mRNAs are unchanged (Hewitt et al., 2022).

4.3. Nonmammalian NAMs

The last category of NAMs often used in toxicity testing are non-mammalian animal models. Examples of alternative animal models include zebrafish, C. elegans (nematode), fish embryos, or *Drosophila* (fruit fly). These model organisms have well defined genomes and developmental patterns that are useful in adding a layer of organ and organism complexity during toxicity testing. Studies have used these for large screening exercises and POD derivations with success (Song et al., 2023; Thomas et al., 2019). These model systems can be coupled with many of the same methods used in *in vitro* settings (e.g., -omics, reporter gene measures) to derive PODs or they can be used to study apical endpoints such as neurobehavioral outcomes or developmental malformations.

5. Approach for Deriving a TRV from Transcriptomic Data

The same general approach is used for deriving a TRV from the various types of NAMs discussed in Section 4, although the details vary. Specifically, a POD is derived from an *in vivo* or *in vitro* model using transcriptomic data. *In vitro to in vivo* extrapolation (IVIVE) needs to be applied to derive an oral administered dose when the POD is from an *in vitro* study. Such extrapolation is not necessary for transcriptomic data from an *in vivo* model. Finally, for both *in vivo* and *in vitro data*, uncertainty factors are applied based on the characteristics of the given experiment. The general conclusion we reached both in this report and in the prior Subtask 2A report was that derivation of the POD from transcriptomic data was the best possible approach as it allowed for comparability between model types (i.e., *in vivo* mammalian, *in vivo* non-mammalian, and *in vitro* systems). This is because the outputs from transcriptomic studies are consistent measures of gene counts (Section 6.6), rather than relative measures. This means that transcriptomic data are more comparable between labs than are data from other NAMs, such as reporter gene assays or behavioral assays (Liao et al., 2014; Tangri et al., 2013). Additionally, it is

important to note that transcriptomic data can also be used to address hazard characterization. To this end, we modeled the transcriptomic data and performed IVIVE in order to determine a POD, as described below.

5.1. Approach to Identify the POD from NAMs Data

Briefly, the approach used in the current assessment was based on the framework developed by Reardon et al. (2023), which utilizes pathway level gene sets to identify transcriptomic PODs (tPODs) based on the lowest consistent response dose (LCRD). In this process, tPODs are calculated on a gene-by-gene basis for the benchmark response (BMR: a pre-defined level of effect) at a one control standard deviation (SD) change . In the current work, the BMR was defined as a relative deviation of 25% (i.e., a change by 25% of the control mean), as noted in Section 7.2. Relative deviation was chosen rather than a change in the mean of one control standard deviation (SD). This choice was made because relative deviation better amplifies the signal of dose responsive genes relative to noise. Consideration of the signal to noise ratio is important in accounting for the inherent variability in gene expression data between biological replicates (Auerbach, 2022). A relative deviation of 25% was used for the BMR rather than the more common 10% as it has been suggested as a more robust measure of true difference by those in the transcriptomics field (Auerbach, 2022). After the tPODs were derived, the genes are grouped based on their biological function by GO biological processes. Next, the LCRD method was applied to each grouping of genes to identify consistent response groups of benchmark concentrations (BMCs). This framework addresses a key need for biological plausibility because it uses an approach that is dependent upon consistency of dose-responsiveness of genes rather than focusing on readouts of a single endpoint (Reardon et al., 2023). Additional details of the modeling and identification of the *in vitro* POD are in Section 7.2.

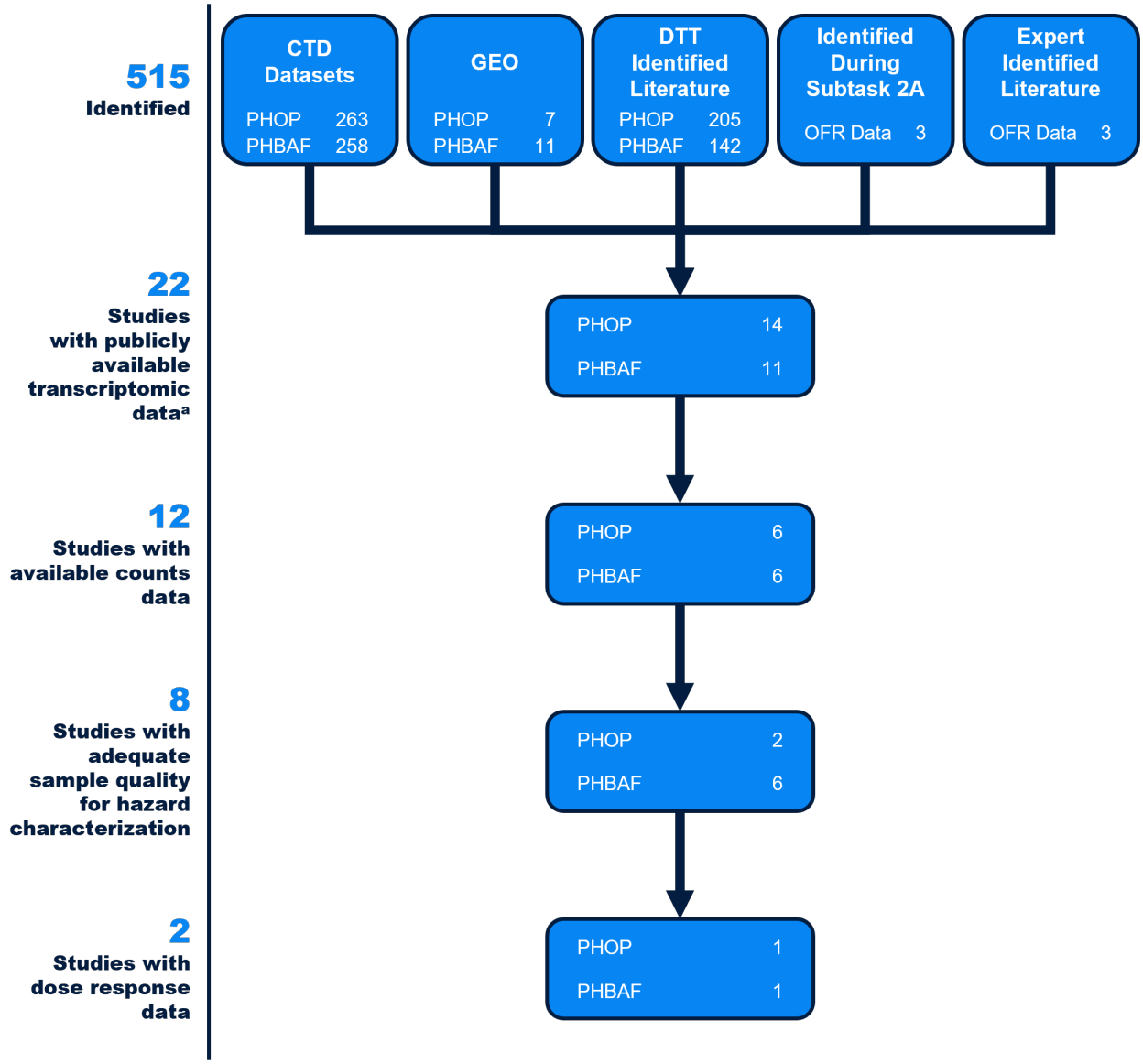
5.2. In Vitro to In Vivo Extrapolation (IVIVE)

For *in vitro* models, it is necessary to extrapolate doses from the concentration tested *in vitro* to an *in vivo* administered dose. The high-throughput toxicokinetics (httk) computational package is commonly used to perform IVIVE (Wambaugh et al., 2019). This package is both part of the NIEHS Integrated Chemical Environment (ICE) and available as a stand-alone R package (Section 7.2). The httk package estimates the *in vivo* dose based on a generic PBPK model and two parameters: plasma protein binding activity and hepatocyte-mediated clearance determined *in vitro* (i.e., the intrinsic clearance). The software incorporates empirical data from 389 chemicals and uses a QSAR model to predict these parameters for several thousand additional chemicals. User defined inputs for these parameters are also allowed, making it possible to apply the model to a broader range of chemicals. IVIVE extrapolations using httk can be used for a variety of applications, including chemical risk prioritizations and TRV derivation (Wambaugh et al., 2019). We describe this process in depth in Section 7.2 for the exact parameters used.

6. Identifying OFR Available Datasets

To identify PHOP and PHBAF chemicals that have publicly accessible transcriptomic datasets, five different resources were reviewed. The Comparative Toxicogenomic Database (CTD) was queried for PHOP and PHBAF chemicals while the Gene Expression Omnibus (GEO) database was queried for PHOP and PHBAF chemicals as well as PHOP metabolites identified in CO-1, Call Order 61320622F2011, “Guidance Document for Conducting Class-based Hazard Assessment of Organohalogen Flame Retardants” under CPSC Blank Purchase Agreement (BPA 61320622A0005). PHBAF metabolites were not available from CO-1 to conduct this additional search. Studies identified during the CPSC co-sponsored scoping review project with the National Institute of Environmental Health Sciences (NIEHS) Division of Translational Toxicology (DTT) that were tagged to either the PHOP or PHBAF subclass were reviewed to determine if they contained transcriptomic data. Finally, three studies each that were identified during Subtask 2A and by experts during this Subtask were reviewed to determine if they contained transcriptomic data. The process for querying each database, reviewing identified literature, and the total number of relevant datasets identified are described below and illustrated in Figure 2.

Figure 2. PHOP and PHBAF Transcriptomic Datasets



[a] The total number of studies include three studies that contained data for both a PHOP and PHBAF chemical.

6.1. Comparative Toxicogenomic Database (CTD)

CTD is a manually curated database that aggregates information across publications related to chemical-gene/protein interactions. To identify studies that contained transcriptomic data, we performed a bulk query by CAS RN for all PHOPs, and PHBAFs (Appendix C). These were then validated by chemical names in single query format. The query yielded a list of studies affiliated with each gene changed in response to the chemical of interest. Studies were first deduplicated by the PubMed identifier number (PMID), which yielded 263 PHOP affiliated studies and 258 PHBAF affiliated studies (Figure 2). Of these studies, 19 included both PHOP and PHBAF data, leading to a total of 502 unique studies (Attachment 2).

Subject matter experts (SMEs) were engaged to perform title/abstract screening to identify studies that had whole transcriptomic data. Specifically, screeners reviewed the title and abstract for the use of the words “array”, “sequencing”, “RNA-seq”, “gene expression”, “mRNA”, “RNA”, “transcripts”, “transcriptomes”, or

"transcriptomics". Studies that were excluded described the use of "qPCR" and "reporter gene" as we were focused on finding datasets that measured changes in the whole transcriptome. If a study was identified as likely to contain transcriptomic data, full text PDFs were obtained and screened for (1) GSE accession numbers that are housed in the Gene Expression Omnibus (GEO) database, (2) BioProject IDs that are housed in the BioProject database, or (3) GenBank Accession (SRA) IDs that are housed in the GenBank database. These three databases house publicly accessible gene expression datasets.

6.2. Gene Expression Omnibus (GEO)

GEO is the largest publicly accessible database of gene expression data. Unlike the GenBank and BioProject databases, where data are entered by users as desired, GEO requires data to be deposited with annotations that are screened by a group of scientists employed by GEO. This increases the likelihood of adequate quality data for analysis. We chose to query GEO to ensure we had captured as many available PHOP and PHBAF datasets as possible, as GEO datasets link to publications that are not necessarily captured in CTD. For this search, we manually queried GEO for all PHOPs, PHBAFs, and PHOP metabolites using both the full chemical names and common acronyms for the PHOP and PHBAF chemicals (Appendix C). CAS RNs were not used as there is no tagging system for CAS RN in GEO. We included PHOP metabolites for this search given that our preliminary review for publicly available transcriptomic datasets in CTD were fewer than expected, and because we expect that dose-response information for shared metabolites may be useful for the class-based dose-response analysis. For the GEO query, GSE accession numbers and PMIDs were recorded. This search yielded 7 PHOP affiliated studies (i.e., chemical and metabolite) and 11 PHBAF affiliated studies (i.e., chemical) (Figure 2). Three of these studies included both PHOP and PHBAF data leading to a total of 15 unique studies (Attachment 2).

6.3. Division of Translational Toxicology (DTT) Identified Studies

As reported in the Subtask 1B report, there were 2,474 studies from the DTT screening project. This initial file was used to identify 205 PHOP affiliated studies (i.e., chemical) and 142 PHBAF affiliated studies (i.e., chemical) (Figure 2). Sixteen of these studies included both PHOP and PHBAF data leading to a total of 331 unique studies (Attachment 2). For these studies, title/abstract and full text screening were conducted as described in Section 6.1 to identify studies that had whole transcriptomic data.

6.4. Studies Identified During Subtask 2A

During the full text review conducted during Subtask 2A, three references were tagged as ones containing data related to OFRs (Figure 2, Attachment 2). For these studies, title/abstract and full text screening were conducted as described in Section 6.1 to identify studies that had whole transcriptomic data.

6.5. Expert Identified Studies

CPSC identified three *in silico* studies as likely containing data for OFRs (Figure 2, Attachment 2). Full text screening was conducted as described in Section 6.1 to identify whether these studies had whole transcriptomic data for the relevant chemicals of interest. Based on the review of the supplement data, it does not appear that there are OFR relevant data contained within these studies.

6.6. Dataset Curation

There were 515 unique studies identified to review for containing whole transcriptomic data. Of these, multiple studies were identified as duplicates via various sources, including 2 studies through the CTD search and GEO search; 1 study through the CTD search, GEO search, and Subtask 2A search; 327 studies through the CTD search and DTT identified studies; and 4 studies through the CTD search, GEO search, and DTT identified studies (Attachment 2).

These 515 unique studies were reviewed to identify publicly available transcriptomic datasets. Following title/abstract review and search of the GEO database (Section 6.2), 66 studies were identified as possibly

containing transcriptomic data. The PDFs for these papers were obtained and reviewed by SMEs. Of these 66 studies, 22 were identified as containing publicly available transcriptomic datasets (Figure 2, Attachment 2). A dataset provides the data for a test condition, which we define as a unique combination of chemical, test condition, and publication. Thus, a study that tested five chemicals under two different test conditions would produce 10 datasets. Three of these studies contained transcriptomic data for both a PHOP and PHBAF. For these studies, a data accession number was recorded (Attachment 2).

These 22 studies were then reviewed to determine if there were available gene counts data (i.e., when sequencing reads have been processed by alignment to the genome and summarized as counts data). Gene counts are the (1) literal counts of a transcript fragment generated during the sequencing process or (2) the intensity of fluorescence for each probe following the process of hybridization of input sample to a microarray. These two measures have been demonstrated to be comparable (Martin et al., 2023). Of the available studies, 6 PHOP and PHBAF studies each (12 total) had gene counts data (Figure 2, Attachment 2).

These 12 studies were further reviewed to identify datasets with adequate sample quality. Reasons for dataset removal were the absence of appropriate annotations (n=2), not depositing actual relevant treated samples (n=1), and inadequate sample number (n=1). This resulted in identification of 2 PHOPs datasets and 6 PHBAF datasets (Figure 2, Attachment 2).

The last step was to identify dose-responsive datasets (i.e., datasets with more than one dose and control). We identified 1 PHOP dataset, covering 2 chemicals (TCIPP and TCEP), and one PHBAF dataset, covering 1 chemical (TBBPA-DBPE).

In an attempt to maximize the use of the available data, we performed differential gene expression analysis on both PHOP datasets of adequate quality and all 6 PHBAF datasets, without regard to whether multiple doses were tested. The purpose of this analysis was to perform hazard characterization and identify common genes and pathways that are enriched across chemicals and chemical classes. This process furthers the goal of class-based risk assessment as it provides insight into the biological underpinnings of how PHOPs and PHBAFs may affect health.

As the goal of this report was to derive TRVs from NAMs data, we performed POD derivation for each of the datasets that were dose responsive for PHBAFs and PHOPs (Section 7.2). This was coupled to IVIVE through the httk package to derive a human relevant oral dose prior to the application of uncertainty factors.

7. Summary of Findings for OFR NAMs Data

7.1. Hazard Characterization

As described above, we identified a total of 8 datasets (2 PHOP and 6 PHBAF) suitable for differential gene expression analysis (Figure 2, Table 1, Attachment 2). These studies covered four major test systems: (1) exposure during development (2-6 hours post fertilization) for zebrafish (n=2), (2) exposure of human embryonic stem cells coupled with differentiation (n=4), (3) exposure of a human liver cell line, HepG2 (n=1), and (4) exposure of rats via drinking water and transcriptomic measurements in the liver (n=1). The studies analyzing PHOPs covered three chemicals: TDCPP, TCEP, and TCIPP. The studies analyzing PHBAFs analyzed TBBPA (n=5), TCBPA (n=4), TBBPA-DBPE (n=2; also called TBBPA-BDBPE), TBBPA-BAE (n=1), and TBBPA-BHEE (n=1). There were 3 studies that utilized microarray technology and 5 studies that utilized high-throughput sequencing. Overall, data were obtained for 19 test conditions that corresponded to one chemical in one test system. The results of the hazard characterization for each chemical and test condition are described in depth in Appendix D.

The microarray and high-throughput sequencing data were pushed through a data analysis pipeline for hazard characterization as demonstrated in Figure 3.

Table 1. Available Transcriptomic Datasets

GSEset	Data Type	Model	Species	Cell line/Tissue Type	Exposure duration	Chemical (s)	Dose-Responsive Dataset
GSE106875	Array	*In vivo*	Danio rerio (T5D (Tropical 5D) WT)	Whole Embryo	Exposure for 2 hours postfertilization or 6 hours postfertilization	TDCPP	No
GSE121336	Array	*In vivo*	Danio rerio (WT)	Whole Embryo	Exposure for 0-5 days postfertilization	TBBPA	No
GSE153363	Array	*In vivo*	Rattus norvegicus (Sprague Dawley)	Liver	5 days post exposure	TBBPA-DBPE	Yes
GSE109463	High Throughput Sequencing	*In vitro*	Homo sapiens	HepG2 Cell Line	72 hours of exposure	TCEP, TCIPP	Yes
GSE147560	High Throughput Sequencing	*In vitro*	Homo sapiens	HeLA Cell Line	2 hours or 24 hours of exposure	TDCPP	No
GSE148518	High Throughput Sequencing	*In vitro*	Homo sapiens	Human Embryonic Stem Cells	4 days or 12 days	TBBPA, TCBPA, TBBPS	No
GSE193176	High Throughput Sequencing	*In vitro*	Homo sapiens	Human Embryonic Stem Cells	2 days, 8 days, 16 days	TBBPA, TCBPA, TBBPA-BAE, TBBPS, TBBPA-BDBPE, TBBPA-BHEE	No
GSE128431	High Throughput Sequencing	*In vitro*	Homo sapiens	Human Embryonic Stem Cells	Not Reported	TBBPA, TCBPA	No
GSE129412	High Throughput Sequencing	*In vitro*	Homo sapiens	Human Embryonic Stem Cells	Not Reported	TBBPA, TCBPA	No

Datasets that were publicly accessible and included in the hazard characterization and/or dose-response assessment. Information includes measurement type (i.e., Array or High Throughput Sequencing), model type (*in vivo* or *in vitro*), species, cell line/tissue type, exposure duration, chemical(s), and whether the dataset was dose-responsive.

Figure 3. Data Analysis Pipeline

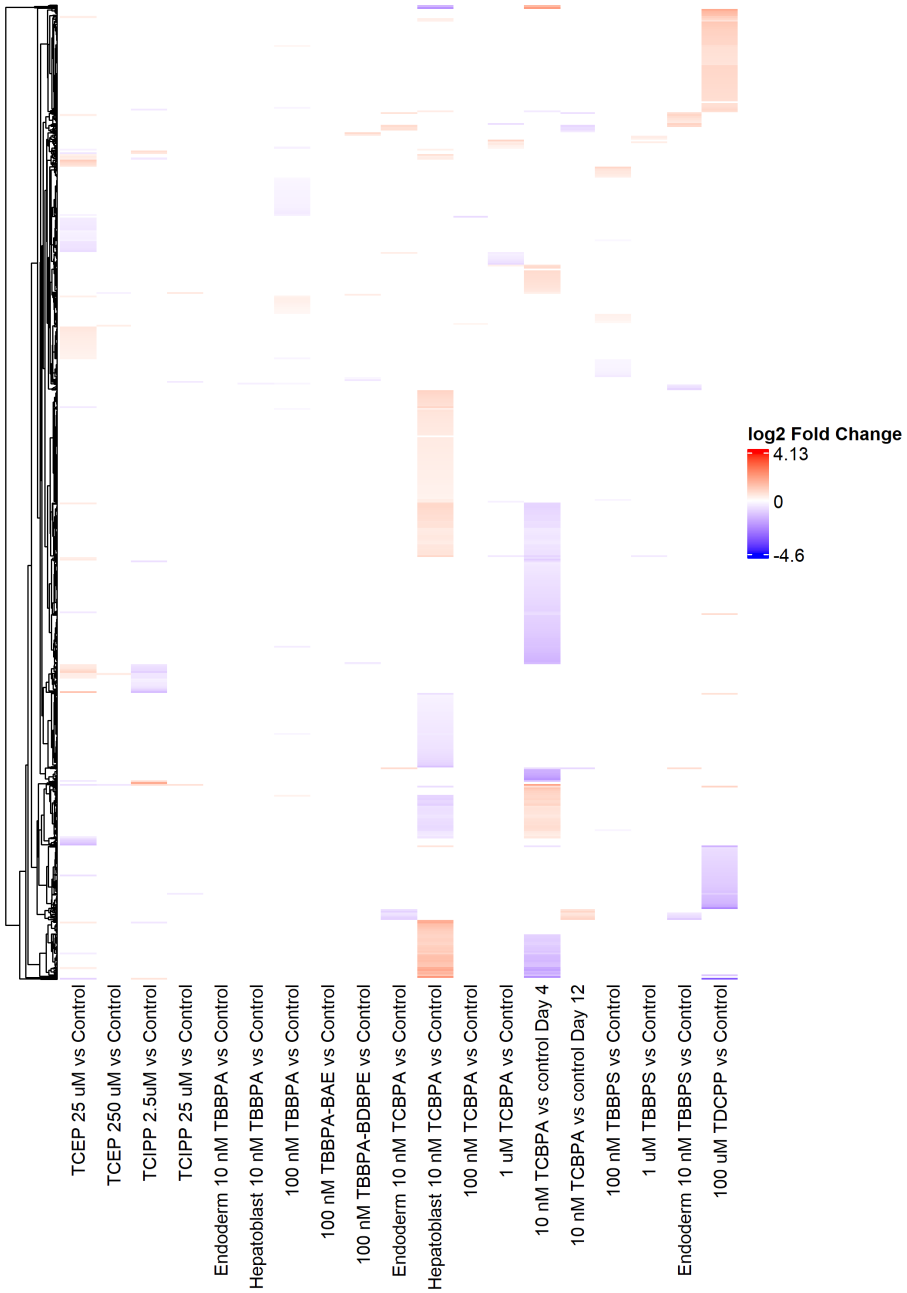
Datasets were accessed directly from GEO. Once accessed, data underwent differential gene expression analysis dependent upon the data source (arrays were processed by limma while RNA-sequencing data was processed by DESeq2). All datasets were then carried forward if they had appropriate sample quality and more than one dose group to BMD Modeling using BMDExpress 3.0.

7.1.1. Differential Gene Expression Analysis

To identify differentially expressed genes to be used for hazard characterization, we either used limma (Linear Models for Microarray) for array data or DESeq2 for high-throughput sequencing data (Figure 3). A given study could test a number of chemicals using several different test systems. Thus, from this point onwards in the report, we describe "test conditions" to refer to a unique combination of chemical, test system, and study. The major results of interest are a comparison across all datasets. Overall, we found a total of 2,971 genes activated in at least one test condition (Figure 4, Attachment 3). The heatmap of genes shows the log2 fold change relative to control for each chemical across various conditions. The chemicals are grouped together along the X-axis. Strikingly, the majority of gene expression changes are unique to each experimental set up: 2,155 (72%) are unique to one test condition (i.e., one chemical, one experiment). The most commonly shared gene was fibronectin

leucine rich transmembrane protein 3 (FLRT3), which was changed in 6 experiments across TCEP (n=1), TCBPA (n=4), and TBBPS (n=1). Generally, it showed an increase in expression. The protein FLRT3 is involved in cell adhesion and receptor signaling particularly during cranio-facial development. However, given that the test systems did not represent relevant tissues (i.e., tissues in which health effects are seen *in vivo*, specifically most effects are kidney related), it is unclear what these alterations might mean for disease at this time. Seven genes were changed in 5 test conditions, 20 genes in 4 test conditions, 118 in 3 test conditions, and 670 in one test condition. When looking at chemical specific changes, TCBPA stood out as it displayed the highest amount of transcriptional activity with most test conditions tested showing a high degree of change. Interestingly, TBBPA showed the least amount of transcriptional activity, with the majority of test conditions seeing almost no change in gene expression patterns. Specifically for studies that assessed TBBPA and TCBPA at 10 nM concentrations, gene expression was significantly altered in those that assessed TCBPA while not for those that assessed TBBPA. Furthermore, looking at a single dataset which tested the impacts of each chemical specifically (Appendix D.2.2.2), there is a clear difference in the activity of these compounds.

Figure 4. PHOP and PHBAF Hazard Characterization Differentially Expressed Gene Results



The heatmap displays the log2 fold change for each gene. White indicates gene expression was not changed compared to controls, blue indicates that there was a decrease in gene expression compared to controls, and red indicates there was an increase in gene expression compared to controls. Overall, we see the most transcriptional activity associated with TCBPA. Also evident in this heatmap is the limited amount of overlap between genes that are differentially regulated by each chemical as indicated by the significant amount of white space.

