EPA Public Access

Author manuscript

Comput Toxicol. Author manuscript; available in PMC 2019 May 01.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Comput Toxicol. 2018 May ; 6: 71–83. doi:10.1016/j.comtox.2017.10.003.

A Workflow for Identifying Metabolically Active Chemicals to Complement *in vitro* Toxicity Screening

Jeremy A. Leonard^a, Caroline Stevens^b, Kamel Mansouri^{a,c,d}, Daniel Chang^e, Harish Pudukodu^f, Sherrie Smith^f, and Yu-Mei Tan^{f,*}

^aOak Ridge Institute for Science and Education, Oak Ridge, TN, USA;

^bNational Exposure Research Laboratory, United States Environmental Protection Agency, Athens, GA, USA;

^cNational Center for Computational Toxicology, United States Environmental Protection Agency, Research Triangle Park, NC, USA;

^dScitoVation LLC, Research Triangle Park, NC, USA;

^eOffice of Pollution and Prevention of Toxics, United States Environmental Protection Agency, Washington, D.C., USA;

^fNational Exposure Research Laboratory, United States Environmental Protection Agency, Research Triangle Park, NC, USA

Abstract

The new paradigm of toxicity testing approaches involves rapid screening of thousands of chemicals across hundreds of biological targets through use of *in vitro* assays. Such assays may lead to false negatives when the complex metabolic processes that render a chemical bioactive in a living system are unable to be replicated in an *in vitro* environment. In the current study, a workflow is presented for complementing *in vitro* testing results with *in silico* and *in vitro* techniques to identify inactive parents that may produce active metabolites. A case study applying this workflow involved investigating the influence of metabolism for over 1,400 chemicals considered inactive across 18 *in vitro* assays related to the estrogen receptor (ER) pathway. Over 7,500 first-generation and second-generation metabolites were generated for these *in vitro* inactive chemicals using an *in silico* software program. Next, a consensus model comprised of four individual quantitative structure activity relationship (QSAR) models was used to predict ER-binding activity for each of the metabolites. Binding activity was predicted for ~8–10% of metabolites in each generation, with these metabolites linked to 259 *in vitro* inactive parent chemicals. Metabolites were enriched in substructures consisting of alcohol, aromatic, and phenol bonds relative to their inactive parent chemicals, suggesting these features are potentially favorable for ER-binding. The workflow presented here can be used to identify parent chemicals that can be potentially bioactive, to aid confidence in high throughput risk screening.

*To whom correspondence should be addressed: 109 T.W. Alexander Dr. Mail Code E205-01, Research Triangle Park, NC 27709; tan.cecilia@epa.gov, leonard.jeremy@epa.gov, stevens.caroline@epa.gov, kmansouri@scitovation.com, chang.daniel@epa.gov.

Publisher's Disclaimer: Disclaimer: The U.S. Environmental Protection Agency has provided administrative review and has approved this paper for publication. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the U.S. Environmental Protection Agency.

Keywords

metabolite prediction; QSAR; ER; *in vitro* screening

1. Introduction

The traditional toxicity testing paradigm involving use of animal studies requires a large investment in time and resources [1,2]. Therefore, the thousands of environmental chemicals currently in commerce [3] and the hundreds of chemicals that are registered on an annual basis [4] render such a low-throughput chemical-by-chemical approach unsustainable. Moreover, extrapolating toxicity results from animals dosed with a specific chemical at high concentrations to relevant human health outcomes arising from much lower concentrations often involves many uncertainties [1,2].

Recognizing these issues, the National Toxicology Program proposed a new “roadmap” for toxicity testing in the 21st century that involved development of rapid profiling strategies to reduce or refine the use of animal studies while remaining scientifically sound and promoting human and animal welfare [5]. High-throughput (HT) *in vitro* screening assays arose from this initiative to act as a means for rapidly investigating the effects of thousands of chemicals across hundreds of biological endpoints linked to disease outcomes relevant to both human and ecosystem health [6,7]. These screening results from HT assays can be used to prioritize chemicals for more extensive *in vivo* testing [8,9].

While *in vitro* assays have significantly expanded the field of toxicity testing, interpreting the results of such studies is not without challenges. False positives may arise for chemicals that are active in *in vitro* assays through means other than direct binding to an intended technological target, or by altering pathways that are not specifically related to that target (e.g., dyes that interfere with an assay relying upon fluorescent signal for detection) [10–12]. A false positive signal may also stem from non-selective cell cytotoxicity, rather than from chemical interaction with the intended enzyme or receptor targets [13]. A high incidence in the number of false positives resulting from *in vitro* testing can lead to an unnecessary waste of time and cost, even though some risk managers may be less concerned with false positives in order to avoid incorrectly labeling some chemicals as safe [14].

Much of the activity caused by endocrine disrupting chemicals (EDCs) can be tested within numerous *in vitro* assays that are part of the US Environmental Protection Agency’s Toxicity Forecaster (ToxCast) HT screening program [15]. EDCs, such as estrogenic chemicals, are of concern due to permanent adverse health outcomes that may arise from perturbations in signaling processes during early-life exposures, at a period when proper communication among cells is critical for normal development [16,17]. Adverse outcomes resulting from perturbations on pathways related to estrogen receptor (ER) binding can be investigated for each of the relevant individual ToxCast assays, but interpreting results from multiple assays as a whole may help screen out false positives due to assay interference or certain characteristics in each assay’s technological design. A recent study attempted to reduce such false positives by developing an orthogonal model to integrate results across a total of 18 assays that are related to perturbations on the ER pathway [18].

Of greater concern, however, is the chance that some chemicals may register as negatives in *in vitro* assays while possessing the ability to induce adverse health outcomes under *in vivo* conditions [19]. False negatives, like false positives, can be the result of assay design or interference, as chemicals that are active *in vivo* may have difficulty interacting with their intended *in vitro* technological target [20]. False negatives can also result from pharmacokinetic behaviors that are unable to be replicated in most *in vitro* assays. Specifically, a chemical found inactive in an *in vitro* assay may undergo bioactivation in an *in vivo* environment to a metabolite capable of perturbing an *in vivo* target [21].

Thousands of putative metabolites, as well as their bioactivities and chemical properties or features mediating those bioactivities, can be predicted through use of computational methods in order to avoid the cost, time, and difficulty associated with deriving such information through *in vitro* or *in vivo* means [22,23]. Though numerous existing studies employ cheminformatics and docking approaches to create models for predicting chemical binding activity, only few consider the binding affinities of those chemicals' metabolites as well [24–27]. Here, we propose an approach that integrates *in silico* model predictions of metabolites with quantitative structure activity relationship (QSAR) models predicting activity for a specific biological endpoint, in order to identify *in vitro* inactive parents that may have *in vivo* active metabolites. The utility of this approach is demonstrated in a case study that first predicts metabolites for all chemicals considered inactive in the aforementioned orthogonal model for 18 ER-related assays [18], followed by use of a consensus QSAR model that identifies chemical structural features influencing interactions with the ER and predicts ER binding potential. Metabolites that are predicted to be active are then linked to their respective inactive parents, and these parents may in turn undergo further *in vitro* analyses to confirm generation of their respective active metabolites, followed by evaluation of those metabolites' bioactivities. This workflow is proposed to complement HT *in vitro* toxicity testing methods through its ability to identify chemicals with the potential for bioactivation in a living system, for more accurate prioritization of such chemicals during risk screening and chemical prioritization.

2. Methods

2.1 Workflow Overview

Our workflow (Figure 1) to investigate the potential influence of metabolism on interpretation of *in vitro* toxicity testing results begins by identifying chemicals that show no activity in *in vitro* assays. Next, metabolites are predicted *in silico* for these inactive parent compounds, followed by *in silico* predictions of metabolite activity for the biological endpoint of interest. From this point forward in the current study, for simplicity, the metabolites will be labeled as “active” or “inactive”, though it should be understood that both the identities and activities of metabolites were predicted using *in silico* methods only. Active metabolites can then be linked to their respective *in vitro* inactive parents in order to identify those parent chemicals that may potentially become bioactive in a living system. A qualitative screening approach proposed previously [28] can be used to prioritize those *in vitro* inactive parents that possess the necessary exposure and absorption potentials that allow them to reach a site of metabolism and generate active metabolites capable of

perturbing the designated molecular target. As it is likely that several predicted metabolites are also commercially available, the activity of these metabolites can be confirmed through *in vitro* methods. For those metabolites that are not commercially available, the probability of being endogenously generated by their inactive parents may be confirmed through *in vitro* metabolism studies. Those metabolites that are generated in the metabolism studies can then be analyzed in *in vitro* assays in order to confirm those metabolites' suspected activity. While the totality of this workflow (Fig. 1) should allow for a more definitive identification of *in vitro* inactive parents capable of being bioactivated, our case study progresses only through *in silico* prediction of metabolite activity due to the limited resources available to us for conducting the *in vitro* studies necessary to confirm the production and activity of those metabolites. Further details of the approaches specific to this case study are provided below.

