Title

An Integrated Approach Using Publicly Available Resources for Identifying and Characterizing Chemicals of Potential Toxicity Concern: Proof-of-Concept With Chemicals That Affect Cancer Pathways.

Journal

Toxicological Sciences, 169(1)

ISSN

1096-6080

Authors

Iyer, Shoba
Pham, Nathalie
Marty, Melanie
et al.

Publication Date

2019-05-01

DOI

10.1093/toxsci/kfz017

Peer reviewed

OXFORD

SOT | Society of Toxicology
www.toxsci.oxfordjournals.org

TOXICOLOGICAL SCIENCES, 2019, 1–11

doi: 10.1093/toxsci/kfz017
Advance Access Publication Date: January 12, 2019
Research Article

An Integrated Approach Using Publicly Available Resources for Identifying and Characterizing Chemicals of Potential Toxicity Concern: Proof-of-Concept With Chemicals That Affect Cancer Pathways

Shoba Iyer,*,[1] Nathalie Pham,† Melanie Marty,† Martha Sandy,* Gina Solomon,‡ and Lauren Zeise*

*Office of Environmental Health Hazard Assessment (OEHHA), California Environmental Protection Agency's (CalEPA's), Oakland, California 94612; †Office of Environmental Health Hazard Assessment (OEHHA), California Environmental Protection Agency's (CalEPA's), Sacramento, California 95814; and ‡Department of Medicine, University of California, San Francisco (UCSF), San Francisco, California 94143

[1]To whom correspondence should be addressed at Office of Environmental Health Hazard Assessment (OEHHA), California Environmental Protection Agency's (CalEPA's), Oakland, CA 94612. Fax: (510) 622-3211; E-mail: shoba.iyer@oehha.ca.gov.

Disclaimer: The views expressed are those of the authors and do not necessarily represent those of the Office of Environmental Health Hazard Assessment, the California Environmental Protection Agency, or the State of California.

ABSTRACT

We developed an integrated, modular approach to predicting chemical toxicity relying on in vitro assay data, linkage of molecular targets to disease categories, and software for ranking chemical activity and examining structural features (chemotypes). We evaluate our approach in a proof-of-concept exercise to identify and prioritize chemicals of potential carcinogenicity concern. We identified 137 cancer pathway-related assays from a subset of U.S. EPA's ToxCast platforms. We mapped these assays to key characteristics of carcinogens and found they collectively assess 5 of 10 characteristics. We ranked all 1061 chemicals screened in Phases I and II of ToxCast by their activity in the selected cancer pathway-related assays using Toxicological Prioritization Index software. More chemicals used as biologically active agents (eg, pharmaceuticals) ranked in the upper 50% versus lower 50%. Twenty-three chemotypes are enriched in the top 5% ($n = 54$) of chemicals; these features may be important for their activity in cancer pathway-related assays. The biological coverage of the ToxCast assays related to cancer pathways is limited and short-term assays may not capture the biology of some key characteristics. Metabolism is also minimal in the assays. The ability of our approach to identify chemicals with cancer hazard is limited with the current input data, but we expect that our approach can be applied with future iterations of ToxCast and other data for improved chemical prioritization and characterization. The novel approach and proof-of-concept exercise described here for ranking chemicals for potential carcinogenicity concern is modular, adaptable, and amenable to evolving data streams.

Key words: carcinogen; chemical prioritization; chemotype; new approach methodologies (NAMs); ToxCast; ToxPi.

Over a decade ago, the National Academy of Science (NAS) noted the substantial time and cost required to conduct traditional chemical toxicity testing in animal models and recommended increased use of in vitro methods to more quickly obtain mechanistic chemical information while reducing the reliance on whole animal testing (NAS, 2007). More

recently, the NAS made recommendations to link pathways and disease components to hazard traits using existing knowledge and current research, and to complement data integration with visualization tools (NAS, 2017). These NAS recommendations motivated our present proof-of-concept exercise, which integrates multiple publicly available resources to characterize and rank chemicals of potential toxicity concern. This modular approach relies on three components: (1) existing in vitro assay data; (2) a strategy for linking the molecular targets examined in the assays to disease categories; and (3) software for ranking relevant chemical activity and examining chemical structural features.

We evaluate our approach by demonstrating how it could be used to identify and prioritize over 1000 chemicals for potential carcinogenicity. Several hundred chemicals have been identified as carcinogens based largely on data from human epidemiology and animal toxicology (IARC, 2018; NTP, 2016; OEHHA, 2018; U.S. EPA, 2017), yet thousands of chemicals currently in commerce have never been evaluated for carcinogenic potential, signifying the need for a method to screen and prioritize chemicals of potential concern.

For this evaluation exercise, we selected a convenience subset of the screening platforms used by the U.S. EPA Toxicity Forecaster (ToxCast) program. ToxCast uses over 700 high-throughput assays on more than 15 commercial or federal government platforms to screen chemicals for a wide variety of biological effects (Judson *et al.*, 2010; Kavlock *et al.*, 2012; U.S. EPA, 2018).

Specifically, we used publicly available resources to:

- Select ToxCast assays that are cancer pathway related, within the selected subset of platforms.
- Map these assays to key characteristics of carcinogens, such as "is genotoxic" or "induces chronic inflammation" (described in Guyton *et al.*, 2018; Smith *et al.*, 2016), based on the endpoints measured.
- Rank and visually depict 1061 chemicals based on activity in these assays using the Toxicological Prioritization Index (ToxPi) software (Marvel *et al.*, 2018; Reif *et al.*, 2010, 2013).
- Compare the use categories or classes of chemicals represented in the upper 50% of chemicals, ranked by activity, with the categories or classes in the lower 50%.
- Identify chemotypes, or structural features, enriched in the top ranked 5% of chemicals using the ChemoTyper application (Yang *et al.*, 2015).

This proof-of-concept approach integrates multiple information sources and software as an example of how chemicals could be ranked or prioritized for carcinogenicity concern. This approach is highly adaptable and amenable to evolving data streams. Our approach could also be tailored to examine other toxicological endpoints and other methods of ranking chemicals.

METHODS

Selecting assays and chemicals from U.S. EPA's ToxCast program. We evaluated a total of 236 assays from the ACEA, Apredica, and BioSeek platforms that were included in U.S. EPA's Interactive Chemical Safety for Sustainability (iCSS) ToxCast Dashboard at the time of our data export on November 30, 2015. Although there are a dozen additional platforms in the continuously evolving ToxCast program (U.S. EPA, 2018), we selected a convenience subset of three assay platforms to demonstrate proof of concept while maintaining manageability. Within the three selected assay platforms, we identified 137 assays related to cancer pathways. Selection of the 137 assays was based on: (1) expert judgement and scientific literature ($n = 61$); and (2) curated associations of molecular targets with cancer in the Comparative Toxicogenomics Database (CTD) ($n = 76$).

The cancer pathway-related assays identified by expert judgement included assays evaluating: altered cell proliferation ($n = 21$), increased mitochondrial mass ($n = 3$), mitotic arrest ($n = 3$), and altered cell protein content ($n = 16$). Scientific literature was used to support cancer pathway associations for 18 assays, such as assays evaluating protein modifications associated with genotoxicity. This supporting literature was identified through PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) and Google Scholar (https://scholar.google.com/) searches, using the search term "cancer" and terms relevant to the assay endpoints measured. See Supplementary Table 1 for further details.

The CTD is maintained by the North Carolina State University's NIEHS Environmental Health Science Center, and includes information on gene-disease interactions manually gathered from peer-reviewed scientific literature (Davis *et al.*, 2017; Wiegers *et al.*, 2014). The CTD includes curated associations as a means of indicating that there is peer-reviewed scientific literature supporting a link between a molecular target and diseases belonging to the category of "cancer," which can include site-specific conditions (eg, glioblastoma, hepatocellular carcinoma, lung neoplasms), as well as more generalized conditions (eg, adenocarcinoma, carcinoma, neoplasm metastasis). The 76 cancer pathway-related assays identified using the CTD-interrogated protein targets were all part of the BioSeek platform, and each of the protein targets in the 76 selected BioSeek assays had a minimum of three curated associations with cancer in the CTD as of the August 24, 2017, data update. The cutoff of three curated associations (ie, associations with at least three diseases within the "cancer" category) was chosen to ensure a fairly high level of specificity by excluding targets with limited or no evidence in the scientific literature of being cancer-pathway related. We selected assays that evaluated both increased and decreased expression of the molecular targets. See Supplementary Table 2 for CTD cancer-curated associations. Overall cancer pathway-related assay selection is depicted in Figure 1.