7.1.2. Pathway Enrichment Analysis

To identify whether genes could be mapped to common pathways, we performed Gene Set Enrichment Analysis (GSEA) using the tool ClusterProfiler (Figure 3). GSEA uses Gene Ontology, a standardized vocabulary used by bioinformaticians to catalogue and classify biological pathways. Each unique pathway maps to 1 of 3 different

Gene Ontology classifications: Biological Process (BP), Cellular Component (CC), and Molecular Function (MF). Pathways that map to BP are involved in broader biological processes that require the coordinated efforts of multiple molecular activities, such as *DNA repair* or *signal transduction*. Pathways that map to CC describe cellular locations or structures, such as *plasma membrane* or *nucleus*. Pathways that map to MF represent biological processes that are directly performed by gene products (e.g., proteins, small non-coding regulatory microRNA), such as *catalytic activity* or *transporter activity*. These classifications can be used to separate out and group data in a more specific way and are especially helpful when visualizing GSEA results with many pathways.

Overall, we found a total of 1,597 pathways activated or suppressed in at least 1 test condition across both chemical classes. These 1,597 pathways were visualized separately by their Gene Ontology classifications where 1,204 pathways mapped to BP, 189 pathways mapped to CC, and 204 pathways mapped to MF. Of these 1,597 pathways, 504 of them were shared across 2 or more test conditions. Of note, there was a larger than expected number of changes in pathway level hits (1,597) based upon the number of differential gene expression hits (2,971). This is unexpected as typically the number of pathways is substantially reduced relative to the number of genes.

Across the PHOP data, we found 621 pathways were activated or suppressed, with 49 of these pathways shared between 2 or more test conditions. Of the PHOP data analyzed, 2 pathways were shared between TCEP test conditions (n = 2), and 0 pathways were shared between TCIPP test conditions (n = 2). The 1 pathway shared across the most PHOP test conditions (n = 3) was *membrane* (GO:0016020). *Membrane* maps to CC and, therefore, represents a cellular component likely involved in response to PHOP exposure.

Across the PHBAF data, we found 1,171 pathways were activated or suppressed, with 356 of these pathways shared between 2 or more test conditions. Of the PHBAF data analyzed, 299 pathways were shared with TCBPA test conditions (n = 5), and 2 pathways were shared between TBBPS test conditions (n = 2). The 2 pathways shared across the most PHBAF test conditions (n = 4) were *cellular response to endogenous stimulus* (GO:0071495) and *plasma membrane* (GO:0005886). *Cellular response to endogenous stimulus* maps to BP and, therefore, represents a broader biological process where the cell responds to stimuli that arise from within the organism. This could suggest that PHBAF chemicals could be perturbing target cells in a way that mimics an endogenous compound or affects pathways that are essential to a cell's normal function. *Plasma membrane* maps to CC and, therefore, represents a cellular component likely involved in response to PHBAF exposure.

7.2. Derivation of Point of Departure (POD) with BMDExpress 3.0

Of the studies identified for modeling, only two passed quality control check and had accessible datasets (Figure 2, Attachment 2). These two studies covered a total of three chemicals: TCEP, TCIPP and TBBPA-DBPE. These data were pushed through the data analysis pipeline to derive a POD as demonstrated in Figure 3. To derive BMDs, BMDExpress 3.0 was utilized. First, sequencing data was normalized to transcript per million (TPM) which allows for sequencing depth normalization (i.e., normalization to the total number of transcripts sequenced). For array data, quantile normalization was performed as it scales all data so that samples are comparable to each other by intensity. Next, unexpressed genes (defined as genes with less than 1 read per sample for sequencing, or with beta-values equaling 0 for arrays) were removed. Data were then log2 normalized to calculate fold change based upon platform recommendations (NTP, 2024; Phillips et al., 2019; Yang et al., 2007) prior to importing into BMDExpress 3.0.

To derive the POD, the remaining transcripts were then pre-filtered by the William's test for trend to identify those genes that were in fact dose responsive and then filtered again by the curve fit prefilter test to identify those genes with a monotonic dose-response. While non-monotonic dose-response relationships do occur, with such a small dataset it is difficult to verify and likely to be an artifact of the data rather than a true biological phenomenon. In larger datasets, this supposition could be removed. After this final filtering step, BMDs were calculated using the EPA BMDS algorithm with the following parameters: BMR Type - relative deviation, Variance - Constant, BMR Factor - 25%, and BMDU/L Estimation Method - Profile Likelihood. All continuous models were run and the best fitting model was selected. Next, the genes were grouped by GO biological pathway. The median BMD of each GO biological grouping was plotted against the cumulative number of genes with the median BMD or lower. This functionality is a part of BMDExpress. For the BMD derivation, a lowest consistent response dose (LCRD) threshold

of 1.67 was applied to results displayed in the accumulation plot, whereby the median BMD between two gene sets was >1.67. A threshold of 1.67 was used because this represents the spacing of doses where at least 2 genes would have BMDs between 2 adjacent dose levels assuming half log dose spacing was used (Crizer et al., 2021). The results of these analyses are described below.

For data gathered from an *in vitro* system, IVIVE was performed using httk. The *in vitro* bioactive concentration (i.e., the LCRD) was used to compute an oral administered equivalent dose through the calc_mc_oral_equiv function of httk_2.3.1 R package (Pearce et al., 2017). The method is based on the construction of a linear relationship between the steady state concentration *in vitro* and the administered equivalent dose *in vivo*. Once the linear model is built, a scaling factor is calculated and used for conversion. For chemicals that exist in the httk library, calc_mc_oral_equiv function requires bioactive *in vitro* concentration as an input. In order to add a new molecule to the library, fraction unbound in plasma (F_{up}) and intrinsic hepatic clearance (CL_{int}) of the molecule are also required. OPERA 2.9 was used to predict values for those two parameters for molecules for which experimental values were not available (Mansouri et al., 2018). The Monte Carlo method was utilized for simulation of variability and uncertainty in the estimate of these parameters. The administered doses are computed for the 0.95 percentile of the Monte Carlo distribution. 'Honda1' flag was used for IVIVE (Honda et al., 2019). The rest of the parameters were the defaults of the httk package. Thus, we performed IVIVE using httk on the PODs calculated following the LCRD method in the same computational environment to derive an *in vivo* administered dose from the *in vitro* concentration.

7.2.1. PHOPs: GSE109463

GSE109463 is the dataset affiliated with Krivoshiev et al. (2018). Briefly, this study compared the effects of TCEP and TCIPP in HepG2 cells, a human hepatocellular carcinoma cell line. RNA was extracted 72 hours post exposure for DMSO controls, and from cells exposed to 2.5 µM TCIPP, 25 µM TCIPP, 25 µM TCEP, or 250 µM TCEP (Appendix D.1.2.1). For the dose response derivation, data for TCEP and for TCIPP was analyzed separately. For these datasets, TPM gene count data was first filtered for expressed genes (16,160 of 39,376 transcripts) followed by log2 normalization.

TCEP

Log2 normalized expressed TPM count data for 16,160 transcripts were imported into BMDExpress 3.0 with Ensembl annotations. Analysis of the first two principal components showed some separation by dose (Figure 5A). The Williams Trend test identified 415 genes that were likely dose-responsive. In the curve fit pre-filtering tests, 403 genes were identified that could be reasonably fit to a trend. These 403 genes were then brought forward to BMD modeling. Of these genes, the majority (71%) were best fit to a linear model, followed by 27.8% best fit to a Hill model (Figure 5B). Next, GO pathway analysis of gene function was conducted to identify groups of genes in which transcriptional changes occur. The results of this analysis are displayed as an accumulation plot. The accumulation plot shows the BMD median of the group (x axis) versus the cumulative number of genes with the BMD median or lower (y axis) (Figure 5C). Each dot in an accumulation plot accounts for the dots prior to it. This figure demonstrates that more genes are activated as the BMD increases and shows the rate at which additional genes are activated. To identify the POD, the LCRD is identified as the lowest of the BMDs at which a consistent gene response is observed. To do this, the BMD of each step is divided by the BMD of the prior step. The threshold for this response is 1.67 as it represents 2 adjacent dose levels where half log dose spacing is used. For TCEP, the dose this is observed at (i.e., the POD) is 0.047 µM.

Figure 5. TCEP Dose-Response Modeling Outputs

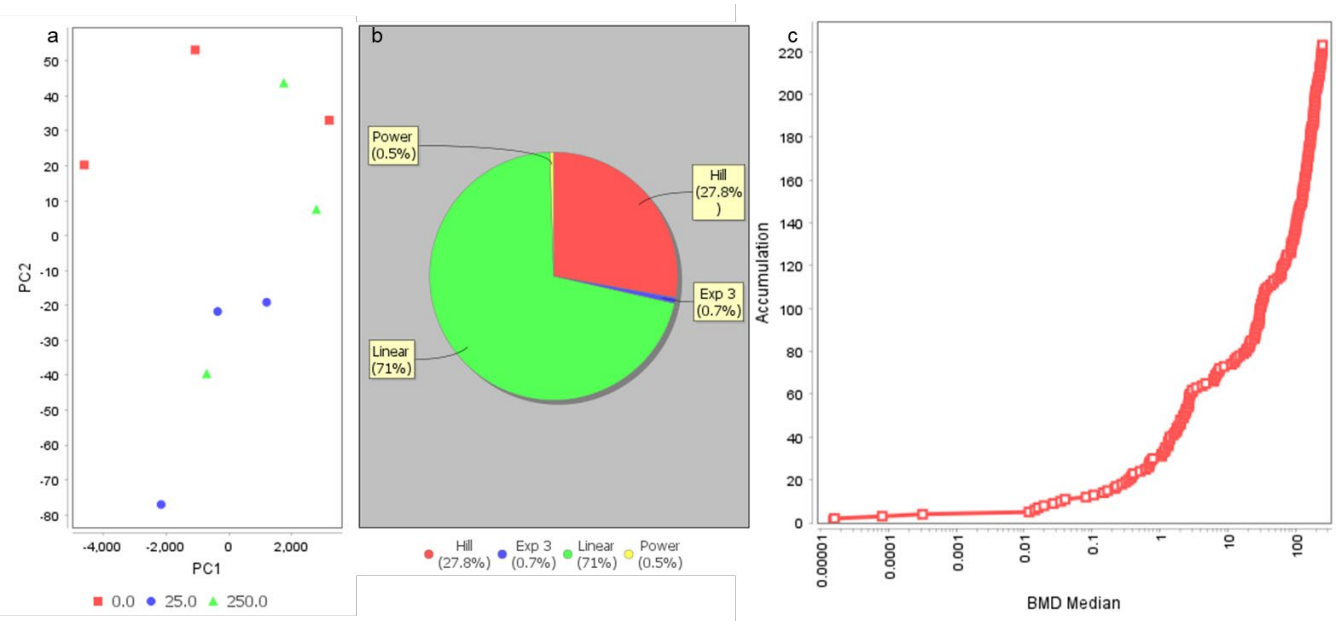


Figure 5 shows (a) separation of data by principal component, (b) best fitting model proportions for each gene, and (c) BMD median accumulation following aggregation by pathway analysis for TCEP exposure at 2 concentrations.

TCIPP

Log2 normalized expressed TPM count data for 16,160 transcripts were imported into BMDExpress 3.0 with Ensembl annotations. Analysis of the first two principal components showed some separation by dose with 2.5 µM being clearly separated from control (Figure 6A). The Williams Trend test identified 417 genes that were likely dose-responsive. During the curve fit pre-filtering tests, 411 genes were identified that could be reasonably fit to a trend. These 411 genes were then brought forward to BMD modeling. Of these genes, the majority (65%) were best fit to a linear model, followed by 33.8% that were best fit to a Hill model (Figure 6B). Using the specified BMR definition and application of the LCRD method (Figure 6C) yielded a POD of 0.081 µM.

Figure 6. TCIPP Dose-Response Modeling Outputs

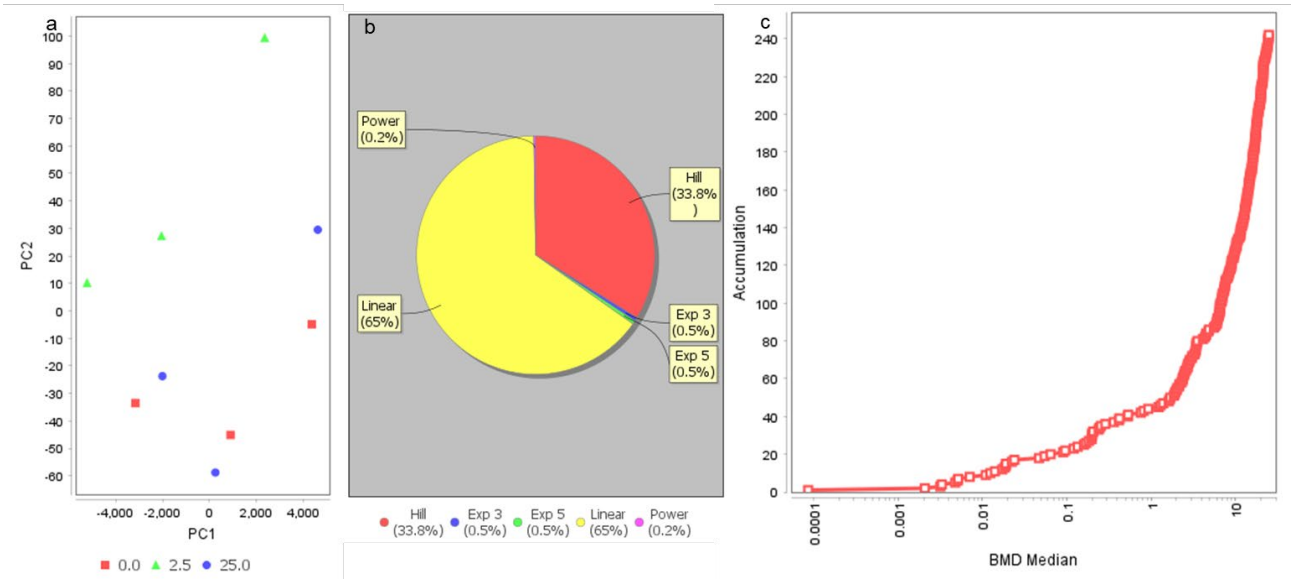


Figure 6 shows (a) separation of data by principal component, (b) best fitting model proportions for each gene, and (c) BMD median accumulation following aggregation by pathway analysis for TCIPP exposure at 2 concentrations.

7.2.2. PHBAF: GSE153363

GSE153363 is the dataset affiliated with Shockley et al. (2020b). Briefly, this study exposed Harlan Sprague Dawley rats to various brominated flame retardants or vehicle via gavage in corn oil for 5 days. For TBBPA-DBPE, animals were exposed to 0.1, 0.94, 9.4, 94.3 or 943 mg/kg/day (Appendix D.1.2.1). Following exposure, RNA was extracted from liver samples and analyzed on the Affymetrix Rat Genome 230 2.0 Arrays.

Log2 normalized RNA processed intensity values for 31,099 probes were imported into BMDExpress. Analysis of the first two principal components showed no separation by dose (Figure 7A). The Williams Trend test identified 122 genes that were likely dose-responsive. Based on the curve fit pre-filtering tests, only 19 genes were identified that could be reasonably fit to a trend. These 19 genes were then brought forward to BMD modeling. Of these genes, the majority (68.4%) were best fit to a Linear model, followed by 15.8% that were best fit to a Hill model (Figure 7B). These genes were then analyzed for the best BMD by individual gene functional analysis as not enough genes were present for pathway analysis (Figure 7C). Again, the LCRD method was used to analyze the genes following individual level gene analysis which yielded a POD of 11.7 mg/kg/day.

Figure 7. TBBPA-DBPE Dose-Response Modeling Outputs

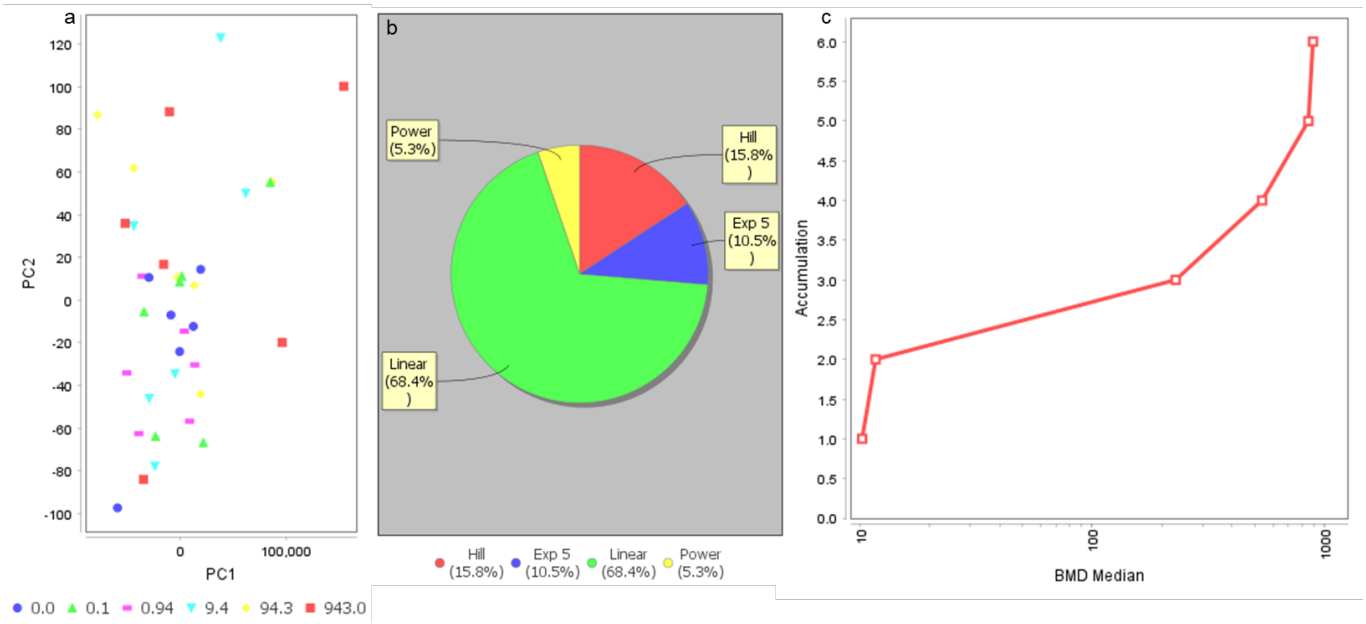


Figure 7 shows (a) separation of data by principal component, (b) best fitting model proportions for each gene, and (c) BMD median accumulation following individual gene analysis for TBBPA-BPDE exposure at 5 concentrations.

8. Deriving TRVs from Analyzed Data

8.1. httk Application to Derive Value

As described in Section .2, the httk_2.3.1 R package was used to conduct IVIVE and convert the *in vitro* PODs calculated in Section 7.2.1 to an *in vivo* dose. TCEP (CASRN 115-96-8) exists in the httk library. The *in vitro* concentration (0.047 µM) was used as an input for calc_mc_oral_equiv function to derive an oral equivalent dose of 0.014 mg/kg/day. For this project, TCIPP (CASRN 13674-84-5) was added to the httk library as a new molecule. CL_{int} and F_{up} values predicted by OPERA default settings were 10.82 µl/min/10^6 cells and 0.2, respectively. An *in vitro* concentration of 0.081 µM was used as an input to calc_mc_oral_equiv function that resulted in an oral equivalent dose of 0.030 mg/kg/day. Note the calc_mc_oral_equiv function includes some Monte Carlo simulation, and thus efforts to replicate these oral equivalent doses will result in slightly different values. Oral equivalent doses predicted by the calc_mc_oral_equiv function represent the 95th percentile.

The POD for TBBPA-DBPE, calculated in Section 7.2.2 was based on an *in vivo* study, and so no IVIVE was needed.

9. Comparing TRVs from Analyzed Chemicals

9.1. Subtask 1A Values

For the three chemicals with quantitative NAMs data (i.e., TCEP, TCIPP, TBBPA-DBPE) in the present subtask, Table 2 summarizes the TRVs based on *in vivo* data that were identified in Subtask 1A from authoritative assessments or derived in the current assessment and identifies in bold the overall TRV recommended in the current assessment. The rest of this section summarizes the key elements of the TRV derivations, comparing the assessments from the various organizations, and presents our rationale for the choice of an overall TRV for each

chemical. The goal of the overall TRV is to have a benchmark for comparison with the *in vitro*-based TRV. It is recognized that a different endpoint or TRV may ultimately be used in the class-based assessment.

Table 22. Summary of TRVs from Authoritative Reviews (TCEP and TCIPP) and TRVs from Task 1B (TBBPA-DBPE)

Organization	Principal Study/Species and Strain (sex)	Exposure Scenario	Critical Effect	Point of Departure Type/Value (mg/kg/day)	Uncertainty Factors or Factors for MOE	TRV (mg/kg/day)
TCEP						
U.S. EPA, 2009	Matthews et al., 1990; NTP, 1991/ Rat, F344/N (female)	Dosed by gavage 5 days/week for 16 weeks.	Increased relative kidney weight	BMDL1SD[ADJ]/ 6.9	UFA = 10 UFH = 10 UFD = 3 UFL = 1 UFS = 3 (total 1,000)	0.007[1]
ATSDR, 2009	NTP, 1991/ Rat, F344/N (female)	Dosed by gavage 5 days/week for 104 weeks.	Renal tubular epithelial hyperplasia	BMDL10[ADJ]/ 23.44	UFA = 10 UFH = 10 (total 100)	0.2
Health Canada, 2009	NTP, 1991; Matthews et al., 1993/ Rats (both)	NR	Renal tubular hyperplasia	LOAEL/ 44	NR	NR (0.4)[2]
ECHA, 2009	Takada et al., 1989/ Mice, Scl:ddY (both)	Exposed daily via feed for 18 months.	Renal tubule epithelium hyperplasia and hypertrophy with nuclei enlargement	LOAEL/ 12	NR	NR (0.1)[2]
U.S. EPA, 2023 (DRAFT)	**Chen et al., 2015/ Mice, ICR (male)**	**Exposed via feed for 35 days.**	**Decreased seminiferous tubules**	**BMDL5[HED]/ 2.73**	**UFA = 3 UFH = 10**	**0.09[3]**
U.S. EPA, 2023 (DRAFT)	NTP, 1991/ Rat, F344/N (female)	Dosed by gavage 5 days/week for 104 weeks.	Renal tubular epithelial hyperplasia	BMDL10[HED] 5.49 (alternative potential POD)	UFA = 3 UFH = 10	0.2[3]
TCIPP						
U.S. EPA, 2012	NTP, 2011/ Mice, B6C3F1 (male)	Exposed daily via feed for 14 weeks.	Increased incidence of hepatocyte hypertrophy	BMDL10/ 138	UFA = 10 UFH = 10 UFD = 10 UFL = 1 UFS = 10	0.01
Health Canada, 2016 (DRAFT)	**TNO Quality of Life, 2007/ Rats, Wistar (both)**	**2 gen study Exposed daily via diet for at least 10 weeks during**	**Increased number of runts**	**LOAEL/ 99**	**NR**	**NR (0.1)[2]**

		premating and mating, throughout gestation and lactation until sacrifice.				
Health Canada, 2016 (DRAFT)	Stauffer Chemical Co., 1981/ Rats, Sprague Dawley (male)	Exposed daily via feed for 13 weeks.	Increased liver weight	LOAEL/ 52	NR	NR (0.05)[2]
ECHA, 2008 (DRAFT)	Stauffer Chemical Company, 1981/ Rats, Sprague Dawley (male)	Exposed daily via feed for 13 weeks.	Increased liver weight	LOAEL/ 52	UFA = 10 UFH = 10 UFL = 1[4]	0.5[3]
ECHA, 2008 (DRAFT)	TNO Quality of Life, 2007/ Rats, Wistar (both)	2 gen study Exposed daily via feed for at least 10 weeks throughout gestation and lactation until sacrifice.	Increased number of runts	LOAEL/ 99	UFA = 10 UFH = 10 UFL = 3 (Total = 300)	0.3[3]
TBBPA-DBPE (Subtask 1B)						
This assessment	Shockley et al., 2020/ Rats, Sprague-Dawley (male)	Gavage in corn oil for 5 days	Decreased T4 levels	NOAEL/94[5]	UFA = 10 UFH = 10 UFD = 10 UFS = 3-10? (Total = 3,000)	0.03

[1]This TRV appears to be overconservative based on both the total UF and the choice of POD, and so was not used for the overall TRV.
[2]For assessments using MOE approach that did not identify the appropriate MOE, the TRV that would be derived using the default values of UFA = 10 and UFH = 10 and UFL = 10 when applicable is shown in parentheses to aid with comparisons.
[3]Implied TRV; assessment reported only POD and information on acceptable MOE.
[4]A factor of 1 was used for LOAEL to NOAEL because the critical effect was "not considered particularly toxicologically significant and the LOAEL is probably quite close to the NOAEL." No additional factor was used for extrapolation to chronic duration because similar effects were seen at similar doses in the 28 and 90-day studies.
[5]NOAEL was erroneously identified as 0.94 mg/kg/day in the Subtask 1B preliminary assessment. More complete analysis for the current report identified the NOAEL as 94 mg/kg/day. See text for rationale.

