Screening for neurotoxic potential of 15 flame retardants using freshwater planarians

Siqi Zhang[a,*], **Danielle Ireland**[b,e,*], **Nisha S. Sipes**[c], **Mamta Behl**[c], **Eva-Maria S. Collins**[b,d,e]

[a]Department of Bioengineering, University of California San Diego, La Jolla, CA 92093, USA

[b]Division of Cell and Developmental Biology, University of California San Diego, La Jolla, CA 92093, USA

[c]Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, 27709, USA

[d]Department of Physics, University of California San Diego, La Jolla, CA 92093, USA

[e]Department of Biology, Swarthmore College Swarthmore, PA 19081, USA

Abstract

Asexual freshwater planarians are an attractive invertebrate model for high-throughput neurotoxicity screening, because they possess multiple quantifiable behaviors to assess distinct neuronal functions. Planarians uniquely allow direct comparisons between developing and adult animals to distinguish developmentally selective effects from general neurotoxicity. In this study, we used our automated planarian screening platform to compare the neurotoxicity of 15 flame retardants (FRs), consisting of representative phased-out brominated (BFRs) and replacement organophosphorus FRs (OPFRs). OPFRs have emerged as a proposed safer alternative to BFRs; however, limited information is available on their health effects. We found 11 of the 15 FRs (3/6 BFRs, 7/8 OPFRs, and Firemaster 550) caused adverse effects in both adult and developing planarians with similar nominal lowest-effect-levels for BFRs and OPFRs. This suggests that replacement OPFRs are comparably neurotoxic to the phased-out compounds. BFRs were primarily systemically toxic, whereas OPFRs, except Tris(2-chloroethyl) phosphate, shared a behavioral phenotype in response to noxious heat at sublethal concentrations, indicating specific neurotoxic effects. We found this behavioral phenotype was correlated with cholinesterase inhibition, thus linking behavioral outcomes to molecular targets. By directly comparing effects on adult and developing planarians, we further found that one BFR (3,3',5,5'-Tetrabromobisphenol A) caused a developmental selective defect. Together, these results demonstrate that our planarian screening platform yields high content data from various behavioral and morphological endpoints, allowing us to distinguish selective neurotoxic effects and effects specific to the developing nervous system. Ten of these 11 bioactive FRs were previously found to be bioactive in other

Correspondence should be addressed to Eva-Maria S. Collins; ecollin3@swarthmore.edu; Mailing address: Biology Department, Swarthmore College, 500 College Ave, Swarthmore, PA 19081, Telephone: 610-690-5380; Fax: (610) 328-8663.
[*]These authors contributed equally to this work.

models, including cell culture and alternative animal models (nematodes and zebrafish). This level of concordance across different platforms emphasizes the urgent need for further evaluation of OPFRs in mammalian systems.

Keywords

alternative animal models; planarian; flame retardants; OPFRs; polybrominated diphenyl ethers; developmental neurotoxicity

1. Introduction

Traditional hazard assessment using mammalian models cannot keep up with the speed of chemical development. The Toxic Substances Control Act Inventory lists approximately 68,000 chemicals currently in use in the United States (https://www.epa.gov/tsca-inventory), but <2% have been evaluated for developmental neurotoxicity (DNT) (Smirnova et al., 2014). *In vitro* testing combined with bioinformatic analyses have emerged as a powerful tool to achieve more rapid chemical testing (Fritsche et al., 2018), but is limited in predicting adverse effects on neuronal function (Smirnova et al., 2014; Tohyama, 2016). A promising alternative solution for efficient DNT screening are non-mammalian animal models, amenable to high-throughput, cost-effective screening, since they allow for automated behavioral assays to quantitatively evaluate different neuronal functions (Bal-Price et al., 2012).

Flame retardants (FRs), widely added to commercial products for fire safety (Costa and Giordano, 2007; Salimi et al., 2017), are a class of chemicals which only have limited mammalian data for safety assessment. Prior to 2005, brominated FRs (BFRs) such as polybrominated diphenyl ether (PBDE) mixtures were the primary FRs used in the United States (Hale et al., 2003). Since then, many were voluntarily phased-out due to growing evidence of their associations with impaired neurodevelopment and decreased fertility, exacerbated by their persistence in the environment and ability to bioaccumulate (Darnerud, 2003; Herbstman et al., 2010; Stapleton et al., 2011; Talsness, 2008). Organophosphorus FRs (OPFRs) have gradually become replacements for PBDEs over the last decade (Stapleton et al., 2012; Stapleton et al., 2014). Concerns exist that OPFRs, similar to PBDEs, may persist in the environment and bioaccumulate, but little is known about their potential toxicity (Hendriks and Westerink, 2015; Meeker and Stapleton, 2010; Stapleton et al., 2009; Van Der Veen and De Boer, 2012). Particularly concerning is the potential DNT of OPFRs, as they share structural similarities to organophosphorus pesticides, which adversely affect neurodevelopment and cause neurobehavioral impairments (González-Alzaga et al., 2014; Muñoz-Quezada et al., 2013; Ricceri et al., 2006; Slotkin et al., 2006). Previous work using a screening battery of *in vitro* and alternative animal models (zebrafish and nematodes) found that several OPFRs had comparable toxicity and potency to BFRs, including phased-out BDE-47 (Behl et al., 2016, 2015; Jarema et al., 2015), indicating that OPFRs may not be the safer alternative to BFRs that they have been marketed to be.

Because these existing studies focused on a few FRs and/or developmental endpoints, we screened 15 FRs, including 8 OPFRs, in the freshwater planarian *Dugesia japonica*

using a custom high-throughput screening (HTS) platform to assess neurotoxicity and DNT. This planaria model is particularly well-suited for DNT studies, because it allows for interrogation of multiple endpoints of distinct neuronal functions and differentiation of neurotoxicity from DNT in parallel with the same assays in a single screen (Zhang et al., 2019). Therefore, the results obtained herein enrich the toxicity profiles of these chemicals by providing complementary insight into potential DNT. Additionally, we previously characterized two putative genes responsible for cholinesterase activity and demonstrated that *D. japonica* is well-suited to study organophosphorus pesticide neurotoxicity linked to cholinesterase inhibition (Hagstrom et al., 2018, 2017; Zhang et al., 2019). Because of the structural similarities of OPFRs to organophosphorus pesticides, the planarian model therefore promises to provide insights into the neurotoxicity mechanisms of these FRs.

Because *D. japonica* is a fairly new model for neurotoxicity studies, we briefly summarize its key features to familiarize the reader and refer to our recent review paper (Hagstrom et al., 2016) for an in-depth review of planarians as an alternative model for DNT. In asexual *D. japonica* planarians, neurodevelopment is solely achieved through neuroregeneration following asexual reproduction by fission. This allows neurodevelopment to be induced by head amputation. In addition, because adult and regenerating/developing planarians are of similar size, both worm types can be screened in parallel with the same assays to differentiate DNT from neurotoxicity (Hagstrom et al., 2016, 2015; Zhang et al., 2019). Although the planarian central nervous system is morphologically simple, consisting of a bilobed cephalic ganglion and two ventral nerve cords, it is molecularly complex and highly compartmentalized. It shares the same major neurotransmitters as the vertebrate brain, including dopamine, serotonin, octopamine, GABA, and acetylcholine (Cebrià, 2007; Cebrià et al., 2002a, 2002b; Ross et al., 2017). Furthermore, planarians possess a large repertoire of diverse quantifiable behaviors, such as gliding and swimming, phototaxis, thermotaxis, and scrunching. The molecular mediators of some of these behaviors have been characterized, such as serotonergic neurons in gliding or transient receptor potential channels in thermotaxis (Birkholz and Beane, 2017; Inoue et al., 2014; Nishimura et al., 2010). Thus, these behaviors can be used to assess specific neuronal functions and provide insight into the cellular effects of neurotoxicity. In summary, by testing diverse behaviors in both adult and regenerating/developing planarians, this platform uniquely allows differentiation of selective neurotoxic effects and effects specific to the developing nervous system.

The 15 FRs tested herein include 3 phased-out PBDEs (BDE-47, BDE-99, and BDE-153), 3 currently in-use BFRs (TBB, TBBPA, and TBPH), 8 in-use OPFRs (BPDP, EHDP, IDDP, IPP, TCEP, TCPP, TMPP, and TPHP), and one BFR mixture (FM550). These chemicals comprise a subset of an 87-compound library from the National Toxicology Program (NTP) (Behl et al., 2019; Hagstrom et al., 2019; Zhang et al., 2019), which we had previously screened. However, in this previous study we had not evaluated these compounds in detail, because we focused on the comparison of compound classes (Zhang et al., 2019) and concordance with other systems (Hagstrom et al., 2019). Moreover, we re-screened these 15 FRs independently in a second screen and used the two independent screens to evaluate the robustness of our planarian screening platform. Thus, the present study serves a dual purpose: To more closely examine the potential neurotoxicity of FRs and to quantify screen

robustness. To contextualize our results, we evaluated concordance to published data from zebrafish, nematode, and human, mouse, and rat cell-culture models (Behl et al., 2016, 2015; Jarema et al., 2015; Noyes et al., 2015). There was high concordance among the different alternative systems showing comparable DNT of OPFRs to BFRs, warranting further investigation into OPFR toxicity.

2. Materials and methods

2.1. Chemical library

18 chemicals consisting of 6 BFRs (including phased-out BDE-47 and BDE-99), 8 OPFRs, 1 BFR mixture, 2 negative controls, and 1 duplicate (TPHP) were screened in this study. Acetaminophen and L-ascorbic acid were chosen as negative controls because we previously demonstrated they showed no effects in our planarian screening platform (Zhang et al., 2019). Table 1 lists the 17 unique chemicals with their name, chemical ID, type, supplier, structure, and purity information provided by the NTP (either as determined by the NTP or, if not available, from the chemical supplier's certificate of analysis). The chemicals were provided by the NTP as ~20mM stocks in dimethyl sulfoxide (DMSO, Gaylord Chemical Company, Slidell, LA), with the exception of 2,2',4,4',5,5'-Hexabromodiphenyl ether (BDE-153), which was provided at 10mM.