2.2 Chemical Selection

Chemicals selected for the current case study were those 1,677 ToxCast compounds determined by the orthogonal ER model to be inactive [18]. Briefly, the orthogonal ER model incorporates the activity results of the chemicals within each of the 18 ToxCast ER-related assays (e.g., receptor binding, dimerization, DNA binding, and protein production or cell proliferation), while also taking into account the possibility of cytotoxicity contributing to false positives. The model was validated using 45 positive and negative reference chemicals. Further details regarding the development and results of the orthogonal ER model are provided elsewhere [18]. Out of the 1,677 ToxCast chemicals tested across the 18 ER-related *in vitro* assays, 1,440 (~86%) were found to be inactive using the scoring algorithm for the model. These 1,440 inactive parents were then used in the next step to predict their potential metabolites.

2.3 Metabolite Predictions

The structures of the selected chemicals were processed through removal of salts, energy minimization, and removal of inorganic counter-ions prior to inputting these chemical's simplified molecular input line-entry system (SMILES) into ChemAxon's Metabolizer application (version 16.9.5.0; ChemAxon LLC, Cambridge, MA USA) to predict two generations of metabolic products. This software was selected due to its ability to eliminate metabolite candidates considered unstable or with slow generation rates, by estimating transformation rates across a large number of potential reaction schemes. Additionally, the core function of the Metabolizer makes use of a library of 159 reaction schemes that encode how structural fragments within a molecule are subject to phase I metabolism in humans, whereas other software often utilize non-human enzymatic reactions to predict metabolites. The library contains nearly 900 examples of its given reaction schemes from literature. The ChemAxon Metabolizer outputs four metrics for each chemical entered: degradation rate, formation rate, production ratio, and an accumulation index in order to determine the transformation rates for that chemical. The transformation rate of each reaction scheme is compared to the rates of other schemes existing within the library to assign a rank between one and seven. Briefly, a formation rate is derived from its exponential relationship with the rank and a given constant according to the equation:

$$F = \exp^{(R \times 1.94591)} - 1$$

where F is the rate of formation and R is the rank for the reaction scheme of interest. Rank 1 is the most minor (slowest) reaction scheme and rank 7 is the most major (fastest) scheme. Degradation rate of the parent compound is considered to be the sum of formation rates for each transformation product. Production is the ratio of formation and degradation calculated for each transformation product and aids in identifying the transformation pathway(s) generating the most metabolite products. The accumulation index is a metric that describes the difference between the formation of a metabolite and its degradation to a second transformation product and indicates the metabolites likely to be formed and to remain stable.

ChemAxon's Metabolizer includes a built-in function that limits the number of predicted metabolites by focusing on the major metabolic pathways that are most likely to occur. This approach involved allowing the ChemAxon Metabolizer to be executed with the option to ignore a product for which the estimated production rate was less than 10% of the maximum production rate (i.e., the production rate of the major metabolite) of all products within that generation of metabolism. Additionally, we removed those products that formed from the slowest reaction schemes (i.e., those with a rank of 1 or 2) from the set of generated metabolites in our case study. It was assumed that the metabolites produced by these two slower reaction schemes, relative to the parent compound's residence time in the body, would be insignificant.

2.4 Activity Predictions

The structures of first- and second-generation metabolites were made QSAR ready through a standardized process. Briefly, (1) the integrity of the structural information required to render the molecules was verified, and invalid entries were corrected by retrieving structures from online chemical databases (e.g., PubChem, ChemSpider, etc.) or removed if ambiguous; (2) those structures with a molecular weight greater than 1,000 g/mol were removed to facilitate descriptor calculations; (3) tautomers were transformed, standardized, and neutralized by removing stereochemistry; and (4) explicit hydrogen atoms were added. Further details regarding processing of a chemical to render it QSAR-ready can be found elsewhere [29]. QSAR-ready duplicate structures were found and removed by assigning each structure an IUPAC (International Union of Pure and Applied Chemistry) InChI (International Chemical Identifier) code that acts as an unequivocal unique identifier for that metabolite.

The ER binding activity of first- and second-generation metabolites was predicted using four QSAR models that were available to us. These four models were obtained from the *Collaborative Estrogen Receptor Activity Prediction Project* (CERAPP), in which 48 continuous and categorical models were developed by 17 institutes for predicting ER binding, agonist, and antagonist activity for over 30,000 chemicals [29]. The four QSAR models selected for our case study were developed by Lockheed Martin, by the National Center for Computational Toxicology (NCCT), by the Laboratory of Cheminformatics at the

University of Strasbourg, France (UNISTRA; <http://infochim.u-strasbg.fr/webserv/VSEngine.html>), and with the publicly available Online Chemical Database with Modeling Environment (OCHEM) web platform (<https://ochem.eu/home/show.do>). The training set for these models consisted of the results from the orthogonal ER model, and included 45 positive and negative reference ER compounds and a variety of chemicals with varying degrees of estrogen-like activity [18]. The CERAPP models were then evaluated using a wide array of literature data consisting of ~44,600 experimental entries from *in vitro* studies for over 7,500 chemicals, in addition to inclusion of the original ToxCast chemicals used for the training set. Briefly, summary scores for evaluation were calculated using the following equation:

$$Score = \frac{1}{2} (BA_{ToxCast} + BA_{Corrected}) \quad 2$$

where $BA_{ToxCast}$ is the balanced accuracy of the models when predicting binding for chemicals in the evaluation set based on the ToxCast chemicals, and $BA_{Corrected}$ is the balanced accuracy of the models when predicting binding for chemicals in the evaluation set after removing those chemicals (1) with only one literature source; (2) with very weak activity; and (3) that fall outside of the applicability domain of the training set. Further details regarding evaluation of these activity models and their performance can be found elsewhere [29], where authors calculated summary scores based on balanced accuracy (i.e., average of sensitivity and specificity for each model). Sensitivity (i.e., the true positive rate) can be determined using the equation:

$$TPR = TP/(TP + FN) \quad 3$$

where TP is the number of true positives (e.g., active chemicals predicted to be active) and FN is the number of false negatives (e.g., active chemicals predicted to lack activity). Specificity (i.e., the true negative rate) can be determined using the equation:

$$TNR = TN/(TN + FP) \quad 4$$

where TN is the number of true negatives (e.g., inactive chemicals predicted to be inactive) and FP is the number of false positives (e.g., inactive chemicals predicted to have activity). Scores ranging from 0.69 to 0.85 indicated high predictive capabilities for the CERAPP models, and the summary scores of the four models used in the current study ranged from 0.70 to 0.78 (Figure S1). It should be noted that these four models, as used in our case study, only predicted a qualitative designation of ER binding (i.e., yes/no), rather than a more detailed estimate of agonist or antagonist activity. We developed a consensus QSAR model, which employed a majority-rules approach, to determine whether a metabolite might exhibit ER-binding activity based on the predictions from each of the four individual models. Using our consensus model, each predicted metabolite was assigned to one of the following three binding activity categories: (1) metabolites whose binding activities were predicted by at least three models; (2) metabolites predicted to have no binding activity by at least three

models; and (3) metabolites whose binding activities are uncertain due to an equal distribution among predictions (e.g., two models predicted binding while two others did not).

2.5 Chemotype Analysis

The application Chemotyper (version 1.0; Molecular Networks; Erlangen, Germany) was used to identify chemical substructures (i.e., chemotypes) that contribute to ER binding capability for 202 ToxCast chemicals that were determined to be active both *in vitro* [18] and *in silico* [29]. The SMILES strings of these 202 *in vitro* and *in silico* active chemicals were entered into the Chemotyper for comparison against a library of ToxPrint chemotypes. The 729 ToxPrint chemotypes are a dataset developed by the US Food and Drug Administration, Center for Food Safety and Applied Nutrition, and Office of Food Additive Safety, and are publically available at <https://www.toxprint.org>. The chemotypes common to a majority of the 202 *in vitro* and *in silico* active parents (i.e., in at least 35% of those parent chemicals) were identified to establish chemotype enrichment. These enriched chemotypes were then queried within four groups of chemicals to determine the number of chemicals within each group that possessed such chemotypes. These four groups were (1) *in vitro* inactive parents without active metabolites; (2) *in vitro* inactive parents with active metabolites; (3) active first-generation metabolites; and (4) active second-generation metabolites.