Five noncancer TRVs were identified for TCEP based on authoritative sources in Subtask 1A. Three of the noncancer TRVs (ATSDR, 2012; Health Canada, 2009; U.S. EPA, 2009) were based on kidney effects (i.e., renal tubular hyperplasia, increased relative kidney weight) in a chronic rat study (NTP, 1991). ECHA (ECB, 2009) also based its TRV on kidney effects (i.e., renal tubular epithelial hyperplasia and hypertrophy), but used an 18-month mouse study (Takada et al., 1989) as the principal study. The corresponding PODs ranged from 6.9 to 44 mg/kg/day, depending on how the BMR was defined, and whether the POD was a BMDL or LOAEL. Not all organizations calculated a TRV, and the total UFs varied from 100 to 1,000; the range of final TRVs was 0.007 to 0.2 mg/kg/day (high end of 0.4 mg/kg/day using the Health Canada (2009) POD and a factor of 100 for the MOE).

U.S. EPA (2023), in a draft assessment, determined that the most sensitive endpoint for a chronic assessment was decreased seminiferous tubules in mice exposed in feed for 35 days (Chen et al., 2015). The POD (human equivalent dose, HED) based on a BMDL5(HED), was 2.73 mg/kg/day. The 2023 EPA assessment also identified a potential chronic POD of 5.49 mg/kg/day based on a BMDL10(HED) for the same study (NTP, 1991) and critical effect (renal tubule hyperplasia) as used by Health Canada, and ATSDR in earlier assessments. Applying a factor of 10 for human variability and a factor of 3 for interspecies differences when using an HED, the implied TRV based on the decreased seminiferous tubules is 0.09 mg/kg/day (note: the draft risk evaluation document identified UFs but used a margin of exposure (MOE) approach, rather than calculating noncancer TRVs.) The U.S. EPA (2009) document calculated a lower TRV, based on the kidney effects, but this is because a larger UF was used (Table 1). **The noncancer TRV of 0.09 mg/kg/day** implied by U.S. EPA (2023) can be considered the best TRV for TCEP, because it is based on a recent evaluation of the data, with an updated assessment reflecting a study published after earlier authoritative reviews. The corresponding value from U.S. EPA (2023) based on kidney effects is 0.2 mg/kg/day, within a factor of two of the most sensitive TRV.

The kidney was also the target for cancer effects of TCEP. IARC (1999) considered TCEP not classifiable as to human carcinogenicity but U.S. EPA (2009, 2023) considered TCEP likely to be carcinogenic to humans based on renal tubule adenomas or carcinomas in rats (NTP, 1991). These assessments calculated an oral slope factor of 0.02 $(mg/kg/day)^{-1}$.

For TCIPP, Health Canada (2016) and ECHA (ECB, 2009) both identified two PODs. A LOAEL of 99 mg/kg/day was identified for reproductive and developmental effects, based on an increased number of runts in a 2-generation study in rats (TNO Quality of Life, 2007). Both organizations also identified a repeated dose LOAEL of 52 mg/kg/day, based on increased liver weight in a 13-week study in rats (Stauffer Chemical Co., 1981). Health Canada (2016) noted that liver effects were not observed at higher doses in the 2-generation study but stated that there were uncertainties regarding the purity of the TCIPP in both of the principal studies. Health Canada (2016) and ECHA (ECB, 2009) used a MOE approach but only ECHA (ECB, 2009) explicitly identified a minimum MOE. U.S. EPA (2012) also considered the critical effect to be in the liver but derived a BMDL10 of 138 mg/kg/day based on a subchronic study in mice (NTP, 2011). U.S. EPA (2012) applied a total UF of 10,000, resulting in a screening-level TRV of 0.01 mg/kg/day. Full factors of 10 were used for interspecies extrapolation, human variability, and for use of a subchronic study. A factor of 10 was also used for database uncertainty, based on the availability of only one developmental toxicity study (the authors were apparently unaware of the TNO [The Netherlands Organisation for Applied Scientific Research] study), and concerns about developmental neurotoxicity. The TRV was characterized as "screening level," because the evaluation was based on the NTP (2011) study abstract and summary tables, but the full study report was not yet published. The Health Canada (2016) assessment came after the U.S. EPA (2012) one but Health Canada apparently chose not to base an assessment on the summary tables.

Several issues with the U.S. EPA value should be noted. First, because it is unlikely for a chemical to be at the high end of the UF distribution for multiple UFs, U.S. EPA practice is generally that four full factors of 10 result in a composite UF of 3,000 not 10,000 (U.S. EPA, 2002b). Second, even though several organizations identified the critical effect as hepatocellular hypertrophy or increased liver weight, this is no longer considered best practice for identifying a POD. In the absence of other signs of liver toxicity, these changes are considered adaptive, not adverse (Hall et al., 2012; U.S. EPA, 2002a). Therefore, the most appropriate critical effect for TCIPP is the increase in the number of runts in the 2-generation study at 99 mg/kg/day (TNO Quality of Life, 2007). ECHA (ECB, 2009) identified a minimum MOE of 300 based on factors of 10 each for interspecies extrapolation, intraspecies variability, and its default factor of 3 for extrapolation from a LOAEL. Health Canada did not identify a minimum MOE but use of the default factor of 10 for use of a LOAEL would result in an overall factor of 1,000. Because the critical effect was a developmental effect in a 2-generation study, no UF is needed for duration extrapolation. **The resulting TRV is 0.1 mg/kg/day.**

Health Canada (2016) noted concern about the potential carcinogenicity of TCIPP in light of the structural similarity to TCEP and TDCPP, which were carcinogenic in rodents, but there were no completed 2-year bioassays of TCIPP at the time of the Health Canada assessment. NTP (2023) reported the results of the 2-year bioassays in rats and mice. That study concluded that there was *some evidence* of carcinogenic activity in male rats, based on the increased incidence of hepatocellular adenoma or carcinoma (combined), and *some evidence* in female rats based on the increased incidence of uterine adenoma or adenocarcinoma (combined), noting that the marginal increase

in the incidence of hepatocellular adenoma in female rats may have been related to exposure. There was also *some evidence* of carcinogenic activity in male mice, based on increased incidence of hepatocellular carcinoma, and *clear evidence* of carcinogenic activity in female mice, based on the increased incidences of hepatocellular adenoma, hepatocellular carcinoma, and hepatocellular adenoma or carcinoma (combined). In light of these prior assessments, and in the absence of consideration of MOA, the Subtask 1B report calculated slope factors based on the tumors seen in the NTP (2023) study. However, as described in the next paragraph, further consideration of the MOA data suggests that these slope factors are not appropriate.

NTP (2023) discussed the carcinogenic MOA of TCIPP. The report stated that TCIPP has little or no genotoxic potential, suggesting a nongenotoxic MOA. In high-throughput screening assays, TCIPP is a weak activator of the constitutive androstane receptor (CAR) and pregnane X receptor (PXR), and toxicogenomic data suggest that TCIPP activates transcripts related to CAR, PXR, and the peroxisome proliferator-activated receptor alpha (PPARα) signaling pathways. NTP (2023) did not address the human relevance of the rodent liver tumors, but the general scientific consensus is that rodent liver tumors that act through these pathways are not relevant to humans (Felter et al., 2018). The potential human relevance of the uterine tumors in rats has not been addressed, but "some evidence" would likely result in a categorization as a *possible human carcinogen*, for which cancer quantification is generally not done.

It is noted that the NTP (2023) study also had data on nonneoplastic effects of TCIPP that could be used to derive an updated noncancer TRV. However, Subtask 1A only included published TRVs, and the searching for Subtask 1B was not conducted for chemicals that had existing TRVs.

9.2. Subtask 1B Values

For TBBPA-DBPE (also known as TBBPA-BDBPE), only one potentially relevant endpoint was identified in Subtask 1B. Sporadic decreased serum T4 concentrations were observed in male Sprague Dawley rats treated by gavage with daily doses of 0.1-943 mg/kg/day (Shockley et al., 2020a, 2020b). Subtask 1B identified the NOAEL as 0.94 mg/kg/day (the second lowest dose), but it is unclear whether the change at this dose was biologically meaningful, since higher T4 levels were observed at 94 mg/kg/day compared to the 0.94 mg/kg/day dose group. It is also noted that the control T4 level was substantially higher than control levels for all of the other tested chemicals. Based on these considerations, it appears that it may be more appropriate to identify 94.3 mg/kg/day as the NOAEL and 943 mg/kg/day as the LOAEL. BMD modeling of this dataset was not successful, due to the nonmonotonic nature of the dose-response. In addition to the thyroid hormone measurements, this study evaluated hematology, clinical serum chemistry, histopathology of the liver and thyroid, and microarray analysis of liver tissue. Thus, the study evaluated a range of sensitive endpoints, but had incomplete histopathology. Transcriptional analysis was conducted only for the liver, and the authors reported "no significant liver disease-related transcript changes." In contrast, our analysis of the liver transcriptomic data identified a few differentially expressed genes, for which a POD of 11.7 mg/kg/day was calculated (Section 7.2.2).

Due to the acute exposure duration, no chronic TRV was derived from this study (Shockley et al., 2020a, 2020b) in Subtask 1B. However, for the purposes of the current assessment, one can consider the implications of the NOAEL for T4 changes of 94.3 mg/kg/day and the absence of biologically meaningful transcriptomic changes in the liver. Default UFs of 10 for intraspecies variability and 10 for interspecies extrapolation in the absence of an HED would be applied. A UF of 10 for database deficiencies would be applied, reflecting the absence of other *in vivo* studies. This UF of 10 is also consistent with the database UF used in the development of EPA transcriptomic assessment products (ETAPs, U.S. EPA, 2024). A factor of 3 or 10 might be used for duration extrapolation. This factor takes into account the absence of transcriptomic changes in the liver, the standard ETAP approach of conducting transcriptomics in a much broader variety of tissues, and the finding that no duration UF is needed when the POD is based on the GO biological process class with the lowest median BMD value using the ETAP approach. Overall, a total UF of 3,000 might result. Applying this to the NOAEL of 94 mg/kg/day results in **an approximate TRV of 0.03 mg/kg/day.**

9.3. Subtask 1C Values

As noted above, there is substantial uncertainty in the derivation of TRVs from the transcriptomic data, due to the small number of transcriptomic datasets identified for the PHOPs and PHBAFs. This uncertainty is increased due to the mismatch between the tissues evaluated in the transcriptomic studies and the critical effects identified in Sections 9.1 and 9.2. Specifically, the transcriptomic data for both TCEP and TCIPP were obtained from HepG2 cells, a human liver cell line. The primary targets of TCEP exposure were the kidney and, in a 35-day study, the male reproductive tract. For TCIPP, the TRV was based on a developmental effect, increased number of runts. Increased hepatocyte hypertrophy and increased liver weight were seen at somewhat lower doses, but these changes alone are not considered adverse, and thus not appropriate as the basis for the TRV.

Despite this mismatch in targets, the transcriptomic data might still constitute an appropriate POD, based on the supposition that effects on cellular processes begin to be affected at similar doses even in different organs. However, there is also a substantial mismatch between the tPODs and apical PODs. Since the transcriptomic data are not derived from the primary target tissue, one might expect the *in vitro* POD (after IVIVE) to be somewhat *higher* than the *in vivo* POD. Instead, the calculated transcriptomic PODs (tPODs) from the *in vitro* data are orders of magnitude *lower* than the PODs based on apical endpoints. Thus, the tPOD for TCEP is 0.014 mg/kg/day, compared with the POD of 2.73 mg/kg/day based on decreased seminiferous tubules. Similarly, the TCIPP tPOD was 0.03 mg/kg/day, while the POD based on an increased number of runts was a LOAEL of 99 mg/kg/day. Even if one were to use the liver changes as the POD for TCIPP, the result would be a (minimal) LOAEL of 52 mg/kg/day. While the approach for the tPOD is designed to be conservative, that is, to be comparable to or somewhat lower than the *in vivo* POD, the goal is for there to be a maximum of about 10-fold difference between the tPOD and apical POD. The reason for the apparently extremely low tPODs in this case is unclear.

An advantage with the data for TBBPA-DBPE is that the transcriptomic data and *in vivo* POD come from the same 5-day study. On the other hand, it is difficult to use the *in vivo* data to benchmark a TRV, in the absence of longer duration studies, or any reproductive or developmental toxicity studies. Nonetheless, it is reassuring that the tPOD for TBBPA-DBPE of 11.7 mg/kg/day is lower than the *in vivo* POD of 94 mg/kg/day, and within an order of magnitude.

In light of the uncertainties associated with the tPODs, the discussion of potential UFs here will be brief.

- As discussed in the Subtask 2A report, the default UF of 10 would be appropriate for intraspecies variability.
- No UF for interspecies variability would be needed for the PHOPs, since the tPODs are from a human cell line. For TBBPA-DBPE, a UF of 3 could be used for interspecies variability if the dose were converted to an HED; otherwise, a factor of 10 is used.
- No LOAEL to NOAEL factor is needed because the tPODs are based on BMD modeling.
- As noted, there is considerable uncertainty associated with the tPODs because quantitative transcriptional data are available only for the liver, raising questions about the adequacy of the toxicological coverage of all potential effects. A UF of 10 appears appropriate here for both the *in vitro* and *in vivo* tPODs, consistent with the approach of both Health Canada (2021) and U.S. EPA (2024).
- The Subtask 2A report discussed other potential UFs for the *in vitro* tPODs, including a subchronic to chronic factor, a bioactivity UF, and a UF for addressing IVIVE *(*note that these UF areas may include substantial overlap). A UF for these areas of uncertainty does not appear to be needed for TCEP or TCIPP in light of the very low tPODs. Evaluation of the need for an additional UF for TBBPA-DBPE is challenging in the absence of an authoritative TRV based on traditional methods. However, the finding that the tPOD (11.7 mg/kg/day) is almost a full order of magnitude lower than the *in vivo* 5-day POD may support the decision that no additional UF is
  needed.

10. Discussion

The purpose of this report was to (1) summarize methods that are currently used to derive TRVs from NAMs data and (2) derive a TRV from one stream of NAMs data for the PHOPs and PHBAFs. Overall, we identified computational, *in vitro,* and *in vivo* non-mammalian models that are currently being used to derive PODs for chemicals with the goal of finding alternative methods to identify PODs from toxicity testing. Based on our findings in this report, coupled with those presented in Subtask 2A, we undertook the task of deriving TRV values from transcriptomic data. This approach was chosen because data used in these analyses are deposited in public repositories, and transcriptomics data suffers from fewer issues surrounding inter-lab and inter-experiment variability than other data types. This concept was explored more fully in the Subtask 2A report.

To derive the values presented in this report, we queried three databases, reviewed studies identified in Subtask 2A, and assessed expert identified papers. Despite initially identifying 22 possible datasets, only two met the criteria for derivation of a TRV. Issues identified during this process were (1) experimental design, (2) data accessibility, and (3) data annotation. The most common issue we identified was that the majority of studies used only one dose group when measuring molecular endpoints. This made these studies unsuitable for dose-response modeling, although they were used for hazard characterization. Additionally, not all data is publicly accessible. For example, we identified transcriptomic data for TCIPP within the 2023 NTP report on TCIPP (NTP, 2023). However, this data is not currently accessible to the public. This lack of deposition of data was a major issue that we encountered in our attempts to perform TRV derivation; the majority of the identified studies lacked accessible transcriptomic data. This is a major area that should be addressed as deposition of the data into publicly accessible databases could lead to a significant increase in chemical coverage. The last major issue we identified was data annotation. We lost 33% of the accessible counts data due to poor annotation or inadequate sample design. Specific issues included mismatch between sample ids and the identified chemical treatments, inappropriate sample sizes (n=1/exposure group was unusable), and deposition of data that did not correspond with the experiment described. These issues all further reduced the number of datasets available for analysis. Because data availability and curation were the major factors limiting our ability to derive TRVs, efforts to ensure deposition of results in data repositories and appropriate curation should be undertaken by funding and regulatory agencies. This will help to facilitate the use of NAMs data for both hazard characterization and TRV derivation.

In addition to limited deposition of data, we found limitations in the model systems studied. For TCEP, based on *in vivo* data, the key targets are kidney and reproductive outcomes based on *in vivo* data. For TCIPP, the key target *in vivo* is reproductive outcomes. However, the only available data were from Hep2G (liver) cells. Similarly, for TBBPA- DBPE, the transcriptional data are similarly from the liver from animals exposed for 5 days. The only POD available for TBBPA-DBPE is for thyroid hormone levels in a 5-day study. Further, the toxicity database is not sufficient to confirm that the thyroid is the main target. This is an important caveat to the TRVs derived in this report- the tissue in which they were derived is not a target organ of interest. Alignment of target tissues that represent apical endpoints of interest would be a worthwhile exercise as each tissue and cell type have their own unique baseline transcriptome (Fagerberg et al., 2014). As such, the TRVs derived here are of limited regulatory utility. Similar problems plague the hazard characterization findings: the majority of studies are conducted in embryonic stem cells, whole embryos, or embryo bodies. These data thus provide limited information about non-developmental impacts of OFRs. Additionally, this narrow focus could help to explain the limited transcriptomic effects if these are not the tissues of interest. Greater emphasis on a variety of different cell types and organs should be the focus of future research as it will enable better characterization of potential hazards identified during studies of apical endpoints, as well as providing data for anchoring NAMs-based assessments relative to apical endpoints.

In conclusion, we identified methods to derive TRVs from various NAMs data, and derived TRVs from transcriptomic data for PHOPs and PHBAFs. Data accessibility, quality, and experimental design were factors that limited the utility of this exercise. To improve this process and move it towards regulatory standards, guidelines regarding study design, accessibility of data and curation of data should all be developed and/or implemented.

Appendix A. Literature Search String and Strategy

A.1. Peer Database Search

The peer database search was executed on 12/04/2023.

A.1.1. Approaches for Deriving a Toxicological Reference Value (TRV) using New Alternative Methods (NAMs) Data

A.1.1.1. PubMed

Search String:

(“Toxicological Reference Value” OR "Toxicological Reference Values" OR "Toxicity Reference Value" OR “Toxicity Reference Values” OR “Reference Dose” OR “Reference Doses” OR “Reference Dosage” OR “Reference Concentration” OR “Reference Concentrations” OR “AccepTable Daily Concentration” OR “AccepTable Daily Concentrations” OR “AccepTable Daily Intake” OR “AccepTable Daily Intakes” OR “Minimal risk level” OR “Minimal risk levels” OR “Cancer potency estimate” OR “Cancer potency estimates” OR “Cancer Potency Estimation” OR “Cancer Potency Estimations” OR “slope factor” OR “slope factors” OR “relative hazard potency” OR “relative hazard potential” OR “margin of exposure” OR “margins of exposures” OR “threshold of toxicological concern” OR “Tolerable Daily Intake” OR “Tolerable Daily Intakes” OR “Inhalation Unit Risk” OR (TRVs AND Toxicol*[tiab]))

AND

("New Approach" OR “New Approaches” OR "Alternative method" OR "Alternative methods" OR "alternative methodology" OR "Novel Method" OR "novel methods" OR "novel methodology" OR "New method" OR "new methods" OR "new methodology" OR NAMs[ti] OR "Alternative data" OR "Novel data" OR "unique data" OR "alternate data" OR "in vitro data" OR "mechanistic"[tiab] OR "computational method" OR "computational methods" OR "computational methodology" OR “Read Across” OR transcriptomic* OR QIVIVE OR “Quantitative in vitro to in vivo extrapolation” OR IVIVE OR “in vitro to in vivo extrapolation”)

Limits: None

A.1.1.2. Web of Science

Search String:

TS=(“Toxicological Reference Value” OR "Toxicological Reference Values" OR "Toxicity Reference Value" OR “Toxicity Reference Values” OR “Reference Dose” OR “Reference Doses” OR “Reference Dosage” OR “Reference Concentration” OR “Reference Concentrations” OR “AccepTable Daily Concentration” OR “AccepTable Daily Concentrations” OR “AccepTable Daily Intake” OR “AccepTable Daily Intakes” OR “Minimal risk level” OR “Minimal risk levels” OR “Cancer potency estimate” OR “Cancer potency estimates” OR “Cancer Potency Estimation” OR “Cancer Potency Estimations” OR “slope factor” OR “slope factors” OR “relative hazard potency” OR “relative hazard potential” OR “margin of exposure” OR “margins of exposures” OR “threshold of toxicological concern” OR “Tolerable Daily Intake” OR “Tolerable Daily Intakes” OR “Inhalation Unit Risk” OR (TRVs AND Toxicol*))

AND

TS=("New Approach" OR “New Approaches” OR "Alternative method" OR "Alternative methods" OR "alternative methodology" OR "Novel Method" OR "novel methods" OR "novel methodology" OR "New method" OR "new methods" OR "new methodology" OR "Alternative data" OR "Novel data" OR "unique data" OR "alternate data" OR "in vitro data" OR "mechanistic" OR "computational method" OR "computational methods" OR "computational

methodology" OR “Read Across” OR transcriptomic* OR QIVIVE OR “Quantitative in vitro to in vivo extrapolation” OR IVIVE OR “in vitro to in vivo extrapolation”)

Limits: None

A.2. Grey Literature Google Web Scraping

All grey literature searches were executed on 12/04/2023. Table A-1 lists the grey literature sources that were searched using the Google web scraping tool.

Table A-1. Grey Literature Sources

Source	Type	URL Domain
US Environmental Protection Agency (EPA)	Federal, Regulatory	epa.gov
National Institute of Environmental Health Sciences (NIEHS)	Federal, Research	niehs.nih.gov
Centers for Disease Control and Prevention (CDC)	Federal, Regulatory	cdc.gov
National Technical Reports Library (NTRL)	Federal, Repository	ntrl.ntis.gov
National Academies of Sciences, Engineering, and Medicine (NASEM)	NGO, Research	nationalacademies.org
National Institute for Public Health and the Environment (RIVM)	International, Regulatory	rivm.nl
European Chemicals Agency (ECHA)	International, Regulatory	echa.europa.eu
European Food Safety Authority (EFSA)	International, Regulatory	efsa.europa.eu
Organisation for Economic Co-operation and Development (OECD)	International, Research	oecd.org
Health Canada (HC)	International, Regulatory	canada.ca
Japanese Ministry of Health	International, Regulatory	mhlw.go.jp

A.2.1.1. Approaches for Deriving a Toxicological Reference Value (TRV) using New Alternative Methods (NAMs) Data

The below search string was used. Due to strict character limits in Google searching, the string was edited from the peer database string.

(“Toxicological Reference Value” OR “Reference Dose” OR “Daily Concentration” OR “Minimal Risk Level” OR “Cancer Potency” OR “Slope factor”) AND (“New Approach” OR “Alternative Methods” OR “Alternative Data” OR “In vitro data” OR “mechanistic” OR “computational methods”)

Appendix B. Generative Artificial Intelligence (GenAI) for Full Text Screening

B.1. Prompts Used for Full Text Screening

A set of 10 prompts were developed to categorize 10 expert curated PDFs according to topic. Subject matter experts graded these prompts for accuracy against the seed studies and confirmed they were sufficient for this subject matter topic. These 10 prompts were used to review the 356 peer reviewed papers identified as containing potential approaches to deriving a TRV from NAMs data.

Table B-1 provides an overview of the 10 prompts that were used for testing. Table B-2 lists the various parameters used for the GenAI.

Table B-1. Final Prompt Questions

Topic	Prompt Question
Framework	Does this paper present a framework, method, decision tree, toolbox or workflow for deriving toxicity reference values (TRV) for chemicals using in vitro data, New Approaches Methods (NAMs), toxicogenomics, early mechanistic endpoints, or nonmammalian data?
Applicable to OFR Chemicals	Does this paper present data on flame retardants?
Surrogate	Does this paper discuss methods or criteria for choosing a surrogate or anchor chemical for a class-based assessment or grouping or read-across?
BMR	Does this paper discuss methods or issues related to the establishment of a point of departure (POD), benchmark response (BMR), AC50 (50% activity concentration, or concentration at which 50% of maximum activity is observed), ACC (activity concentration at cutoff), or administered equivalent dose (AED)?
AOP	Does this paper discuss methods or issues related to filling in quantitative Adverse Outcome Pathways (AOPs) with New Approach Methods (NAMs) data (or otherwise using AOPs and NAMs data to inform deriving a toxicity reference value (TRV)?
IVIVE	Does this paper discuss methods or issues related to in vitro to in vivo extrapolation (IVIVE) or administered equivalent dose (AED)?
NAMs UF	Does this paper discuss methods or issues related to deriving uncertainty factors for New Approach Methods (NAMs) data (duration, use of NAMs data, etc.)?
NAMs RA	Does this paper discuss methods or issues related to applying New Approach Methods (NAMs) in risk assessment?
NAM Type	What types of New Approach Methods (NAMs) are used in this paper?
RPF TEF	Does the paper contain anything about the application of relative potency factors (RPFs) or toxicity equivalent factors (TEFs), or any other similar method of comparing potency of related chemicals?

Table B-2. Final Parameters Used for GenAI Full Text Screening

Parameter	Final Details
GenAI	Claude version 3.0

Parameter	Final Details
Preamble	Please act as a toxicologist. Wrap each response in an XML tag based on the prefix 'response' concatenated with the question number.
Post-amble	If yes, justify your response with quotes. If you don't know, say "I don't know."

Appendix C. List of PHOP and PHBAF Chemicals, as well as PHOP Metabolites

A list of PHOP and PHBAF chemicals as well as PHOP metabolites are listed in Table C-1. The metabolite list for PHOPs was refined during CO-1, Call Order 61320622F2011, "Guidance Document for Conducting Class-based Hazard Assessment of Organohalogen Flame Retardants" under CPSC Blank Purchase Agreement (BPA 61320622A0005). For the metabolites, only the first three metabolites for each PHOP chemical were searched. The PHOP and PHBAF chemicals were searched for in both the Comparative Toxicogenomic Database (CTD) and the Gene Expression Omnibus (GEO) database. The top three metabolites for each PHOP chemical were also searched in the GEO database.