The tested FRs were a subset of the NTP 87-compound library, which we previously reported on (FR screen 1; (Zhang et al., 2019)). The NTP later provided us with another set of the library (now including 4 duplicate chemicals for a total of 91 chemicals), in which we only screened the subset of chemicals mentioned in this study (FR screen 2). Chemicals in the new library were provided as individual vials, stored at 4°C and used within 2 months. For all but two chemicals (IPP and TMPP), the same supplier and mostly (14/17) the same chemical lot was used for the chemicals in both FR screen 1 and 2. TPHP was tested in duplicate in FR screen 2, using 2 separate stock vials from the same supplier and lot. The stocks of the negative control L-ascorbic acid were from different lots but were from the same supplier with the same concentration and purities. Five of the chemical stocks differed slightly in concentration between the two libraries, but by at most 0.3mM. Additional information on the NTP library can be found in (Behl et al., 2019).

The chemicals were tested at 0.01, 0.1, 1, 10, and 100μM (0.005, 0.05, 0.5, 5, and 50μM for BDE-153), with a final DMSO concentration of 0.5% for all concentrations. Solvent control populations were exposed to 0.5% DMSO. Chemical dilutions for both libraries were done as described in (Zhang et al., 2019).

2.2. FR screen in the planarian system

Asexual *D. japonica*, originally obtained from Shanghai University (Shanghai, China) and cultivated in our lab for > 5 years, were used for all experiments. We screened the library with a custom, fully automated HTS platform using adult (intact) and regenerating, i.e. tail pieces regenerating a new head, planarians in parallel with the same assays, assessing morphological and behavioral endpoints at day 7 and day 12 (Figure 1). Two independent screens, the original NTP 87-compound library screen ("FR screen 1") and the second,

FR-only screen ("FR screen 2"), were performed with 3 replicates (3 independent runs on different days) for each screen. For each replicate, n=8 regenerating and adult planarians were each screened for every chemical concentration. One adult or regenerating animal was loaded into each well of a 48-well plate (Genesee Scientific, San Diego, CA). Static exposure began on day 1, within 3 hours of amputation of the regenerating planarians, by adding the appropriate chemical concentration to the planarians' aquatic environment. A solvent control population (n=8) exposed to 0.5% DMSO was included in every plate. This DMSO concentration has no effect on planarian morphology or behavior (Hagstrom et al., 2015). The planarians were kept in the dark in the sealed 48-well plates for 12 days, with screening on days 7 and 12 (Figure 1). By day 12, regenerating planarians are considered fully regenerated and behave like adults (Hagstrom et al., 2015; Zhang et al., 2019), thus the two screening days allow us to distinguish regeneration delays from regeneration defects.

A detailed description of our automated planarian screening platform and statistical workflow can be found in (Zhang et al., 2019). Briefly, adult and regenerating planarians were assessed using morphological and behavioral endpoints at day 7 and day 12 (Figure 1). Since dead worms generally disintegrate, lethality was quantified by determination of the presence or absence of a body in each well. Eye regeneration was quantified by determining the percentage of planarians with 2 eyes regenerated (normal regeneration). For behavioral endpoints, we used computational image analysis to quantify animal behavior to assay unstimulated locomotion, negative phototactic behavior in response to blue light, the ability to react to a heat gradient in thermotaxis, and the ability to scrunch (musculature-driven, oscillatory escape gait (Cochet-Escartin et al., 2015)) in response to noxious heat. For statistical analysis, data from the replicate runs were compiled and each chemical concentration was compared with its own in-plate DMSO controls. Statistical significance was determined using either a one-tailed Fisher's exact test for lethality, eye regeneration, phototaxis, and scrunching endpoints; Mann Whitney U-test for thermotaxis and unstimulated behavior; or two-tailed t-test for unstimulated behavior (depending on normality of the sample) using a significance level of 0.05. To extract biologically significant results from the statistically significant hits, "biological relevancy" cutoffs, based on the distribution of all control populations, were used to exclude hits that fell within the variability of the DMSO controls. Lastly, hits generated from inconsistent data from the different replicates were also excluded. For each endpoint, the lowest concentration which showed a significant effect (statistically and biologically) was defined as the lowest-effect-level (LEL). For each chemical, the overall LEL was determined as the most sensitive LEL when considering all endpoints. All image analyses and statistical analyses were performed in MATLAB.

The same experimental procedures and data analysis methods as described in detail in (Zhang et al., 2019) were used, except for the following differences in FR screen 2, building upon our experiences from FR screen 1: Firstly, since we found that food quality affects planarian fitness and sensitivity to chemicals (Zhang et al., 2019), we fed planarians used in FR screen 2 commercial freeze-dried organic chicken liver (Amazon, Seattle, WA) to better control food quality and thus minimize animal fitness variability. Secondly, "biological relevance cutoffs", used to minimize false positives by accounting for variability of the solvent controls in each endpoint (Zhang et al., 2019), were updated by using the combined

DMSO control samples from the previous NTP 87-compound screen (n=87 control samples) and the later FR-only screen (n=18 control samples) for a total control population of n=105 samples, with each sample consisting of n=24 planarians. Thirdly, the background noise speed threshold used in the phototaxis assay was recalculated based on the mean speed in the unstimulated behavioral assay of the expanded DMSO controls. Hits due to hyper- (rather than hypo-) activity in unstimulated behavior or hits which were concentration-independent were not included, because they were not conserved when compiling all 6 runs together, suggesting they may be artifacts.

2.3. Robustness analysis of planarian screening platform

The robustness of our planarian screening platform was evaluated by comparing the results obtained from six replicate runs (n=8 planarians per condition per replicate) from the two independent FR screens, FR screen 1 and 2. It is reasonable to assume that using all 6 replicates provides the most accurate results given the large sample size (n=48). We therefore compared the results of compiling different numbers of replicates (3, 4, or 5) with the results from 6 replicates. As described previously (Zhang et al., 2019), we used a 3-step (statistical test, biological relevance cutoff, and inconsistency check) statistical workflow to determine whether a chemical showed bioactivity, defined as a significant adverse effect, in any endpoint. Herein, we refer to the determination of bioactivity (i.e. bioactive or inactive) in each endpoint as a "readout" to distinguish it from a bioactive "hit", which refers to having an adverse effect at an endpoint. Since the biological relevance cutoffs are meant as a means to distinguish normal wild-type noise levels, biological relevance cutoffs were determined based on the total control population of n=105 control samples (see Section 2.2), and were applied equally to these different sets of compiled data (i.e. the same cutoffs, determined from this total control population, were used for 3, 4, 5, or 6 replicates). For the inconsistency check, instances were excluded where more than half of a set of replicates (i.e. >1 out of 3 replicates, >2 out of 4 replicates, >2 out of 5 replicates, or >3 out of 6 replicates) were within the biological relevancy cutoffs. Although the majority of the experimental procedures were consistent between the two screens, some differences existed which could affect the reproducibility of the screens. For example, varying food quality (between different home-made batches and between home-made and commercial sources) affects animal health and thus affects the animals' sensitivity to chemical exposure (Zhang et al., 2019). Different chemical batches can also cause variability due to differences in stock concentrations and potential variability in the serial dilution process. Therefore, we chose replicates (different plates from different days) with the most similar conditions (i.e. in the same screen) to minimize experimental variability and to focus on the robustness of the system itself. To compile 3 replicates, we chose the 3 replicates from FR screen 2. To compile 4 (or 5) replicates, we chose 1 (or 2) replicate(s) from FR screen 1 in addition to the 3 replicates from FR screen 2. Firstly, overall chemical bioactivity (i.e. whether the chemical was bioactive in any endpoint or inactive in all endpoints) in different sets of replicates (3, 4, 5, or 6) was compared. Secondly, concordance of the bioactivity of the readouts (i.e. whether the chemical was deemed bioactive or inactive in a specific endpoint) between the data using 3, 4, or 5 replicates and the data using 6 replicates were compared. Lastly, the concordance of the sensitivity of the concordant bioactive readouts (i.e. LEL of the bioactive readout) was compared between the data using 3, 4, or 5 replicates and the data using 6 replicates.

2.4. Comparison with other alternative models

To directly compare our results with published data from other alternative models, we consulted the literature (Behl et al., 2016, 2015; Jarema et al., 2015; Noyes et al., 2015) and focused on studies which had screened the majority of the same FRs. Studies which only tested one or a few of these FRs were not included in the comparison. We extracted the points-of-departures (PODs) and LELs reported in these papers for relevant developmental and developmental neurotoxic endpoints and focused only on the concentration-dependent hits for LELs to facilitate cross-system comparisons. Specifically, we extracted PODs for developmental toxicity and developmental neurotoxicity reported in (Behl et al., 2015), LELs for developmental behavior in (Jarema et al., 2015), LELs for larval development in (Behl et al., 2016), and LELs for all endpoints reported in (Noyes et al., 2015). Only data for regenerating planarians was used for this analysis as the comparative studies used developing models.

To quantitatively determine the agreement between the different alternative animal studies, Cohen's kappa (κ) was calculated for each pairwise comparison where κ = (observed agreement - probability of chance agreement)/(1- probability of chance agreement) (McHugh, 2012). The 95% confidence intervals were quantified as $\kappa \pm$ (1.96 x the standard error). Of note, the Noyes et al. study showed a $\kappa=0$ for all comparisons because it had no inactives in this set of 9 FRs, leading to a high probability of agreement due to chance.

2.5. Cholinesterase activity staining in OPFR treated planarians

To determine whether OPFR toxicity was correlated with inhibition of acetylcholinesterase activity, adult planarians were incubated in either 0.5% DMSO, 10 μM EHDP (an active OPFR in planarians) or 100 μM TCEP (an inactive OPFR in planarians) for 12 days and then fixed and stained to visualize acetylthiocholine catalysis as previously described (Hagstrom et al., 2018; Zheng et al., 2011), with a staining incubation of 3.5 hours. The staining was performed in duplicate on separate days on a total of 6 worms per chemical (3/day) and 12 DMSO controls (6/day).

3. Results

A library of 15 FRs and 2 negative controls (acetaminophen and L-ascorbic acid) (Table 1), was screened in adult and regenerating planarians to assess their effects on mortality, eye regeneration, and behavior. These chemicals were screened twice, once as part of the NTP 87-compound library screen (FR screen 1) (Zhang et al., 2019) and once as a new, independent screen of the FR-specific subset of a newly obtained version of the library (FR screen 2), with 3 replicate runs in each screen. Therefore, we compared the results obtained from these two independent screens using the same screening methodology to evaluate the robustness and reproducibility of our planarian screening platform. Moreover, as our previous study of the NTP 87-compound library focused on the screening methodology and major trends seen in the screen, herein we provide more in-depth insight of the toxicological profiles induced by this discrete subset of chemicals. The raw and analyzed data are available in Mendeley Data and Supplementary File 1, respectively.