3. Results

3.1 Metabolite Predictions

As noted above, the first step within the proposed workflow (Fig. 1) involves *in silico* prediction of metabolites for those chemicals found to be inactive *in vitro*. Detailed information regarding predicted metabolites for chemicals used in the current case study and the respective transformation pathways leading to those structures is available upon request. From the 1,440 *in vitro* inactive parent structures entered into the ChemAxon Metabolizer, a total of 15,369 first- and second-generation metabolites were predicted, a quarter of which were first-generation metabolites. Over 200 first-generation metabolites and over 500 second-generation metabolites were excluded from further analysis due to their production through slow reaction schemes (i.e., those schemes with a rank of 1 or 2). After eliminating duplicate entries based on the unique InChI codes, 7,536 unique metabolites remained. Again, just over one quarter were first-generation metabolites, with the remainder being second-generation metabolites. Also, out of the 7,536 metabolites, 753 (~10%) were found in both the first and second generations for either the same parent compound or for different parents.

3.2 Activity Predictions

The next step of the proposed workflow (Fig. 1) involves prediction of metabolite activity, and the predictions of metabolite activity by the consensus ER binding model are also available upon request. Over 75% of the generated metabolites were predicted to have no binding activity (5,941), while less than 10% were predicted to be potential ER binders (719), and a definitive conclusion for the remainder was unable to be made (876). Generally,

the number of second-generation metabolites within each of the three binding activity categories was 60–68% of the total for that category, with the number of first-generation metabolites comprising 24–28% for that category, and the remaining 6–10% were metabolites found across both generations (Fig. 2). Of the 2,101 first generation-only metabolites, 173 were active (~8%); of the 4,682 second-generation metabolites, 480 were active (~10%). Of the 756 metabolites that were found within both generations, 66 were active (~8%).

It was found that 55 active metabolites were also in the original test set of the 30,000 environmental compounds analyzed by the CERAPP models [29]. Presence of these 55 metabolites across both studies allowed for an opportunity to evaluate the performance of our consensus model that was based on four individual models by comparing its predictions to those outputted by the CERAPP consensus model that was based on 48 individual models (Table 1). Of the 20 first-generation-only active metabolites identified using our consensus model, four were predicted by CERAPP to be inactive but labeled as “suspicious”, while the remaining 80% showed agreement between the two models regarding predicted activity. Similarly, of the 24 second-generation-only active metabolites identified using our consensus model, four were predicted by CERAPP to be inactive but three of them were labeled as “suspicious”, and the predicted activity of the remaining 83% showed agreement between the two models. For the 11 active metabolites that are in both the first and second generations, four were predicted by CERAPP to be inactive but labeled as “suspicious”, and predicted activity of the remaining 64% agreed between the two models. At first glance, the average error rate of our consensus model, which is a subset extracted from the larger number of CERAPP models, is ~24% when compared to the consensus estimate assigned using the full set of CERAPP models. However, in light of CERAPP’s labeling of all of these contradicting predictions but one as suspicious, the resulting error rate is reduced to ~5% (i.e., only three predictions from our consensus model disagreed with CERAPP suspicious consensus predictions aligning with literature information of no binding). Furthermore, when comparing our consensus predictions with only the 44 definitive CERAPP consensus binding predictions (e.g., weak, moderate, inactive) and removing the suspicious entries from consideration, there was nearly complete agreement, at 98%, between the two models.

3.3 Predicted Transformation Pathways for Metabolites

After linking the 719 active metabolites to their respective parents, it was found that 259 of the 1,440 *in vitro* inactive parent chemicals have the potential to generate at least one active metabolite. Among those 259 parents, 23 have only first-generation active metabolites, while 127 have only second-generation active metabolites. The rest of the 109 parents have both first- and second-generation active metabolites (Fig. 2). The identities of the 259 parent compounds, along with their respective active metabolites in each generation are available upon request. First-generation active metabolites were formed through 33 unique transformation pathways in ChemAxon’s Metabolizer, while 42 unique transformation pathways were used to form second-generation active metabolites.

The dominant pathways within the first generation were phenyl hydroxylation (formation rank of 5), ketone reduction (formation rank of 6), phenol hydroxylation (formation rank of 3), benzylic hydroxylation (formation rank of 6), and O-dealkylation (formation rank of 5). The percentage of metabolites predicted to be formed through these pathways was 22%, 11%, 9%, 8%, and 8%, respectively (Fig. 3). The dominant pathways within the second-generation were similar, with a few exceptions. These pathways, in order of importance, included phenyl hydroxylation for 19% of the metabolites, phenol hydroxylation (formation rank of 3) for 14%, O-dealkylation for 7%, dehydration (formation rank of 5) for 6%, and ketone reduction for 6%. Other important transformation pathways within the second generation included carbonyl alpha hydroxylation (formation rank of 5), 1-hydroxylation (formation rank of 3), primary alcohol oxidation (formation rank of 6), hydroxy alpha hydroxylation and hydroxy beta hydroxylation (both with formation ranks of 3), and epoxide rearrangement (formation rank of 7). Although a higher number of metabolites were predicted to be formed through these additional pathways, compared to benzylic hydroxylation (only 3%), those pathways are primarily responsible for second-generation products, and so only benzylic hydroxylation is shown in figure 3 due to its overlap between generations.

3.4 Activity of Parent Compounds

A summary of the number of ToxCast parent compounds separated into active and inactive categories determined by either *in vitro* [18] or *in silico* [29] means, as well as the number of *in vitro* inactive chemicals predicted to have active or inactive metabolites, is presented below (Fig. 4). Among the 237 *in vitro* active parents, the CERAPP model predicted 202 to be active and 35 to be inactive (15% disagreement). It was these 202 chemicals that were used to determine chemotype enrichment contributing to ER binding activity. Among the 1,440 *in vitro* inactive parents, the CERAPP model predicted 1,400 to be inactive and 40 to be active (3% disagreement). This set of 1,440 chemicals included both the 1,181 *in vitro* inactive parents with inactive metabolites and the 259 *in vitro* inactive parents with active metabolites. Interestingly, the CERAPP model predictions of probable activity disagreed with *in vitro* activity estimates of 15 chemicals for the first group ($15/1,181 = 1\%$ disagreement) and *in vitro* activity estimates of 25 chemicals for the second group ($25/259 = 10\%$ disagreement), respectively. The higher chance of the CERAPP model to estimate binding activity for an inactive parent with an active metabolite suggests that these parent chemicals are comprised of structural features favorable for ER-binding, which are captured by the CERAPP model.

Literature information concerning *in vivo* estrogen-like activity, including weak ER binding and cell proliferation [30], and *in vitro* studies other than ToxCast was available for 60 of the 259 *in vitro* inactive parent chemicals with active metabolites (Supplemental File 1). It should be noted, however, that for some measured *in vivo* endpoints, it was difficult to determine whether a chemical causes direct estrogenic effects or is acting through upstream mechanisms of action (e.g., aromatase inhibition) that might induce adverse reproductive effects in general. In the second case, a chemical may perturb an adverse outcome pathway (AOP) that leads to the same adverse reproductive effects as AOPs that involve ER binding. Other *in vivo* endpoints included changes in physiological characteristics often ascribed to

estrogenic compounds, such as increase in uterine weight [31]. In these cases, estrogen-like activity was assumed to occur. Of the 60 *in vitro* inactive parent chemicals with available literature information, 27 have *in vitro* data only, 16 have *in vivo* data only, and the remaining 17 have both *in vitro* and *in vivo* data. Thirteen chemicals out of 27 (48%) that have other *in vitro* data (i.e., results from toxicity tests separate from the ToxCast screening program) were found to be active in at least one study, while only 3 out of 17 (18%) having *in vivo* data were found to be active in their respective *in vivo* studies (Table 2). Two chemicals were observed to be inactive in both *in vitro* and *in vivo* studies- methyl salicylate and 1,2,4-trichlorobenzene; Seven other chemicals were observed to be active in both *in vitro* and *in vivo* studies. The remaining 8 chemicals were found to be active in other *in vitro* studies but lack activity *in vivo*.

Literature observations were available for eight of the 25 *in vitro* inactive parent chemicals that were both predicted by the CERAPP consensus model to be active and to have active metabolites. Five of these chemicals were found to be active in other *in vitro* studies, while dodecyl gallate and 4-ethylphenol were found to be inactive in other *in vitro* studies, and lack of activity was observed for amiodarone hydrochloride *in vivo* (Table 2). Finally, 15 *in vitro* inactive parent chemicals with inactive metabolites were themselves predicted by the CERAPP consensus model to be active, and literature information was available for 7 of these (Supplemental File 1). Three of these parent chemicals were observed as active in separate *in vitro* studies, while an additional two were observed as active *in vivo*. This result indicates that potential errors may exist in orthogonal ER model designation of these five chemicals as being inactive[18], and/or in our current *in silico* approach that predicted no active metabolites.