Activity data for all 1061 chemicals tested in Phases 1 and 2 of ToxCast were exported from the iCSS ToxCast Dashboard on November 30, 2015. We examined the activities of all 1061 chemicals in the selected 137 cancer pathway-related assays.

Linking cancer pathway-related assays to key characteristics of carcinogens. In order to evaluate the biological coverage of the cancer pathway-related assays, we independently mapped each of the selected 137 assays to the 10 key characteristics of carcinogens described by IARC (Guyton *et al.*, 2018; Smith *et al.*, 2016), and compared our results to their similar mapping effort (Chiu *et al.*, 2018; IARC, 2017). Unlike IARC, we mapped some assays to two key characteristics. See Supplementary Spreadsheet 1 for a full list of the 137 assays mapped to key characteristics of carcinogens. The comparison of our mapping results for assays from three ToxCast platforms with IARC's independent mapping of 265 cancer pathway-related assays they identified from seven ToxCast platforms is displayed in Table 1.

Ranking chemical activity in cancer pathway-related assays. We used the ToxPi software, version 2.0 (available at: http://toxpi.org/) to rank each chemical for activity in the 137 cancer pathway-related assays. The software integrates different data streams

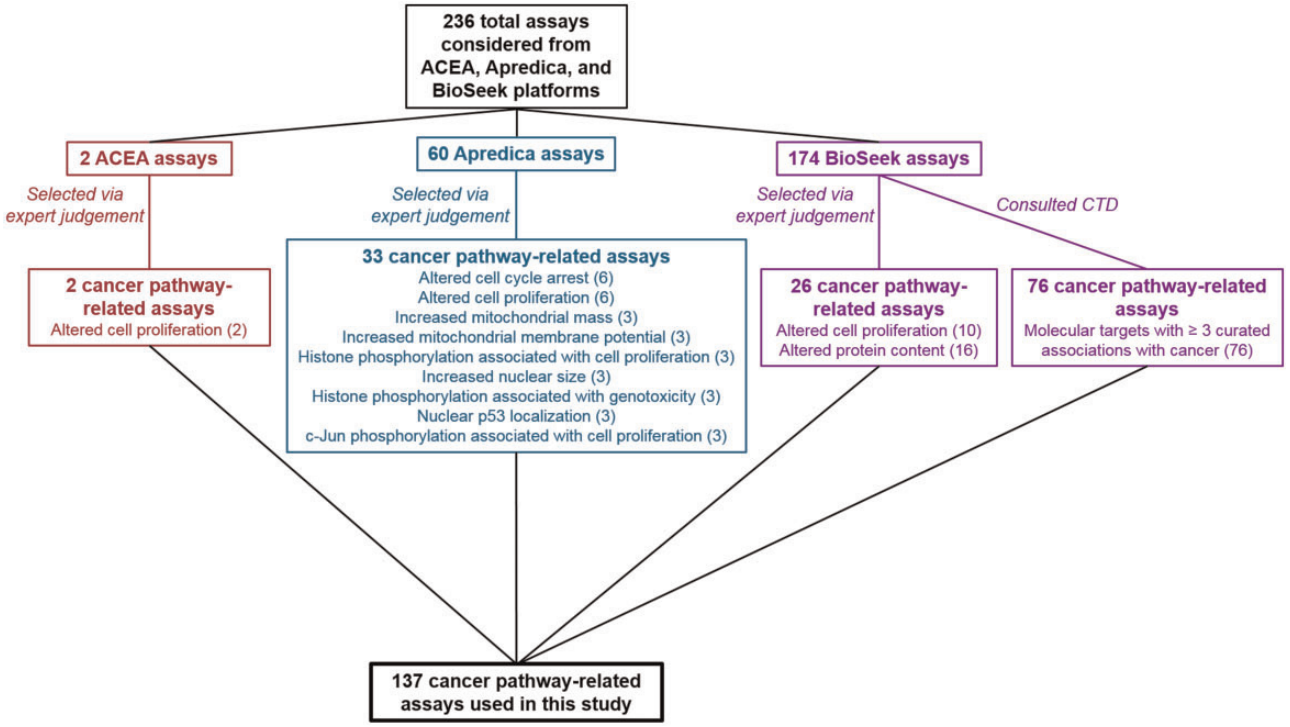

Figure 1. Selection of cancer pathway-related assays in the ACEA, Apredica, and BioSeek platforms. The number of assays evaluating each type of endpoint is displayed in parentheses.

Table 1. Biological Coverage of Key Characteristics of Carcinogens in Selected Cancer Pathway-Related Assay Endpoints

Key Characteristic of Carcinogens	Percentage of Cancer Pathway-Related Assays in a Subset of ToxCast Platforms (ACEA, Apredica, and BioSeek) Linked to Characteristic[a]	Percentage of Cancer Pathway-Related Assays in Seven ToxCast Platforms Linked to Characteristic, Identified by IARC[b]
1. Is electrophilic or can be metabolically activated	0% (0/137)	12% (31/265)
2. Is genotoxic	4% (6/137)	0% (0/265)
3. Alters DNA repair or causes genomic instability	0% (0/137)	0% (0/265)
4. Induces epigenetic alterations	4% (6/137)	4% (11/265)
5. Induces oxidative stress	4% (6/137)	7% (18/265)
6. Induces chronic inflammation	34% (46/137)	17% (45/265)
7. Is immunosuppressive	0% (0/137)	0% (0/265)
8. Modulates receptor-mediated effects	0% (0/137)	35% (92/265)
9. Causes immortalization	0% (0/137)	0% (0/265)
10. Alters cell proliferation, cell death, or nutrient supply	67% (92/137)	26% (68/265)

[a]The 137 cancer pathway-related assays identified in this study were mapped to the key characteristics of carcinogens described by IARC (IARC, 2017; Smith *et al.*, 2016). Some of these assays were mapped to two characteristics. See Supplementary Spreadsheet 1 for a full list of these assays mapped to key characteristics of carcinogens.

[b]IARC mapped 265 cancer pathway-related assays from seven of the ToxCast platforms (ACEA, Apredica, Attagene, BioSeek, NovaScreen, Odyssey Thera, and Tox21) to the key characteristics of carcinogens. IARC mapped each assay to one characteristic. Additional details on IARC's mapping can be found in the Excel file "Section 4.3 Spreadsheet" available at: https://monographs.iarc.fr/iarc-monographs-on-the-evaluation-of-carcinogenic-risks-to-humans-3/.

and calculates a unitless ToxPi score that represents a relative ranking of biological activity across multiple assays. The ToxPi output can be used for rank ordering chemicals and informing prioritization. The software also generates a visual ToxPi image for each chemical, reflective of its relative biological activity. Details about the ToxPi software and algorithm have been previously described (Marvel *et al.*, 2018; Reif *et al.*, 2010, 2013). In this analysis, each ToxPi slice represents the set of ToxCast assays that we mapped to each key characteristic of carcinogens. Because the 137 assays considered here mapped to only 5 of the 10 key characteristics, each ToxPi is composed of 5 corresponding slices. The input data used for ToxPi were: AC_{50} values (the concentrations inducing a half-maximal assay response) for the chemicals active in assays, and an assigned value of 10^6 for chemicals inactive in assays. The ToxPi scaling type $-\log_{10}(AC_{50}) + 6$ was used. These are the intended inputs and

scaling type for ranking ToxCast data, as described in the original ToxPi GUI User Manual (UNC, 2009). For this work, we did not adjust the data to account for cytotoxicity. Because all the cytotoxicity assays were mapped to the "alters cell proliferation, cell death, or nutrient supply" key characteristic, it is challenging to adjust for this potential confounder without diluting a key cancer pathway-related effect. As a consequence of the scaling type used, each ToxPi slice length is proportional to the normalized potency of the assay value ($-\log_{10}(AC_{50}) + 6$) of the component assays included in that slice. For the ToxPi analysis conducted in this study, weighting was applied based on the number of "assay component names" making up the slice. The "assay component names" provided in the iCSS ToxCast Dashboard are short names containing the assay and a component readout, such as "APR_HepG2_CellCycleArrest_1hr." Distinct from the "assay component endpoint name," the "assay component name" excludes the direction of the assay signal (eg, "up" or "down"). There were:

- Fifty-five assay component names for the "alters cell proliferation, cell death, or nutrient supply" key characteristic ToxPi slice
- Three assay component names for the "induces oxidative stress" key characteristic slice
- Twenty-three assay component names for the "induces chronic inflammation" key characteristic slice
- Six assay component names for the "is genotoxic" key characteristic slice
- Six assay component names for the "induces epigenetic alterations" key characteristic slice

To correspond with the variable numbers of assay component names making up each ToxPi slice, weights of 18, 1, 8, 2, and 2, were applied for each of these slices, respectively. The ToxPi input data file is Supplementary Spreadsheet 2. Using this approach with ToxPi, all 1061 chemicals were ranked for activity in the cancer pathway-related assays.