Table C-1. PHOP and PHBAF Chemicals and Metabolites

Chemical	Common Acronym	CASRN	Searched in CTD	Searched in GEO
PHOP Chemicals				
2,3-Dibromopropylphosphate	BDBPP	5324-12-9	X	X
2,2-Bis(chloromethyl)-1,3-propanediyl tetrakis(1-chloro-2-propanyl) bis(phosphate)	N/A	1047637-37-5	X	X
Tris(3-chloropropyl)phosphate	N/A	1067-98-7	X	X
Tris(2-chloroethyl) phosphate	TCEP	115-96-8	X	X
Bis(2-chloroethyl) vinylphosphonate	N/A	115-98-0	X	X
Tris(2,3-dibromopropyl) phosphate	TDBPP	126-72-7	X	X
Tris(2-chloroisopropyl)phosphate	TCIPP	13674-84-5	X	X
Tris(1,3-dichloro-2-propyl) phosphate	TDCIPP	13674-87-8	X	X
Tris(2-chloroethyl) phosphite	N/A	140-08-9	X	X
Tris(tribromoneopentyl)phosphate	TTBNPP	19186-97-1	X	X
Tris(dichloropropyl) phosphate	N/A	26604-51-3	X	X
Ethanol, 2-bromo-, phosphate (3:1)	N/A	27568-90-7	X	X
Phosphoric acid, 1,2-ethanediyl tetrakis(2-chloroethyl) ester	N/A	33125-86-9	X	X
Tetrakis(1-chloropropan-2-yl) ethane-1,2-diyl bis(phosphate)	N/A	34621-99-3	X	X
Phosphoric acid, 2,2-bis(chloromethyl)-1,3-propanediyl tetrakis(2-chloroethyl) ester	BCMP-BCEP; V6	38051-10-4	X	X
Phosphonic acid, P-[1-[[(2-chloroethoxy)(2-chloroethyl)phosphinyl]oxy]ethyl]-, 1-[bis(2-chloroethoxy)phosphinyl]ethyl 2-chloroethyl ester	N/A	4351-70-6	X	X

Chemical	Common Acronym	CASRN	Searched in CTD	Searched in GEO
Diethylene glycol bis[bis(2-chloroethyl)phosphate]	N/A	53461-82-8	X	X
Bis(2,3-dibromopropyl) hydrogen phosphate	DDBPP	5412-25-9	X	X
2,4,8,10-Tetraoxa-3,9-diphosphaspiro[5.5]undecane, 3,9-bis[3-bromo-2,2-bis(bromomethyl)propoxy]-, 3,9-dioxide	N/A	61090-89-9	X	X
Tris(2-chloropropyl) phosphate	N/A	6145-73-9	X	X
Bis(2-chloroethyl) 2-chloroethylphosphonate	N/A	6294-34-4	X	X
2,2-Bis(bromomethyl)-3-chloropropyl bis[2-chloro-1-(chloromethyl)ethyl] phosphate	N/A	66108-37-0	X	X
Tris(1,3-dichloropropan-2-yl) phosphite	N/A	6749-73-1	X	X
Bis(1,3-dichloropropan-2-yl) hydrogen phosphate	N/A	72236-72-7	X	X
Bis(2-chloro-1-methylethyl) 2-chloropropyl phosphate	N/A	76025-08-6	X	X
(2-Chloro-1-methylethyl) bis(2-chloropropyl) phosphate	N/A	76649-15-5	X	X
Tris(2,3-dichloropropyl)phosphate	TDCPP	78-43-3	X	X
2-Bromoethyl 5-bromopentyl 2-chloroethyl phosphate	N/A	84282-27-9	X	X
4-Bromo-2-chlorobutyl 3-bromo-2,2-dimethylpropyl phosphate	N/A	98923-48-9	X	X
PHOP Metabolites				
Epibromohydrin 1-Bromo-2,3-epoxypropane 2-(Bromomethyl)oxirane	NA	NA	NS	X
Chloroacetone 1-Chloropropan-2-one Chloropropanone	NA	NA	NS	X
2-Chloropropionic acid 2-CHLOROPROPANOIC ACID Propanoic acid, 2-chloro-	NA	NA	NS	X
2-Chloropropan-1-ol 2-CHLORO-1-PROPANOL 2-Chloropropanol	NA	NA	NS	X
1-Chloro-2-propanol 1-Chloropropan-2-ol 2-PROPANOL, 1-CHLORO-	NA	NA	NS	X
1,3-Dichloroacetone 1,3-dichloropropan-2-one 1,3-Dichloro-2-propanone	NA	NA	NS	X
2,3-DIBROMOPROPIONIC ACID 2,3-Dibromopropanoic acid	NA	NA	NS	X

Chemical	Common Acronym	CASRN	Searched in CTD	Searched in GEO
Propanoic acid, 2,3-dibromo-				
3-bromo-2,2-bis(bromomethyl)propanoic Acid 2,2-Bis(bromomethyl)-3-bromopropanoic acid Propanoic acid, 3-bromo-2,2-bis(bromomethyl)-	NA	NA	NS	X
2,3-DICHLOROPROPIONIC ACID 2,3-Dichloropropanoic acid Propanoic acid, 2,3-dichloro-	NA	NA	NS	X
3-CHLOROPROPIONIC ACID 3-Chloropropanoic acid Propanoic acid, 3-chloro-	NA	NA	NS	X
chloroacetic acid Monochloroacetic acid 2-chloroacetic acid	NA	NA	NS	X
Diglycolic acid, 2,2-OXYDIACETIC ACID, Acetic acid, 2,2-oxybis-	NA	NA	NS	X
2,3-Dibromopropanal, Acrolein dibromide, Propanal, 2,3-dibromo-	NA	NA	NS	X
2,3-Dichloropropanal 2,3-Dichloropropionaldehyde Propanal, 2,3-dichloro-	NA	NA	NS	X
Bromoacetaldehyde 2-Bromoacetaldehyde Acetaldehyde, bromo-	NA	NA	NS	X
CHLOROACETALDEHYDE, 2-chloroacetaldehyde Chloroethanal	NA	NA	NS	X
ETHEPHON 2-Chloroethylphosphonic acid (2-Chloroethyl)phosphonic acid	NA	NA	NS	X
1,3-DICHLORO-2-PROPANOL 1,3-Dichloropropan-2-ol Dichlorohydrin,	NA	NA	NS	X
2,4,6-tribromophenol Tribromophenol Bromol	NA	NA	NS	X
2,4-DIBROMOPHENOL Phenol, 2,4-dibromo- 2,4-Dibromo-phenol	NA	NA	NS	X
2,3-Dibromo-1-propanol 2,3-Dibromopropan-1-ol 2,3-DIBROMOPROPANOL	NA	NA	NS	X
2,3-Dichloro-1-propanol 2,3-Dichloropropan-1-ol 1-Propanol, 2,3-dichloro-	NA	NA	NS	X

Chemical	Common Acronym	CASRN	Searched in CTD	Searched in GEO
2-BROMOETHANOL Ethylene bromohydrin Ethanol, 2-bromo-,	NA	NA	NS	X
3-CHLORO-1-PROPANOL 3-Chloropropan-1-ol 3-Chloropropanol	NA	NA	NS	X
2-chloroethanol Ethylene chlorohydrin Ethanol, 2-chloro-	NA	NA	NS	X
PHBAF Chemicals				
(Propane-2,2-diyl)bis(2,6-dibromo-4,1-phenylene) diprop-2-enoate	TBBPA-BA	55205-38-4	X	X
3,3',5,5'-Tetrabromobisphenol A	TBBPA	37419-42-4	X	X
Bis(p-acryloxyethoxy)tetrabromobisphenol A	TBBPA-BHEEBA	66710-97-2	X	X
Tetrabromobisphenol A-bis(2,3-dibromopropyl ether)	TBBPA-BDBPE	21850-44-2	X	X
Tetrabromobisphenol A diallyl ether	TBBPA-BAE	25327-89-3	X	X
Tetrabromobisphenol A bis(2-hydroxyethyl) ether	TBBPA-BHEE; TBBPA-OHEE	4162-45-2	X	X
Phenol, 4,4'-(1-methylethylidene)bis[2,6-dibromo-, 1,1'-diacetate	TBBPA-BOAc	33798-02-6	X	X
Tetrabromobisphenol A dimethyl ether	TBBPA-BME	37853-61-5	X	X
1,1'-(Isopropylidene)bis(3,5-dibromo-4-(2,3-dibromo-2-methylpropoxy)benzene)	N/A	97416-84-7	X	X
2,2'-[(1-Methylethylidene)bis[(2,6-dibromo-4,1-phenylene)oxymethylene]]bis[oxirane]	TBBPA-BGE	3072-84-2	X	X
3,3',5,5'-Tetrabromobisphenol A bispropionate	TBBPA-BP	37419-42-4	X	X
2,2',6,6'-Tetrachlorobisphenol A	TCBPA	79-95-8	X	X
1,1'-sulfonylbis(3,5-dibromo-4-methoxybenzene)	TBBPS-BME	70156-79-5	X	X
4,4'(propane2,2diyl)bis(2bromophenol)	N/A	29426-78-6	X	X

X = searched in the database; NS = not searched in the database; N/A = not applicable as some chemicals and metabolites do not have common abbreviations or CASRNs affiliated with them. For PHOP metabolites, the first three chemical names were searched in GEO to ensure coverage. Each metabolite is separated by a line break. For PHBAFs, no metabolites were identified in CO-1 and thus were not searched.

Appendix D. Use of Transcriptomic Data in Hazard Analysis

D.1. Polyhalogenated Organophosphate (PHOP)

D.1.1. Tris(1,3-dichloro-2-propyl) Phosphate (TDCIPP)

D.1.1.1. Dataset Number – GSE106875

Data Type: Array

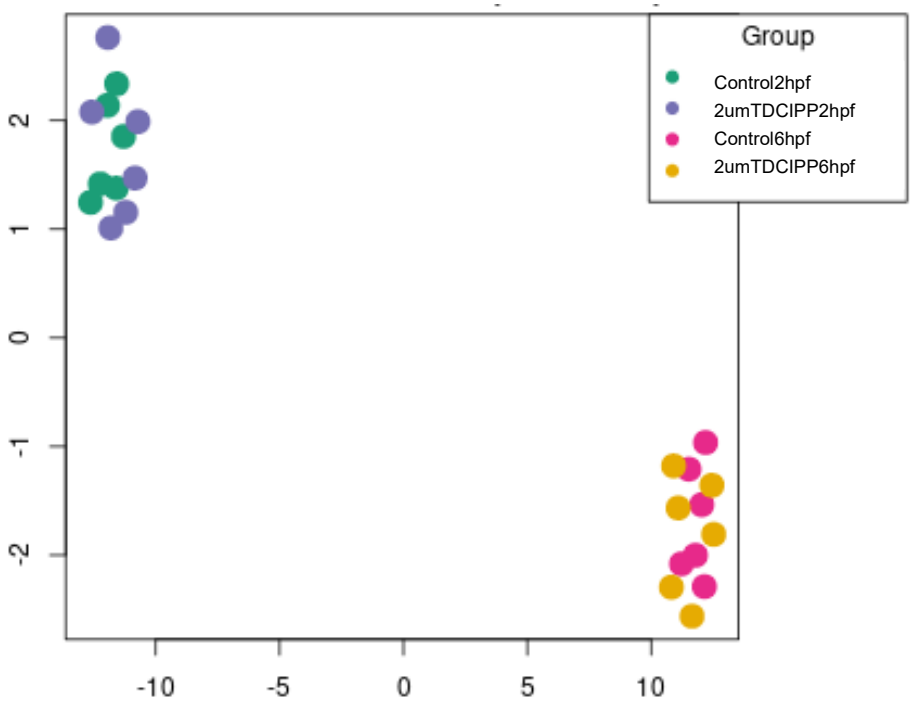
Model Type: *In vivo*

Species: Zebrafish *(Danio rerio)*

Exposure Paradigm: Zebrafish embryos were exposed to 2 µM TDCIPP for 2 or 6 hours post fertilization (hpf).

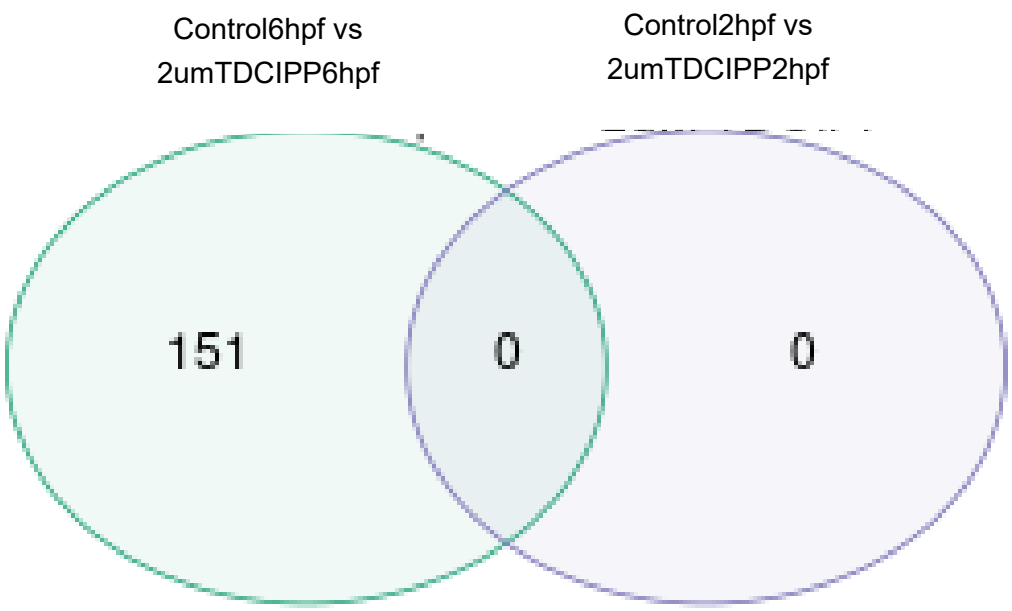
Note: Gene set enrichment analysis was run but no significant pathways were identified.

Figure D-1. Uniform Manifold Approximation and Projection (UMAP) Plot for All Data Contained Within the GSE106875 Dataset



Based on this UMAP plot, hours post fertilization (hpf) are the main drivers of difference, with limited separation by exposure to TDCIPP.

Figure D-2. A Venn Diagram Indicating the Number of Genes Differentially Expressed Between Control and Exposed Embryos for All Data Contained Within the GSE106875 Dataset



TDCIPP affects a limited gene set at a longer duration of exposure (6 hours post fertilization) as indicated by no changes in gene expression at the 2-hour time point and 151 changes in gene expression at the 6 hour time point.

D.1.1.2. Dataset Number – GSE147560

Data Type: High-Throughput Screening

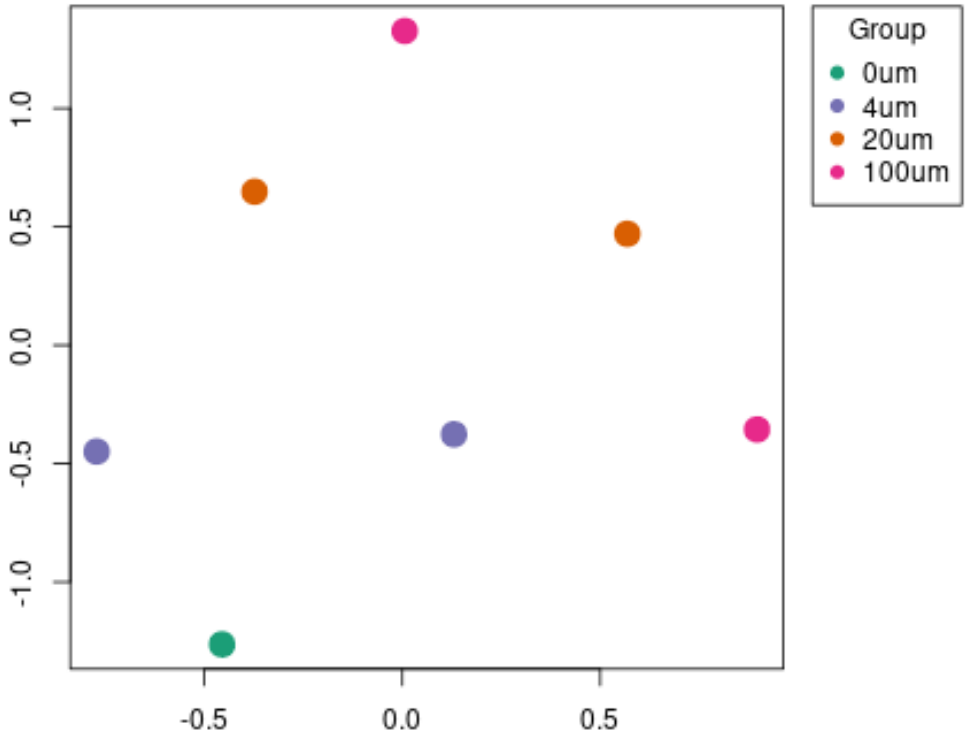
Model Type: *In vitro*

Cell Type: HeLa cells

Exposure paradigm: HeLa cells were exposed to TDCIPP for 2 and 24 hours at concentrations of 4 µM, 20 µM, and 100 µM.

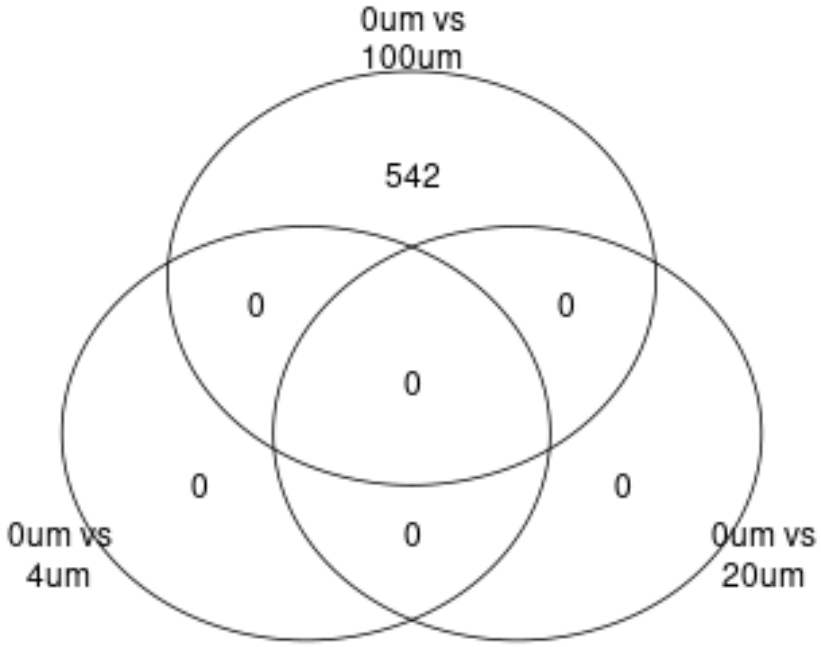
Note: Gene set enrichment analysis could not be conducted due to n=1 per condition. Since the data quality was insufficient, it was not taken forward for BMD modeling.

Figure D-3. Uniform Manifold Approximation and Projection (UMAP) Plot for All Data Contained Within the GSE147560 Dataset



Based on this UMAP plot, there is limited separation by TDCIPP exposure concentration.

Figure D-4. A Venn Diagram Indicating the Number of Genes Differentially Expressed Between the Control and Exposed HeLa Cells for All Data Contained Within the GSE147560 Dataset



The high dose of TDCIPP (100 µM) induced 542 changes in gene expression compared to the control dose while no changes in gene expression were noted following exposure to the low (4 µM) and medium (20 µM) dose of TDCIPP.

D.1.2. Tris(2-chloroethyl) Phosphate (TCEP) and Tris(2-chloroisopropyl) Phosphate (TCIPP)

D.1.2.1. Dataset Number – GSE109463

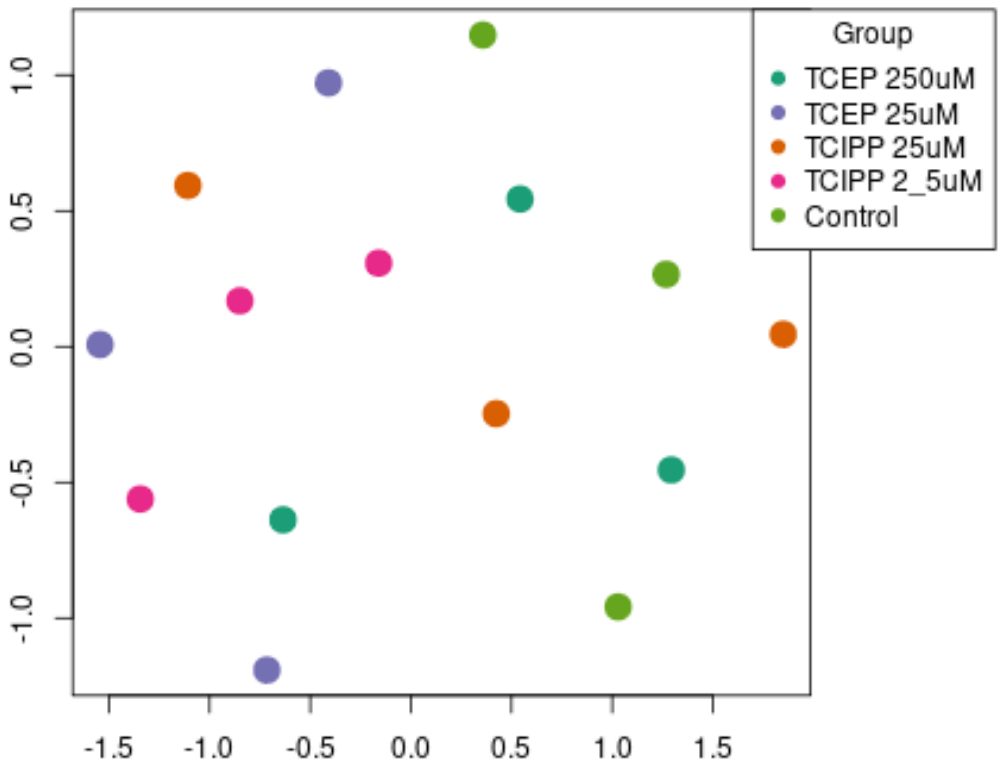
Data Type: High-Throughput Screening

Model Type: *In vitro*

Cell Type: HepG2 cells

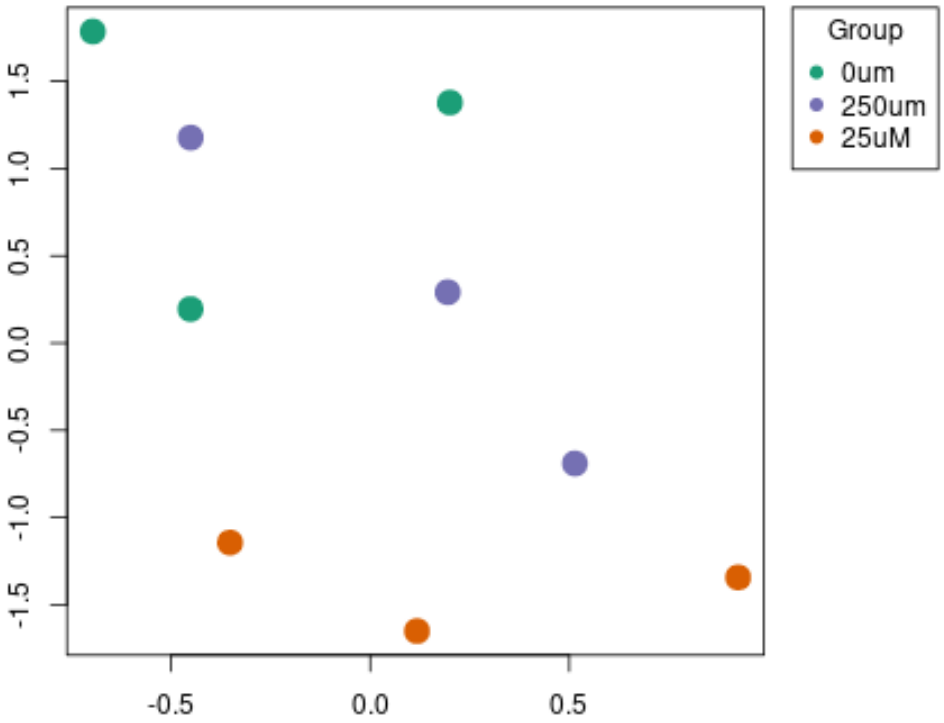
Exposure paradigm: HepG2 cells were exposed to TCEP at concentrations of 25 µM or 250 µM and TCIPP at concentrations of 2.5 µM or 25 µM for 72 hours

Figure D-5. Uniform Manifold Approximation and Projection (UMAP) Plot for All Data Contained Within the GSE109463 Dataset



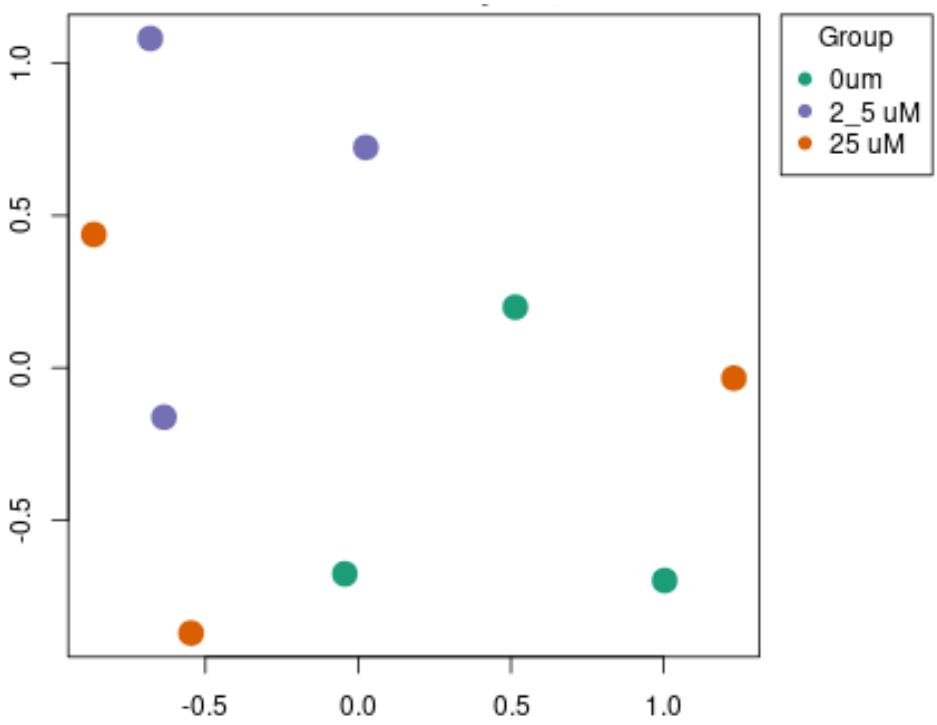
Based on this UMAP plot, there is limited separation between concentrations, chemicals and control.