3.5 Chemotype Analysis

The set of 202 *in vitro* and *in silico* active ToxCast chemicals was comprised of 303 ToxPrint chemotypes of the 729 investigated. Seven chemotypes most frequently found in the structures of these active parents were (1) an aromatic benzene ring (86%); (2) a generic alcohol bond (58%); (3) a carbon-carbon bond at C1 of the benzene ring (56%); (4) a generic carbonyl bond (51%); (5) an alcohol aromatic bond (45%); (6) an alcohol aromatic phenol bond (45%), and (7) long chain carbon bonds (43%) (Fig. 5). Along with the five groups of chemicals listed in the Methods section (i.e., the *in vitro* active set, the two *in vitro* inactive sets, and two metabolite sets), the 25 *in vitro* inactive chemicals predicted by the CERAPP consensus model to be active and having active metabolites were also included to examine whether specific chemotypes are enriched in certain groups of chemicals. There were no apparent differences among chemical sets for chemotype 4, other than slightly lower enrichment in both the *in vitro* active parent set and set of 25 chemicals (Figure 5). Chemotypes 1 and 3 were also similar among all chemical sets besides a lower enrichment in the set containing *in vitro* inactive chemicals with inactive metabolites. Chemotype 7 was most enriched in the set of 25 chemicals. There was a high degree of enrichment in chemotype 2 in the sets of both first- and second-generation active metabolites, followed by the set of 25 chemicals and set of *in vitro* active parents. Chemotypes 5 and 6 were enriched in a similar manner for both first- and second-generation active metabolites and the set of 25 chemicals, followed by the set of *in vitro* active parents. Relative to the other chemical sets,

the sets of *in vitro* inactive parents with active metabolites and *in vitro* inactive parents with inactive metabolites were the least enriched in the alcohol-related chemotypes (i.e., chemotypes 2, 5, and 6). The set of *in vitro* inactive parents with inactive metabolites was generally the least enriched among all sets in the aromatic-related chemotypes (i.e., chemotypes 1, 3, and 6). Out of interest, an ER structural alert filter [32] was applied to the six chemical sets to identify those chemicals possessing any of the ~24 structural alert features indicating estrogenic activity, for comparison against the enriched chemotypes. Of the seven enriched chemotypes, three were present in a majority of the chemicals having structural alerts related to phenol, carbonyl, and alcohol bond substructures.

4. Discussion

4.1 A Workflow for Identifying Potentially Bioactive Compounds

As toxicity testing shifts away from *in vivo* studies and more towards new approach methods such as high-throughput *in vitro* screening assays, integration of data related to a chemical's exposure, pharmacokinetics, biological pathways, and dose-response is critical for providing a mechanistic understanding of potential health effects than can arise from real-world exposures to that chemical. It is often the case that complex metabolic processes cannot be captured within many *in vitro* assays, and removing some *in vitro* inactive parents from consideration of further testing may pose an unexpected risk when those parent chemicals are capable of becoming bioactivated in a living system. Now more than ever, as chemical inventories continue to expand, there is emphasis on development of international standardized guidelines for inclusion of metabolic and toxicokinetic methodologies in *in vitro* tests [21,33]. In seeking to address this challenge, the current paper presents a workflow for identifying those chemicals with the potential of generating active metabolites, using a combination of *in vitro* and *in silico* approaches (Fig. 1).

4.2 Limitations of *in silico* and *in vitro* Approaches

There are often discrepancies that exist between *in vitro* or *in silico* results and *in vivo* outcomes. One reason is that *in vitro* assays and *in silico* models often are capable of evaluating chemical effects on only a single molecular target at a time. The concept of AOPs arose as a means for interpretation of *in vitro* screening results in the context of the molecular initiating event (MIE) of an AOP [34,35]. The MIE relevant to the current case study, binding of chemicals to the ER, can lead to any number of adverse outcomes at the individual or population level, such as reduced female fecundity, malformations in gonad development, reduced sexual behavior, and population decline [36–38]. On the other hand, these outcomes may be linked to AOPs that do not include ER binding due to the existence of AOP networks. The chemical-agnostic nature of AOPs allows their key components to be reusable, so that many biological processes (i.e., key events) and the relationships among those processes are common throughout the network itself, while certain processes are unique to a specific AOP. As such, and because AOPs represent events occurring in living systems, one adverse outcome may be the result of different MIEs. Although a chemical does not bind to the ER, it can progress through a separate AOP and result in a similar outcome, such as reduced female fecundity. Therefore, some chemicals may cause

reproductive effects *in vivo*, though they truly are inactive both *in vitro* and *in silico* in regards to ER binding.

Another reason that *in vitro* and *in silico* results may disagree with those from *in vivo* studies involves problems with assay or model design. Problems with inter-assay variability can lead to erroneous interpretation of *in vitro* results, though the orthogonal assay design approach, such as the ER model used in the current study [18], should limit this error. The predictive ability of individual *in silico* models, such as those developed by members of the CERAPP project, may be impacted by data availability for training and evaluation of those models, or even limitations in the modeling approaches themselves. The purpose of the consensus CERAPP model was to overcome the limitations of single models, and although the members of the CERAPP project observed similar performances among their individual models, those models with a narrower domain of applicability performed slightly better [29]. Unfortunately, not all 48 models developed for the CERAPP project were available to us for predicting activity of metabolites generated *in silico*. Thus, the scoring algorithm used to determine consensus among individual models in the CERAPP project [29] was not applicable for the small subset of four models used in the current study, due to bias that could result from one or two of the models having a much lower or higher score relative to the others. While the majority-rules approach taken for our analysis is a reasonable alternative, one major limitation with only four models is that activity is inconclusive when there is an even split among model predictions (i.e., two models predicting activity and two models predicting inactivity).

The resulting uncertainty in activity predictions for nearly 900 metabolites might lead to their parents being missed when selecting chemicals for additional testing, though it is acknowledged that the conservative approach of retaining these parents as well could result in the unnecessary use of resources. To address this issue, the full set of CERAPP models could be used in future investigations to increase confidence in predictions and to provide a more definitive determination of activity for these predicted metabolites. However, the overall agreement in predictions between the small model subset used in the current study and the CERAPP consensus model suggests that the former is sufficient for demonstrating the utility of activity predictions within the workflow.

In addition to possible limitations of the four individual models used in the current study, the particular types of metabolic schemes used to generate predicted metabolites may have also influenced those metabolites whose activity was difficult to predict. Thus, the metabolic schemes predicted to generate inconclusive and inactive metabolites were compared to those that were predicted to generate active metabolites. The number of schemes generating inconclusive metabolites were only slightly more than those generating active metabolites (37 vs. 33 and 52 vs. 42 for first- and second-generation metabolites, respectively). The types of metabolic schemes generating inconclusive and active metabolites were similar, with the exception of schemes such as aromatic hydroxylation, primary alcohol oxidation, and ester hydrolysis generating a greater percentage of inconclusive metabolites than did phenyl or phenol hydroxylation. These last two schemes appeared to generate metabolites that were more likely to be active. Differences were even more pronounced with the number of metabolic schemes generating first-generation inactive metabolites (89) and second-

generation inactive metabolites (100). Additionally, the schemes generating the majority of these inactive metabolites included oxidation of primary and secondary alcohols, dehydration, and ester hydrolysis, with less than 1% of inactive metabolites generated through phenyl oxidation or phenol oxidation.

The hydroxylated phenyl and phenol groups that are produced through the primary metabolic schemes generating active metabolites (Fig. 3) can form strong hydrophobic bonds with active site residues of the ER, in addition to a potent hydrogen bond with aspartic acid 353 [23]. It is such metabolic transformations that likely led the consensus QSAR model comprised of the four individual models to predict these metabolites as active. This assumption is further supported by the enrichment of aromatic alcohols and phenols, and especially generic alcohol bonds, in the two generations of active metabolites, compared to the lower presence of these chemotypes in their respective parents. In contrast, the metabolic schemes such as primary and secondary alcohol oxidation form aldehydes and ketones, whose carbonyl groups might decrease binding affinity by reducing hydrophobic contacts [23]. Schemes such as dehydration remove the hydroxyl group, and replacement with alkyl groups such as a methyl or ethyl group can also reduce binding affinity to the ER [39].