3.1. Robustness of the planarian FR screen

To evaluate screen robustness and reproducibility, we compared the results obtained from compiling different numbers of replicates from the two independent screens (FR screen 1 and FR screen 2) (Figure 2 and Supplementary Figure 1). Compiled results for the individual replicates and different sets of replicates can be found in Supplemental File 1. Of the 18 screened chemicals, IPP and TMPP were provided by two different suppliers for the two screens (Table 1). Therefore, given potential unknown variability between the different chemical batches, IPP and TMPP were excluded from the robustness evaluation. Notably, a recent screen with developing zebrafish found significant readout differences between IPP stocks from three different suppliers (Noyes et al., 2015), confirming potential supplier variability.

To determine the robustness of our data, we evaluated the overall agreement across different numbers of compiled replicates on bioactivity identification and potency (LELs) for any endpoint (overall determination of toxicity) and for each endpoint (specific effects). When considering all endpoints collectively, the same 10/16 chemicals, including the duplicate TPHP, caused adverse effects in both adult and regenerating planarians in all numbers of replicates. Thus, the overall chemical hit-call was unchanged regardless of the number of replicates used. Sensitivity was generally conserved as the overall LELs, i.e. the most sensitive LEL from any assay endpoint for each chemical, were the same for 9/10 bioactive chemicals in all numbers of replicates. Only BDE-47 had different LELs depending on the number of replicates analyzed (Supplementary Information).

Because one of the strengths of our planarian screening platform is the ability to distinguish effects on different neuronal functions using various behavioral readouts, we also compared the bioactivity of each endpoint readout for 3, 4, or 5 replicates with all 6 replicates (Figure 2). Bioactivity was 96% concordant for 3, 4, and 5 replicates compared to all 6 replicates (Figure 2A) and sensitivity was 97% concordant (Figure 2B). Details on the few differences that were observed can be found in Supplementary Figure 1 and Supplementary Information. Together, these data show the high reproducibility among the different replicates in both endpoint-specific bioactivity and potency. TPHP was screened as a duplicate in FR screen 2 (FR-specific library screen). The results of the duplicates were consistent in any set of compiled data, being bioactive in the same endpoints at the same LELs (Supplementary Figure 1). Thus, our current screening strategy using 3 replicates is sufficiently robust.

3.2. OPFRS are more potent than BFRs in planarians

After confirming our data is robust and reproducible, we used the compiled data from all 6 replicates to delve into the details of FR toxicity. Eleven of the 15 unique FRs caused adverse effects in both adult and regenerating planarians in at least 1 endpoint (Figure 3). The duplicate chemical, TPHP, was consistent between batches and thus the results are only shown once. The bioactive FRs consisted of 3 of the 6 BFRs (BDE-99, BDE-47, TBBPA), 7 of the 8 OPFRs (IDDP, TCPP, EHDP, IPP, BPDP, IPP, TMPP, TPHP), and the FM550 mixture. TBPH, BDE-153, TBB, TCEP, as well as the two negative control chemicals, were inactive in both worm types at the tested concentrations. All chemicals were screened at

5 test concentrations (see Materials and Methods, Section 2.1); however, hits were only observed at the two highest concentrations, 10 and 100 μM. Due to poor solubility, BDE-153 was tested at a maximum concentration of 50 μM, which may explain its inactivity, since the other BDE compounds were only active at 100 μM.

Because our screening platform provides high-content data consisting of effects on planarian morphology and a variety of diverse behaviors, we determined whether any specific toxicities were induced by the different FRs. Eight FRs in regenerating planarians (BDE-99, TBBPA, TCPP, EHDP, IPP, BPDP, TMPP, and TPHP) and 6 FRs in adult planarians (TCPP, EHDP, IPP, BPDP, TMPP, and TPHP) caused adverse outcomes on morphological and/or behavioral endpoints at concentrations where animal viability was not affected, suggesting specific non-systemic toxicity at these concentrations. Two BFRs, BDE-99 and TBBPA, caused sublethal effects in regenerating planarians but only showed toxicity concomitant with lethality in adult planarians. This was due to increased sensitivity to lethality in the adult worms, which is a trend we have previously observed for some other chemicals (Hagstrom et al., 2015; Zhang et al., 2019).

Next, we further examined which specific endpoints were affected by the sublethal concentrations of the different FRs. Scrunching, a planarian musculature-driven gait used as an escape response to specific adverse stimuli such as noxious heat (Cochet-Escartin et al., 2015), was the most sensitive endpoint with 7 and 6 FRs causing scrunching defects at sublethal concentrations in regenerating and adult planarians, respectively. In regenerating planarians, 6/7 of these FRs showed scrunching defects in the absence of eye regeneration defects. Similarly, in 6/7 and 4/6 of the sublethal, active FRs in regenerating and adult planarians, respectively, scrunching defects were observed in the absence of defects on unstimulated behavior/locomotion. Together, these indicate that inability to scrunch in these cases is likely due to specific neurotoxicity and is not a result of general morphological or locomotive phenotypes. To a lesser extent, sublethal effects were also seen on unstimulated behavior (regenerating planarians in TBBPA and adult planarians in IPP) and eye regeneration (TBBPA and TCPP). Interestingly, in addition to effects on eye regeneration, TCPP affected almost all behavioral endpoints (with the exception of day 7 phototaxis in regenerating planarians) in both worm types in the absence of lethality. Together, these toxicological profiles demonstrate the power of the large number of diverse endpoints to discern specific toxic effects from systemic toxicity and pinpoint effects on specific neuronal functions, such as neuromuscular communication.

Our planarian HTS platform has the unique advantage of allowing direct comparisons of chemical effects on adult and developing/regenerating animals. At 10 μM, TBBPA caused defects on day 12 unstimulated behavior in regenerating but not adult planarians, indicating a developmental selective defect (Figure 3). Of note, this concentration did induce lethality in 24% of adult planarians at day 12 (see Supplemental File 1). However, this low level of lethality, although statistically significant, may suggest that any overt systemic toxicity at this concentration would not be potent enough to mask effects in other endpoints. Furthermore, TBBPA and TCPP caused eye regeneration defects at a sublethal concentration. Thus, while most compounds adversely affected both adult and developing animals, some compounds displayed additional developmental-specific toxicities.

Lastly, to ask whether OPFRs are safer alternatives to BFRs, we compared the toxicological profiles induced by the two types of FRs. The OPFRs were overall more potent than the BFRs with LELs of 10μM in 5 out of 7 bioactive OPFRs, whereas the bioactive BFRs, except for TBBPA, had LELs of 100μM. TBBPA had a LEL of 10μM in both adult and regenerating planarians, with the most endpoints affected at this concentration, compared to the other FRs. Taken together, the bioactive OPFRs were at least as potent as BDE-47, which was phased-out due to concerns about its toxicity (Stapleton et al., 2014), emphasizing the importance of evaluating the toxicity of these supposedly safer replacement FRs.

Almost all bioactive OPFRs (6/7 and 5/6 for regenerating and adult planarians, respectively) induced specific sublethal effects on scrunching. This was not seen to the same extent in the BFRs, where only 1 bioactive BFR (BDE-99) showed sublethal effects on scrunching and only in regenerating planarians. Because scrunching is a muscular-driven gait and motor activity in planarians is controlled by cholingeric neurotransmission (Nishimura et al., 2012), we hypothesized that scrunching defects caused by the OPFRs were correlated with inhibition of cholinesterase activity. We tested this hypothesis using a cholinesterase activity staining method previously validated in planarians (Hagstrom et al., 2018) and found an almost complete loss of cholinesterase activity in day 12 adult planarians exposed to 10 μM EHDP, which causes scrunching defects. In contrast, cholinesterase activity in day 12 adult planarians exposed to 100 μM TCEP, which was the only inactive OFPR in our planarian screen, was indistinguishable from controls (Figure 4). Thus, by assaying a variety of different neuronal functions in planarians, we were able to differentiate toxicities elicited by different chemical domains, suggesting that our system provides specificity.

3.3. Concordance of bioactive hits between planarians and other alternative models

While several of the 15 FRs that we screened have been previously studied in other systems, including rodents (Cope et al., 2015; EFSA, 2011; Moser et al., 2015; Nakajima et al., 2009; National Toxicology Program, 1991; Viberg and Eriksson, 2011) and other alternative models (Alzualde et al., 2018; Bailey and Levin, 2015; Cano-Sancho et al., 2017; Glazer et al., 2018; Oliveri et al., 2015; Slotkin et al., 2017; Usenko et al., 2016), only a handful of studies have performed direct multi-FR comparisons as provided here. Fourteen of the 15 unique FRs tested in this screen have been previously studied in developing zebrafish, nematodes, and *in vitro* cell-based (mouse embryonic stem cell, human neural stem cell, and rat neuron) systems with some FRs showing developmental toxicity or DNT. Thus, to contextualize our findings, we compared our planarian data with these published results (Behl et al., 2016, 2015; Jarema et al., 2015; Noyes et al., 2015) (Table 2, Figure 5). The FM550 mixture was excluded from this comparison because it was not screened in any of these previous studies. Comparisons were made with the overall regenerating planarian LEL, i.e. most sensitive LEL from any endpoint, determined from the compiled data of 6 replicates from FR screens 1 and 2 (see Methods Section 2.4, Table 2 and Figure 5).

First, we compared the chemical hit calls between the different systems. Thirteen of the 14 shared FRs were bioactive in at least one other model; however, not all 14 FRs were tested in all systems or studies. IPP was the only FR to be bioactive in all tested systems. TBB was

the only FR which was inactive in all screens; however, it was only tested in our study and the (Noyes et al., 2015) zebrafish study. Generally, the alternative animal models detected more bioactive FRs than the *in vitro* models (Table 2). Of the 9 FRs tested in the *in vitro* models, 8 were bioactive in any model, with at most 7 detected in any one model (human stem cell neurite outgrowth). Most effects caused by the FRs in the *in vitro* models occurred concomitantly with cytotoxicity, suggesting non-specific effects.

Due to the increased specificity and higher similarity to our planarian screening platform, we focused our concordance analysis on the alternative developing/regenerating animal models (planarians, nematodes, and zebrafish). The chemical hit calls for the 14 FRs tested in these models were highly concordant. The two invertebrate systems, nematodes and planarians, had perfect hit concordance for the 10 FRs tested in both systems. These 10 FRs were also tested in at least one zebrafish study and showed high concordance among the three models (considering all studies together) with the same 9 chemicals eliciting adverse effects in at least one study from each system (Table 2, Figure 5A). Of these 10 FRs, only TCEP had discordant results between the different systems, as it was bioactive in two of the zebrafish studies, but inactive in the third zebrafish study and in nematodes and planarians.