Figure D-6. Uniform Manifold Approximation and Projection (UMAP) Plot for TCEP Contained Within the GSE109463 Dataset



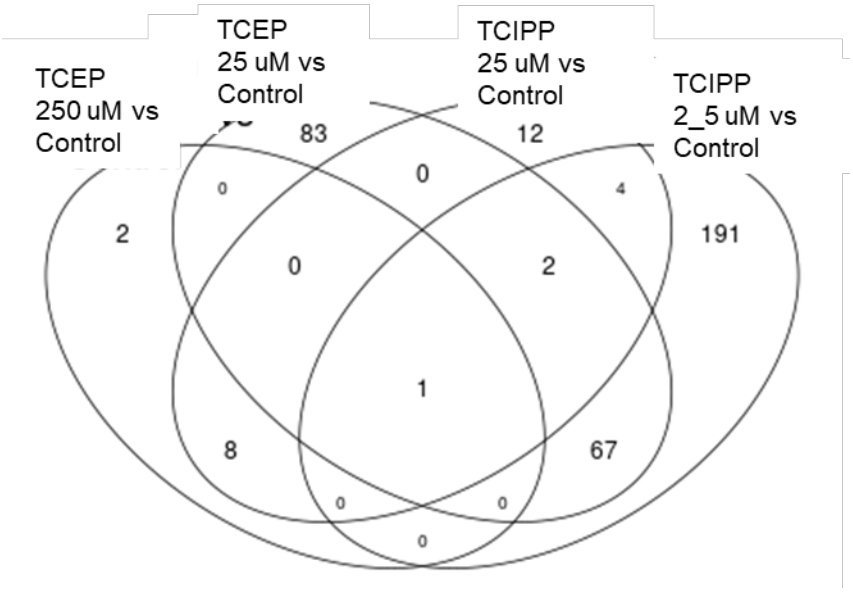
Based on this UMAP plot, there is limited separation by TCEP exposure concentrations.

Figure D-7. Uniform Manifold Approximation and Projection (UMAP) Plot for TCIPP Contained Within the GSE109463 Dataset



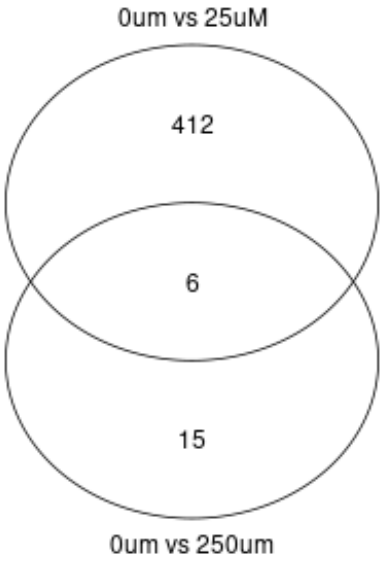
Based on this UMAP plot, there is limited separation by TCIPP exposure concentrations.

Figure D-8. A Venn diagram Indicating the Number of Genes Differentially Expressed Between HepG2 Cells Exposed to TCEP or TCIPP for All Data Contained Within the GSE109463 Dataset



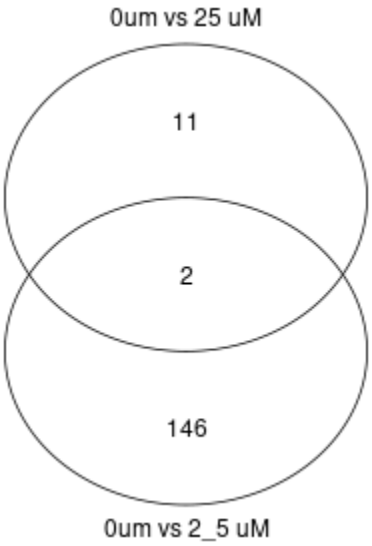
TCEP and TCIPP affect a limited gene set with increased exposure concentration.

Figure D-9. A Venn diagram Indicating the Number of Genes Differentially Expressed Between HepG2 Cells Exposed to TCEP for All Data Contained Within the GSE109463 Dataset



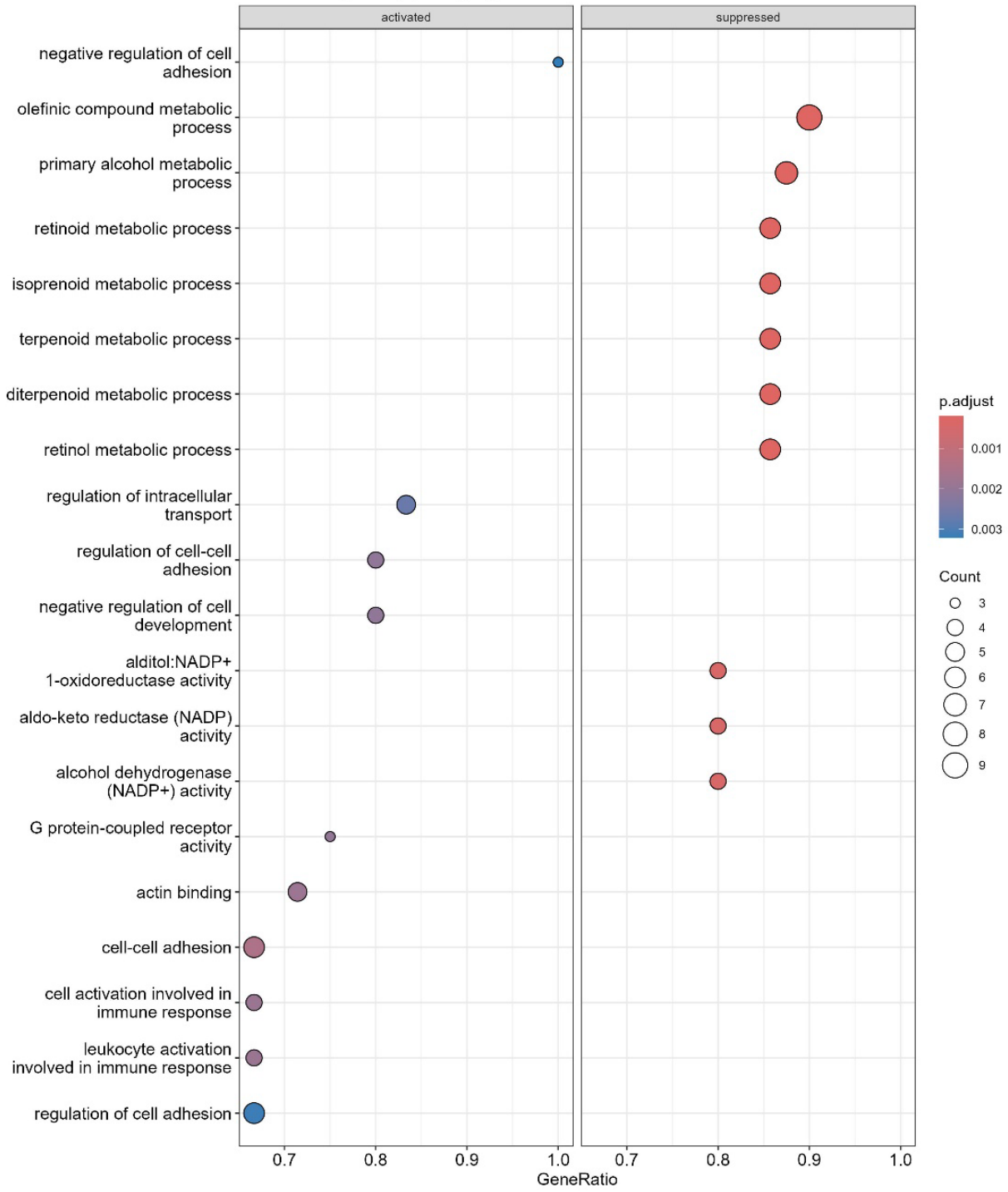
TCEP affects a more genes at the lower dose and fewer genes at the higher dose. This is striking as is the small overlap between test conditions.

Figure D-10. A Venn diagram Indicating the Number of Genes Differentially Expressed Between HepG2 Cells Exposed to TCIPP for All Data Contained Within the GSE109463 Dataset



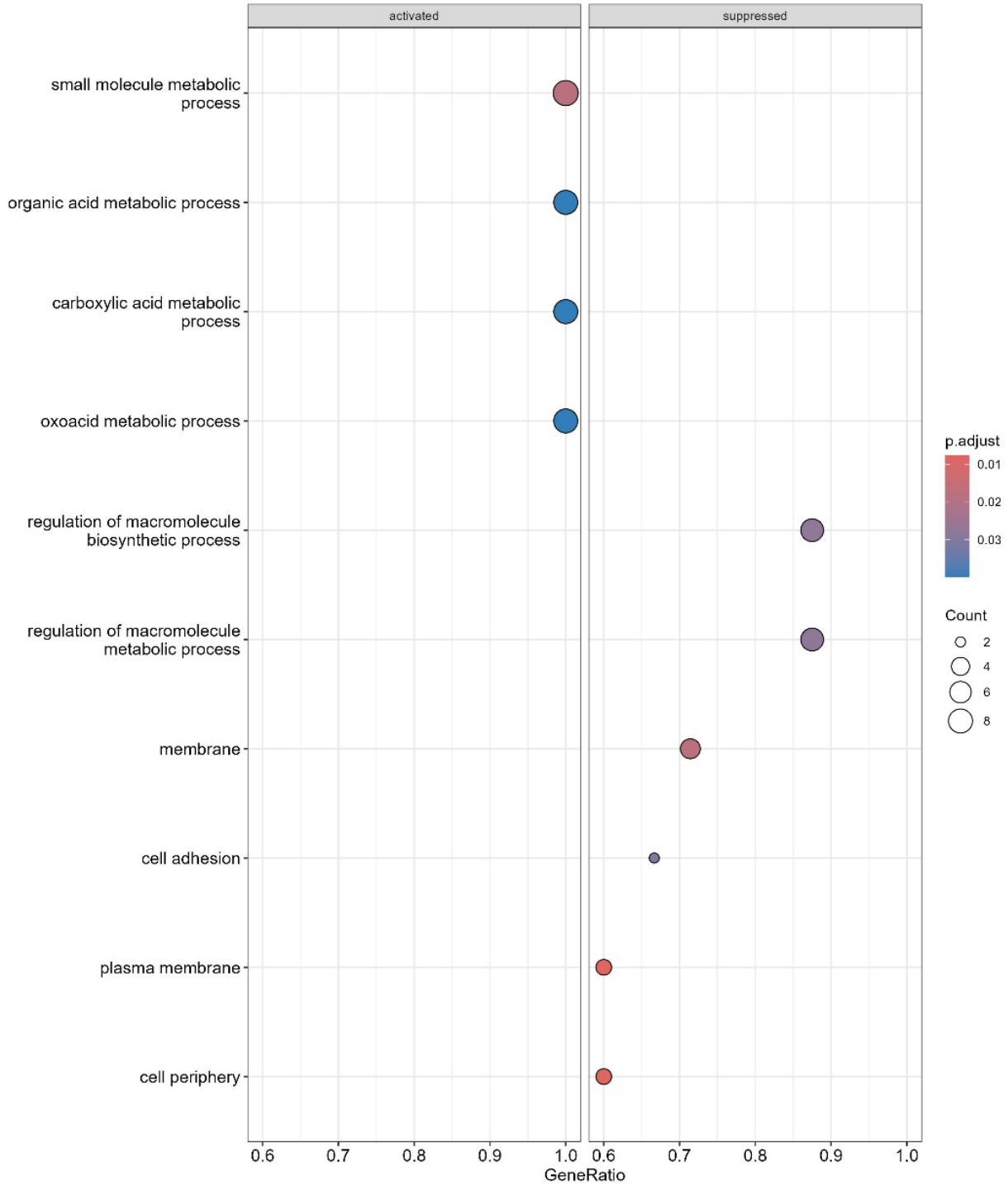
TCIPP affects a more genes at the lower dose and fewer genes at the higher dose. This is striking as is the small overlap between test conditions.

Figure D-11. Gene Set Enrichment Analysis (GSEA) from the GSE109463 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in HepG2 Cells Exposed to 25 µM TCEP Versus Control Cells



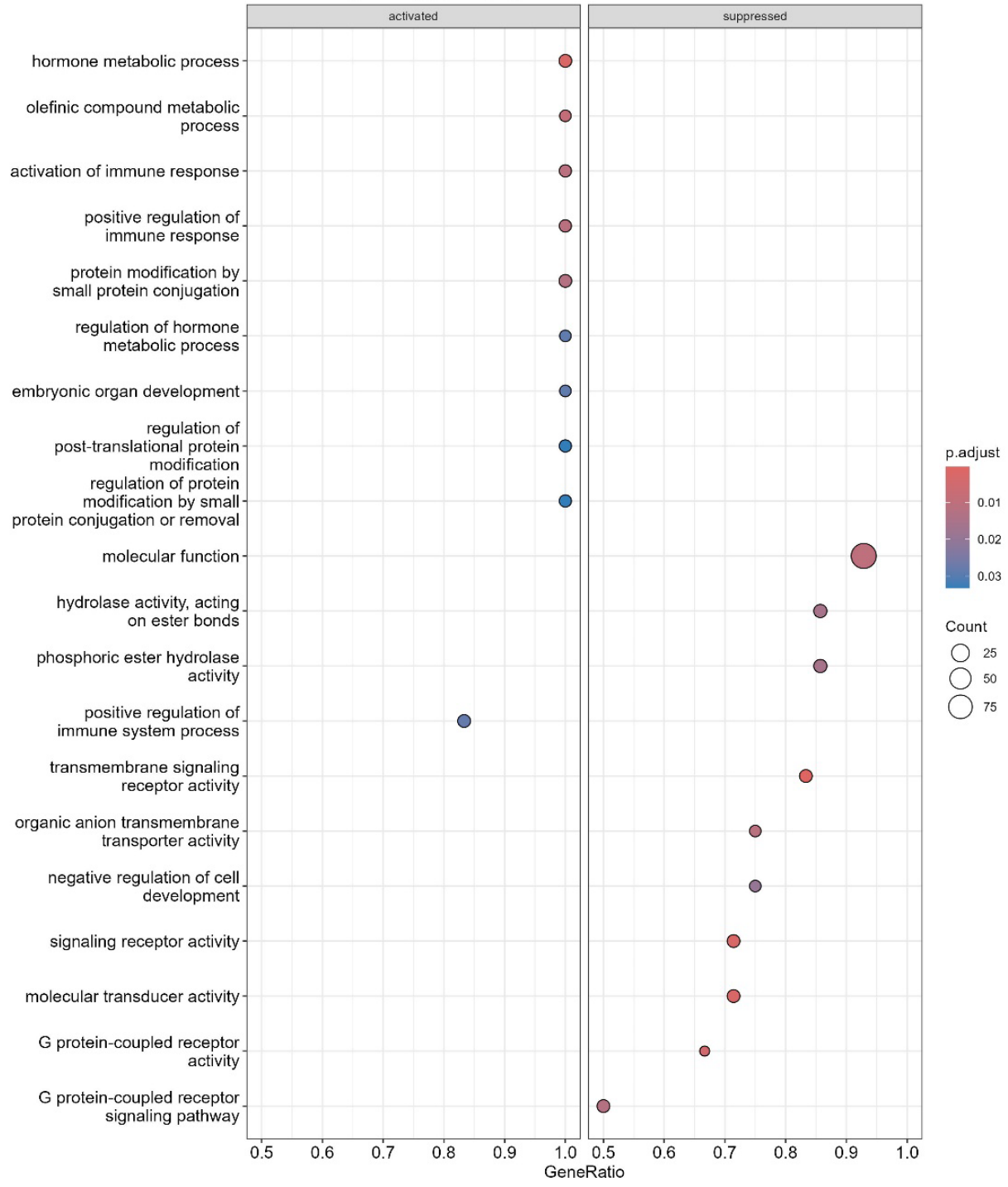
Top enriched pathways (either suppressed or activated as determined by log2 fold change directions) include olefinic compound metabolic process, primary alcohol metabolic process, and retinoid metabolic process as indicated by the low adjusted p-value (red circles) and gene count numbers (circle size).

Figure D-12. Gene Set Enrichment Analysis (GSEA) from the GSE109463 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in HepG2 Cells Exposed to 250 µM TCEP Versus Control Cells



Top enriched pathways (either suppressed or activated as determined by log2 fold change directions) include small molecule metabolic processes, plasma membrane, and cell periphery as indicated by the low adjusted p-value (red circles) and gene count numbers (circle size).

Figure D-13. Gene Set Enrichment Analysis (GSEA) from the GSE109463 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in HepG2 Cells Exposed to 2.5 µM TCIPP Versus Control Cells



Top enriched pathways (either suppressed or activated as determined by log2 fold change directions) include molecular function, hormone metabolic function, and transmembrane signaling receptor activity as indicated by the low adjusted p-value (red circles) and gene count numbers (circle size).

Figure D-14. Gene Set Enrichment Analysis (GSEA) from the GSE109463 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in HepG2 Cells Exposed to 25 µM TCIPP Versus Control Cells

The top enriched pathway (either suppressed or activated as determined by log2 fold change directions) includes cytosol, negative regulation of biological process, and negative regulation of cellular process as indicated by the low adjusted p-value (red and blue circles) and gene count numbers (circle size).

D.2. Polyhalogenated Bisphenol Aliphatics and Functionalized (PHBAF)

D.2.1. Tetrabromobisphenol A (TBBPA)

D.2.1.1. Dataset Number – GSE121336

Data Type: Array

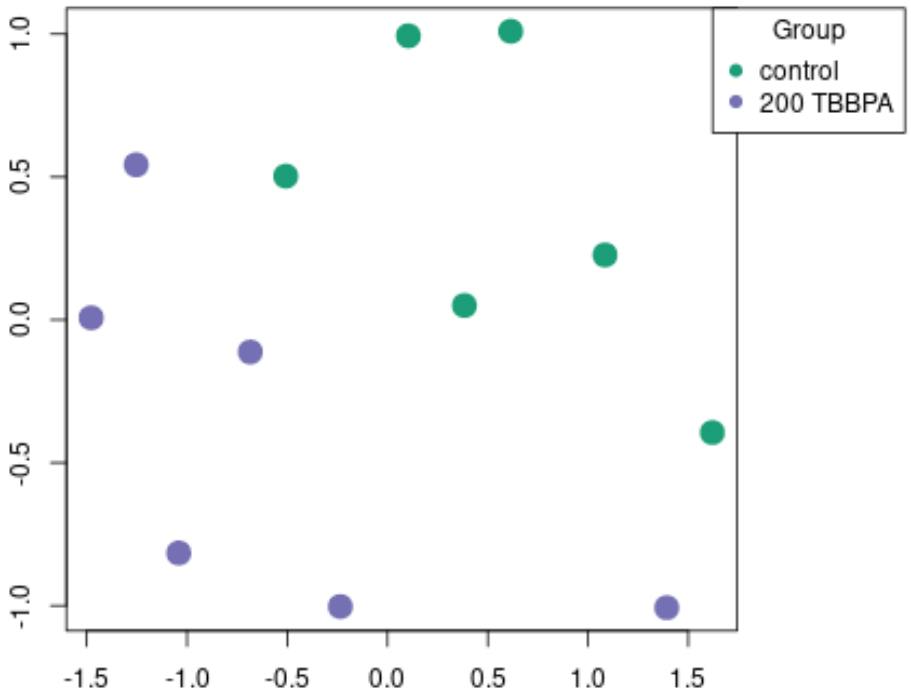
Model Type: *In vivo*

Species: Zebrafish *(Danio rerio)*

Exposure Paradigm: Zebrafish embryos were exposed to 200 µg/L TBBPA from 0 to 5 days post fertilization (dpf).

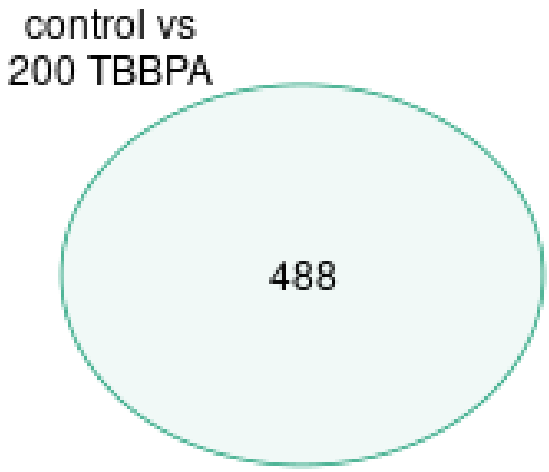
Note: Gene set enrichment analysis could not be conducted due to lack of annotations.

Figure D-15. Uniform Manifold Approximation and Projection (UMAP) Plot for All Data Contained Within the GSE121336 Dataset



Based on this UMAP plot, samples are separated by exposure to TBBPA.

Figure D-16. A Venn Diagram Indicating the Number of Genes Differentially Expressed Between Control and Exposed Embryos for All Data Contained Within the GSE121336 Dataset



TBBPA exposure led to 488 changes in gene expression compared to the controls.

D.2.2. Tetrabromobisphenol A (TBBPA) and Tetrachlorobisphenol A (TCBPA)

D.2.2.1. Dataset Number – GSE128431

Data Type: High-Throughput Screening

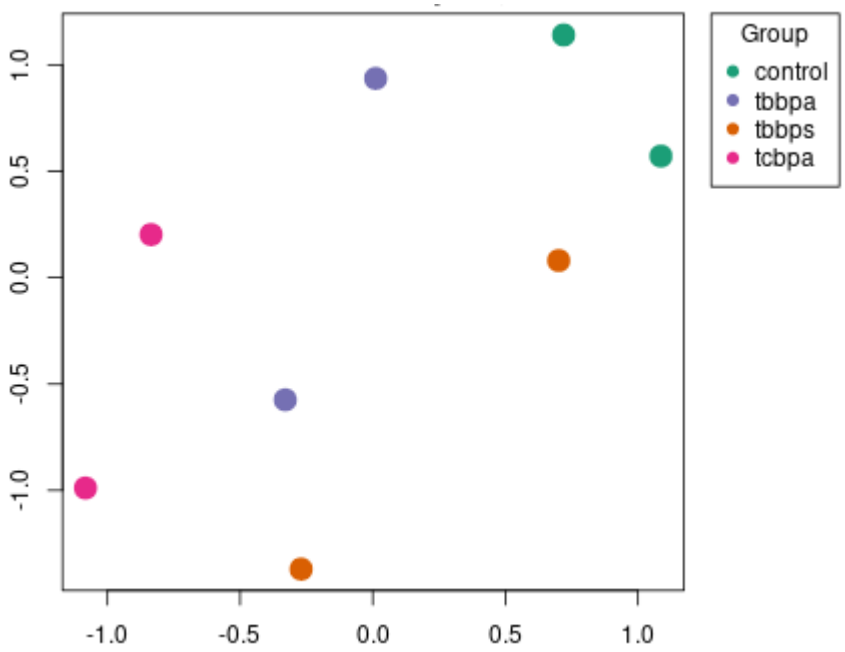
Model Type: *In vitro*

Cell Type: Human Embryonic Stem Cells

Exposure Paradigm: Human embryonic stem cells were exposed to 1 µM of TBBPA, TBBPS, or TCBPA.

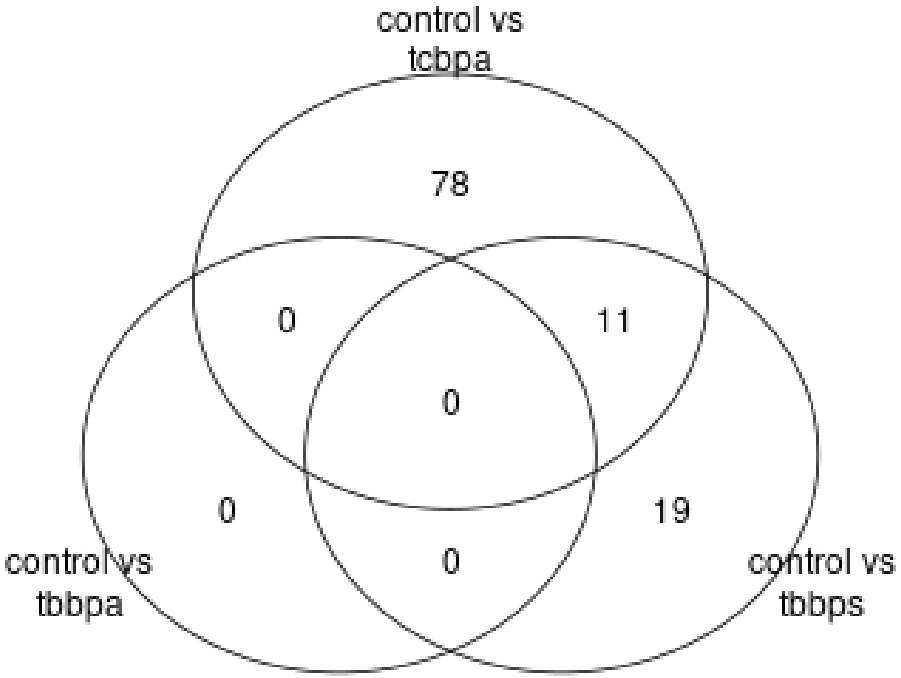
Note: Gene set enrichment analysis for TBBPA was run but no significant pathways were identified.