The orthogonal ER model and the CERAPP model evaluate consensus among multiple *in vitro* and *in silico* results, respectively [18,29]. The abundance of literature information for estrogenic activity of chemicals and the variety of ER-related *in vitro* assays contributed significantly to the development of the two aforementioned models. Often, such a luxury is not available when investigating other toxicological endpoints. The value of the workflow presented in this work lies in its ability to remain applicable in cases where only one to a few *in vitro* assays are available or where information from literature is scarce for a chemical set or toxicological endpoint of interest. Though it is acknowledged that confidence in results will be lower in such cases, our workflow facilitates rapid identification of chemicals capable of triggering an MIE after bioactivation, for preliminary screening.

A recent publication by Pinto and colleagues attempted to address the issues of metabolism in *in vitro* studies in a similar manner as the current study [27]. The purpose of that publication was to quantitatively predict ER binding activity, using the relationship between scoring results in the orthogonal ER model [18] and CERAPP model predictions, of metabolites for chemicals determined in literature to produce active metabolites. As such, authors in that study used a much smaller and defined chemical set of 50 parent chemicals, 38 of which are known to produce estrogenic metabolites [27], for generating metabolites *in silico* and predicting activity for those metabolites through use of a subset of CERAPP models. Although the models used in that study were developed by the same CERAPP institutions as those that developed the models in the current study (i.e., OCHEM, UNISTRA, and Lockheed Martin), the former are models based on agonist activity [27], while the ones used here are based on binding activity. Additionally, in the Pinto study, a different software program (ADMET Predictor) was used to predict metabolites [27]. As 33 chemicals overlap both studies, comparisons can be made regarding the performance of our current model compared to predicted and measured activity provided by the alternate study.

Of the 33 parent chemicals considered positive for bioactivation *in vivo* and predicted to have estrogenic metabolites in the Pinto et al. study [27], only two disagreed with activity predictions in the current study. In the current study, azobenzene, which is often used in the creation of dyes [40], was placed into the set of *in vitro* inactive chemicals with inactive metabolites because the active metabolite that it generates, 4-hydroxyazobenzene, had a formation rank below 3 predicted by the ChemAxon Metabolizer, suggesting its low production or rapid elimination. Azobenzene was shown to be active in at least one rat uterotrophic assay in the literature [41], and thus even though the ChemAxon Metabolizer does not predict an active metabolite to reside in the body long, it does so in a sufficient manner to increase uterine weight in rats *in vivo*. Alternatively, another metabolite not predicted by the ChemAxon Metabolizer may be generated *in vivo* and lead to the observed activity in the rat uterotrophic assay. In our study, only four chemicals were assigned to the set of *in vitro* inactive parents with inactive metabolites due to the low formation ranks of their respective metabolites, and so this error was likely small.

The other chemical showing differences between the current study and the Pinto et al. study was the polycyclic aromatic hydrocarbon, phenanthrene. In our study, this chemical was placed into the set of *in vitro* inactive chemicals with inactive metabolites because of definitive non-binding predicted for three possible metabolites and inconclusive binding for four others. The predicted scores for phenanthrene and its metabolites suggest that it is only a very weak ER agonist *in vitro* but generates metabolites with at least weak to moderate agonist activity [27]. This, in addition to the potential for bioactivation in rat uterotrophic assays [41], reveals the limitation of having inconclusive outcomes from the 4-model approach that might result in the omission of potentially active metabolites. To investigate the reasoning behind the discrepancies for phenanthrene between the current study and the Pinto et al. study [27], the SMILES for all predicted metabolites of phenanthrene generated by the ChemAxon Metabolizer were inputted into the open-source name-structure site Chemical Entities of Biological Interest (ChEBI; <https://www.ebi.ac.uk/chebi/>) to determine those metabolites' identities. Two of the metabolites returned were 9-phenanthrol and 2-phenanthrol, which were the two predicted by Pinto and colleagues to have some agonist activity [27], but were predicted in the current study as being inconclusive. A consensus model containing a larger number of models for predictions could resolve this issue. A greater ability to identify active metabolites of phenanthrene is especially important due to its many uses, such as an additive to dyes, drugs, and plastics, as well as its presence in cigarette smoke [42].

4.3 Challenges in Predicting Metabolites

Most metabolite prediction programs use all possible reaction schemes of interest or known to the developer to predict all possible metabolites, since it is difficult to discern whether a specific metabolite will be definitively generated within a living organism. This limitation continues to result in an expanding search for optimal methodologies and tools capable of generating the most likely metabolites for a specific chemical. As such, applications other than the ChemAxon Metabolizer currently exist. According to the developer website, ChemAxon is no longer available as a stand-alone application as of July 2016 (<https://www.chemaxon.com/products/metabolizer/>), and users are directed towards accessing its

metabolite predicting capabilities through a sister program named JChem Suite. Some alternative platforms for predicting metabolites include the proprietary software Meteor Nexus (Lhasa Limited) and ADMET Predictor (Simulations Plus, Inc.), as well as the open-source program OECD QSAR Toolbox (Organization for Economic Cooperation and Development). The QSAR Toolbox utilizes information taken from literature and databases for thousands of chemicals to predict metabolites of a user-entered chemical set, sometimes even quantitative values of formation, but primarily uses rat, non-human, or microbial transformation pathways. ADMET Predictor not only predicts metabolites of inputted parent chemicals, but also their kinetic metabolism values (e.g., intrinsic clearance) based on the assumed site of metabolic enzyme activity. However, as of this time, metabolite predictions are available for only five major cytochrome P450 Phase I enzyme pathways, and these too are trained using rat enzymes. Meteor Nexus predicts products generated through both Phase I and Phase II metabolism and so can generate a large number of potential structures. However, the ability to identify the most likely metabolites from such a large library of potential candidates remains a difficult task. While the ChemAxon Metabolizer was applied to our proposed workflow in the current study, the many software options available can be used at the discretion of other investigators, though the advantages and disadvantages of each should be taken into consideration.

4.4 Application of the Workflow in High Throughput Risk Screening

The goal of high throughput risk screening is to identify those chemicals that have the potential of causing the most harm to human and ecosystem health in as little time as possible and through efficient use of resources. Evaluating the estrogenic potential of the 259 *in vitro* inactive parents and their respective metabolites may place additional pressure on such resource use and cost, and so those chemicals can be prioritized further. For example, the set of 25 *in vitro* inactive parents predicted to be active *in silico* and with active metabolites are likely of high concern due to structural similarities that are similar between themselves and their respective metabolites and that indicate the potential for ER-binding. Because of the low number of chemicals within this set, their activity might be confirmed in *in vitro* assays outside of those already run for ToxCast. If these chemicals are found to be active using other *in vitro* assays, this additional weight of evidence may suggest that these parents can be active themselves, thus removing the need to confirm metabolite production and activity. Additionally, those parents found to be active in *in vitro* studies separate from ToxCast assays might also be considered of concern, especially if evidence also points to their potential for estrogenic activity *in vivo*. In contrast, those parent compounds considered to be inactive *in vitro*, predicted to be inactive *in silico*, and predicted to have active metabolites, may be of high concern due to the potential of such activity being captured only in an *in vivo* environment upon metabolism. There is likely little value in administering alternative *in vitro* assays to this set of 209 parent chemicals, and rather analyzing their metabolites through ToxCast or alternative *in vitro* assays may be of more interest. Prioritizing the parent chemicals and their active metabolites in our case study in such a manner will allow for complementing *in vitro* testing results with *in silico* model predictions for improving confidence in screening of environmental chemicals.

5. Conclusions

This study presents a workflow for identifying parent chemicals that may be biotransformed in an *in vivo* system into metabolites capable of perturbing a molecular target, and we have demonstrated how this workflow may be applied in a case study involving the ER pathway. As the number of environmental chemicals in commerce continues to expand, rapid determination of their potential for harming human or ecosystem health becomes critical. Our proposed workflow facilitates detection of those chemicals that may exhibit only weak activity on their own but that may produce adverse outcomes upon metabolic activation. Used in conjunction with approaches for limiting the number of false positives in *in vitro* assays, our workflow, which provides the ability to also screen for false negatives, can aid in more reliable high throughput risk screening.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank Drs. Sehan Lee, Jeremy Fitzpatrick, Peter Egeghy, and Tom Pierce of the US EPA for their review of a draft of this manuscript. The authors also thank Drs. Ann Richard, Grace Patlewicz, and Chris Grulke for their advice on the study design. Funding for Dr. Leonard and Dr. Mansouri was provided by the Oak Ridge Institute for Science and Education Research Participation Program at the US EPA.