As this example illustrates, some discrepancies were found across different studies using the same model. Detailed comparisons are shown in Table 2 and Figure 5A. We used Cohen's kappa, which also considers agreement due to chance (McHugh, 2012), to quantify concordance across the different studies for the 9 shared FRs (Figure 5B). As expected, we found perfect agreement between the two nematode studies and our planarian study. However, all other comparisons yielded substantial agreement or only agreement due to chance, even when comparing different studies in the same system, demonstrating that interlab methodological differences can be as impactful as interspecies differences. This emphasizes the importance of transparency, so results from different studies can be adequately compared and compiled to leverage the strength of multi-system testing.

We also compared the sensitivity (i.e. chemical potency) of our regenerating planarian screening data to those of the comparative alternative model studies (Figure 5C, Table 2). Notably, some of these studies, including ours herein, evaluated toxicity potency using the nominal test concentrations (i.e. LEL) (Behl et al., 2016; Jarema et al., 2015; Noyes et al., 2015) while others used modeling approaches to calculate a point-of-departure (POD) (Behl et al., 2015). Eight of the 10 FRs causing adverse effects in planarians (BDE-47, TBBPA, BPDP, EHDP, IPP, TMPP, TPHP, and TCPP) had overall LELs (from any endpoint) in the regenerating planarian system within an order of magnitude of the effect concentration (LEL or POD) in at least one of the other studies using zebrafish, nematode, and *in vitro* cell-based models (Figure 5C, Table 2).

4. Discussion

4.1. Robustness of planarian screening platform

Robustness of screening data is indispensable for building confidence in a new testing platform. In previous proof-of-concept work (Hagstrom et al., 2015), we used post-hoc statistical power analysis to determine that n=24 specimen per condition were necessary

to achieve sufficient statistical power. Here, we experimentally verified the validity of our current test strategy, testing in 3 replicate runs with n=8 each, through direct comparison of the reproducibility of 3 (n=24), 4 (n=32), or 5 replicates (n=40) with 6 replicates (n=48). We found that our planarian platform is robust and that 3 replicates were sufficient to identify all chemical hit-calls, distinguishing the same bioactive chemicals as 6 replicates. Moreover, readout bioactivity and sensitivity (i.e. LELs) of the different endpoints were highly concordant among the different sets of replicates. Loss of bioactive readouts with less replicates was only found for morphological or behavioral endpoints at lethal concentrations. Since all endpoints, except lethality, were quantified from the data from living animals, high lethality led to a small sample size and limited the power to detect a statistically significant effect. Moreover, lethality suggests overt systemic toxicity at these concentrations which may have masked effects at more specific endpoint readouts. Given the overt toxicity at lethal concentrations, many labs only evaluate more sensitive morphological or behavioral effects at sublethal concentrations, disregarding effects at lethal concentrations (Alzualde et al., 2018; Jarema et al., 2015; Truong et al., 2014). All together, these data indicate that different number of replicates (from 3–6) yield very similar results, validating that our current screening strategy using 3 replicates can robustly identify chemical bioactivity and different toxicities.

4.2. The role of in vitro HTS for DNT

Few mammalian guideline studies have been conducted to evaluate the DNT of chemicals, including the FRs tested herein. High-throughput *in vitro* assays, coupled with bioinformatic analysis, have emerged to fill this gap by quickly and efficiently evaluating potential chemical toxicity. However, our comparison of several human and rodent cell culture assays with non-mammalian whole animal models found that the alternative animal models were better at identifying specific developmental or developmental neurotoxic effects at sublethal concentrations (Table 2). When considering all compared *in vitro* assays, 8/9 FRs were found to be bioactive; however, 6 of these bioactive FRs were only identified in 1 or 2 assays (primarily the human stem cell systems). The two rodent cell culture assays only yielded 2 hits each, both of which were also detected by the human stem cell assays. Most of the *in vitro* effects were found concomitantly with cytotoxicity, while selective sublethal effects were more apparent in the alternative animal models. A recent study evaluating a small number of BFRs and OPFRs in rat neural stem cells and neuronotypic PC12 cells was able to distinguish effects on neurodifferentiation from cytotoxicity (Slotkin et al., 2017), suggesting that certain types of cell culture assays may be better suited for DNT studies. Thus, while current *in vitro* techniques are extremely powerful for rapid screening of toxicants on other organ systems, they may have limited value for detecting potential developmental toxicity or DNT, because they inherently lack the necessary complexity for detecting functional perturbations that are independent of systemic toxicity. Alternative whole animal systems present an opportunity to fill this gap between *in vitro* and mammalian animal models.

4.3. Need for transparency in screening methodology and analysis

As shown in Figure 5B, intraspecies differences in chemical hit-calls between different studies were as big as interspecies differences. Additionally, even when a chemical

was identified as bioactive in two studies from the same model, sensitivity could vary dramatically. For example, the zebrafish data from (Noyes et al., 2015) shows increased sensitivity (up to 3 orders of magnitude difference) to 4 FRs (TBBPA, BPDP, IDDP, and IPP) when compared to the other published zebrafish studies (Table 2).

Important methodological differences in experimental design and data analysis may explain these differences. For example, we tested the FRs in log-steps, whereas several of the other studies used smaller concentration steps, thus the large dose gaps in our study limits precise comparisons of potency. In addition, how significant effects are determined can vary dramatically between laboratories. Some use statistics alone, while we and others incorporate additional stringency tests including biological relevancy cutoffs and inconsistency checks (Zhang et al., 2019). Our statistical pipeline is based on empirical observations to reduce the false positive rate but may limit sensitivity (Supplementary Table 1). Variability can also arise due to using LELs vs PODs to determine potency. LELs are limited to the tested concentrations, and thus the respective POD could lie somewhere within the range of the LEL and the next lower tested concentration level. Therefore, because the different experimental approaches and statistical criteria used to evaluate bioactivity and potency are as important as differences between species, transparency of methodology is crucial to make system comparisons meaningful and build confidence in these new approaches for complementing mammalian toxicity testing.

4.4. Comparisons of different models emphasizes the value of a battery approach

To contextualize the results of our planarian screen with what is known for FR toxicity, we found that our results were highly concordant with available data from other alternative, non-mammalian animal models. Eight of 10 common FRs tested in planarians, developing zebrafish, and nematodes were bioactive in at least one study in all three systems (Table 2 and Figure 5A). Thus, these shared bioactive FRs affected multiple models and multiple endpoints, spanning effects on development, neurodevelopment, and general neuroactivity, indicating multiple mechanisms may be involved in their toxicity. Battery approaches such as these are powerful because they allow parallel interrogation of different species and endpoints, from the cellular to organismal level, to integrate multiple lines of evidence and build confidence in preliminary hazard assessment.

For all systems comparisons, it is imperative to understand the differences between nominal water concentrations and internal doses, particularly at the site of action. The toxicity in the alternative animal models discussed herein are based on nominal water/media concentrations and not doses. Since the effective internal concentrations in the animals are unknown, strong conclusions of comparative sensitivity and their relevancy to human exposure are difficult to make.

4.5. Replacement FRs show comparable neurotoxicity to phased out PBDEs

In this screen, we evaluated the neurotoxicity and DNT of 15 FRs (6 BFRs, 8 OPFRs, and 1 BFR mixture) and found that 11 (3 BFRs, 7 OPFRs, and the BFR mixture) caused adverse effects in both adult and regenerating planarians. It should be noted that the 3 inactive BFRs, BDE-153, TBPH, and TBB, have relatively high molecular weights (550–706 g/mol).

Moreover, TBPH and BDE-153 had solubility issues at the highest tested concentration (100 and 50 μM, respectively) (Zhang et al., 2019). Thus, these issues could have contributed to their inactivity in our planarian screen. TCEP, the only OPFR inactive in our planarian screen, was also inactive in all studies (*in vivo* and *in vitro*) except for 2 zebrafish studies (Table 2) (Jarema et al., 2015; Noyes et al., 2015). Only the Jarema et al and Noyes et al studies found TCEP to be bioactive in zebrafish developmental behavioral endpoints, albeit at relatively higher concentrations than the other tested FRs, suggesting relatively low potency. Inactivity in nematodes was likely not an issue of limited test concentrations as TCEP was tested at up to 1mM without any observed effects. In agreement with the results in worms, studies with Long-Evans rats found no significant DNT was induced by oral exposure (gavage) of up to 90 mg/kg/day TCEP from gestational day 10 to weaning (Moser et al., 2015).

Several currently used FRs, including one in-use BFR, were found to be similarly or more toxic than the phased-out PBDEs in planarians. TBBPA, a BFR which is still currently widely in use, was found to cause specific, sublethal effects on unstimulated behavior (day 7 and 12) and eye regeneration in regenerating planarians (Figure 3) and was more potent than the phased-out BDE-47 and −99. In addition, effects on day 12 unstimulated behavior were selective to regenerating planarians. Together, these results suggest TBBPA caused selective adverse effects on planarian development. Although guideline animal studies have not demonstrated any teratogenic effects of TBBPA (EFSA, 2011), developmental defects were observed in several alternative models compared here, including planarians, zebrafish (Noyes et al., 2015), and nematodes (Behl et al., 2016, 2015) at concentrations of 10 μM and below, suggesting conserved mechanisms of toxicity which may need to be re-examined for human risk assessment. In addition, in mice, TBBPA has been reported to cause neurobehavioral defects after acute exposure and to act as a thyroid endocrine disruptor (Nakajima et al., 2009; Viberg and Eriksson, 2011).

Similarly, almost all the tested OPFRs, which were marketed as safer alternatives to BFRs, were found to be at least as toxic as the phased-out BDEs in planarians. Our results resonate with previous studies in nematode, zebrafish and human, mouse, and rat cell culture, which similarly found that most OPFRs had comparable activity to BFRs (Alzualde et al., 2018; Behl et al., 2016, 2015; Jarema et al., 2015; Noyes et al., 2015). Thus, by assaying various neuronal functions through a range of behavioral endpoints, our planarian screen provides novel insight into the distinct toxicities of OPFRs compared to BFRs. OPFRs generally caused more sublethal, specific neurotoxic effects whereas BFRs were primarily systemically toxic, as discussed in detail in section 4.6.