Figure D-17. Uniform Manifold Approximation and Projection (UMAP) Plot for All Data Contained Within the GSE128431 Dataset



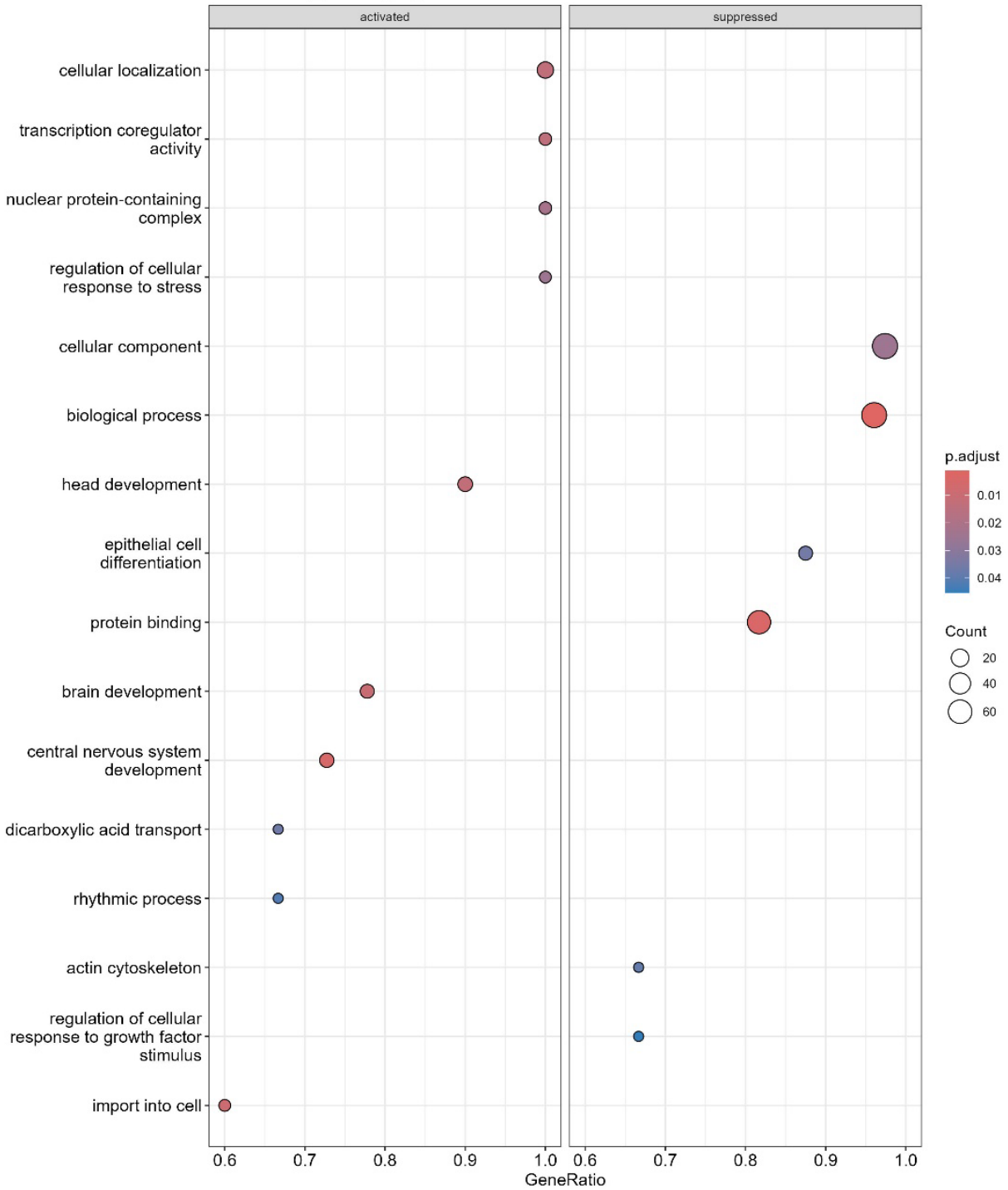
Based on this UMAP plot, there is separation by exposure to TBBPA, TBBPS, and TCBPA.

Figure D-18. A Venn Diagram Indicating the Number of Genes Differentially Expressed Between the Control and Exposed Human Embryonic Stem Cells for All Data Contained Within the GSE128431 Dataset



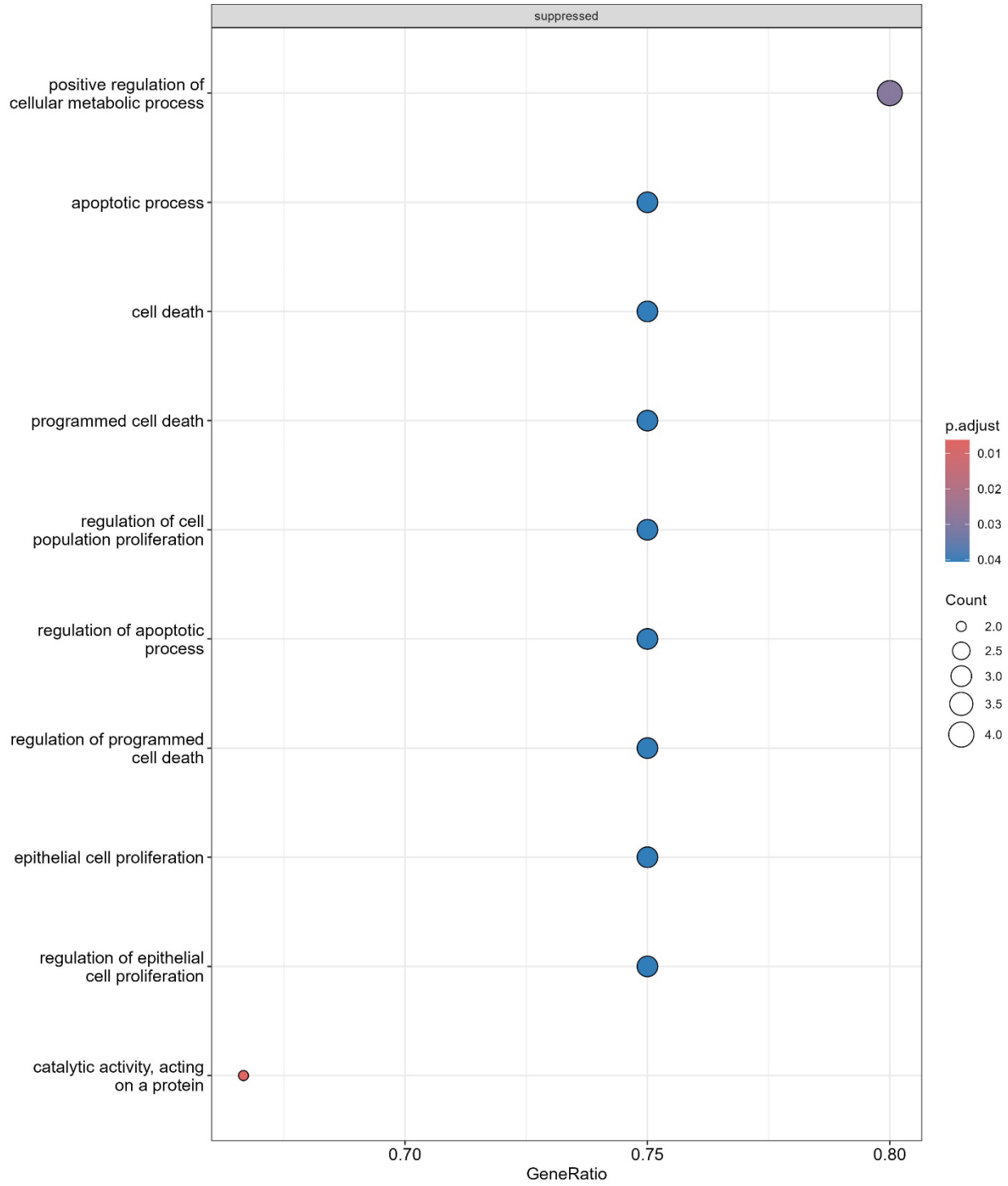
Exposure to TCBPA affected a limited gene set as indicated by 89 changes in gene expression whereas exposure to TBBPA led to no changes in gene expression. Exposure to TBBPS affected a limited gene set as indicated by 30 changes in gene expression.

Figure D-19. Gene Set Enrichment Analysis (GSEA) from the GSE128431 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in Human Embryonic Stem Cells Exposed to 1 µM TCBPA Versus Control Cells



Top enriched pathways (either suppressed or activated as determined by log2 fold change directions) include biological process and protein binding as indicated by the low adjusted p-value (red circles) and gene count numbers (circle size).

Figure D-20. Gene Set Enrichment Analysis (GSEA) from the GSE128431 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in Human Embryonic Stem Cells Exposed to 1 µM TBBPS Versus Control Cells



Top enriched pathways (either suppressed or activated as determined by log2 fold change directions) include positive regulation of cellular metabolic process and catalytic activity as indicated by the low adjusted p-value (red circles) and gene count numbers (circle size).

D.2.2.2. Dataset Number – GSE129412

Data Type: High-Throughput Screening

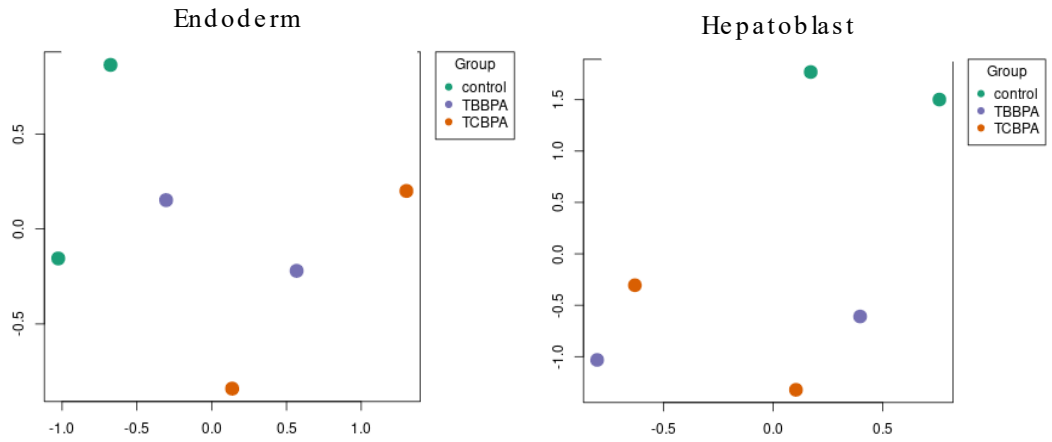
Model Type: *In vitro*

Cell Type: Human Embryonic Stem Cells

Exposure Paradigm: Human embryonic stem cells that were in definitive endoderm or hepatoblast specification form were exposed to 10 nM TBBPA and 10 nM TCBPA.

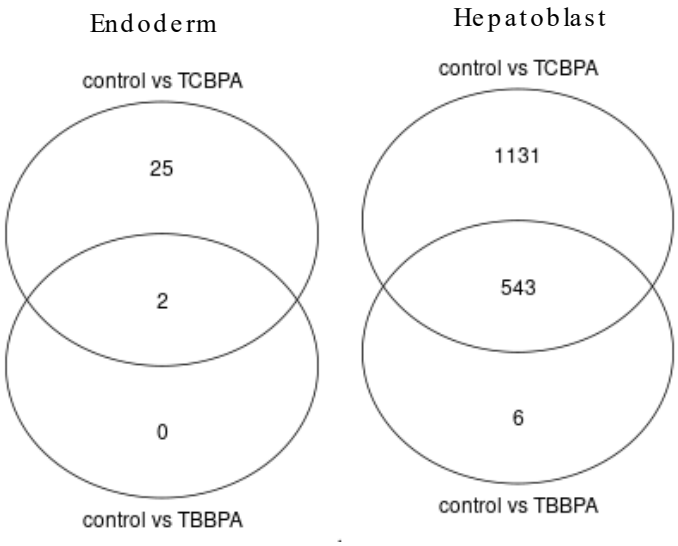
Note: Gene set enrichment analysis for TBBPA was run but no significant pathways were identified.

Figure D-21. Uniform Manifold Approximation and Projection (UMAP) Plot for All Data Contained Within the GSE129412 Dataset



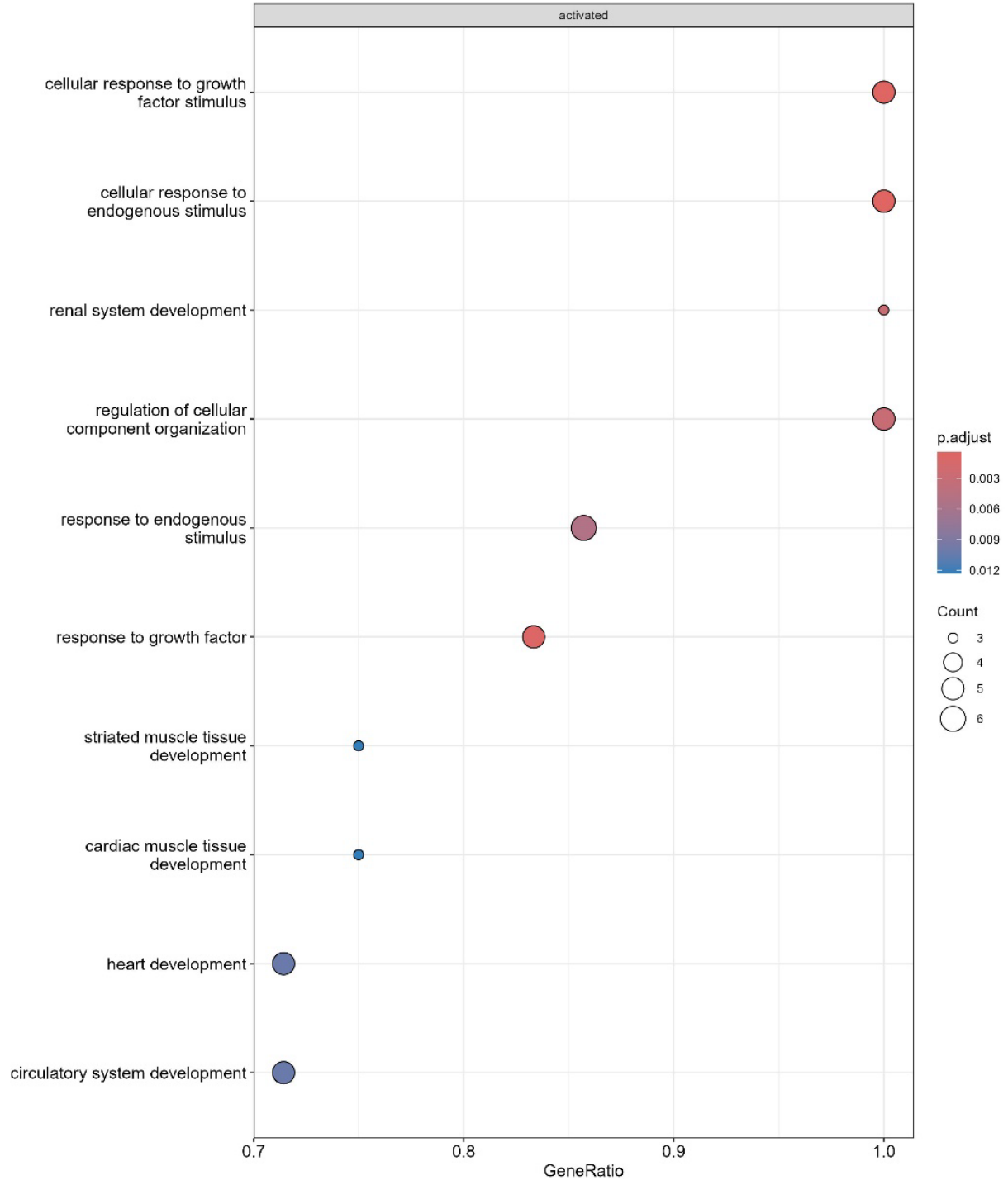
Based on these UMAP plots, there is limited separation by exposure to TBBPA and TCBPA.

Figure D-22. A Venn Diagram Indicating the Number of Genes Differentially Expressed Between the Control and Exposed Human Embryonic Stem Cells in the Endoderm/Hepatoblast Form for All Data Contained Within the GSE129412 Dataset



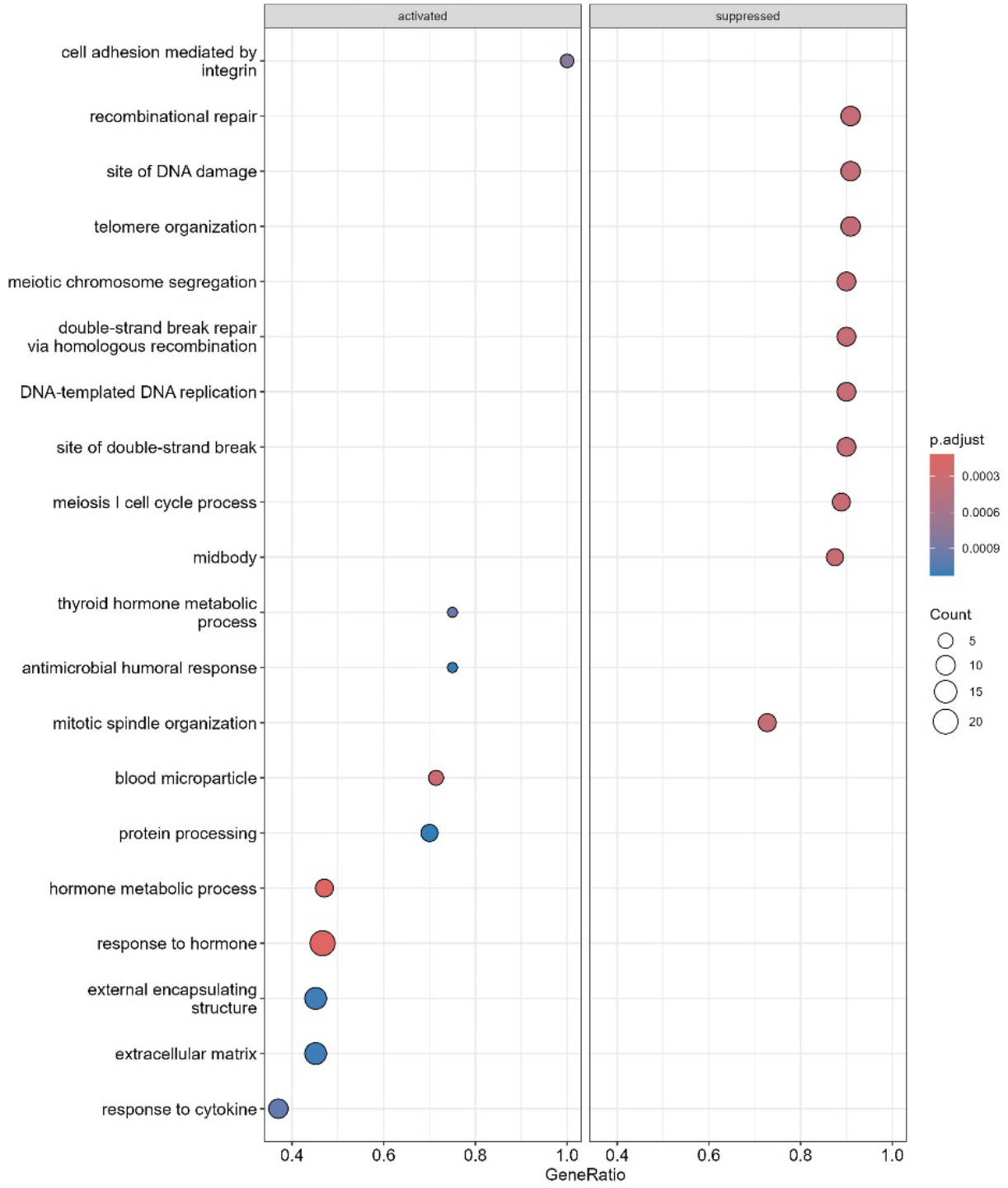
Exposure to TCBPA affected a limited gene set as indicated by 27 changes in gene expression whereas exposure to TBBPA led to 2 changes in gene expression in endoderm. In contrast, exposure to TCBPA affected a large gene set as indicated by 1,674 changes in gene expression compared to exposure to TBBPA which led to 549 changes in gene expression in hepatoblasts.

Figure D-23. Gene Set Enrichment Analysis (GSEA) from the GSE129412 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in Human Embryonic Stem Cells in Endoderm Form Exposed to 1 µM TCBPA Versus Control Cells



Top enriched pathways (either suppressed or activated as determined by log2 fold change directions) include cellular response to growth factor stimulus, cellular response to endogenous stimulus, and response to growth factor as indicated by the low adjusted p-value (red circles) and gene count numbers (circle size).

Figure D-24. Gene Set Enrichment Analysis (GSEA) from the GSE129412 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in Human Embryonic Stem Cells in Hepatoblast Form Exposed to 1 µM TCBPA Versus Control Cells



Top enriched pathways (either suppressed or activated as determined by log2 fold change directions) include response to hormone, recombinational repair, and site of DNA damage as indicated by the low adjusted p-value (red circles) and gene count numbers (circle size).

D.2.2.3. Dataset Number – GSE148518

Data Type: High-Throughput Screening

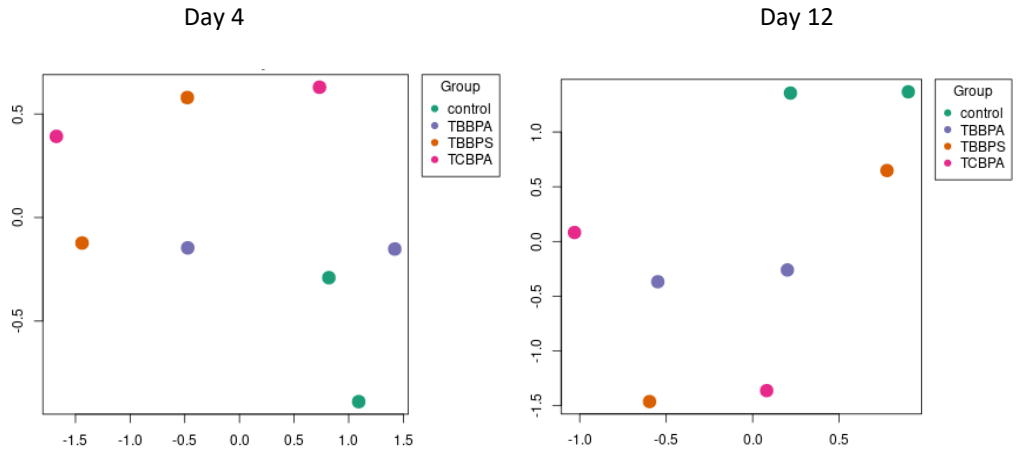
Model Type: *In vitro*

Cell Type: Human Embryonic Stem Cells

Exposure Paradigm: Human embryonic stem cells were exposed to 10 nM of TBBPA, TBBPS, or TCBPA for 4 or 12 days.

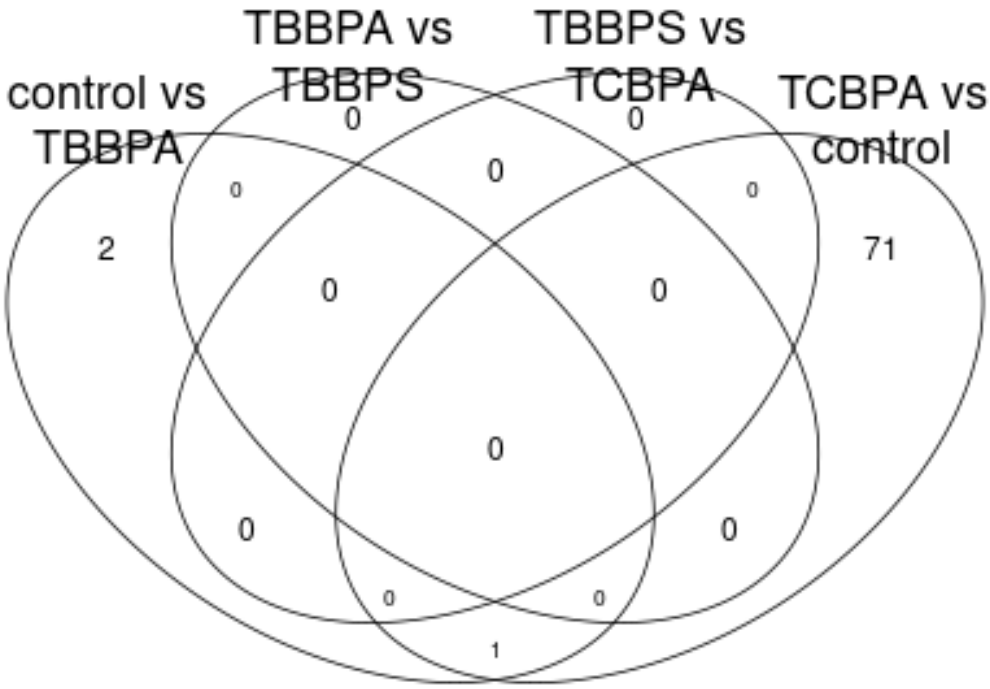
Note: Gene set enrichment analysis for TBBPA was run but no significant pathways were identified.