Assigning descriptive chemical categories. To further examine the ToxPi rankings, we assigned a descriptive category type to each chemical. For many chemicals, the category type assigned was simply the use category information manually obtained from the iCSS ToxCast Dashboard. In cases where the iCSS ToxCast Dashboard assigned more than one use category to a single chemical, we conducted further research using PubChem (https://pubchem.ncbi.nlm.nih.gov/) and Google (https://www.google.com/) searching to assign a single predominant descriptive category. If a single predominant descriptive category could not be determined for a chemical, it was assigned to the "other" category. Occasionally, we created more specific descriptive categories for chemical groups of known carcinogenicity concern (eg, nitrosamines and polycyclic aromatic hydrocarbons). We chose to compare the chemical categories of the upper 50% ($n = 531$) of ranked chemicals with the categories represented in the lower 50% ($n = 530$). Other users of the approach presented here could choose to examine other quantiles of ranked chemicals, as suitable for their study objectives.

Evaluating chemotypes. We arbitrarily selected the top 5% of chemicals ranked with ToxPi for further exploration and characterization. The percentage of chemicals examined can be customized to suit a user's study objectives. We examined the chemotypes, or structural features, represented in the top 5% of chemicals with the ChemoTyper application (available at: https://chemotyper.org/), developed by Molecular Networks GmbH and Altamira LLC (Yang *et al.*, 2015), and the "TOXCST_v4a_1892_20Mar2012.sdf" and "toxprint_v2.0_r212.xml" files. We determined the number of chemotypes represented in the 5% ($n = 54$) most active chemicals by searching across "any" of them in the ChemoTyper application. The search included chemotypes that are Ashby Tennant structural alerts for DNA reactivity (Ashby and Tennant, 1991) and cancer threshold of toxicological concern (TTC) structures (Kroes *et al.*, 2004).

One-tailed two-proportion Z-tests were conducted using the Bonferroni correction to compare the proportion of each chemotype ($n = 201$) in the 54 most active chemicals with the proportion in the remaining 1007 chemicals. The null hypothesis was that the proportion of each chemotype in the 54 most highly ranked chemicals was not different from the proportion of each chemotype in the remainder of the chemicals; the null was rejected when $p < 2.5 \times 10^{-4}$, which was the significance cutoff after applying the Bonferroni correction (0.05/201). This allowed us to identify the enriched chemotypes in the top 5% of chemicals active in the cancer pathway-related assays.

RESULTS

Figure 1 shows the breakdown of assays in the ACEA, Apredica, and BioSeek platforms that we determined to be cancer pathway-related either via expert judgement or via consultation with the CTD. (Additional assay selection details can be found in Supplementary Tables 1 and 2.) In total, we identified 137 assays as cancer pathway-related.

We then linked these 137 cancer pathway-related assays to IARC's 10 key characteristics of carcinogens, as a way to evaluate the biological coverage of the ToxCast assay subset and compare our findings with those of IARC (Chiu *et al.*, 2018; IARC, 2017). As shown in Table 1, the cancer pathway-related assays from the three platforms considered in this study primarily evaluate endpoints related to the "alters cell proliferation, cell death, or nutrient supply" ($n = 92$), and "induces chronic inflammation" ($n = 46$) characteristics. Some of the assays we determined to be cancer pathway-related evaluate the characteristics "is genotoxic" ($n = 6$), "induces epigenetic alterations" ($n = 6$), and "induces oxidative stress" ($n = 6$). See Supplementary Spreadsheet 1 for a full list of assays mapped to key characteristics, as well as a list of the assays considered not cancer pathway-related in this study. We did not identify any assays that evaluate the remaining five characteristics of carcinogens; thus according to our analyses, no assays in the ACEA, Apredica, or BioSeek platforms of ToxCast evaluate the characteristics "is electrophilic or can be metabolically activated," "alters DNA repair or causes genomic instability," "is immunosuppressive," "modulates receptor-mediated effects," or "causes immortalization."

For comparison purposes, the mapping results of the cancer pathway-related assays selected by IARC from seven ToxCast platforms is shown alongside our mapping results in Table 1. Based on IARC's work, there are ToxCast assays in other platforms that evaluate the characteristics "is electrophilic or can be metabolically activated," and "modulates receptor-mediated effects." Neither our mapping nor IARC's mapping of four additional platforms identified any ToxCast assays that evaluate the key characteristics, "alters DNA repair or causes genomic instability," "is immunosuppressive," or "causes immortalization."

Our mapping was generally concordant with that of Chiu *et al.* (2018) and IARC (2017). Among the three platforms evaluated in this study, we mapped many assays to the same key characteristics as did IARC (eg, many assays mapped to

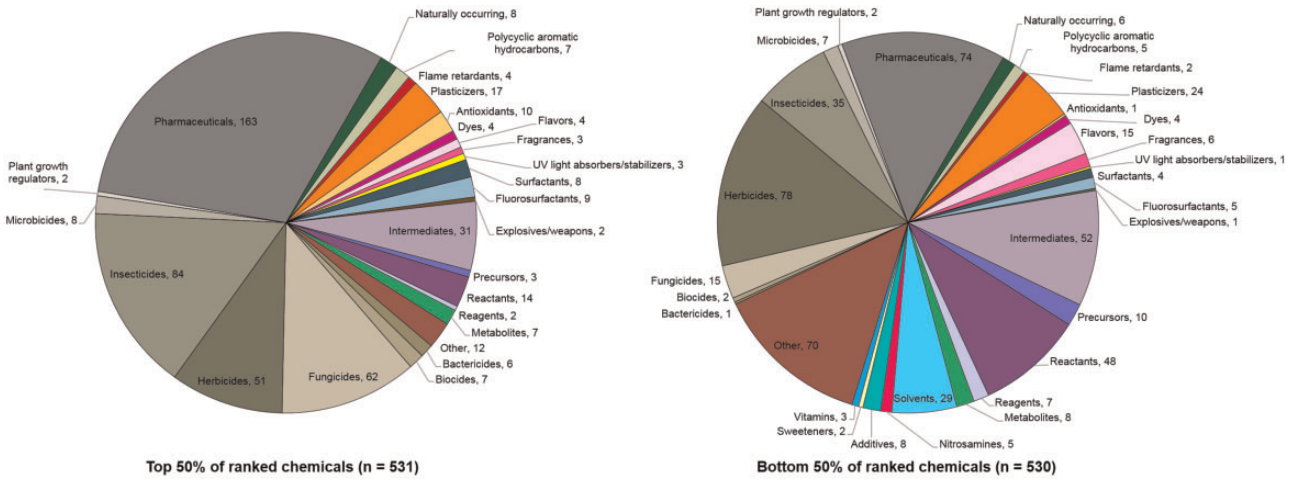

Figure 2. Pie charts depicting category types of the upper 50% ($n = 531$) and lower 50% ($n = 530$) of chemicals ranked based on activity in cancer pathway-related assays. Pharmaceutical chemicals are represented in the gray pie slice, chemicals in pesticide categories (bactericides, biocides, fungicides, herbicides, insecticides, microbicides, and plant growth regulators) are represented in the bronze pie slices, and the remaining categories are shown in the pie slices of other colors.