4.6. Behavioral profiling in planarians adds screening specificity

The majority (5/7) of the bioactive OPFRs in regenerating and adult planarians showed specific, sublethal effects on scrunching, which were not seen to the same extent in the BFRs. Scrunching is an oscillatory, musculature-driven escape gait in planarians that is induced by specific adverse stimuli such as noxious heat (Cochet-Escartin et al., 2015), but is not a generic response to all noxious environments (Cochet-Escartin et al., 2016). Although scrunching shares qualitative features, such as muscle-based motion and increased

mucus secretion, with other muscle-driven body shape changes, such as peristalsis, it can be quantitatively distinguished using four characteristic parameters: frequency and asymmetry of body length oscillations, speed, and maximum elongation (Cochet-Escartin et al., 2015). Scrunching is also distinctively different from other body undulations described in the literature, including snake-like behaviors and seizures (Ross et al., 2018).

Scrunching defects, as tested in our HTS platform, therefore point to specific adverse effects on one or multiple molecular targets controlling the underlying signaling cascade from stimulus sensation to muscle contraction, mucus secretion, and locomotion. Importantly, some of the molecular targets that mediate scrunching have been previously identified. For instance, it has been shown that the Transient Receptor Potential Channel Ankyrin 1 (TRPA1) is required for proper scrunching in response to tail amputation (Arenas et al., 2017). TRPA1 is also an indirect sensor of noxious heat in planarians (Arenas et al., 2017) and we have recently confirmed that TRPA1 activation is required for heat-induced scrunching (unpublished results). Additionally, cholinergic neurotransmission, which is responsible for coordinating planarian motor functions (Nishimura et al., 2012), is likely important for scrunching as this gait relies on muscle contraction. Organophosphorus compounds share the ability to inhibit cholinesterase, though their potency can vary due to structural and pharmacokinetic differences (Taylor, 2018). Because comparable concentrations of organophosphorus pesticides, such as parathion and chlorpyrifos (Zhang et al., 2019), also cause scrunching defects, this phenotype is likely representative of the shared cholinergic toxicity. Consistent with this idea, we found loss of cholinesterase activity in day 12 adult planarians was correlated with scrunching defects by comparing activity after exposure to 10 μM EHDP or 100 μM TCEP, the only inactive OFPR in our planarian screen (Figure 4). Similarly, DNT of 4 OPFRs tested in medaka was correlated with changes in acetylcholinesterase activity. For example, TPHP caused developmental and behavioral defects concomitant with significant acetylcholinesterase inhibition whereas TCEP, which had no significant effect on acetylcholinesterase activity, had much milder toxic effects (Sun et al., 2016a). Lack of significant effects of TCEP on acetylcholinesterase activity or transcription were also found in developing zebrafish (Sun et al., 2016b). Together, these studies indicate that the inactivity of TCEP in these different models may result from its inability to significantly inhibit acetylcholinesterase at the tested concentrations.

Our behavior-rich planarian system can be a powerful complement to existing developing zebrafish and nematode models. While these existing alternative animal models provide morphological breadth and genetic tractability, respectively, the planarian model and specifically our HTS platform provide an unmet diversity of behavioral endpoints. As demonstrated with this screen, we can phenotypically and mechanistically distinguish different chemical classes by interrogating different neuronal functions in an automated, HT manner. Additionally, because we can directly compare toxicities in regenerating and adult planarians, we are able to distinguish toxic from developmental toxic effects, which is not directly possible in these other systems. For example, we found developmentally selective effects on day 12 unstimulated behavior following exposure to 10μM TBBPA in regenerating but not adult planarians, suggesting TBBPA may act on specific developmental processes or on processes which are more sensitive during development. The shared effects and potencies of the other bioactive FRs on both regenerating and adult planarians suggest

a potential shared mechanism of toxicity, with effects on overall neural function rather than neurodevelopment directly. Together, these examples illustrate that, in a single screen, the planarian platform is uniquely able to differentiate DNT from neurotoxicity and distinguish effects on specific neuronal functions.

5. Conclusions

We showed, using the concrete example of screening 15 FRs, that our planarian HTS platform generates robust data and allows us to distinguish systemic toxicity from neurotoxicity and developmentally selective effects from general neurotoxicity. Additionally, the diverse behaviors tested in this planarian system allowed us to differentiate class-dictated phenotypic signatures, particularly evidenced by the shared effect of OPFRs, but not BFRs, on planarian scrunching. We showed that this behavioral phenotype is correlated with decreased cholinesterase activity, thus linking molecular mechanisms to organismal outcomes. Together, these data provide novel insight into the toxicity profiles of these compounds. Concordance between this planarian study and previous studies in cell culture, nematodes and developing zebrafish, demonstrated the comparable toxicity of replacement OPFRs with BFRs and emphasizes the urgent need for further evaluation of OPFRs in mammalian systems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

8. Funding information

Abbreviations

DNT	developmental neurotoxicity
FR	flame retardant
BFR	brominated flame retardant
PBDE	Polybrominated Diphenyl Ether
OPFR	organophosphorus flame retardant
NTP	National Toxicology Program
DMSO	dimethyl sulfoxide
HTS	high-throughput screening

| LEL | lowest effect level |
| HTTK | high-throughput toxicokinetics |
| POD | point of departure |

9. References

Alzualde A, Behl M, Sipes NS, Hsieh J-H, Alday A, Tice RR, Paules RS, Muriana A, Quevedo C, 2018. Toxicity profiling of flame retardants in zebrafish embryos using a battery of assays for teratogenicity, behavior, cardiotoxicity, and hepatotoxicity toward human relevance. Neurotoxicol. Teratol 10.1016/j.ntt.2018.10.002

Arenas OM, Zaharieva EE, Para A, Vásquez-Doorman C, Petersen CP, Gallio M, 2017. Activation of planarian TRPA1 by reactive oxygen species reveals a conserved mechanism for animal nociception. Nat. Neurosci 20, 1686–1693. 10.1038/s41593-017-0005-0 [PubMed: 29184198]

Bailey JM, Levin ED, 2015. Neurotoxicity of FireMaster 550® in zebrafish (Danio rerio): Chronic developmental and acute adolescent exposures. Neurotoxicol. Teratol 52, 210–219. 10.1016/j.ntt.2015.07.001 [PubMed: 26239867]

Bal-Price AK, Coecke S, Costa L, Crofton KM, Fritsche E, Goldberg A, Grandjean P, Lein PJ, Li A, Lucchini R, Mundy WR, Padilla S, Persico AM, Seiler AEM, Kreysa J, 2012. Advancing the science of developmental neurotoxicity (DNT): testing for better safety evaluation. ALTEX 29, 202–15. [PubMed: 22892558]

Behl M, Hsieh JH, Shafer TJ, Mundy WR, Rice JR, Boyd WA, Freedman JH, Hunter ES, Jarema KA, Padilla S, Tice RR, 2015. Use of alternative assays to identify and prioritize organophosphorus flame retardants for potential developmental and neurotoxicity. Neurotoxicol. Teratol 10.1016/j.ntt.2015.09.003

Behl M, Rice JR, Smith MV, Co CA, Bridge MF, Hsieh J-H, Freedman JH, Boyd WA, 2016. Comparative toxicity of organophosphate flame retardants and polybrominated diphenyl ethers to Caenorhabditis elegans. Toxicol. Sci 154, 241–252. 10.1093/toxsci/kfw162 [PubMed: 27566445]

Behl M, Ryan K, Hsieh J-H, Parham F, Shapiro AJ, Collins BJ, Sipes NS, Birnbaum LS, Bucher JR, Foster PMD, Walker NJ, Paules RS, Tice RR, 2019. Screening for Developmental Neurotoxicity at the National Toxicology Program: The Future Is Here. Toxicol. Sci 167, 6–14. 10.1093/toxsci/kfy278 [PubMed: 30496580]

Birkholz TR, Beane WS, 2017. The planarian TRPA1 homolog mediates extraocular behavioral responses to near-ultraviolet light. J. Exp. Biol 220, 2616–2625. 10.1242/jeb.152298 [PubMed: 28495872]

Cano-Sancho G, Smith A, La Merrill MA, 2017. Triphenyl phosphate enhances adipogenic differentiation, glucose uptake and lipolysis via endocrine and noradrenergic mechanisms. Toxicol. Vitr 40, 280–288. 10.1016/j.tiv.2017.01.021

Cebrià F, 2007. Regenerating the central nervous system: how easy for planarians! Dev. Genes Evol 217, 733–748. 10.1007/s00427-007-0188-6 [PubMed: 17999079]

Cebrià F, Kudome T, Nakazawa M, Mineta K, Ikeo K, Gojobori T, Agata K, 2002a. The expression of neural-specific genes reveals the structural and molecular complexity of the planarian central nervous system. Mech. Dev 116, 199–204. 10.1016/S0925-4773(02)00134-X [PubMed: 12128224]

Cebrià F, Nakazawa M, Mineta K, Ikeo K, Gojobori T, Agata K, 2002b. Dissecting planarian central nervous system regeneration by the expression of neural-specific genes. Dev. Growth Differ 44, 135–46. [PubMed: 11940100]

Cochet-Escartin O, Carter JA, Chakraverti-Wuerthwein M, Sinha J, Collins EMS, 2016. Slo1 regulates ethanol-induced scrunching in freshwater planarians. Phys. Biol 13, 1–12. 10.1088/1478-3975/13/5/055001

Cochet-Escartin O, Mickolajczk KJ, Collins E-MS, 2015. Scrunching: a novel escape gait in planarians. Phys. Biol 12, 055001. 10.1088/1478-3975/12/5/056010 [PubMed: 26228740]

Cope RB, Kacew S, Dourson M, 2015. A reproductive, developmental and neurobehavioral study following oral exposure of tetrabromobisphenol A on Sprague-Dawley rats. Toxicology 329, 49–59. 10.1016/J.TOX.2014.12.013 [PubMed: 25523853]

EFSA, 2011. European Food Safety Authority (EFSA) panel on contaminants in the food chain (CONTAM): Scientific opinion on tetrabromobisphenol A (TBBPA) and its derivatives in food. EFSA J 10.2903/j.efsa.2011.2477.

Fritsche E, Grandjean P, Crofton KM, Aschner M, Goldberg A, Heinonen T, Hessel EVS, Hogberg HT, Bennekou SH, Lein PJ, Leist M, Mundy WR, Paparella M, Piersma AH, Sachana M, Schmuck G, Solecki R, Terron A, Monnet-Tschudi F, Wilks MF, Witters H, Zurich M-G, Bal-Price A, 2018. Consensus statement on the need for innovation, transition and implementation of developmental neurotoxicity (DNT) testing for regulatory purposes. Toxicol. Appl. Pharmacol 354, 3–6. 10.1016/j.taap.2018.02.004 [PubMed: 29447839]

Glazer L, Wells CN, Drastal M, Odamah KA, Galat RE, Behl M, Levin ED, 2018. Developmental exposure to low concentrations of two brominated flame retardants, BDE-47 and BDE-99, causes life-long behavioral alterations in zebrafish. Neurotoxicology 66, 221–232. 10.1016/j.neuro.2017.09.007 [PubMed: 28935585]

Hagstrom D, Cochet-Escartin O, Collins E-MS, 2016. Planarian brain regeneration as a model system for developmental neurotoxicology. Regeneration 3, 65–77. 10.1002/reg2.52 [PubMed: 27499880]

Hagstrom D, Cochet-Escartin O, Zhang S, Khuu C, Collins E-MS, 2015. Freshwater planarians as an alternative animal model for neurotoxicology. Toxicol. Sci 147, 270–285. 10.1093/toxsci/kfv129 [PubMed: 26116028]

Hagstrom D, Hirokawa H, Zhang L, Radi Z, Taylor P, Collins E-MS, 2017. Planarian cholinesterase: in vitro characterization of an evolutionarily ancient enzyme to study organophosphorus pesticide toxicity and reactivation. Arch. Toxicol 91, 2837–2847. 10.1007/s00204-016-1908-3 [PubMed: 27990564]

Hagstrom D, Truong L, Zhang S, Tanguay R, Collins E-MS, 2019. Comparative analysis of zebrafish and planarian model systems for developmental neurotoxicity screens using an 87-compound library. Toxicol. Sci 167, 15–25. 10.1093/toxsci/kfy180 [PubMed: 30011007]

Hagstrom D, Zhang S, Ho A, Tsai ES, Radi Z, Jahromi A, Kaj KJ, He Y, Taylor P, Collins EMS, 2018. Planarian cholinesterase: molecular and functional characterization of an evolutionarily ancient enzyme to study organophosphorus pesticide toxicity. Arch. Toxicol 92, 1161–1176. 10.1007/s00204-017-2130-7 [PubMed: 29167930]

Hendriks HS, Westerink RHS, 2015. Neurotoxicity and risk assessment of brominated and alternative flame retardants. Neurotoxicol. Teratol 52, 248–269. 10.1016/J.NTT.2015.09.002 [PubMed: 26363216]

Inoue T, Yamashita T, Agata K, 2014. Thermosensory signaling by TRPM is processed by brain serotonergic neurons to produce planarian thermotaxis. J. Neurosci 34, 15701–14. 10.1523/JNEUROSCI.5379-13.2014 [PubMed: 25411498]

Jarema KA, Hunter DL, Shaffer RM, Behl M, Padilla S, 2015. Acute and developmental behavioral effects of flame retardants and related chemicals in zebrafish. Neurotoxicol. Teratol 10.1016/j.ntt.2015.08.010

Masjosthusmann S, Barenys M, El- Gamal M, Geerts L, Gerosa L, Gorreja A, Kühne B, Marchetti N, Tigges J, Viviani B, Witters H, Fritsche E, 2018. Literature review and appraisal on alternative neurotoxicity testing methods. EFSA Support. Publ. 15 10.2903/sp.efsa.2018.EN-1410

McHugh ML, 2012. Interrater reliability: the kappa statistic. Biochem. medica 22, 276–82.

Meeker JD, Stapleton HM, 2010. House dust concentrations of organophosphate flame retardants in relation to hormone levels and semen quality parameters. Environ. Health Perspect 118, 318–23. 10.1289/ehp.0901332 [PubMed: 20194068]

Moser VC, Phillips PM, Hedge JM, McDaniel KL, 2015. Neurotoxicological and thyroid evaluations of rats developmentally exposed to tris(1,3-dichloro-2-propyl)phosphate (TDCIPP) and tris(2-chloro-2-ethyl)phosphate (TCEP). Neurotoxicol. Teratol 52, 236–247. 10.1016/J.NTT.2015.08.004 [PubMed: 26300399]

Nakajima A, Saigusa D, Tetsu N, Yamakuni T, Tomioka Y, Hishinuma T, 2009. Neurobehavioral effects of tetrabromobisphenol A, a brominated flame retardant, in mice. Toxicol. Lett 189, 78–83. 10.1016/j.toxlet.2009.05.003 [PubMed: 19463927]

National Toxicology Program, 1991. Toxicology and Carcinogenesis Studies of Tris(2-chloroethyl) Phosphate (CAS No. 115–96-8) in F344/N Rats and B6C3F1 Mice (Gavage Studies). Natl. Toxicol. Program Tech. Rep. Ser 391, 1–233. [PubMed: 12637968]

Nishimura K, Kitamura Y, Taniguchi T, Agata K, 2010. Analysis of motor function modulated by cholinergic neurons in planarian Dugesia japonica. Neuroscience 168, 18–30. 10.1016/j.neuroscience.2010.03.038 [PubMed: 20338223]

Nishimura O, Hirao Y, Tarui H, Agata K, 2012. Comparative transcriptome analysis between planarian Dugesia japonica and other platyhelminth species. BMC Genomics 13, 289. [PubMed: 22747887]

Noyes PD, Haggard DE, Gonnerman GD, Tanguay RL, 2015. Advanced morphological-behavioral test platform reveals neurodevelopmental defects in embryonic zebrafish exposed to comprehensive suite of halogenated and organophosphate flame retardants. Toxicol. Sci 145, 177–195. 10.1093/toxsci/kfv044 [PubMed: 25711236]

Oliveri AN, Bailey JM, Levin ED, 2015. Developmental exposure to organophosphate flame retardants causes behavioral effects in larval and adult zebrafish. Neurotoxicol. Teratol 52, 220–227. 10.1016/j.ntt.2015.08.008 [PubMed: 26344674]

Ross KG, Currie KW, Pearson BJ, Zayas RM, 2017. Nervous system development and regeneration in freshwater planarians. Wiley Interdiscip. Rev. Dev. Biol 6, e266. 10.1002/wdev.266

Ross KG, Molinaro AM, Romero C, Dockter B, Cable KL, Gonzalez K, Zhang S, Collins EMS, Pearson BJ, Zayas RM, 2018. SoxB1 Activity Regulates Sensory Neuron Regeneration, Maintenance, and Function in Planarians. Dev. Cell 47, 331–347.e5. 10.1016/j.devcel.2018.10.014 [PubMed: 30399335]

Slotkin TA, Skavicus S, Stapleton HM, Seidler FJ, 2017. Brominated and organophosphate flame retardants target different neurodevelopmental stages, characterized with embryonic neural stem cells and neuronotypic PC12 cells. Toxicology 390, 32–42. 10.1016/j.tox.2017.08.009 [PubMed: 28851516]

Smirnova L, Hogberg HT, Leist M, Hartung T, 2014. Food for Thought …: Developmental Neurotoxicity – Challenges in the 21(st) Century and In Vitro Opportunities. ALTEX 31, 129–156. [PubMed: 24687333]

Stapleton HM, Klosterhaus S, Eagle S, Fuh J, Meeker JD, Blum A, Webster TF, 2009. Detection of Organophosphate Flame Retardants in Furniture Foam and U.S. House Dust. Environ. Sci. Technol 43, 7490–7495. 10.1021/es9014019 [PubMed: 19848166]

Stapleton HM, Misenheimer J, Hoffman K, Webster TF, 2014. Flame retardant associations between children's handwipes and house dust. Chemosphere 116, 54–60. 10.1016/J.CHEMOSPHERE.2013.12.100 [PubMed: 24485814]

Sun L, Tan H, Peng T, Wang S, Xu W, Qian H, Jin Y, Fu Z, 2016a. Developmental neurotoxicity of organophosphate flame retardants in early life stages of Japanese medaka (Oryzias latipes). Environ. Toxicol. Chem 35, 2931–2940. 10.1002/etc.3477 [PubMed: 27146889]

Sun L, Xu W, Peng T, Chen H, Ren L, Tan H, Xiao D, Qian H, Fu Z, 2016b. Developmental exposure of zebrafish larvae to organophosphate flame retardants causes neurotoxicity. Neurotoxicol. Teratol 55, 16–22. 10.1016/J.NTT.2016.03.003 [PubMed: 27018022]

Taylor P, 2018. Anticholinesterase agents, in: Brunton Laurence L (Ed.), Goodman and Gilman's The Pharmacological Basis of Therapeutics McGraw Hill Education, San Francisco, pp. 163–176.

Tohyama C, 2016. Developmental neurotoxicity test guidelines: problems and perspectives. J. Toxicol. Sci 41, SP69–SP79. 10.2131/jts.41.SP69 [PubMed: 28250285]

Truong L, Reif DM, St Mary L, Geier MC, Truong HD, Tanguay RL, 2014. Multidimensional in vivo hazard assessment using zebrafish. Toxicol. Sci 137, 212–33. [PubMed: 24136191]

Tung EWY, Ahmed S, Peshdary V, Atlas E, 2017. Firemaster® 550 and its components isopropylated triphenyl phosphate and triphenyl phosphate enhance adipogenesis and transcriptional activity of peroxisome proliferator activated receptor (Pparγ) on the adipocyte protein 2 (aP2) promoter. PLoS One 12, e0175855. 10.1371/journal.pone.0175855 [PubMed: 28437481]

Usenko C, Abel E, Hopkins A, Martinez G, Tijerina J, Kudela M, Norris N, Joudeh L, Bruce E, 2016. Evaluation of Common Use Brominated Flame Retardant (BFR) Toxicity Using a Zebrafish Embryo Model. Toxics 4, 21. 10.3390/toxics4030021

Van Der Veen I, De Boer J, 2012. Phosphorus flame retardants: Properties, production, environmental occurrence, toxicity and analysis. Chemosphere 88, 1119–1153. 10.1016/j.chemosphere.2012.03.067 [PubMed: 22537891]

Viberg H, Eriksson P, 2011. Differences in neonatal neurotoxicity of brominated flame retardants, PBDE 99 and TBBPA, in mice. Toxicology 289, 59–65. 10.1016/j.tox.2011.07.010 [PubMed: 21820030]