Figure D-25. Uniform Manifold Approximation and Projection (UMAP) Plot for All Data Contained Within the GSE148518 Dataset



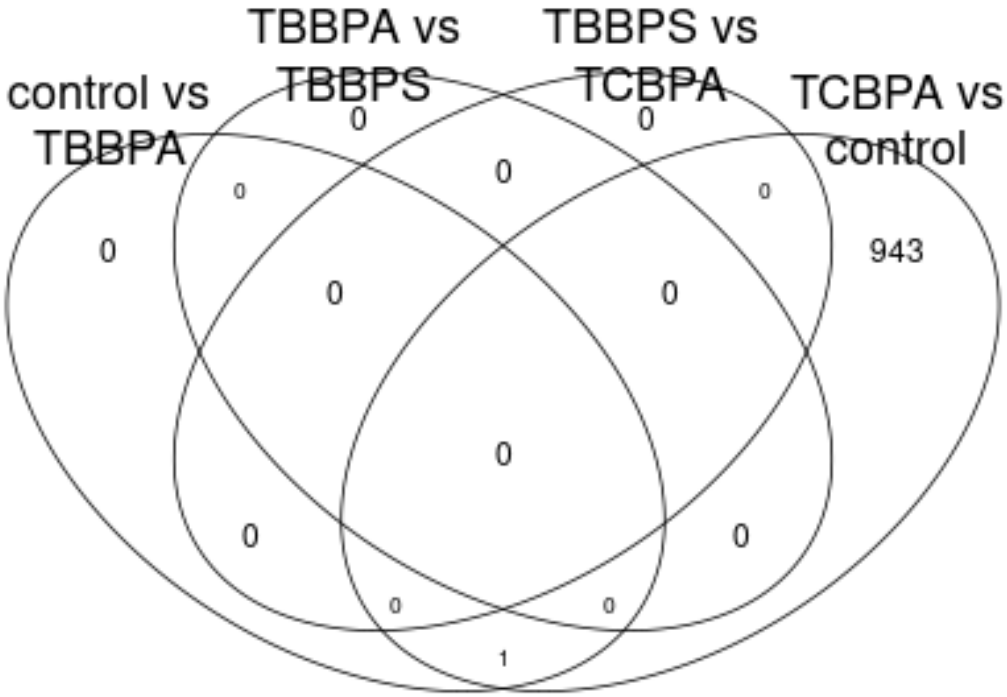
Based on this UMAP plot, there is limited separation by exposure to TBBPA, TBBPS, and TCBPA on days 4 and 12.

Figure D-26. A Venn Diagram Indicating the Number of Genes Differentially Expressed Between the Human Embryonic Stem Cells Exposed to TBBPA, TBBPS, or TCBPA for 4 Days for All Data Contained Within the GSE148518 Dataset



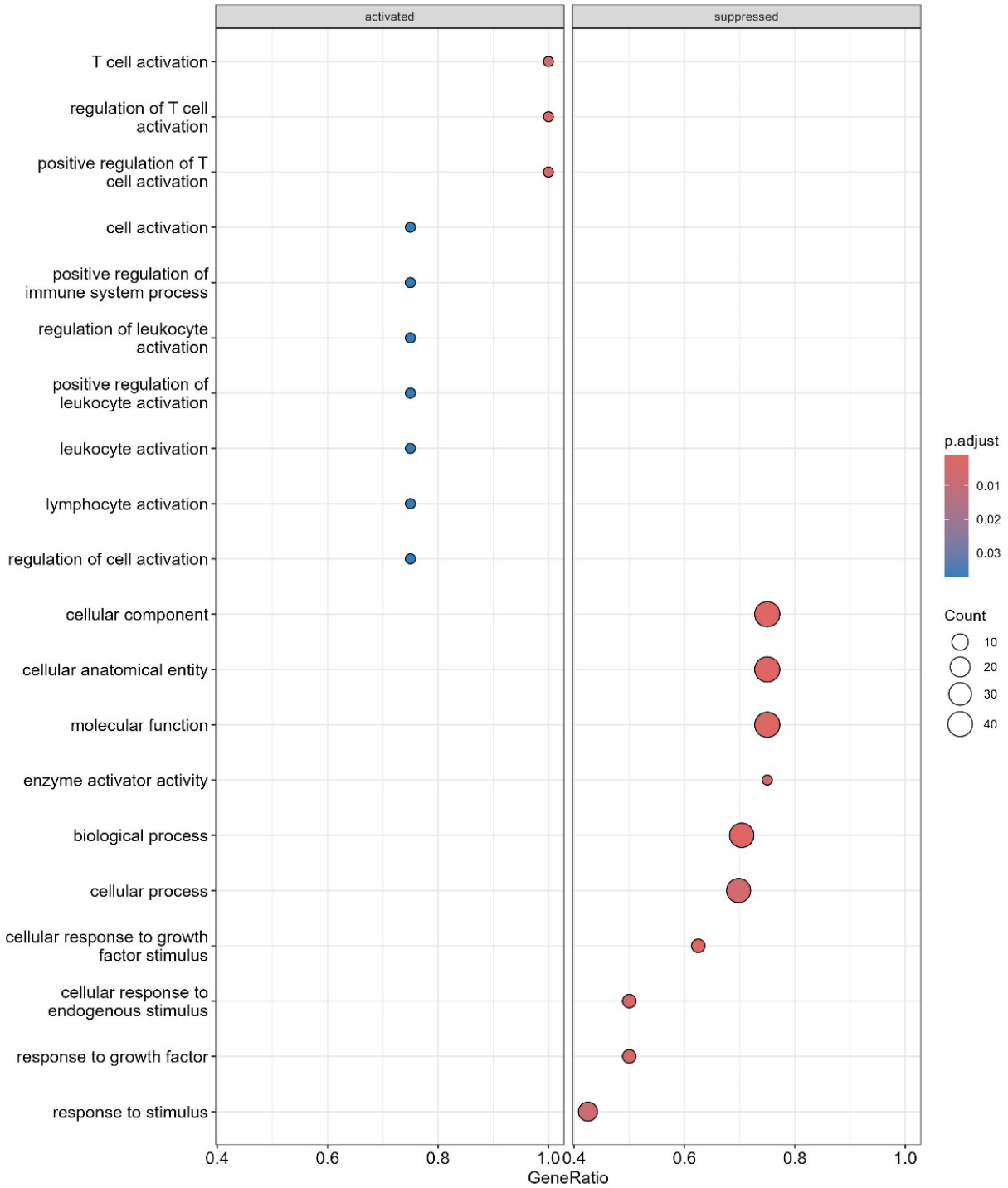
Exposure to TCBPA affected a limited gene set as indicated by 72 changes in gene expression whereas exposure to TBBPA led to 3 changes in gene expression. Similarly, TBBPS had one gene changed in association.

Figure D-27. A Venn Diagram Indicating the Number of Genes Differentially Expressed Between the Human Embryonic Stem Cells Exposed to TBBPA, TBBPS, or TCBPA for 12 Days for All Data Contained Within the GSE148518 Dataset



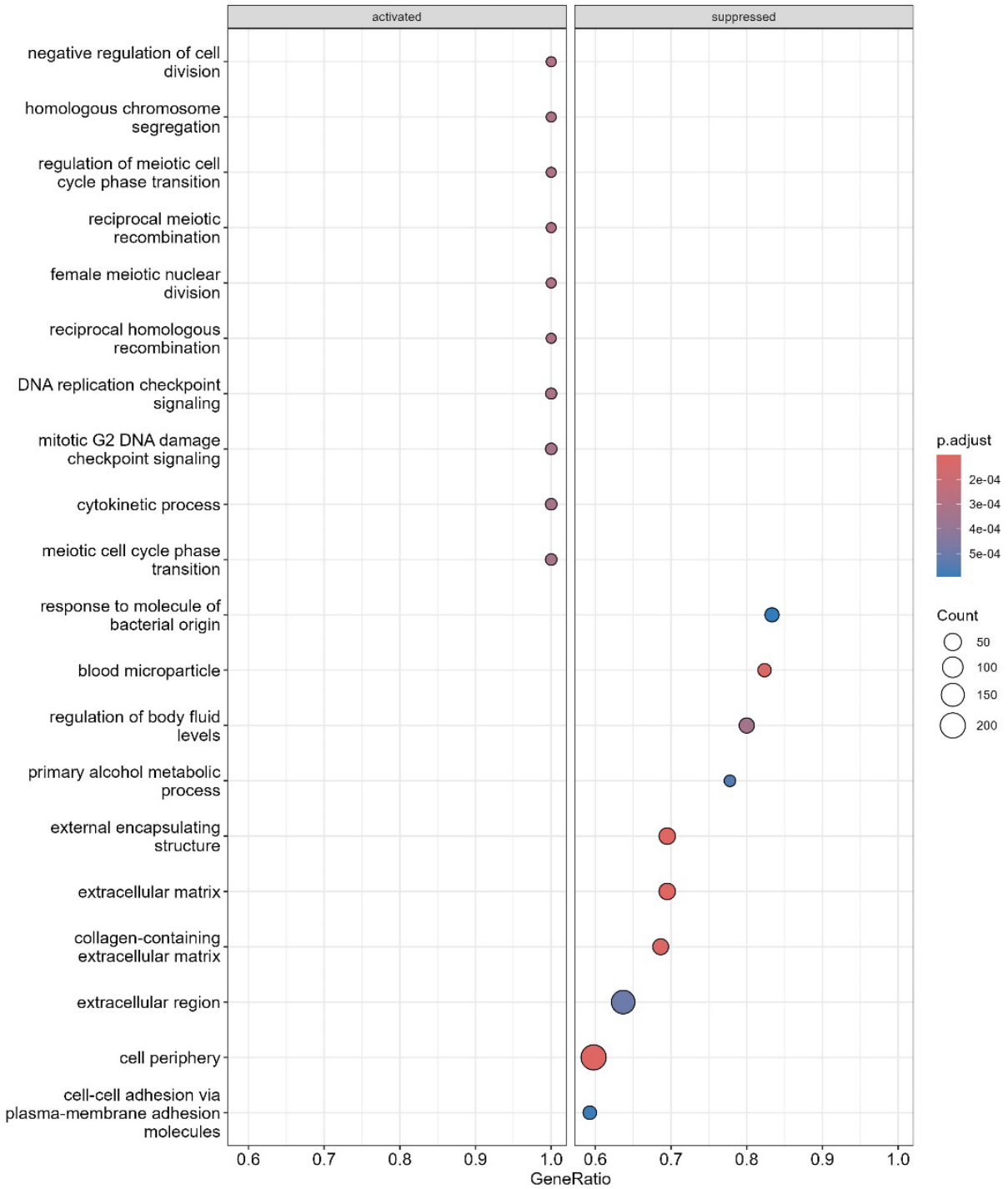
Exposure to TCBPA affected a limited gene set as indicated by 944 changes in gene expression whereas exposure to TBBPA led to 1 change in gene expression and TBBPS lead to a change in 1 gene.

Figure D-28. Gene Set Enrichment Analysis (GSEA) from the GSE148518 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in Human Embryonic Stem Cells Exposed to 10 nM TCBPA for 4 Days Versus Control Cells



Top enriched pathways (either suppressed or activated as determined by log2 fold change directions) include cellular component, cellular anatomical entity, and molecular function as indicated by the low adjusted p-value (red circles) and gene count numbers (circle size).

Figure D-29. Gene Set Enrichment Analysis (GSEA) from the GSE148518 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in Human Embryonic Stem Cells Exposed to 10 nM TCBPA for 12 Days Versus Control Cells



Top enriched pathways (either suppressed or activated as determined by log2 fold change directions) include cell periphery, external encapsulating structure, extracellular matrix, and collagen-containing extracellular matrix as indicated by the low adjusted p-value (red circles) and gene count numbers (circle size).

D.2.3. Tetrabromobisphenol A-bis(2,3-dibromopropyl ether) (TBBPA-DBPE)

D.2.3.1. Dataset Number – GSE153363

Data Type: Array

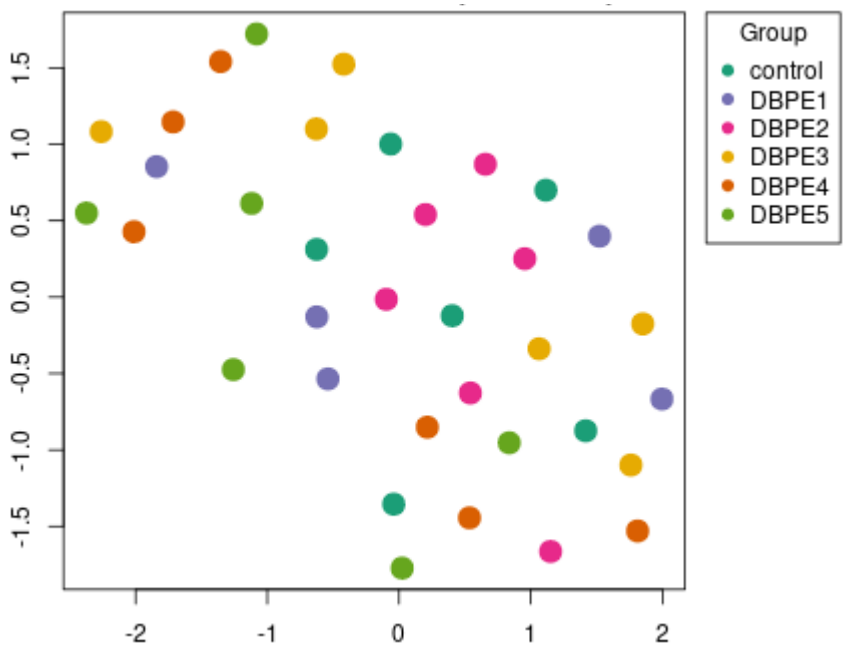
Model Type: *In vivo*

Species: Sprague Dawley Rat (*Rattus norvegicus*)

Exposure paradigm: Male Harlan Sprague Dawley rats were exposed to concentrations of 0, 0.1, 0.94, 9.4, 94.3 or 943 mg/kg TBBPA-DBPE. After five days of exposure, RNA was collected to identify differentially expressed liver transcripts.

Note: No significant genes were identified.

Figure D-30. Uniform Manifold Approximation and Projection (UMAP) Plot for All Data Contained Within the GSE153363 Dataset



Based on this UMAP plot, there is no separation by exposure to TBBPA-DBPE.

D.2.4. Tetrabromobisphenol A (TBBPA), Tetrachlorobisphenol A (TCBPA), Tetrabromobisphenol A bis (allyl ether) (TBBPA-BAE), Tetrabromobisphenol S (TBBPS), TBBPA-bis(2,3-dibromopropyl ether) (TBBPA-BDBPE), Tetrabromobisphenol A bis(2-hydroxyethyl ether) (TBBPA-BHEE)

D.2.4.1. Dataset Number – GSE193176

Data Type: High-Throughput Screening

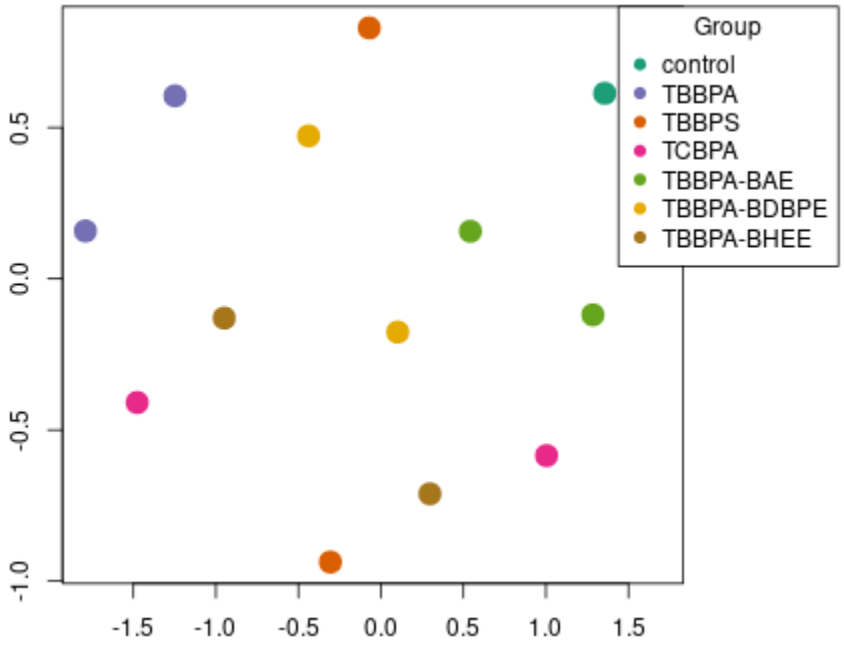
Model Type: *In vitro*

Cell Type: Human Embryonic Stem Cells

Exposure paradigm: Human embryonic stem cells were exposed to 1 µM of TBBPA, TCBPA, TBBPA-BAE, TBBPS, TBBPA-BDBPE, or TBBPA-BHEE.

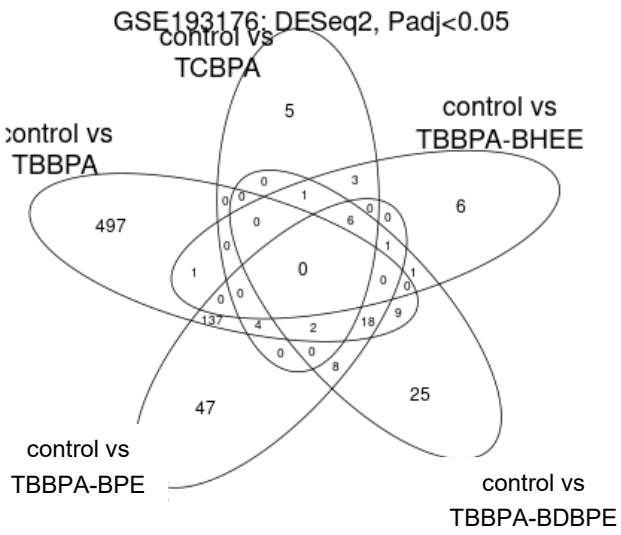
Note: Lack of significant genes resulted in no pathway level responses for TCBPA, TBBPA-BAE, or TBBPA-BHEE.

Figure D-31. Uniform Manifold Approximation and Projection (UMAP) Plot for All Data Contained Within the GSE193176 Dataset



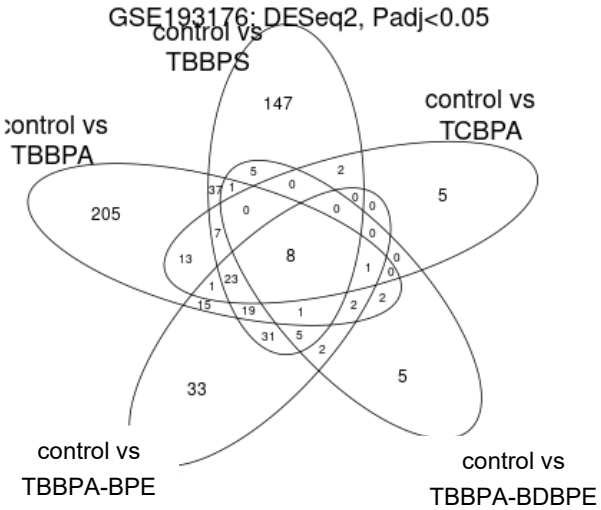
Based on this UMAP plot, there is reasonable separation between chemicals and the control condition.

Figure D-32. A Venn Diagram Indicating the Number of Genes Differentially Expressed Between the Control and Exposed Human Embryonic Stem Cells for All Data Contained Within the GSE193176 Dataset



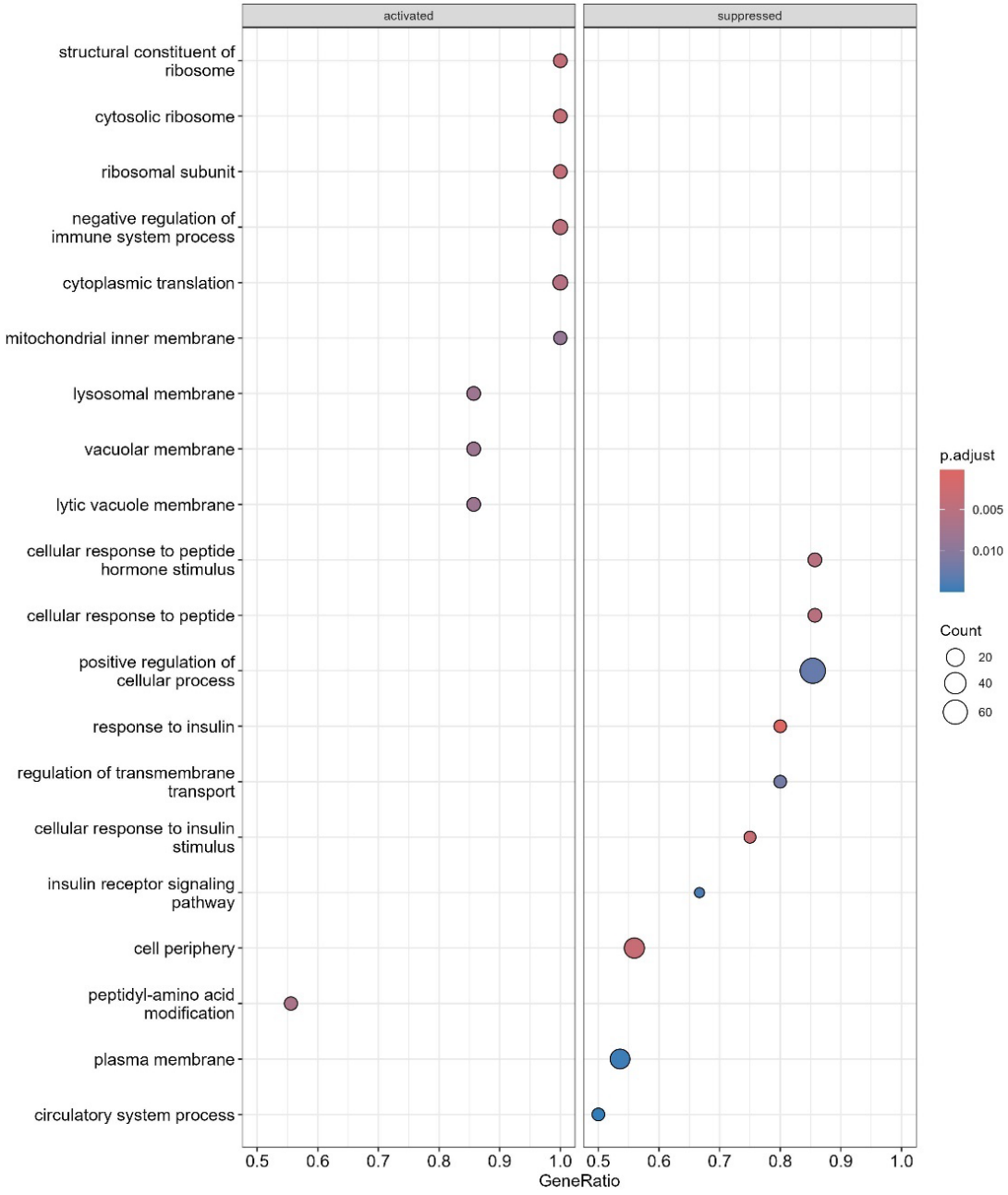
Exposure to TCBPA affected a limited gene set as indicated by 21 changes in gene expression whereas exposure to TBBPA led to 668 changes, TBBPA-BAE led to 215, TBBPA-BDBPE led to 71, and TBBPA-BHEE led to 19 changes in gene expression.