"induces chronic inflammation" and "alters cell proliferation, cell death, or nutrient supply"). However, there were several distinctions between our approach and IARC's. Our initial prescreen to select cancer-related assays in the CTD could have excluded some assays relevant to the key characteristics of carcinogens. However, on inspection, our approach included some assay targets that IARC excluded (eg, uPAR, and assays on the BioSeek platform evaluating decreased expression of molecular targets), and we excluded some assay targets that IARC included (eg, CD38 and HLA-DR, which did not have any curated cancer associations in the CTD). Some of our mappings also differed. For example, IARC mapped the ACEA_T47D_80hr_Positive assay to the "modulates receptor-mediated effects" characteristic, presumably because the gene target for this assay on the iCSS ToxCast Dashboard is estrogen receptor 1, based on the linear growth response that estrogen receptor agonists induce in these cells (Rotroff *et al.*, 2013). However, we mapped this ACEA assay to the "alters cell proliferation, cell death, or nutrient supply" characteristic to be inclusive of the multiple mechanisms that may be responsible for the cell proliferation effects measured in the assay. In addition, we mapped some assays to two different characteristics. For example, the APR_HepG2_MitoticArrest assays evaluate histone H3 phosphorylation at serine 10 (K. Houck, e-mail communication, December 30, 2015), and this endpoint is indicative of high mitotic activity tissue regions in various human tumors (Bossard *et al.*, 2006; Colman *et al.*, 2006; Kim *et al.*, 2007; Ribalta *et al.*, 2004; Scott *et al.*, 2005; Skaland *et al.*, 2007). Therefore, we mapped these assays to both the "induces epigenetic alterations" and "alters cell proliferation, cell death, or nutrient supply" characteristics, whereas IARC mapped them to just the latter characteristic.

We used the ToxPi software to rank all 1061 chemicals screened in ToxCast Phases I and II, based on activity in the 137 identified cancer pathway-related assays. Each ToxPi was composed of five slices, with each slice corresponding to a key characteristic of carcinogens and weighted to reflect the number of mapped assay components. Supplementary Spreadsheet 3 contains the complete ToxPi results output file from this analysis.

The chemical categories for the upper 50% ($n = 531$) and lower 50% ($n = 530$) of ranked chemicals are reflected in Figure 2. About 58% ($n = 309$) of the higher ranked chemicals are pharmaceuticals, insecticides, and fungicides, whereas only about 23% ($n = 124$) of the lower ranked chemicals belong to these categories. The higher ranked chemicals include more antioxidants, surfactants, and fluorosurfactants, whereas the lower ranked chemicals include more herbicides, flavors, fragrances, intermediates, and reactants. The lower ranked chemicals include categories not represented among the higher ranked chemicals: solvents, nitrosamines, additives, sweeteners, and vitamins. The lower ranked chemicals also include many that were represented by more than one descriptive chemical category and thus are captured under the "other" category.

To focus further on the most highly active chemicals, we identified the top 5% ($n = 54$) of total chemicals ranked. They are listed in ranked order in Table 2. Nine of these top 54 chemicals are identified as known to cause cancer under California's Proposition 65 (Title 27, California Code of Regulations, § 27001).The ToxPi images of these top 5% of chemicals are depicted in ranked order in Figure 3. The images show the relative chemical activities, both overall and across the slices, based on activity data from the assays mapped to the key characteristics of carcinogens.

We identified 23 chemotypes, or structural features, that are present at significantly higher proportions in the top 5% of chemicals ranked, relative to the remaining 95% lower ranked chemicals. These enriched chemotypes are listed in Table 3. We noted that multiple tin (Sn)-containing chemotypes were included in this list and further examined the 1061 chemicals evaluated here for tin-containing compounds. We found that all three of the tin-containing compounds included in the chemical set we evaluated are within the top 5%, and are in fact the top three ranked chemicals (tributyltin chloride, tributyltin methacrylate, and triphenyltin hydroxide). The enriched chemotypes also include multiple mercury (Hg)-containing chemotypes. One of the two mercury-containing compounds in the full set of chemicals evaluated here is the fourth ranked chemical, phenylmercuric acetate; the other mercury-containing compound, mercuric chloride, is rank number 89. One of the chemotypes in Table 3 is a threshold of toxicological concern (TTC) structure for carcinogenicity. None of the enriched chemotypes are Ashby Tennant structural alerts for DNA reactivity. Figure 4 shows 10 of the 23 enriched chemotypes.

Table 2. Top 5% ($n = 54$) of Chemicals Ranked Based on Activity in Cancer Pathway-Related Assays

1. Tributyltin chloride	28. Octyl gallate
2. Tributyltin methacrylate	29. PharmaGSID_48519[b]
3. Triphenyltin hydroxide[a]	30. Sodium (2-pyridylthio)-*N*-oxide
4. Phenylmercuric acetate	31. SAR115740[b]
5. Chlorothalonil[a]	32. Clomiphene citrate[a,b]
6. Fluazinam	33. UK-337312[b]
7. Gentian violet[a]	34. 2,4-Bis(2-methylbutan-2-yl)phenol
8. Niclosamide[b]	35. 2,4-Bis(1-methyl-1-phenylethyl)phenol
9. Didecyldimethylammonium chloride	36. Cycloheximide
10. Tamoxifen[a,b]	37. 4-(1,1,3,3-Tetramethylbutyl)phenol
11. AVE8923[b]	38. 2-(Thiocyanomethylthio)benzothiazole
12. Octhilinone	39. Abamectin
13. Ziram	40. Milbemectin (mixture of 70% milbemycin A4, 30% milbemycin A3)
14. SR146131[b]	41. SB236057A[b]
15. Triclosan	42. Kepone[a]
16. Thiram	43. PD 0343701[b]
17. Tamoxifen citrate[a,b]	44. Diethylstilbestrol[a,b]
18. 3-Iodo-2-propynyl-*N*-butylcarbamate	45. Farglitazar[b]
19. 9-Phenanthrol[b]	46. SSR241586[b]
20. Emamectin benzoate	47. Dodecyltrimethylammonium chloride
21. Captafol[a]	48. SR271425[b]
22. Disulfiram[b]	49. PharmaGSID_47337[b]
23. PharmaGSID_47315[b]	50. Zoxamide
24. AVE5638[b]	51. HMR1171[b]
25. 4-Nonylphenol, branched	52. Clorophene
26. 1, 2-Benzisothiazolin-3-one	53. 2,4-Di-*tert*-butylphenol
27. AVE6324[b]	54. Clotrimazole

[a]Chemical is listed as known to the state to cause cancer under California's Proposition 65, as of November 23, 2018 (Title 27, California Code of Regulations, § 27001).
[b]Pharmaceutical compound.