Zhang S, Hagstrom D, Hayes P, Graham A, Collins E-MS, 2019. Multi-behavioral endpoint testing of an 87-chemical compound library in freshwater planarians. Toxicol. Sci 167, 26–44. 10.1093/toxsci/kfy145 [PubMed: 29893936]

Zheng D-M, Xie H-Q, Wang A-T, Wu C-C, 2011. The nerve system identificiation by histochemical localization of acetylcholinesterase in planarian Dugesia japonica. Chinese J. Zool 45, 68–75.
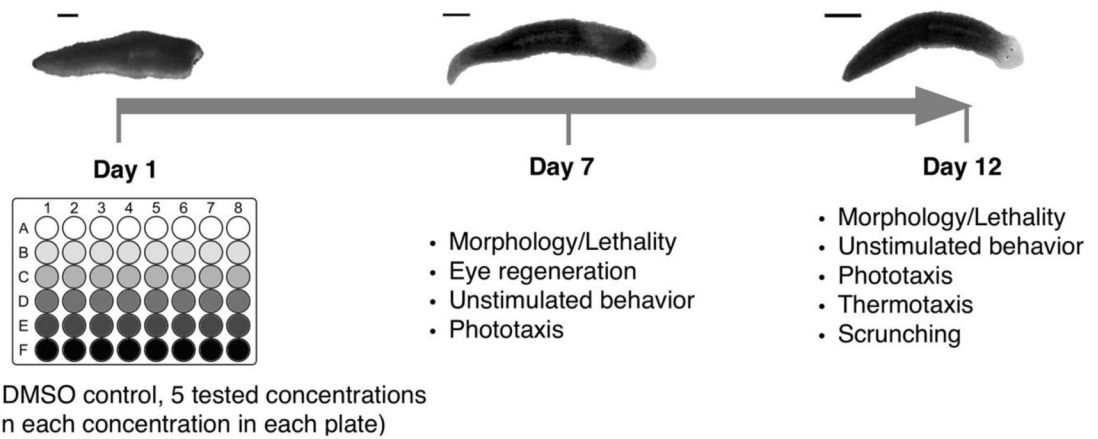
Figure 1.

Schematic of overall screen flow in the planarian system. Exposure lasted 12 days. Amputated planarians (amputation at day 1) and adult planarians are loaded into 48-well plates (one worm per well, and one worm type per plate) with the test chemicals at day 1 and assayed for different morphological and behavioral endpoints at day 7 and day 12. Scale bar: 1mm.
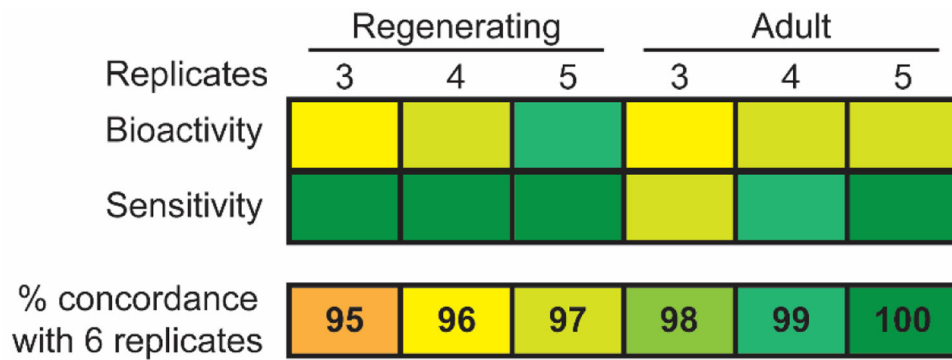
Figure 2.
Concordance analysis of readouts for 3, 4, or 5 replicates compared to 6 replicates. A. Readout bioactivity concordance (top) was determined as the percentage of readouts showing the same bioactivity (bioactive or inactive) as using 6 replicates, out of the total number of readouts (9 for regenerating and 8 for adult planarians) for all 16 chemicals, for a total of 144 and 128 readouts for regenerating and adult planarians, respectively. Sensitivity concordance (bottom) of the concordant bioactive readouts was determined as the number of bioactive readouts with the same LEL as in 6 replicates, out of the total number of concordant bioactive readouts (data from top) for all 16 chemicals in this set of compiled data. IPP and TMPP were excluded from the analysis.
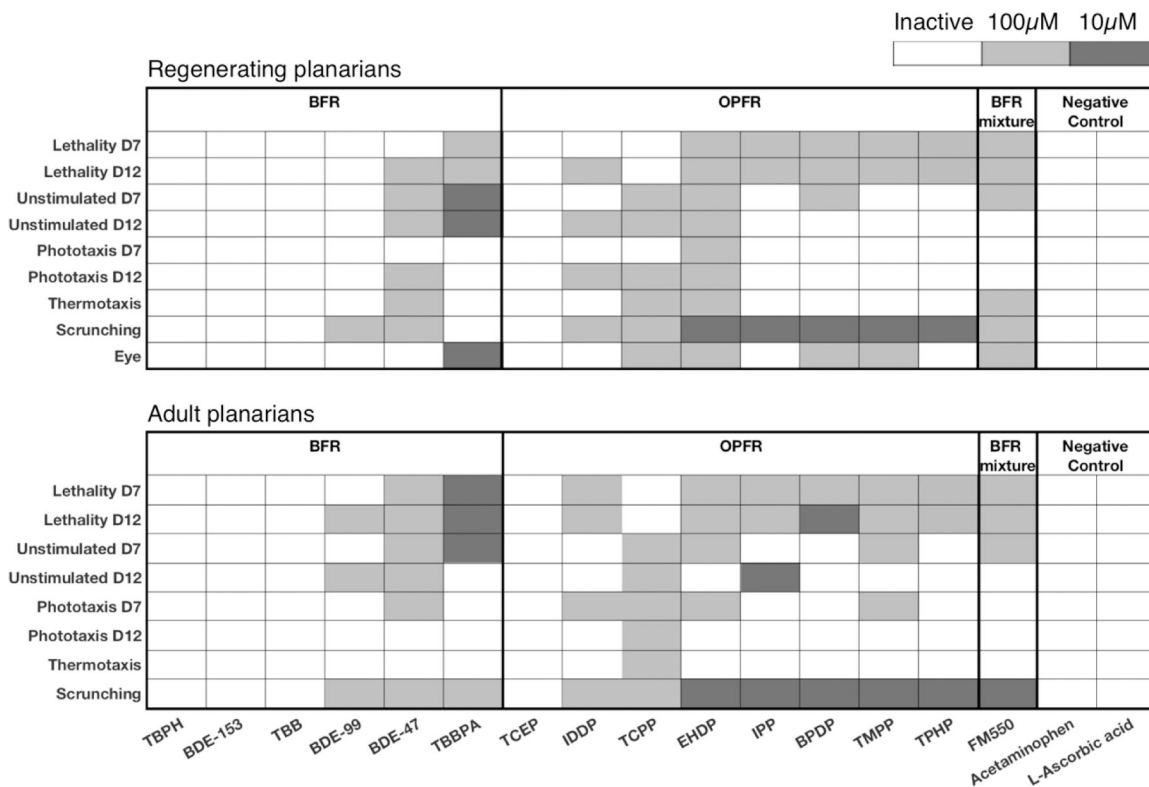
Figure 3.
Overview of the planarian screening data of 15 unique flame retardants and 2 negative controls using 6 replicates. Heat maps of effects of brominated flame retardants (BFRs), organophosphorus flame retardants (OPFRs), a BFR mixture, and negative controls on regenerating (Top) and adult (Bottom) planarians for all endpoints with lowest-effect-level shaded. TPHP was screened as a duplicate in FR screen 2, but only shown once here since the results were consistent. Note that BDE-153 was screened at a maximum concentration of 50 μM.
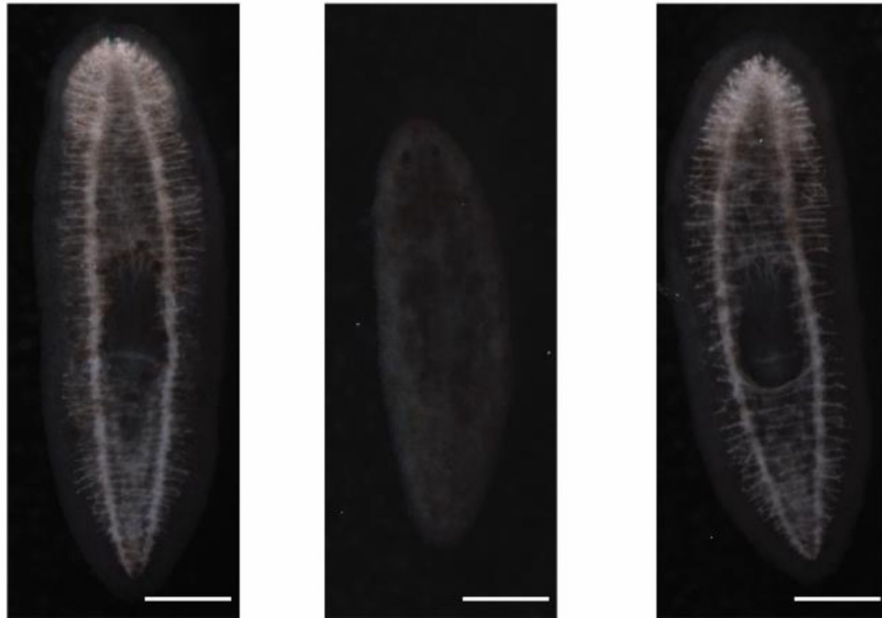
Figure 4.
Significant loss of cholinesterase activity in planarians exposed to EHDP but not TCEP. Representative images (out of 12 or 6 samples for 0.5% DMSO and the FRs, respectively, from two independent experiments) are shown of adult planarians incubated in the respective chemical concentrations for 12 days and then fixed and stained to visualize catalysis of acetylthiocholine (Hagstrom et al., 2018, 2017; Zheng et al., 2011). Note that almost no activity, seen as the light stain in the nervous system, is seen in the EHDP-exposed planarian. Anterior of the worms (brain) is to the top of the image. Scale bars: 0.5 mm.
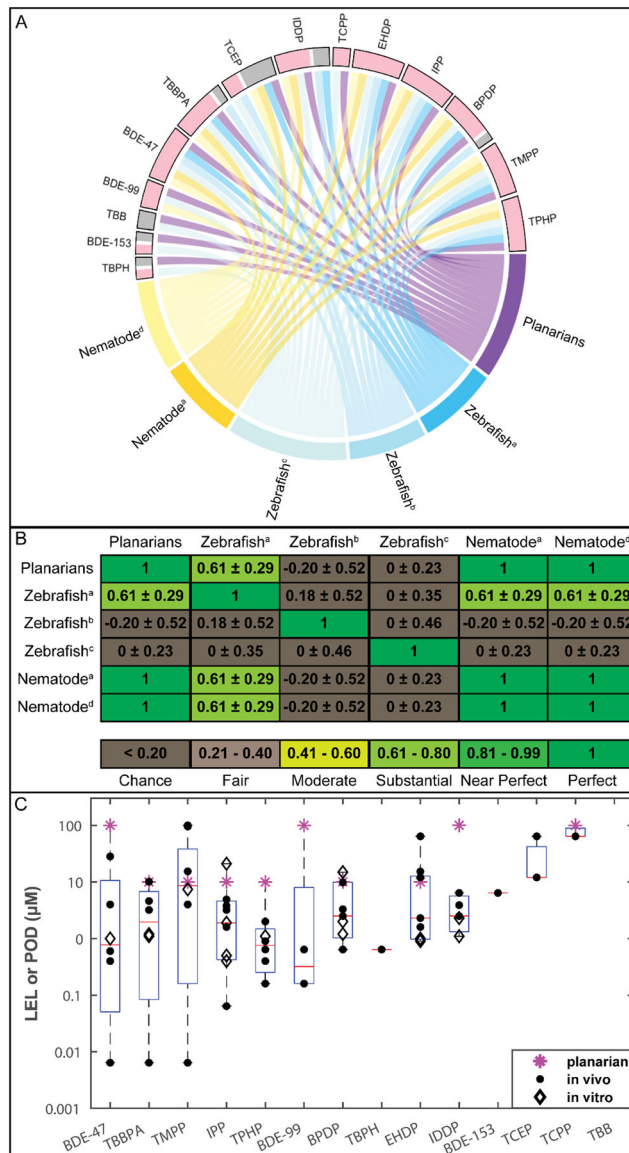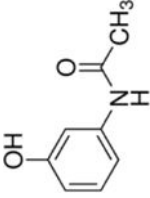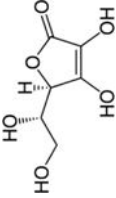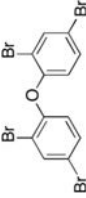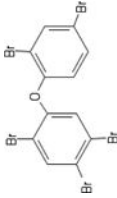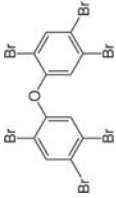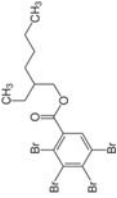
Figure 5.
Concordance between different alternative models. A. Interaction of bioactivity of the
14 flame retardants (FRs) with the different tested alternative animal models. Each box
represents the interactions of this FR with the different models and studies. Active
interactions are in pink. Inactive interactions are in grey. FRs with both active and
inactive interactions showed different effects in different studies/models. B. Quantification
of concordance between alternative animal models for 9 FRs screened in all alternative
animal models using Cohen's kappa (κ), listed as $\kappa \pm 95\%$ confidence intervals (McHugh,
2012). the Noyes et al. study showed a kappa of 0 for all comparisons because it had no
inactives in this set of 9 FRs. For (A) and (B), studies are listed as: a, data from (Behl et
al., 2015); b, data from (Jarema et al., 2015); c, data from (Noyes et al., 2015); d, data from
(Behl et al., 2016). C. Comparison of toxicity potency of 14 FRs in regenerating planarians
and other alternative developing models. The data for the individual studies are overlaid on
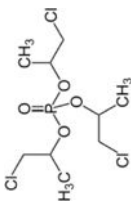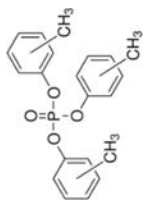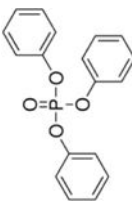
a box and whisker plot compiling the lowest-effect-levels (LELs) and points-of-departure (PODs) for the different FRs in all compared models. Stars represent planarian LELs, filled dots represent LELs/PODs in developing zebrafish and nematode models (Behl et al., 2016, 2015; Jarema et al., 2015; Noyes et al., 2015), and diamonds represent PODs in *in vitro* cell-based models (Behl et al., 2015). The x-axis shows the FRs in decreasing order of potency (lowest effect concentration).