Abbreviations:

AOP	adverse outcome pathway
CERAPP	Collaborative Estrogen Receptor Activity Prediction Project
EDC	endocrine disrupting chemical
ER	estrogen receptor
HT	high-throughput
IUPAC	International Union of Pure and Applied Chemistry
InChI	International Chemical Identifier
MIE	molecular initiating event
NCCT	National Center for Computational Toxicology
OCHEM	Online Chemical Database with Modeling Environment
OECD	Organization for Economic Cooperation and Development
QSAR	quantitative structure activity relationship
SMILES	simplified molecular input line-entry system
UNISTRA	University of Strasbourg, France

References

- [1]. Zurlo J, Rudacille D, Goldberg AM, Animals and Alternatives in Testing: History, Science, and Ethics, Mary Ann Liebert, Inc, Larchmont, NY, 2001 http://caat.jhsph.edu/publications/animal_alternatives/ (accessed September 1, 2015).
- [2]. Phillips RD, Bahadori T, Barry BE, Bus JS, Gant TW, Mostowy JM, Smith C, Willuhn M, Zimmer U, Twenty-first century approaches to toxicity testing, biomonitoring, and risk assessment: perspectives from the global chemical industry, *J Expos Sci Environ Epidemiol* 19 (2009) 536–543. doi:10.1038/jes.2009.38.
- [3]. Egeghy PP, Judson R, Gangwal S, Mosher S, Smith D, Vail J, Cohen Hubal EA, The exposure data landscape for manufactured chemicals, *Sci. Total Environ* 414 (2012) 159–166. doi:10.1016/j.scitotenv.2011.10.046. [PubMed: 22104386]
- [4]. Binetti R, Costamagna FM, Marcello I, Exponential growth of new chemicals and evolution of information relevant to risk control, *Ann. Ist. Super. Sanita* 44 (2008) 13–15. [PubMed: 18469371]
- [5]. Shukla SJ, Huang R, Austin CP, Xia M, The future of toxicity testing: A focus on in vitro methods using a quantitative high throughput screening platform, *Drug Discov Today* 15 (2010) 997–1007. doi:10.1016/j.drudis.2010.07.007. [PubMed: 20708096]
- [6]. Krewski D, Andersen ME, Mantus E, Zeise L, Toxicity testing in the 21st century: Implications for human health risk assessment, *Risk Anal* 29 (2009) 474–479. doi:10.1111/j.1539-6924.2008.01150.x. [PubMed: 19144067]
- [7]. National Research Council (NRC), Toxicity testing in the 21st Century: A vision and a strategy, The National Academies Press, Washington, DC, USA, 2007.
- [8]. Austin C, Kavlock RJ, Tice R, Tox21: Putting a lens on the vision of toxicity testing in the 21st century, *Alttox.org* (2008). <http://alttox.org/tox21-putting-a-lens-on-the-vision-of-toxicity-testing-in-the-21st-century/>. (accessed June 9, 2017).
- [9]. Kavlock RJ, Austin CP, Tice R, Toxicity Testing in the 21st century: Implications for human health risk assessment, *Risk Anal* 29 (2009) 485–497. doi:10.1111/j.1539-6924.2008.01168.x. [PubMed: 19076321]
- [10]. Baell J, Walters MA, Chemistry: Chemical con artists foil drug discovery, *Nature* 513 (2014) 481–483. doi:10.1038/513481a. [PubMed: 25254460]
- [11]. Baell JB, Holloway GA, New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays, *J. Med. Chem* 53 (2010) 2719–2740. doi:10.1021/jm901137j. [PubMed: 20131845]
- [12]. Bruns RF, Watson IA, rules for identifying potentially reactive or promiscuous compounds, *J. Med. Chem* 55 (2012) 9763–9772. doi:10.1021/jm301008n. [PubMed: 23061697]
- [13]. Judson R, Houck K, Martin M, Richard AM, Knudsen TB, Shah I, Little S, Wambaugh J, Woodrow Setzer R, Kothya P, Phuong J, Filer D, Smith D, Reif D, Rotroff D, Kleinstreuer N, Sipes N, Xia M, Huang R, Crofton K, Thomas RS, Editor's highlight: Analysis of the effects of cell stress and cytotoxicity on *in vitro* assay activity across a diverse chemical and assay space, *Toxicol Sci* 152 (2016) 323–339. doi:10.1093/toxsci/kfw092. [PubMed: 27208079]
- [14]. National Resources Defense Council (NRDC), Risk-Based procedural rule for chemical prioritization under TSCA as amended by the Frank R Lautenberg Chemical Safety for the 21st Century Act, 2016 <https://www.nrdc.org/sites/default/files/tsca-amended-risk-based-chemical-prioritization-comments-20160824.pdf> (accessed April 21, 2017).
- [15]. United States Environmental Protection Agency (USEPA), Use of high throughput assays and computational tools; Endocrine Disruptor Screening Program; Notice of Availability and Opportunity for Comment, *Federal Register* (2015). <https://www.federalregister.gov/documents/2015/06/19/2015-15182/use-of-high-throughput-assays-and-computational-tools-endocrine-disruptor-screening-program-notice> (accessed April 19, 2017).
- [16]. Vaiserman A, Early-life Exposure to endocrine disrupting chemicals and later-life health outcomes: An epigenetic bridge?, *Aging Dis* 5 (2014) 419–429. doi:10.14336/AD.2014.0500419. [PubMed: 25489493]

- [17]. Ankley GT, Johnson RD, Small fish models for identifying and assessing the effects of endocrine-disrupting chemicals, *ILAR J* 45 (2004) 469–483. doi:10.1093/ilar.45.4.469. [PubMed: 15454686]
- [18]. Judson RS, Magpantay FM, Chickarmane V, Haskell C, Tania N, Taylor J, Xia M, Huang R, Rotroff DM, Filer DL, Houck KA, Martin MT, Sipes N, Richard AM, Mansouri K, Setzer RW, Knudsen TB, Crofton KM, Thomas RS, Integrated model of chemical perturbations of a biological pathway using 18 *in vitro* high-throughput screening assays for the estrogen receptor, *Toxicol. Sci* 148 (2015) 137–154. doi:10.1093/toxsci/kfv168. [PubMed: 26272952]
- [19]. Bus JS, Becker RA, Toxicity testing in the 21st century: A view from the chemical industry, *Toxicol. Sci* 112 (2009) 297–302. doi:10.1093/toxsci/kfp234. [PubMed: 19805406]
- [20]. Hartung T, Daston G, Are *in vitro* tests suitable for regulatory use?, *Toxicol Sci* 111 (2009) 233–237. doi:10.1093/toxsci/kfp149. [PubMed: 19617452]
- [21]. Jacobs MN, Laws SC, Willett K, Schmieder P, Odum J, Bovee TF, In vitro metabolism and bioavailability tests for endocrine active substances: What is needed next for regulatory purposes?, *ALTEX* 30 (2013) 331–351. [PubMed: 23861078]
- [22]. Kirchmair J, Göller AH, Lang D, Kunze J, Testa B, Wilson ID, Glen RC, Schneider G, Predicting drug metabolism: Experiment and/or computation?, *Nat Rev Drug Discov* 14 (2015) 387–404. doi:10.1038/nrd4581. [PubMed: 25907346]
- [23]. Lee S, Barron MG, Structure-based understanding of binding affinity and mode of estrogen receptor α agonists and antagonists, *PLOS ONE* 12 (2017) e0169607. doi:10.1371/journal.pone.0169607. [PubMed: 28061508]
- [24]. Zhu BT, Han G-Z, Shim J-Y, Wen Y, Jiang X-R, Quantitative structure-activity relationship of various endogenous estrogen metabolites for human estrogen receptor α and β subtypes: Insights into the structural determinants favoring a differential subtype binding, *Endocrinology* 147 (2006) 4132–4150. doi:10.1210/en.2006-0113. [PubMed: 16728493]
- [25]. Wang P, McInnes C, Zhu BT, Structural characterization of the binding interactions of various endogenous estrogen metabolites with human estrogen receptor α and β subtypes: A molecular modeling study, *PLOS ONE* 8 (2013) e74615. doi:10.1371/journal.pone.0074615. [PubMed: 24098659]
- [26]. Harris JB, Eldridge ML, Saylor G, Menn F-M, Layton AC, Baudry J, A computational approach predicting CYP450 metabolism and estrogenic activity of an endocrine disrupting compound (PCB-30), *Environ. Toxicol. Chem* 33 (2014) 1615–1623. doi:10.1002/etc.2595. [PubMed: 24687371]
- [27]. Pinto CL, Mansouri K, Judson R, Browne P, Prediction of estrogenic bioactivity of environmental chemical metabolites, *Chem. Res. Toxicol* 29 (2016) 1410–1427. doi:10.1021/acs.chemrestox.6b00079. [PubMed: 27509301]
- [28]. Phillips MB, Leonard JA, Grulke CM, Chang D, Edwards SW, Brooks R, Goldsmith M-R, El-Masri HA, Tan Y-M, A workflow to investigate exposure and pharmacokinetics influences on high-throughput *in vitro* chemical screening based on Adverse Outcome Pathways, *Environ. Health Perspect* (2016) 53–60. doi: 10.1289/ehp.1409450. [PubMed: 25978103]
- [29]. Mansouri K, Abdelaziz A, Rybacka A, Roncaglioni A, Tropsha A, Varnek A, Zakharov A, Worth A, Richard AM, Grulke CM, Trisciuzzi D, Fourches D, Horvath D, Benfenati E, Muratov E, Wedebye EB, Grisoni F, Mangiatordi GF, Incisivo GM, Hong H, Ng HW, Tetko IV, Balabin I, Kancherla J, Shen J, Burton J, Nicklaus M, Cassotti M, Nikolov NG, Nicolotti O, Andersson PL, Zang Q, Politi R, Beger RD, Todeschini R, Huang R, Farag S, Rosenberg SA, Slavov S, Hu X, Judson RS, CERAPP: Collaborative Estrogen Receptor Activity Prediction Project, *Environ. Health Perspect* 124 (2016) 1023–1033. doi:10.1289/ehp.1510267. [PubMed: 26908244]
- [30]. Biggar RJ, Molecular pathways: Digoxin use and estrogen-sensitive cancers--risks and possible therapeutic implications, *Clin. Cancer Res* 18 (2012) 2133–2137. doi: 10.1158/1078-0432.CCR-11-1389. [PubMed: 22368159]
- [31]. Arya M, Lohiya NK, Estrogenic and antifertility effect of cyproterone acetate in female gerbils, *meriones hurriane Jerdon, Acta Biol. Med. Ger* 36 (1977) 1133–1141. [PubMed: 612093]
- [32]. Mellor CL, Steinmetz FP, Cronin MTD, Using molecular initiating events to develop a structural alert based screening workflow for nuclear receptor ligands associated with hepatic steatosis,