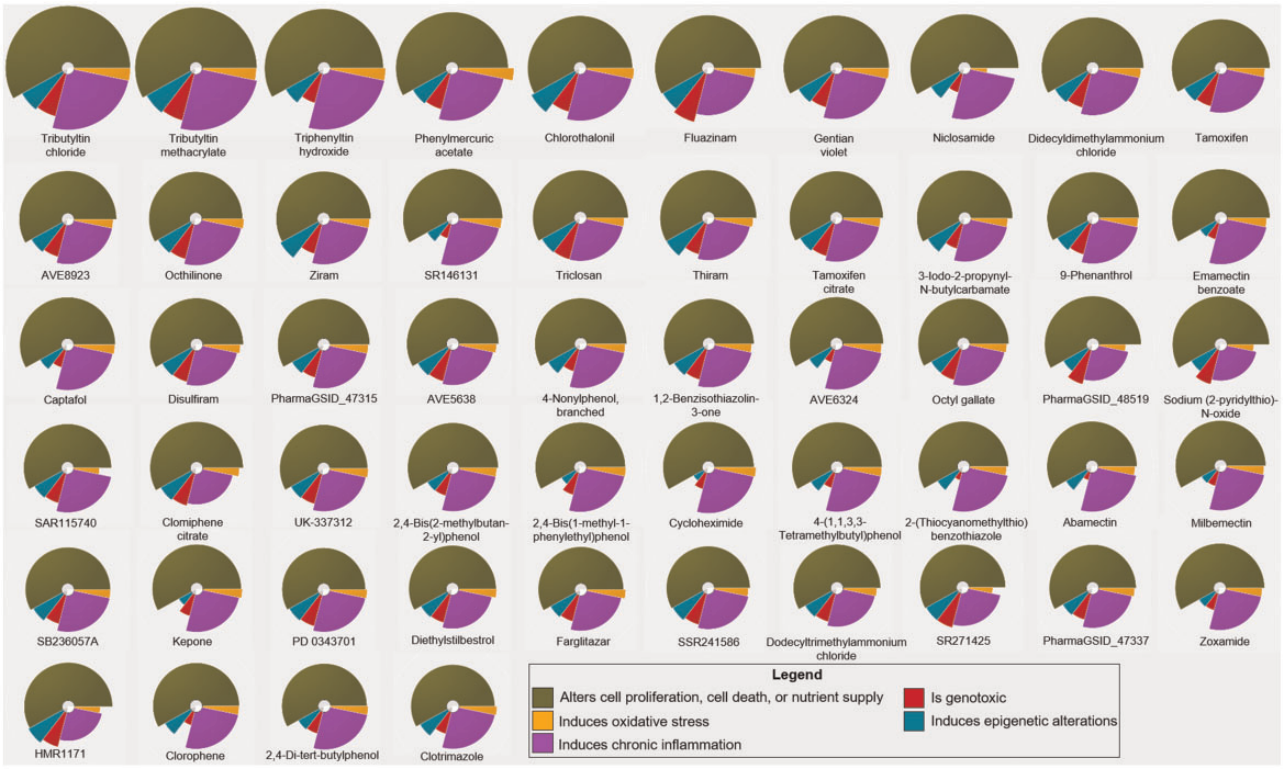

Figure 3. ToxPi images of the top 5% ($n = 54$) of ranked chemicals based on activity in cancer pathway-related assays. Each ToxPi is composed of five slices, each corresponding to a key characteristic of carcinogens, to which the assays considered here were linked. Each ToxPi slice is weighted to reflect the number of assay component names it includes.

Table 3. Enriched Chemotypes in Top 5% ($n = 54$) of Chemicals Ranked Based on Activity in Cancer Pathway-Related Assays

1. bond: metal_group_III_other_Sn_generic
2. bond: metal_group_III_other_Sn_organo
3. atom: element_metal_poor_metal
4. bond: metal_group_III_other_generic
5. bond: metal_group_III_other_generic_oxy
6. bond: metal_group_III_other_Sn_oxy
7. bond: quatN_generic
8. bond: CS_sulfide_di-
9. bond: quatN_alkyl_acyclic
10. chain: aromaticAlkane_Ph-C1-Ph
11. bond: metal_group_III_other_Sn_halide
12. bond: metal_transition_Hg_generic
13. bond: metal_transition_Hg_organo
14. bond: metal_transition_Hg_oxy
15. ring: hetero_[5]_N_S_isothiazole
16. bond: CN_amine_aliphatic_generic
17. bond: C(=O)N_carbamate_dithio[a]
18. chain: aromaticAlkane_Ar-C-Ar
19. group: carbohydrate_hexopyranose_generic
20. ring: hetero_[6]_N_piperidine
21. chain: aromaticAlkene_Ph-C2_acyclic_generic
22. chain: aromaticAlkane_Ph-C1-acyclic_generic
23. ring: hetero_[5]_N_S_thiazole

[a]Threshold of toxicological concern (TTC) structural category for carcinogenicity.

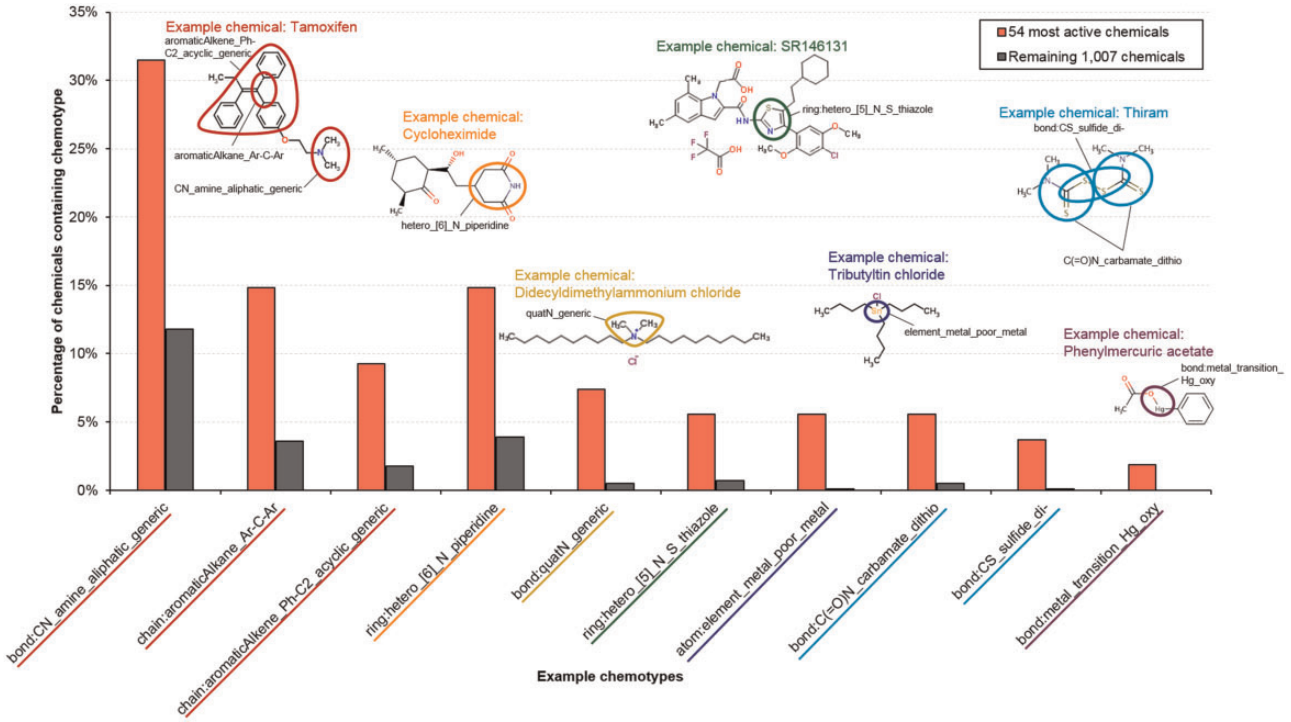

Figure 4. The top 5% of ranked chemicals based on activity in cancer pathway-related assays are enriched for specific chemotypes. One-tailed two-proportion Z tests were conducted with the Bonferroni correction with the 201 chemotypes represented in the top 5% of chemicals ($n = 54$) ranked in our analysis in order to compare the proportion of each chemotype in the most highly ranked chemicals with the proportion in the remaining 1007 chemicals. Ten of these enriched chemotypes, each significant with $p < 2.5 \times 10^{-4}$, are shown here as examples, color coded to link the chemotype name with the example chemical structure.

DISCUSSION

We developed and applied a proof-of-concept exercise to identify and prioritize chemicals of potential carcinogenicity concern. We utilized the iCSS ToxCast Dashboard, a framework for organizing cancer pathway-related assays, and publicly available software in a novel approach to identify chemicals and chemotypes associated with some key characteristics of carcinogens.

Identification of ToxCast assays associated with cancer pathways is consistent with the NAS (2007) recommendations to evaluate chemicals according to their ability to perturb toxicity pathways. Our approach took this recommendation further by independently mapping assays to the key characteristics of carcinogens. This additional step, pioneered by IARC (Chiu *et al.*, 2018; IARC, 2017), is useful because it incorporates more biological details and because it identifies biological coverage gaps in ToxCast assay platforms. It is notable that neither our work nor IARC's identified any ToxCast assays that evaluate the characteristics "alters DNA repair or causes genomic instability," "is immunosuppressive," or "causes immortalization." These biological coverage gaps potentially persist in our approach even after expanding the assay selection to additional ToxCast platforms. The ToxCast program is continuing to refine their assays, which may include expansion of biological coverage across these important key characteristics. Our independent mapping largely validated IARC's, but also identified a few specific areas of divergence that should be addressed before the mapping is applied more broadly.

We found that many of the upper 50% of ranked chemicals belonged, unsurprisingly, to categories of chemicals designed to be bioactive (ie, pharmaceuticals, insecticides, and fungicides). In contrast, other categories made up a majority of the lower 50% of chemicals ranked. We observed unique chemical categories, such as vitamins, solvents, and nitrosamines, in the lower 50% of chemicals. High volatility of some of the solvents may explain their low rankings here because volatile chemicals are difficult to study in vitro. One low-ranking solvent, *N*,*N*-dimethylformamide, is listed as a carcinogen under California's Proposition 65 (Title 27, California Code of Regulations, § 27001), and its mechanism of action requires metabolic activation (Cordeiro and Savarese, 1986; Cross *et al.*, 1990; Gescher, 1993). *N*,*N*-Dimethylformamide's low ranking in our analysis is likely due to none of the assays mapping to the key characteristic "is electrophilic or can be metabolically activated," and the known limited metabolic capacity of ToxCast assays (DeGroot *et al.*, 2018). Nitrosamines were another unique chemical category in the lower 50% of chemicals, and the five nitrosoamines making up this group are listed as known to the state to cause cancer under California's Proposition 65 (Title 27, California Code of Regulations, § 27001). Two of them, *N*-nitrosodibutylamine and *N*-nitrosodipropylamine, were not active in any cancer pathway-related assays and thus received ToxPi scores of zero. Metabolic activation is required for nitrosamine carcinogenicity (Archer, 1989; Montesano and Hall, 1984), so the assays understandably failed to capture these known carcinogens.

We identified 23 chemotypes that were enriched in the top 5% of chemicals, in our examination of the most highly ranked chemicals based on activity in the cancer pathway-related assays. These enriched chemotypes may be important structural features for activity in the cancer pathway-related assays. Multiple enriched chemotypes contained tin, mercury, or metals in general, and the top four ranked chemicals contained tin or mercury. Sipes *et al.* (2013) found that heavy metal-containing chemicals are among the most promiscuous in the Novascreen assay platform, which evaluates receptor binding and enzyme activity. Chemical promiscuity may also explain the high rankings of these compounds in our study.

The approach used here is highly adaptable, allowing incorporation of various data streams, toxicity pathways, and cutoffs that could be tailored to a decision context. This approach could be developed as an initial screen to prioritize chemicals with limited or no toxicity information for further evaluation. Others have published work utilizing similar concepts, or portions of this approach, for such applications as: organizing the ToxCast data mapped to the key characteristics of carcinogens for chemical case studies (Chiu *et al.*, 2018), correlating ToxCast assays with cardio-, hepato-, neuro-, and renal toxicities using information from the CTD (Hu *et al.*, 2015), ranking environmental chemicals for cardiotoxicity hazard using data from induced pluripotent stem cells (Sirenko *et al.*, 2017), and ranking chemicals present at Superfund sites based on such data as pathway information and ToxCast assay target families (Tilley *et al.*, 2017).

Various components of this approach can be further evaluated and improved in future iterations. For example, by excluding assays evaluating molecular targets with only 1–2 curated associations with cancer in the CTD, we may have omitted targets linked with one site-specific effect. Most of the assays we mapped were linked to the "induces chronic inflammation" and "alters cell proliferation, cell death, or nutrient supply" key characteristics. Short-term ToxCast assays may not be the optimal dataset to capture the biology of characteristics requiring chronicity (eg, chronic inflammation, sustained proliferation) to induce cancer. With the dataset used here, the highest ranked chemicals have biologic activity in assays mapped to 5 of the 10 the key characteristics, but carcinogens acting by a different mechanism would not be highly ranked unless the present assay selection were supplemented with data from other sources covering these biological gaps. Another important consideration is the minimal metabolic capacity in many of the current ToxCast assays (DeGroot *et al.*, 2018). Metabolic activation is required for many carcinogens, and the ToxCast assays linked to cancer pathways may not be activated by the parent compound.

For the reasons described above, multiple elements, including the underlying ToxCast data, require refinement before this approach is suitable for applying predictive values to the outcomes. Nine of the top 54 (16.7%) ranked chemicals in our approach are listed as carcinogens under California's Proposition 65. We also separately examined the 223 chemicals that received a ToxPi score of zero in our approach, and 31 of them (13.9%) are carcinogens listed under California's Proposition 65, likely due to the limited biological coverage and metabolic capacity in the present assays. Our approach is therefore not yet capable of predicting suspected carcinogens based on limitations in the current input data, but we expect it will generate stronger predictive outcomes with improved input data and expanded biological coverage.

Becker *et al.* (2017) found that the activities observed in the ToxCast assays that IARC mapped to the key characteristics of carcinogens were unable to predict cancer hazard for pesticides with and without human cancer hazard potential. This study evaluated only pesticides that U.S. EPA has classified for human cancer hazard potential, whereas our analysis examined many more chemicals, some of which are previously untested, for their activities in the mapped assays. Some of the same limitations that we identified above may also explain the conclusion by Becker *et al.* (2017).

Five pesticides within the top 5% of chemicals ranked in our study have been classified by U.S. EPA as either "not likely to be carcinogenic to humans" (3-iodo-2-propynyl-*N*-butylcarbamate, emamectin benzoate, triclosan, and zoxamide) or as having "evidence of noncarcinogenicity for humans" (didecyldimethylammonium chloride), based on lack of carcinogenicity evidence in rodents (U.S. EPA, 2006, 2008, 2011, 2012, 2014, 2017). One explanation is that our method may have identified some false positives in the highly ranked chemicals. Another plausible explanation is that we correctly identified chemicals that display some key characteristics of carcinogens, but that did not produce treatment-related tumors in the traditional rodent tests used to make cancer hazard classifications. Consistent with this explanation, recent studies have independently found that emamectin benzoate induces genotoxicity and cytotoxicity in vitro (Yun *et al.*, 2017; Zhang *et al.*, 2017), that triclosan activates cell migration pathways in vitro (Derouiche *et al.*, 2017; Kim *et al.*, 2015) and is a liver tumor promoter in mice (Yueh *et al.*, 2014), and that didecyldimethylammonium chloride affects oxidative stress and cell growth in vitro (Kwon *et al.*, 2014).

For this study, we input AC_{50} data values as a quantitative measure of chemical activity, but the input values can be customized, given the multiple scaling options available in ToxPi intended to handle various data types. Input values may also be adjusted for cytotoxicity to distinguish between a chemical's specific molecular effects versus general activity reflecting chemical-induced cytotoxicity. A general activity "burst" at cytotoxic chemical concentrations has been observed in the

ToxCast assays (Judson *et al.*, 2016), and accounting for this phenomenon can impact the conclusions drawn from the data (Becker *et al.*, 2017; Fay *et al.*, 2018; Silva *et al.*, 2015). We opted not to adjust the data for cytotoxicity in this study because all the cytotoxicity assays were mapped to the "alters cell proliferation, cell death, or nutrient supply" key characteristic, and making an adjustment could dilute a key cancer pathway-related effect.

The specific cellular and molecular characteristics of carcinogens published by IARC were critical for our analysis because they facilitated mapping of the assay endpoints to key characteristics of carcinogens. Another key element necessary to conduct this work is the frequently updated CTD, which we used to link molecular targets interrogated in the assays with cancer. Cancer is only one of many disease endpoints covered in the CTD; others include cardiovascular disease, male and female urogenital disease, pregnancy complications, and endocrine system disease. If key characteristics were developed for the hazard traits of reproductive toxicity, developmental toxicity, endocrine disruption, or cardiovascular toxicity (OEHHA, 2012), our approach could be easily extended to identify and prioritize chemicals for these toxicities as well.

Identification of enriched chemotypes among highly active chemicals in cancer pathway-related assays could allow data-poor chemicals or chemical classes containing those chemotypes to be flagged for possible carcinogenicity concern and prioritized for further testing. This type of approach offers new possibilities for predictive toxicology by using publicly available tools and data to prioritize chemicals of concern for specific toxicities and further testing while reducing reliance on animal testing.