Table 1.

Summary of the screened chemical library with CAS number, chemical name, ID, type (BFR: brominated flame retardant, OPFR: organophosphorus flame retardant), chemical supplier, structure, and purity. When available, the determined purity (by the NTP) is provided, else the manufacturer supplied purity is provided. NA: not available.

CASRN	Chemical name	ID	Type	Supplier	Structure	Purity
103-90-2	Acetaminophen (4-hydroxyacetanilide)		Negative control	Sigma-Aldrich		100%
50-81-7	L-Ascorbic acid		Negative control	Sigma-Aldrich		100%
5436-43-1	2,2',4,4'-Tetrabromodiphenyl ether	BDE-47	BFR	Cerilliant Corp.		100%
60348-60-9	2,2',4,4',5-Pentabromodiphenyl ether	BDE-99	BFR	Cerilliant Corp.		100%
68631-49-2	2,2',4,4',5,5'-Hexabromodiphenyl ether	BDE-153	BFR	Cerilliant Corp. via Battelle Memorial Institute		100%
183658-27-7	2-ethylhexyl-2,3,4,5-tetrabromobenzoate	TBB	BFR	Toronto Research Chemicals		NA

CASRN	Chemical name	ID	Type	Supplier	Structure	Purity
79-94-7	3,3′,5,5′-Tetrabromobisphenol A	TBBPA	BFR	Sigma-Aldrich		99.77%
26040-51-7	Bis(2-ethylhexyl) 3,4,5,6-tetrabromophthalate	TBPH	BFR	Toronto Research Chemicals		NA
860302-33-6	Firemaster 550	FM550	BFR mixture [†]	Chemtura Corporation		NA
56803-37-3	tert-Butylphenyl diphenyl phosphate	BPDP	OPFR	Ubichem PLC via MRIGlobal		98.22%
1241-94-7	2-ethylhexyl diphenyl phosphate	EHDP	OPFR	TCI America		98.67%
29761-21-5	Isodecyl diphenyl phosphate	IDDP	OPFR	Bayville Chemical Supply Company Inc.		92.19%
68937-41-7	Phenol, isopropylated, phosphate (3:1)	IPP	OPFR	Ameribrom Inc. (screen 1) Amfinecom Inc. (screen 2)		NA
115-96-8	Tris(2-chloroethyl) phosphate	TCEP	OPFR	Sigma-Aldrich		99.07%

CASRN	Chemical name	ID	Type	Supplier	Structure	Purity
13674-84-5	tris(2-Chloroisopropyl)phosphate	TCPP	OPFR	Sigma-Aldrich (screen 1) Acros Organics (screen 2)		99.47%
1330-78-5	Tricresyl phosphate	TMPP	OPFR	Albemarle Corporation via MRIGloba		100%
115-86-6	Triphenyl phosphate *	TPHP	OPFR	Sigma-Aldrich		100%

†Firemaster 550 is a chemical mixture, comprised of 2-ethylhexyl-2,3,4,5-tetrabromobenzoate (TBB), bis(2-ethylhexyl) tetrabromophthalate (TBPH), triphenyl phosphate (TPP), and isopropylated triphenyl phosphate (IPTP) (Tung et al., 2017)

*TPHP was provided as a duplicate in FR screen 2. Duplicate vials were from the same supplier with the same lot numbers.

Table 2.

Comparison of developmental and/or developmentally neurotoxic endpoints used to evaluate FR toxicity between regenerating planarians and different alternative developing models. Lowest-effect-level (LEL) or point-of-departure (POD) for each model are listed in μM. Dash (−) indicates there was no toxicity detected at the tested concentration. An empty cell indicates this chemical was not tested in this model. The number in brackets represents the LELs inclusive of concentration-independent and hyper-active effects. Bold text indicates sublethal effects.

Model	Assay	BFRs							OPFRs						
		TBPH	BDE-153	TBB	BDE-99	BDE-47	TBBPA	TCEP	IDDP	TCPP	EHDP	IPP	BPDP	TMPP	TPHP
Planarian	Morphology and and/or behavior (LEL)	-	-	-	**100**	100	**10**	-	100.5	**100**	**10**	**10.05**	**10.05**	**10**	**9.95**
Zebrafish	Embryonic development (POD)[a]					28.2	4.6	-	-		15.3	**4.9**	**9.8**	**15.4**	**2**
	Developmental behavior (LEL)[b]					**4**	-	**12**	-		**12** [2.1]	**3.8** [1.2]	-	**4**	**0.4**
	Developmental malformations and/or behavior (LEL)[c]	**0.64** [6.4E-4]	**6.4** [0.064]	−[6.4E-4]	**0.64** [0.0064]	0.0064	**0.0064**	**64** [0.0064]	**6.4** [0.0064]	**64** [0.64]	**64** [0.0064]	0.064	**0.64** [0.0064]	**0.0064** [0.0064]	**0.64** [0.0064]
Nematode	Larval development (POD)[a]					**0.6**	**10.1**	-	**3.9**		**2.3**	**3.2**	**3.3**	**95.5**	**0.9**
	Larval development and/or feeding (LEL)[d]				**0.16**	**0.4**	**3.2**	-	**2.5**		**1.6**	**1.6**	**2.5**	**100**	**0.16**
In vitro	Mouse stem cell differentiation (POD)[a]					-	1.1	-	-		-	0.4	-	-	-
	Human stem cell neuroprogenitor proliferation (POD)[a]					-	-	-	2.3		1	1.9	2	7.4	-
	Rat neurite outgrowth (POD)[a]					-	-	-	-		-	21	14.9	-	-

Model	Assay	BFRs						OPFRs							
		TBPH	BDE-153	TBB	BDE-99	BDE-47	TBBPA	TCEP	IDDP	TCPP	EHDP	IPP	BPDP	TMPP	TPHP
	Human stem cell neurite outgrowth (POD) [a]					1	1.2	-	1.1		0.9	0.5	1.2	-	1.1

[a], data from (Behl et al., 2015)

[b], data from (Jarema et al., 2015)

[c], data from (Noyes et al., 2015)

[d], data from (Behl et al., 2016)