Figure D-33. A Venn Diagram Indicating the Number of Genes Differentially Expressed Between the Control and Exposed Human Embryonic Stem Cells for All Data Contained Within the GSE193176 Dataset



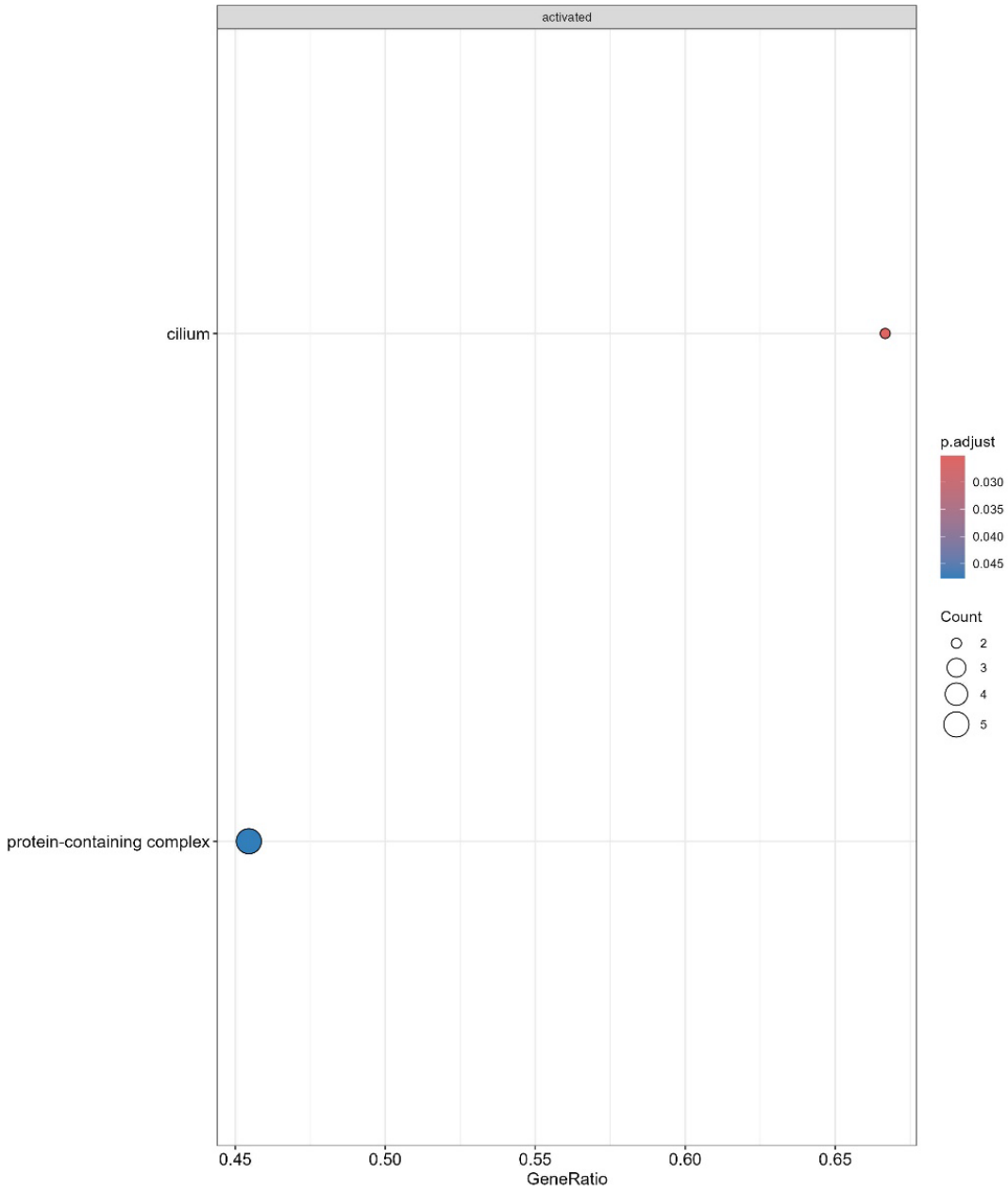
Exposure to TCBPA affected a limited gene set as indicated by 60 changes in gene expression whereas exposure to TBBPA led to 334 changes, TBBPA-BAE led to 32, TBBPA-BDBPE led to 141, and TBBPS led to 286 changes in gene expression.

Figure D-34. Gene Set Enrichment Analysis (GSEA) from the GSE193176 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in Human Embryonic Stem Cells Exposed to 1 µM TBBPA Versus Control Cells



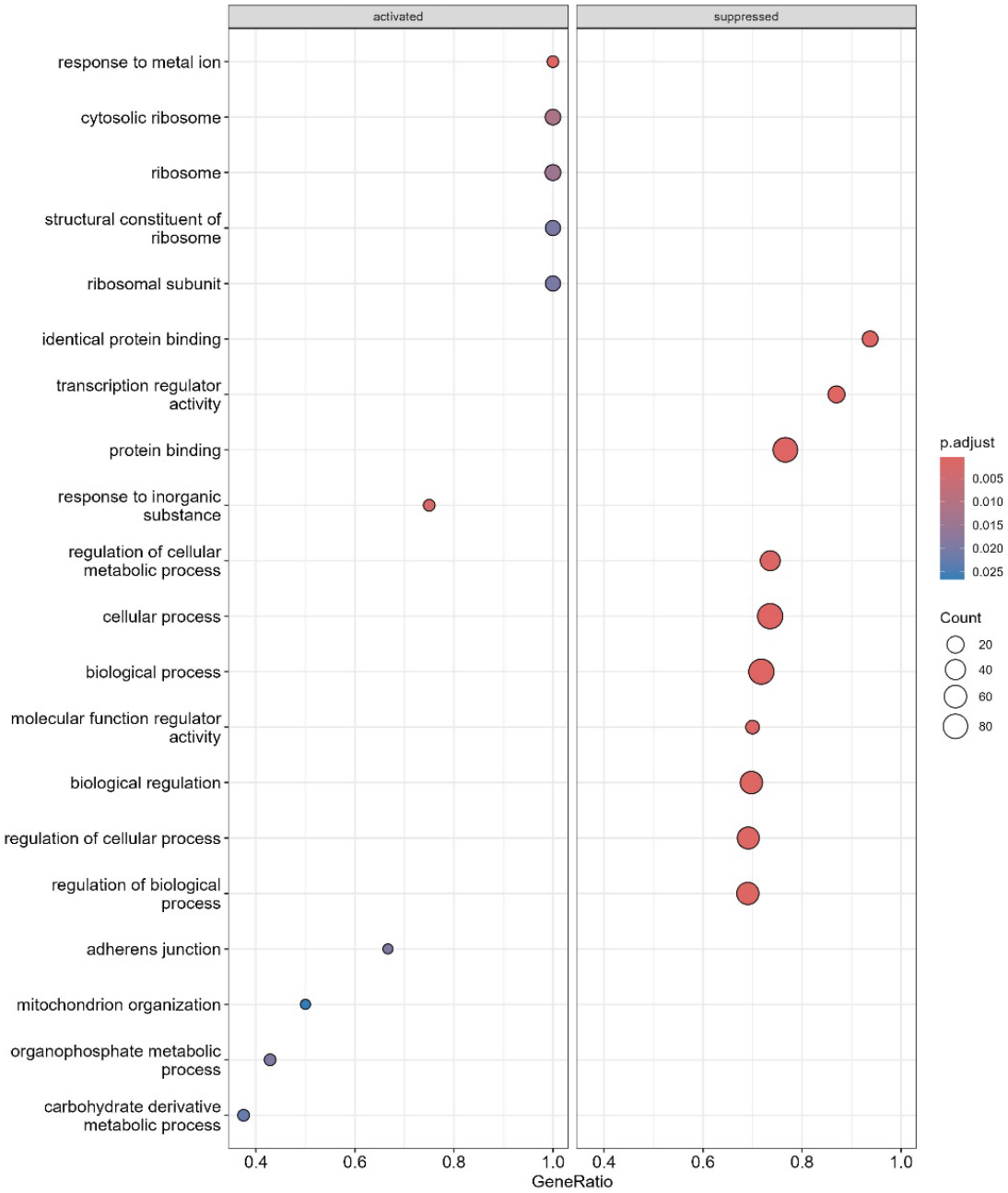
Top enriched pathways (either suppressed or activated as determined by log2 fold change directions) include cell periphery and response to insulin as indicated by the low adjusted p-value (red circles) and gene count numbers (circle size).

Figure D-35. Gene Set Enrichment Analysis (GSEA) from the GSE193176 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in Human Embryonic Stem Cells Exposed to 1 µM TBBPA-BDBPE Versus Control Cells



Top enriched pathways(either suppressed or activated as determined by log2 fold change directions) include cilium and protein-containing complex as indicated by the low adjusted p-value (red circle) and gene count numbers (circle size).

Figure D-36. Gene Set Enrichment Analysis (GSEA) from the GSE193176 Dataset Indicating the Pathways Altered Based on Differentially Expressed Genes in Human Embryonic Stem Cells Exposed to 1 µM TBBPS Versus Control Cells



Top enriched pathways (either suppressed or activated as determined by log2 fold change directions) include protein binding, cellular process, biological process, biological regulation, regulation of cellular process, and regulation of biological process as indicated by the low adjusted p-value (red circles) and gene count numbers (circle size).

Appendix E. References

Agency for Toxic Substances and Disease Registry. (ATSDR). (2012). *Toxicological profile for phosphate ester flame retardants* (PB2013-105796). http://www.atsdr.cdc.gov/toxprofiles/tp202.pdf

Auerbach, S. (2022). US NTP - Nuances of the data analysis pipeline and OECD BMD reporting module. https://www.ecetoc.org/wp-content/uploads/2022/08/7.-Scott-Auerbach.pdf

Black, M. B., Stern, A., Efremenko, A., Mallick, P., Moreau, M., Hartman, J. K., & McMullen, P. D. (2022). Biological system considerations for application of toxicogenomics in next-generation risk assessment and predictive toxicology. *Toxicol In Vitro*, *80*, 105311. https://doi.org/10.1016/j.tiv.2022.105311

Chen, C., Wang, Y., Qian, Y., Zhao, X., & Wang, Q. (2015). The synergistic toxicity of the multiple chemical mixtures: Implications for risk assessment in the terrestrial environment. *Environment International*, *77*, 95-105. https://doi.org/https://doi.org/10.1016/j.envint.2015.01.014

Crizer, D. M., Ramaiahgari, S. C., Ferguson, S. S., Rice, J. R., Dunlap, P. E., Sipes, N. S., Auerbach, S. S., Merrick, B. A., & DeVito, M. J. (2021). Benchmark Concentrations for Untargeted Metabolomics Versus Transcriptomics for Liver Injury Compounds in In Vitro Liver Models. *Toxicol Sci*, *181*(2), 175-186. https://doi.org/10.1093/toxsci/kfab036

Demchuk, E., Ruiz, P., Wilson, J. D., Scinicariello, F., Pohl, H. R., Fay, M., Mumtaz, M. M., Hansen, H., & De Rosa, C. T. (2008). Computational toxicology methods in public health practice. *Toxicol Mech Methods*, *18*(2-3), 119-135. https://doi.org/10.1080/15376510701857148

European Chemicals Bureau. (ECB). (2009). *European Union risk assessment report: Tris(2-chloroethyl) phosphate, TCEP*. https://echa.europa.eu/documents/10162/2663989d-1795-44a1-8f50-153a81133258

Embry, M. R., Belanger, S. E., Braunbeck, T. A., Galay-Burgos, M., Halder, M., Hinton, D. E., Leonard, M. A., Lillicrap, A., Norberg-King, T., & Whale, G. (2010). The fish embryo toxicity test as an animal alternative method in hazard and risk assessment and scientific research. *Aquat Toxicol*, *97*(2), 79-87. https://doi.org/10.1016/j.aquatox.2009.12.008

Fagerberg, L., Hallstrom, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka, M., Tahmasebpoor, S., Danielsson, A., Edlund, K., Asplund, A., Sjostedt, E., Lundberg, E., Szigyarto, C. A., Skogs, M., Takanen, J. O., Berling, H., Tegel, H., Mulder, J.,…Uhlen, M. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics*, *13*(2), 397-406. https://doi.org/10.1074/mcp.M113.035600

Felter, S. P., Foreman, J. E., Boobis, A., Corton, J. C., Doi, A. M., Flowers, L., Goodman, J., Haber, L. T., Jacobs, A., Klaunig, J. E., Lynch, A. M., Moggs, J., & Pandiri, A. (2018). Human relevance of rodent liver tumors: Key insights from a Toxicology Forum workshop on nongenotoxic modes of action. *Regul Toxicol Pharmacol*, *92*, 1-7. https://doi.org/10.1016/j.yrtph.2017.11.003

Gilmour, N., Kern, P. S., Alepee, N., Boisleve, F., Bury, D., Clouet, E., Hirota, M., Hoffmann, S., Kuhnl, J., Lalko, J. F., Mewes, K., Miyazawa, M., Nishida, H., Osmani, A., Petersohn, D., Sekine, S., van Vliet, E., & Klaric, M. (2020). Development of a next generation risk assessment framework for the evaluation of skin sensitisation of cosmetic ingredients. *Regul Toxicol Pharmacol*, *116*, 104721. https://doi.org/10.1016/j.yrtph.2020.104721

Hall, A. P., Elcombe, C. R., Foster, J. R., Harada, T., Kaufmann, W., Knippel, A., Kuttler, K., Malarkey, D. E., Maronpot, R. R., Nishikawa, A., Nolte, T., Schulte, A., Strauss, V., & York, M. J. (2012). Liver

hypertrophy: a review of adaptive (adverse and non-adverse) changes--conclusions from the 3rd International ESTP Expert Workshop. *Toxicol Pathol*, *40*(7), 971-994. https://doi.org/10.1177/0192623312448935

Health Canada. (2009). Ethanol, 2-chloro-, phosphate (3:1) (tris(2-chloroethyl)phosphate) (TCEP): CASRN 115-96-8. https://www.canada.ca/en/health-canada/services/chemical-substances/challenge/batch-5/tcep.html

Health Canada. (2016). Draft screening assessment. Certain organic flame retardants substance groupings. 2-Propanol, 1-chloro-, phosphate (3:1) (TCPP): CASRN 13674-84-5, 2-Propanol, 1,3-dichloro-, phosphate (3:1) (TDCPP): CASRN 13674-87-8. https://www.canada.ca/en/health-canada/services/chemical-substances/challenge/batch-5/tcep.html

Health Canada. (2021). Science approach document. Bioactivity exposure ratio: Application in priority setting and risk assessment. https://www.canada.ca/en/environment-climate-change/services/evaluating-existing-substances/science-approach-document-bioactivity-exposure-ratio-application-priority-setting-risk-assessment.html

Hewitt, N. J., Troutman, J., Przibilla, J., Schepky, A., Ouedraogo, G., Mahony, C., Kenna, G., Varcin, M., Dent, M. P., & Cosmetics Europe. Electronic address, c. c. e. (2022). Use of in vitro metabolism and biokinetics assays to refine predicted in vivo and in vitro internal exposure to the cosmetic ingredient, phenoxyethanol, for use in risk assessment. *Regul Toxicol Pharmacol*, *131*, 105132. https://doi.org/10.1016/j.yrtph.2022.105132

Honda, G. S., Pearce, R. G., Pham, L. L., Setzer, R. W., Wetmore, B. A., Sipes, N. S., Gilbert, J., Franz, B., Thomas, R. S., & Wambaugh, J. F. (2019). Using the concordance of in vitro and in vivo data to evaluate extrapolation assumptions. *PLoS One*, *14*(5), e0217564. https://doi.org/10.1371/journal.pone.0217564

International Agency for Research on Cancer. (IARC). (1999). *Tris(2-chloroethyl) phosphate* (IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Issue. http://monographs.iarc.fr/ENG/Monographs/vol71/mono71.pdf

Krivoshiev, B. V., Beemster, G. T. S., Sprangers, K., Blust, R., & Husson, S. J. (2018). A toxicogenomics approach to screen chlorinated flame retardants tris(2-chloroethyl) phosphate and tris(2-chloroisopropyl) phosphate for potential health effects. *J Appl Toxicol*, *38*(4), 459-470. https://doi.org/10.1002/jat.3553

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923-930. https://doi.org/10.1093/bioinformatics/btt656

Mansouri, K., Grulke, C. M., Judson, R. S., & Williams, A. J. (2018). OPERA models for predicting physicochemical properties and environmental fate endpoints. *J Cheminform*, *10*(1), 10. https://doi.org/10.1186/s13321-018-0263-1

Martin, E. M., Grimm, S. A., Xu, Z., Taylor, J. A., & Wade, P. A. (2023). Beadchip technology to detect DNA methylation in mouse faithfully recapitulates whole-genome bisulfite sequencing. *Epigenomics*, *15*(3), 115-129. https://doi.org/10.2217/epi-2023-0034

National Toxicology Program. (NTP). (1991). *Toxicology and carcinogenesis studies of Tris(2-chloroethyl) Phosphate (CAS No. 115-96-8) in F344/N rats and B6C3F1 mice (gavage studies)*.

National Toxiciology Program. (NTP). (2007). *NTP report on the toxicology studies of dichloroacetic acid (CAS No. 79-43-6) in genetically modified (FVB Tg.AC hemizygous) mice (dermal and drinking water studies) and carcinogenicity studies of dichloroacetic acid in genetically modified [B6.129-Trp53(tm1Brd) (N5) haploinsufficient] mice (drinking water studies)*. https://www.ncbi.nlm.nih.gov/pubmed/18784768

National Toxicology Program. (NTP). (2011). *90 Days dosed feed toxicity study of tris(2 chloroisopropyl)phosphate (TCPP) (CAS No. 13674 84 5) in B6C3F1 Mice*. http://tools.niehs.nih.gov/cebs3/ntpViews/?studyNumber=C20712

National Toxicology Program. (NTP). (2023). *NTP technical report on the toxicology and carcinogenesis studies of an isomeric mixture of tris(chloropropyl) phosphate administered in feed to Sprague Dawley (Hsd:Sprague Dawley® SD®) rats and B6C3F1/N mice* (Technical Report 602). http://dx.doi.org/10.22427/NTP-TR-602

National Toxicology Program. (NTP). (2024). *Comptox tools and resources*. https://ntp.niehs.nih.gov/whatwestudy/tox21/toolbox

Paul Friedman, K., Gagne, M., Loo, L. H., Karamertzanis, P., Netzeva, T., Sobanski, T., Franzosa, J. A., Richard, A. M., Lougee, R. R., Gissi, A., Lee, J. J., Angrish, M., Dorne, J. L., Foster, S., Raffaele, K., Bahadori, T., Gwinn, M. R., Lambert, J., Whelan, M.,…Thomas, R. S. (2020). Utility of In Vitro Bioactivity as a Lower Bound Estimate of In Vivo Adverse Effect Levels and in Risk-Based Prioritization. *Toxicol Sci*, *173*(1), 202-225. https://doi.org/10.1093/toxsci/kfz201

Pearce, R. G., Setzer, R. W., Strope, C. L., Wambaugh, J. F., & Sipes, N. S. (2017). httk: R Package for High-Throughput Toxicokinetics. *J Stat Softw*, *79*(4), 1-26. https://doi.org/10.18637/jss.v079.i04

Phillips, J. R., Svoboda, D. L., Tandon, A., Patel, S., Sedykh, A., Mav, D., Kuo, B., Yauk, C. L., Yang, L., Thomas, R. S., Gift, J. S., Davis, J. A., Olszyk, L., Merrick, B. A., Paules, R. S., Parham, F., Saddler, T., Shah, R. R., & Auerbach, S. S. (2019). BMDExpress 2: enhanced transcriptomic dose-response analysis workflow. *Bioinformatics*, *35*(10), 1780-1782. https://doi.org/10.1093/bioinformatics/bty878

Reardon, A. J. F., Farmahin, R., Williams, A., Meier, M. J., Addicks, G. C., Yauk, C. L., Matteo, G., Atlas, E., Harrill, J., Everett, L. J., Shah, I., Judson, R., Ramaiahgari, S., Ferguson, S. S., & Barton-Maclaren, T. S. (2023). From vision toward best practices: Evaluating in vitro transcriptomic points of departure for application in risk assessment using a uniform workflow. *Front Toxicol*, *5*, 1194895. https://doi.org/10.3389/ftox.2023.1194895

Sharma, A., Banyiova, K., Babica, P., El Yamani, N., Collins, A. R., & Cupr, P. (2017). Different DNA damage response of cis and trans isomers of commonly used UV filter after the exposure on adult human liver stem cells and human lymphoblastoid cells. *Sci Total Environ*, *593-594*, 18-26. https://doi.org/10.1016/j.scitotenv.2017.03.043

Shockley, K. R., Cora, M. C., Malarkey, D. E., Jackson-Humbles, D., Vallant, M., Collins, B. J., Mutlu, E., Robinson, V. G., Waidyanatha, S., Zmarowski, A., Machesky, N., Richey, J., Harbo, S., Cheng, E., Patton, K., Sparrow, B., & Dunnick, J. K. (2020a). Comparative toxicity and liver transcriptomics of legacy and emerging brominated flame retardants following 5-day exposure in the rat. *Toxicol Lett*, *332*, 222-234. https://doi.org/10.1016/j.toxlet.2020.07.016

Shockley, K. R., Cora, M. C., Malarkey, D. E., Jackson-Humbles, D., Vallant, M., Collins, B. J., Mutlu, E., Robinson, V. G., Waidyanatha, S., Zmarowski, A., Machesky, N., Richey, J., Harbo, S., Cheng, E., Patton, K., Sparrow, B., & Dunnick, J. K. (2020b). Transcriptomic data from the rat liver after five days of exposure to legacy or emerging brominated flame retardants. *Data Brief*, *32*, 106136. https://doi.org/10.1016/j.dib.2020.106136

Soeteman-Hernandez, L. G., Fellows, M. D., Johnson, G. E., & Slob, W. (2015). Correlation of In Vivo Versus In Vitro Benchmark Doses (BMDs) Derived From Micronucleus Test Data: A Proof of Concept Study. *Toxicol Sci*, *148*(2), 355-367. https://doi.org/10.1093/toxsci/kfv189

Song, Y., Zheng, K., Brede, D. A., Gomes, T., Xie, L., Kassaye, Y., Salbu, B., & Tollefsen, K. E. (2023). Multiomics Point of Departure (moPOD) Modeling Supports an Adverse Outcome Pathway Network for Ionizing Radiation. *Environ Sci Technol*, *57*(8), 3198-3205. https://doi.org/10.1021/acs.est.2c04917

Stauffer Chemical Co. (1981). A two-year oral toxicity/carcinogenicity study of Fyrol FR-2 rats (Volume I - IV) (Final Reports) with attachments, cover sheets and letter dated 093081. .

Takada, K., Yasuhara, K., Nakaji, Y., Yoshimoto, H., Momma, J., Kurokawa, Y., Aida, Y., & Tobe, M. (1989). CARCINOGENICITY STUDY OF TRIS (2-CHLOROETHYL) PHOSPHATE IN ddY MICE. *Journal of Toxicologic Pathology*, *2*(2), 213-222. https://doi.org/10.1293/tox.2.213

Tangri, S., Vall, H., Kaplan, D., Hoffman, B., Purvis, N., Porwit, A., Hunsberger, B., Shankey, T. V., & on behalf of, I. I. W. G. (2013). Validation of cell-based fluorescence assays: Practice guidelines from the ICSH and ICCS – part III – analytical issues. *Cytometry Part B: Clinical Cytometry*, *84*(5), 291-308. https://doi.org/https://doi.org/10.1002/cyto.b.21106

Thomas, D. G., Shankaran, H., Truong, L., Tanguay, R. L., & Waters, K. M. (2019). Time-dependent behavioral data from zebrafish reveals novel signatures of chemical toxicity using point of departure analysis. *Comput Toxicol*, *9*, 50-60. https://doi.org/10.1016/j.comtox.2018.11.001

TNO Quality of Life. (2007). Oral two-generation reproduction toxicity study (including a dose range finding study) with Tris(2-chloro-1-methylethyl) -phosphate in rats (Unpublished report). Cited in EU RAR 2008a.

U.S. Environmental Protection Agency. (U.S. EPA). (2002a). *Hepatocellular Hypertrophy. HED Guidance Document #G2002.01*.

U.S. Environmental Protection Agency. (U.S. EPA). (2002b). *A review of the reference dose and reference concentration process* (EPA/630/P-02/002F).

U.S. Environmental Protection Agency. (U.S. EPA). (2006). *Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment* (EPA/600/R-05/043F). https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=157668

U.S. Environmental Protection Agency. (U.S. EPA). (2009). *Provisional Peer-Reviewed Toxicity Values for tris(2-chloroethyl)phosphate (TCEP) (CASRN 115-96-8)* (EPA/690/R-09/069F). http://hhpprtv.ornl.gov/issue_papers/Tris2chloroethylphosphate.pdf

U.S. Environmental Protection Agency. (U.S. EPA). (2012). *Provisional Peer-Reviewed Toxicity Values for Tris(1-chloro-2-propyl)phosphate (CASRN 13674-84-5)*. http://hhpprtv.ornl.gov/issue_papers/Tris1chloro2propylphosphate.pdf

U.S. Environmental Protection Agency. (U.S. EPA). (2023). *Draft risk evaluation for tris(2-chloroethyl) phosphate (TCEP): CASRN 115-96-8* (EPA-740-D-23-002).

U.S. Environmental Protection Agency. (U.S. EPA). (2024). *Scientific Studies Supporting Development of Transcriptomic Points of Departure for EPA Transcriptomic Assessment Products (ETAPs)*. https://epa.figshare.com/articles/online_resource/Scientific_Studies_Supporting_Development_of_Transcriptomic_Points_of_Departure_for_EPA_Transcriptomic_Assessment_Products_ETAPs_/25365550

Wambaugh, J. F., Bare, J. C., Carignan, C. C., Dionisio, K. L., Dodson, R. E., Jolliet, O., Liu, X., Meyer, D. E., Newton, S. R., Phillips, K. A., Price, P. S., Ring, C. L., Shin, H.-M., Sobus, J. R., Tal, T., Ulrich, E. M., Vallero, D. A., Wetmore, B. A., & Isaacs, K. K. (2019). New approach methodologies for exposure science. *Current Opinion in Toxicology*, *15*, 76-92. https://doi.org/https://doi.org/10.1016/j.cotox.2019.07.001

Wang, N. C., Jay Zhao, Q., Wesselkamper, S. C., Lambert, J. C., Petersen, D., & Hess-Wilson, J. K. (2012). Application of computational toxicological approaches in human health risk assessment. I. A tiered surrogate approach. *Regul Toxicol Pharmacol*, *63*(1), 10-19. https://doi.org/10.1016/j.yrtph.2012.02.006

Wu, Z., Zhu, M., Kang, Y., Leung, E. L., Lei, T., Shen, C., Jiang, D., Wang, Z., Cao, D., & Hou, T. (2021). Do we need different machine learning algorithms for QSAR modeling? A comprehensive assessment of 16 machine learning algorithms on 14 QSAR data sets. *Brief Bioinform*, *22*(4). https://doi.org/10.1093/bib/bbaa321

Yang, L., Allen, B. C., & Thomas, R. S. (2007). BMDExpress: a software tool for the benchmark dose analyses of genomic data. *BMC Genomics*, *8*, 387. https://doi.org/10.1186/1471-2164-8-387