SUPPLEMENTARY DATA

Supplementary data are available at *Toxicological Sciences* online.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ACKNOWLEDGMENTS

We thank U.S. EPA's ToxCast program staff for helpful discussion and guidance, particularly Dr Keith Houck and Mr Jimmy Phuong for technical discussion about the ToxCast assays, and Ms Monica Linnenbrink for facilitating the logistics of these meetings. We also thank Dr David Reif at North Carolina State University for guidance and helpful recommendations on analyzing this dataset with ToxPi. We appreciate the careful review of this manuscript by Dr Chris Banks and Dr Meng Sun at OEHHA.

REFERENCES

Archer, M. C. (1989). Mechanisms of action of N-nitroso compounds. *Cancer Surv*. **8**, 241–250.

Ashby, J., and Tennant, R. W. (1991). Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the U.S. NTP. *Mutat. Res.* **257**, 229–306.

Becker, R. A., Dreier, D. A., Manibusan, M. K., Cox, L. A. T., Simon, T. W., and Bus, J. S. (2017). How well can carcinogenicity be predicted by high throughput "characteristics of carcinogens" mechanistic data? *Regul. Toxicol. Pharmacol.* **90**, 185–196.

Bossard, C., Jarry, A., Colombeix, C., Bach, N. K., Moreau, A., Loussouarn, D., Mosnier, J. F., and Laboisse, C. L. (2006). Phosphohistone H3 labelling for histoprognostic grading of breast adenocarcinomas and computer-assisted determination of mitotic index. *J. Clin. Pathol.* **59**, 706–710.

Chiu, W. A., Guyton, K. Z., Martin, M. T., Reif, D. M., and Rusyn, I. (2018). Use of high-throughput *in vitro* toxicity screening data in cancer hazard evaluations by IARC Monograph Working Groups. *ALTEX* **35**, 51–64.

Colman, H., Giannini, C., Huang, L., Gonzalez, J., Hess, K., Bruner, J., Fuller, G., Langford, L., Pelloski, C., Aaron, J., *et al.* (2006). Assessment and prognostic significance of mitotic index using the mitosis marker phosphor-histone H3 in low and intermediate-grade infiltrating astrocytomas. *Am. J. Surg. Pathol.* **30**, 657–664.

Cordeiro, R. F., and Savarese, T. M. (1986). Role of glutathione depletion in the mechanism of action of N-methylformamide and N,N-dimethylformamide in a cultured human colon carcinoma cell line. *Cancer Res.* **46**, 1297–1305.

Cross, H., Dayal, R., Hyland, R., and Gescher, A. (1990). N-alkylformamides are metabolized to N-alkylcarbamoylating species by hepatic microsomes from rodents and humans. *Chem. Res. Toxicol.* **3**, 357–362.

Davis, A. P., Grondin, C. J., Johnson, R. J., Sciaky, D., King, B. L., McMorran, R., Wiegers, J., Wiegers, T. C., and Mattingly, C. J. (2017). The comparative toxicogenomics database: Update 2017. *Nucleic Acids Res.* **45**, D972–D978.

DeGroot, D. E., Swank, A., Thomas, R. S., Strynar, M., Lee, M. Y., Carmichael, P. L., and Simmons, S. O. (2018). mRNA transfection retrofits cell-based assays with xenobiotic metabolism. *J. Pharmacol. Toxicol. Methods* **92**, 77–94.

Derouiche, S., Mariot, P., Warnier, M., Vancauwenberghe, E., Bidaux, G., Gosset, P., Mauroy, B., Bonnal, J. L., Slomianny, C., Delcourt, P., *et al.* (2017). Activation of TRPA1 channel by antibacterial agent triclosan induces VEGF secretion in human prostate cancer stromal cells. *Cancer Prev. Res.* **10**, 177–187.

Fay, K. A., Villeneuve, D. L., Swintek, J., Edwards, S., Nelms, M., Blackwell, B. R., and Ankley, G. T. (2018). Differentiating pathway-specific from non-specific effects in high-throughput toxicity data: A foundation for prioritizing adverse outcome pathway development. *Toxicol. Sci.* **163**, 500–515.

Gescher, A. (1993). Metabolism of N,N-dimethylformamide: Key to the understanding of its toxicity. *Chem. Res. Toxicol.* **6**, 245–251.

Guyton, K. Z., Rusyn, I., Chiu, W. A., Corpet, D. E., van den Berg, M., Ross, M. K., Christiani, D. C., Beland, F. A., and Smith, M. T. (2018). Application of the key characteristics of carcinogens in cancer hazard identification. *Carcinogenesis* **39**, 614–622.

Hu, B., Gifford, E., Wang, H., Bailey, W., and Johnson, T. (2015). Analysis of the ToxCast chemical-assay space using the Comparative Toxicogenomics Database. *Chem. Res. Toxicol.* **28**, 2210–2223.

International Agency for Research on Cancer (IARC). (2017). *IARC Monographs*, Instructions to authors, Table 7. Key characteristics of carcinogens. Available at: https://monographs.iarc.fr/ENG/Preamble/previous/Instructions_to_Authors.pdf. last accessed July 30, 2018.

IARC. (2018). *Agents Classified by the IARC Monographs*, Volumes 1-121. http://monographs.iarc.fr/ENG/Classification/index.php. Accessed April 26, 2018.

Judson, R. S., Houck, K. A., Kavlock, R. J., Knudsen, T. B., Martin, M. T., Mortensen, H. M., Reif, D. M., Rotroff, D. M., Shah, I., Richard, A. M., *et al.* (2010). *In vitro* screening of environmental chemicals for targeted testing prioritization: The ToxCast project. *Environ. Health Perspect.* **118**, 485–492.

Judson, R., Houck, K., Martin, M., Richard, A. M., Knudsen, T. B., Shah, I., Little, S., Wambaugh, J., Setzer, R. W., Kothiya, P., *et al.* (2016). Analysis of the effects of cell stress and cytotoxicity on *in vitro* assay activity across a diverse chemical and assay space. *Toxicol. Sci.* **153**, 323–339.

Kavlock, R., Chandler, K., Houck, K., Hunter, S., Judson, R., Kleinstreuer, N., Knudsen, T., Martin, M., Padilla, S., Reif, D., *et al.* (2012). Update on EPA's ToxCast program: Providing high throughput decision support tools for chemical risk management. *Chem. Res. Toxicol.* **25**, 1287–1302.

Kim, S. H., Hwang, K. A., Shim, S. M., and Choi, K. C. (2015). Growth and migration of LNCaP prostate cancer cells are promoted by triclosan and benzophenone-1 via an androgen receptor signaling pathway. *Environ. Toxicol. Pharmacol.* **39**, 568–576.

Kim, Y. J., Ketter, R., Steudel, W. I., and Feiden, W. (2007). Prognostic significance of the mitotic index using the mitosis marker anti-phosphohistone H3 in meningiomas. *Am. J. Clin. Pathol.* **128**, 118–125.

Kwon, J. T., Kim, H. M., Kim, P., and Choi, K. (2014). Didecyldimethylammonium chloride induces oxidative stress and inhibits cell growth in lung epithelial cells. *Mol. Cell. Toxicol.* **10**, 41–45.

Kroes, R., Renwick, A. G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., Schilter, B., Schlatter, J., van Schothorst, F., Vos, J. G., *et al.* (2004). Structure-based thresholds of toxicological concern (TTC): Guidance for application to substances present at low levels in the diet. *Food Chem. Toxicol.* **42**, 65–83.

Marvel, S. M., To, K., Grimm, F. A., Wright, F. A., Rusyn, I., and Reif, D. M. (2018). ToxPi graphical user interface 2.0: Dynamic exploration, visualization, and sharing of integrated data models. *BMC Bioinformatics* **19**, 1–7.

Montesano, R., and Hall, J. (1984). Nitrosamine metabolism and carcinogenesis. In *Mutation, Cancer, and Malformation. Environmental Science Research* (E. H. Y. Chu and W. M. Generoso, Eds.), Vol. **31**, pp. 447–464. Springer, Boston, MA.

NAS (National Academy of Sciences). (2007). Toxicity Testing in the 21st Century: A Vision and a Strategy. National Academies Press, Washington, DC.

NAS. (2017). Using 21st Century Science to Improve Risk-Related Evaluations. National Academies Press, Washington, DC.

NTP (National Toxicology Program). (2016). *Substances Listed in the Fourteenth Report on Carcinogens*. Available at: https://ntp.niehs.nih.gov/ntp/roc/content/listed_substances_508.pdf, last accessed April 27, 2018.

OEHHA. (2012). *Green Chemistry Hazard Traits for California's Toxics Information Clearinghouse*. Title 22, California Code of Regulations, sections 69401-69407.2. Available at: https://oehha.ca.gov/media/downloads/risk-assessment//gcregtext011912.pdf, last accessed July 20, 2018.

OEHHA. (2018). *Chemicals Known to the State to Cause Cancer or Reproductive Toxicity*. Title 27, California Code of Regulations, section 27001. Available at: https://oehha.ca.gov/proposition-65/proposition-65-list, last accessed December 16, 2018.

Reif, D. M., Martin, M. T., Tan, S. W., Houck, K. A., Judson, R. S., Richard, A. M., Knudsen, T. B., Dix, D. J., and Kavlock, R. J. (2010). Endocrine profiling and prioritization of environmental chemicals using ToxCast data. *Environ. Health Perspect.* **118**, 1714–1720.

Reif, D. M., Sypa, M., Lock, E. F., Wright, F. A., Wilson, A., Cathey, T., Judson, R. R., and Rusyn, I. (2013). ToxPi GUI: An interactive visualization tool for transparent integration of data from diverse sources of evidence. *Bioinformatics* **29**, 402–403.

Ribalta, T., McCutcheon, I. E., Aldape, K. D., Bruner, J. M., and Fuller, G. N. (2004). The mitosis-specific antibody anti-phosphohistone-H3 (PHH3) facilitates rapid reliable grading of meningiomas according to WHO 2000 criteria. *Am. J. Surg. Pathol.* **28**, 1532–1536.

Rotroff, D. M., Dix, D. J., Houck, K. A., Kavlock, R. J., Knudsen, T. B., Martin, M. T., Reif, D. M., Richard, A. M., Sipes, N. S., Abassi, Y. A., *et al.* (2013). Real-time growth kinetics measuring hormone mimicry for ToxCast chemicals in T-47D human ductal carcinoma cells. *Chem. Res. Toxicol.* **26**, 1097–1107.

Scott, I. S., Morris, L. S., Rushbrook, S. M., Bird, K., Vowler, S. L., Burnet, N. G., and Coleman, N. (2005). Immunohistochemical estimation of cell cycle entry and phase distribution in astrocytomas: Applications in diagnostic neuropathology. *Neuropathol. Appl. Neurobiol.* **31**, 455–466.

Silva, M., Pham, N., Lewis, C., Iyer, S., Kwok, E., Solomon, G., and Zeise, L. (2015). A comparison of ToxCast test results with *in vivo* and other *in vitro* endpoints for neuro, endocrine, and developmental toxicities: A case study using endosulfan and methidathion. *Birth Defects Res. B Dev. Reprod. Toxicol.* **104**, 71–89.

Sipes, N. S., Martin, M. T., Kothiya, P., Reif, D. M., Judson, R. S., Richard, A. M., Houck, K. A., Dix, D. J., Kavlock, R. J., and Knudsen, T. B. (2013). Profiling 976 ToxCast chemicals across 331 enzymatic and receptor signaling assays. *Chem. Res. Toxicol.* **26**, 878–895.

Sirenko, O., Grimm, F. A., Ryan, K. R., Iwata, Y., Chiu, W. A., Parham, F., Wignall, J. A., Anson, B., Cromwell, E. F., Behl, M., *et al.* (2017). *In vitro* cardiotoxicity assessment of environmental chemicals using an organotypic human induced pluripotent stem cell-derived model. *Toxicol. Appl. Pharmacol.* **322**, 60–74.

Skaland, I., Janssen, E. A., Gudlaugsson, E., Klos, J., Kjellevold, K. H., Søiland, H., and Baak, J. P. (2007). Phosphohistone H3 expression has much stronger prognostic value than classical prognosticators in invasive lymph node-negative breast cancer patients less than 55 years of age. *Mod. Pathol.* **20**, 1307–1315.

Smith, M. T., Guyton, K. Z., Gibbons, C. F., Fritz, J. M., Portier, C. J., Rusyn, I., DeMarini, D. M., Caldwell, J. C., Kavlock, R. J., Lambert, P. F., *et al.* (2016). Key characteristics of carcinogens as a basis for organizing data on mechanisms of carcinogenesis. *Environ. Health Perspect.* **124**, 713–721.

Tilley, S. K., Reif, D. M., and Fry, R. C. (2017). Incorporating ToxCast and Tox21 datasets to rank biological activity of chemicals at Superfund sites in North Carolina. *Environ. Int.* **101**, 19–26.

UNC (University of North Carolina at Chapel Hill). (2009). *ToxPi standalone GUI, Version 1.3 User Manual*. Available at: http://comptox.us/toxpi/User%20Manual%201.3.pdf, last accessed May 27, 2014.

U.S. EPA (United States Environmental Protection Agency). (2006). *Toxicology Disciplinary Chapter for the Re-Registration Eligibility Decision (RED) Risk Assessment*. Active Ingredient: Didecyl

dimethyl ammonium chloride (DDAC). Available at: https://www.regulations.gov/contentStreamer? documentId= EPA-HQ-OPP-2006-0338-0020&contentType=pdf, last accessed June 14, 2018.

U.S. EPA. (2008). *5-Chloro-2-(2, 4-dichlorophenoxy)phenol (Triclosan): Risk Assessment for the Reregistration Eligibility Decision (RED) Document*. Available at: https://www.regulations.gov/contentStreamer?documentId=EPA-HQ-OPP-2007-0513-0035&contentType=pdf, last accessed June 14, 2018.

U.S. EPA. (2011). *Emamectin Benzoate - Human Health Assessment Scoping Document in Support of Registration Review*. Available at: https://www.regulations.gov/contentStreamer?documentId= EPA-HQ-OPP-2011-0483-0004&contentType=pdf, last accessed June 14, 2018.

U.S. EPA. (2012). *Revised Summary of Human Health Effects Data for IPBC Registration Review*. Available at: https://www.regulations.gov/contentStreamer?documentId=EPA-HQ-OPP-2011-0420-0010&contentType=pdf, last accessed June 14, 2018.

U.S. EPA. (2014). *Zoxamide- Human Health Scoping Document in Support of Registration Review*. Available at: https://www.regulations.gov/contentStreamer?documentId=EPA-HQ-OPP-2014-0391-0005&contentType=pdf, last accessed June 14, 2018.

U.S. EPA. (2017). *Chemicals Evaluated for Carcinogenic Potential, Annual Cancer Report 2017*. Available at: http://npic.orst.edu/chemicals_evaluated.pdf, last accessed June 14, 2018.

U.S. EPA. (2018). *ToxCast Owner's Manual*. Available at: https://www.epa.gov/sites/production/files/2018-04/documents/toxcastownermanual4252018.pdf, last accessed December 5, 2018.

Wiegers, T. C., Davis, A. P., and Mattingly, C. J. (2014). Web services-based text-mining demonstrates broad impacts for interoperability and process simplification. *Database* **2014**, 1–16.

Yang, C., Tarkhov, A., Marusczyk, J., Bienfait, B., Gasteiger, J., Kleinoeder, T., Magdziarz, T., Sacher, O., Schwab, C. H., Schwoebel, J., *et al*. (2015). New publicly available chemical query language, CSRML, to support chemotype representations for application to data mining and modeling. *J. Chem. Inf. Model*. **55**, 510–528.

Yueh, M. F., Taniguchi, K., Chen, S., Evans, R. M., Hammock, B. D., Karin, M., and Tukey, R. H. (2014). The commonly used antimicrobial additive triclosan is a liver tumor promoter. *Proc. Natl. Acad. Sci. U.S.A*. **111**, 17200–17205.

Yun, X., Huang, Q., Rao, W., Xiao, C., Zhang, T., Mao, Z., and Wan, Z. (2017). A comparative assessment of cytotoxicity of commonly used agricultural insecticides to human and insect cells. *Ecotoxicol. Environ. Saf*. **137**, 179–185.

Zhang, Z., Zhao, X., and Qin, X. (2017). Potential genotoxic and cytotoxicity of emamectin benzoate in human normal liver cells. *Oncotarget* **8**, 82185–82195.