Chem. Res. Toxicol 29 (2016) 203–212. doi:10.1021/acs.chemrestox.5b00480. [PubMed: 26787004]

- [33]. Wilk-Zasadna I, Bernasconi C, Pelkonen O, Coecke S, Biotransformation in vitro: An essential consideration in the quantitative in vitro-to-in vivo extrapolation (QIVIVE) of toxicity data, *Toxicology* 332 (2015) 8–19. doi:10.1016/j.tox.2014.10.006. [PubMed: 25456264]
- [34]. Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrano JA, Tietge JE, Villeneuve DL, Adverse Outcome Pathways: A conceptual framework to support ecotoxicology research and risk assessment, *Environ. Toxicol. Chem* 29 (2010) 730–741. doi:10.1002/etc.34. [PubMed: 20821501]
- [35]. Edwards SW, Tan Y-M, Villeneuve DL, (Bette) Meek ME, McQueen CA, Adverse Outcome Pathways – Organizing toxicological information to improve decision making, *J Pharmacol Exp Ther* 356 (2016) 170–181. doi:10.1124/jpet.115.228239. [PubMed: 26537250]
- [36]. Kidd KA, Blanchfield PJ, Mills KH, Palace VP, Evans RE, Lazorchak JM, Flick RW, Collapse of a fish population after exposure to a synthetic estrogen, *PNAS* 104 (2007) 8897–8901. doi: 10.1073/pnas.0609568104. [PubMed: 17517636]
- [37]. Reyhanian Caspillo N, Volkova K, Hallgren S, Olsson P-E, Porsch-Hällström I, Short-term treatment of adult male zebrafish (*Danio Rerio*) with 17 α -ethinyl estradiol affects the transcription of genes involved in development and male sex differentiation, *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* 164 (2014) 35–42. doi: 10.1016/j.cbpc.2014.04.003. [PubMed: 24747828]
- [38]. Bugel SM, White LA, Cooper KR, Decreased vitellogenin inducibility and 17 β -estradiol levels correlated with reduced egg production in killifish (*Fundulus heteroclitus*) from Newark Bay, NJ, *Aquat Toxicol* 105 (2011) 1–12. doi:10.1016/j.aquatox.2011.03.013.
- [39]. Blair RM, Fang H, Branham WS, Hass BS, Dial SL, Moland CL, Tong W, Shi L, Perkins R, Sheehan DM, The estrogen receptor relative binding affinities of 188 natural and xenochemicals: Structural diversity of ligands, *Toxicol Sci* 54 (2000) 138–153. doi:10.1093/toxsci/54.1.138. [PubMed: 10746941]
- [40]. Cannizzo C, Amigoni-Gerbier S, Frigoli M, Larpent C, A versatile preparation of azobenzene-dye functionalized colored polymer nanoparticles by surface modification, *J. Polym. Sci. A Polym. Chem* 46 (2008) 3375–3386. doi:10.1002/pola.22679.
- [41]. Kleinstreuer N, Ceger P, Allen D, Strickland J, Chang X, Hamm J, Casey W, A curated database of rodent uterotrophic bioactivity, *Environmental Health Perspectives* 124 (2016) 556–662. doi: 10.1289/ehp.1510183. [PubMed: 26431337]
- [42]. Wang J, Zhong Y, Carmella SG, Hochalter JB, Rauch D, Oliver A, Jensen J, Hatsukami DK, Upadhyaya P, Hecht SS, Zimmerman CL, Phenanthrene metabolism in smokers: Use of a two-step diagnostic plot approach to identify subjects with extensive metabolic activation, *J. Pharmacol. Exp. Ther* 342 (2012) 750–760. doi:10.1124/jpet.112.194118. [PubMed: 22674470]

Figure 1.
Workflow for identifying parent chemicals than can become bioactive in a living system.

Figure 2.

Breakdown of *in vitro* inactive parent chemicals, along with first-generation and second-generation metabolites predicted to be formed *in silico* and their activity as predicted through the consensus *Collaborative Estrogen Receptor Prediction Project* (CERAPP) structure activity relationship model for estrogen receptor binding.

Figure 3.

Dominant transformation pathways for formation of first-generation and second-generation metabolites, along with reactions for each given scheme. The grey circles in the reactions represent the structural feature within the parent chemical being enzymatically acted upon, while the red circles represent the resulting change in structure for that parent chemical's metabolite.

Figure 4. Summary of activity of parent compounds, as determined through *in vitro* analysis and predicted using the consensus *Collaborative Estrogen Receptor Activity Prediction Project* (CERAPP) model, in addition to a breakdown of the 259 *in vitro* inactive parent chemicals predicted to have active metabolites.

Figure 5.

Chemotype enrichment of chemical sets consisting of *in vitro* active parents, *in vitro* inactive parents with active metabolites, *in vitro* inactive parents with inactive metabolites, active first-generation metabolites, active second-generation metabolites, and *in vitro* inactive parents with active metabolites predicted to be active *in silico* and with active metabolites. The chemotypes representing aromatic substructures are provided in the top panel, while the chemotypes representing linear substructures are presented in the lower panel. Illustrations

of each chemotype are provided below and to the side of their respective panels. Chemotype descriptions are provided in the main text.

Table 1.

A comparison of first- and second-generation metabolites predicted to be active by the consensus structure activity relationship (QSAR) model used in the current study with the consensus *Collaborative Estrogen Receptor Activity Prediction Project* (CERAPP) model activity predictions¹ and literature estimates.

Metabolite Name	Generation Formed	CERAPP Binary Prediction	CERAPP Esitmate	Observed Binding
4-hydroxybenzyl alcohol	1,2	1	VeryWeak	NA
Androsterone	1	1	VeryWeak	NA
Hydroquinone	1,2	0	Suspicious	NA
2,4-Dihydroxybenzophenone	1	1	Weak	1
4-tert-Butylphenol	1,2	1	Weak	1
gentisic acid	1,2	1	VeryWeak	NA
4-(2-Methylbutan-2-yl)cyclohexanol	1	1	VeryWeak	0
17alpha-Hydroxyprogesterone	1	1	VeryWeak	NA
4-hydroxyphenethyl alcohol	1	1	VeryWeak	NA
4-Benzylphenol	1	1	Moderate	1
Nonylphenoxypolyethoxyethanol - iodine complex	1	1	VeryWeak	NA
1,2,4-Benzenetriol	1	1	VeryWeak	NA
4-hydroxy-alpha-methylbenzyl alcohol	1,2	1	VeryWeak	NA
4-Hydroxyacetophenone	1,2	1	VeryWeak	1
melengestrol	1	1	Moderate	NA
4-Bromophenol	1,2	1	Weak	NA
4-Phenoxyphenol	1,2	1	Moderate	1
4-Propylphenol	1	1	VeryWeak	NA
4-Butylphenol	1	1	VeryWeak	1
Dimepheptanol	1	0	Suspicious	0
4-tert-Butylcyclohexanol	1,2	0	Suspicious	0
3,4,6-trichlorocatechol	1,2	0	Suspicious	NA
Abietic acid	1	1	VeryWeak	1
4-(2,6,6-trimethyl-2-cyclohexen-1-yl)-3-buten-2-ol	1	0	Suspicious	NA
p-Cresol	1,2	0	Suspicious	NA
1'-Hydroxyestragole	1	0	Suspicious	NA
4-hydroxy-3-phenoxybenzoic acid	1	1	Weak	NA
Cyproterone	1	1	Weak	0
dl-alpha-Tocopherol	1	1	VeryWeak	0
2-hydroxycinnamic acid	1	0	Suspicious	NA
4-Phenylphenol	1	1	Weak	1
2,4,4'-Trihydroxybenzophenone	2	1	Weak	1
4-Androstenediol	2	1	Weak	1
4,4'-Biphenyldiol	2	1	Moderate	1
4-(tert-Pentyl)-cyclohexanone	2	1	VeryWeak	0
4-tert-Butylcatechol	2	1	Weak	0
4-hydroxycinnamic acid	2	1	VeryWeak	NA

Metabolite Name	Generation Formed	CERAPP Binary Prediction	CERAPP Esitmate	Observed Binding
pregna-4,16-diene-3,20-dione	2	1	VeryWeak	NA
Paroxypropione	2	1	VeryWeak	NA
4-Hydroxybenzophenone	2	1	VeryWeak	1
Bis(4-hydroxyphenyl)methane	2	1	VeryWeak	1
4-Nonylphenol	2	1	VeryWeak	1
3-(Trifluoromethyl)phenol	2	1	VeryWeak	NA
p-vinylphenol	2	1	VeryWeak	NA
p-vinylphenol Phenol, 4-ethenyl-, homopolymer	2	1	VeryWeak	NA
P-Hydroxybenzaldehyde	2	1	Weak	1
Hydroxyamphetamine	2	1	VeryWeak	0
1,2-Dihydronaphthalene	2	1	VeryWeak	NA
p-hydroxyphenylacetic acid	2	1	VeryWeak	NA
4-hydroxyphenylpyruvic acid	2	1	VeryWeak	0
3-Hydroxybenzoic acid	2	0	Suspicious	0
androsta-1,4-diene-3,17-dione	2	1	Weak	NA
3,4,5-trichlorocatechol	2	0	Suspicious	NA
1,3,5-Trichloro-2,4,6-trihydroxybenzene	2	0	Suspicious	NA
benzilic acid	2	0	Inactive	NA

¹ Activity predictions are given as binary determinations of binding (1) or non-binding (0)

Table 2.

Suggested activity from alternative sources for *in vitro* inactive parent chemicals predicted to have active metabolites. See supplemental File 1 for complete list of chemicals and the web sources describing their activity.

Parent CASRN ¹	ParentName	In Silico Activity ²	In vitro Activity ³	In vivo Activity
100-51-6	Benzyl alcohol	No	No	N/A
10081-67-1	4-(2-Phenylpropan-2-yl)-N-[4-(2-phenylpropan-2-yl)phenyl]aniline	Yes	N/A	N/A
101-81-5	Diphenylmethane	No	No	N/A
103-69-5	N-Ethylaniline	No	No	N/A
104-51-8	Butylbenzene	No	N/A	No
106-44-5	p-Cresol	No	Yes	N/A
107534-96-3	Tebuconazole	No	No	N/A
114369-43-6	Fenbuconazole	No	Yes	No
1166-52-5	Dodecyl gallate	Yes	No	N/A
118-55-8	Phenyl salicylate	No	Yes	Yes
119-36-8	Methyl salicylate	No	No	No
1194-65-6	Dichlobenil	No	No	N/A
120-82-1	1,2,4-Trichlorobenzene	No	No	No
120-83-2	2,4-Dichlorophenol	No	Yes	N/A
120-95-6	2,4-Bis(2-methylbutan-2-yl)phenol	Yes	N/A	N/A
121-32-4	3-Ethoxy-4-hydroxybenzaldehyde	Yes	N/A	N/A
123-07-9	4-Ethylphenol	Yes	No	N/A
1330-78-5	Tricresyl phosphate	Yes	N/A	N/A
1434-54-4	Pregnenolone carbonitrile	Yes	N/A	N/A
14351-66-7	Sodium abietate	Yes	N/A	N/A
16423-68-0	FD&C Red 3	No	Yes	N/A
1675-54-3	Bisphenol A diglycidyl ether	No	No	N/A
1742-14-9	1,1-Bis(3,4-dimethylphenyl)ethane	Yes	N/A	N/A
1836-75-5	Nitrofen	No	Yes	N/A
1843-05-6	Octabenzene	Yes	N/A	N/A
19774-82-4	Amiodarone hydrochloride	Yes	N/A	No
205-99-2	Benzo(b)fluoranthene	No	N/A	Yes
20830-75-5	Digoxin	No	Yes	Yes
22781-23-3	Bendiocarb	No	N/A	Yes
23593-75-1	Clotrimazole	No	N/A	No
2392-39-4	Dexamethasone sodium phosphate	Yes	N/A	N/A
27344-41-8	Disodium 4,4'-bis(2-sulfostyryl)biphenyl	No	Yes	No
2772-45-4	2,4-Bis(1-methyl-1-phenylethyl)phenol	Yes	N/A	N/A
2919-66-6	Melengestrol acetate	No	N/A	No
302-79-4	trans-Retinoic acid	No	N/A	No
309-00-2	Aldrin	No	Yes	No
3147-75-9	Octrizole	Yes	N/A	N/A

Parent CASRN ^I	ParentName	In Silico Activity ²	In vitro Activity ³	In vivo Activity
3739–38-6	3-Phenoxybenzoic acid	No	N/A	No
41372–20-7	Apomorphine hydrochloride hydrate	Yes	N/A	N/A
427–51-0	Cyproterone acetate	No	Yes	Yes
4449–51-8	Cyclopamine	Yes	N/A	N/A
484–17-3	9-Phenanthrol	Yes	N/A	N/A
50–14-6	Ergocalciferol	Yes	N/A	N/A
50–78-2	Aspirin	No	Yes	No
52–01-7	Spironolactone	No	Yes	Yes
52–86-8	Haloperidol	No	N/A	No
5315–79-7	1-Hydroxypyrene	Yes	Yes	N/A
541–73-1	1,3-Dichlorobenzene	No	Yes	N/A
57–41-0	5,5-Diphenylhydantoin	No	N/A	No
60–57-1	Dieldrin	No	Yes	No
63–05-8	4-Androstene-3,17-dione	Yes	Yes	Yes
630–56-8	Hydroxyprogesterone caproate	No	N/A	No
6344–67-8	3-Hydroxyfluorene	Yes	N/A	N/A
6422–86-2	Bis(2-ethylhexyl) terephthalate	No	N/A	No
65277–42-1	Ketoconazole	No	N/A	No
67747–09-5	Prochloraz	No	Yes	N/A
77–06-5	Gibberellic acid	No	N/A	Yes
79–94-7	3,3',5,5'-Tetrabromobisphenol A	No	Yes	No
8001–54-5	Benzalkonium chloride	No	No	N/A
80844–07-1	Etofenprox	Yes	N/A	N/A
81131–70-6	Pravastatin sodium	No	N/A	No
82657–04-3	Bifenthrin	No	Yes	Yes
838–85-7	Diphenyl phosphate	No	No	N/A
83834–59-7	2-Ethylhexyl trans-4-methoxycinnamate	No	Yes	Yes
87–61-6	1,2,3-Trichlorobenzene	No	No	N/A
88–06-2	2,4,6-Trichlorophenol	No	No	N/A
88–30-2	3-Trifluoromethyl-4-nitrophenol	No	Yes	N/A
89–04-3	Trioctyl trimellitate	No	Yes	N/A
89–83-8	Thymol	No	No	N/A
90–43-7	2-Phenylphenol	Yes	Yes	N/A
92–52-4	Biphenyl	Yes	Yes	No
93–76-5	2,4,5-Trichlorophenoxyacetic acid	No	Yes	N/A
95–95-4	2,4,5-Trichlorophenol	No	No	N/A
95737–68-1	Pyriproxyfen	No	Yes	N/A
96–76-4	2,4-Di-tert-butylphenol	No	Yes	No
98319–26-7	Finasteride	No	N/A	No
99–93-4	4-Hydroxyacetophenone	Yes	Yes	N/A

^I Chemical Abstract Service Registry number

²Predicted using consensus *Collaborative Estrogen Receptor Activity Prediction Project* (CERAPP) model

³Measured in assays other than those considered by the estrogen receptor orthogonal model

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript