PART I

GENERAL ASPECTS

1

ToxCast: PREDICTING TOXICITY POTENTIAL THROUGH HIGH-THROUGHPUT BIOACTIVITY PROFILING

KEITH A. HOUCK, ANN M. RICHARD, RICHARD S. JUDSON, MATTHEW T. MARTIN, DAVID M. REIF, AND IMRAN SHAH

1.1 INTRODUCTION

Chemical safety assessment has long relied on exposing a few species of laboratory animals to high doses of chemicals and observing adverse effects. These results are extrapolated to humans by applying safety factors (uncertainty factors) to account for species differences, susceptible sub-populations, establishing no observed adverse effect levels (NOAEL) from the lowest observed adverse effect levels, and data gaps yielding theoretically safe exposure limits. This approach is often criticized for lack of relevance to human health effects due to the many demonstrated differences in physiology, metabolism, and toxicological effects between humans and rodents or other laboratory animals [1]. Such criticism exists mainly due to the lack of knowledge of specific mechanisms of toxicity and whether these are relevant to humans. Toxicological modes of action (MOA) have been elucidated for only a limited number of chemicals; even fewer chemicals have had their specific molecular mechanisms of action determined. Having such detailed knowledge would facilitate higher confidence in species extrapolation and setting of exposure limits. However, tens of thousands of chemicals currently in commerce and with some potential for human exposure lack even traditional toxicity testing and much less elucidated modes or mechanisms of toxicity [2]. Understanding mechanisms of toxicity usually results

High-Throughput Screening Methods in Toxicity Testing, First Edition. Edited by Pablo Steinberg.

from decades-long research dedicated to single chemicals of interest, a model unsuitable for such vast numbers of chemicals. Even with dedicated research, such efforts are not guaranteed to succeed; the extended focus on understanding the mechanism of toxicity of 2,3,7,8-tetrachlorodibenzodioxin (TCDD) is an example [3]. Traditional animal testing, in addition to the criticisms discussed above, is not appropriate or feasible for the large numbers of untested chemicals due to the high costs and number of animals required [1].

One major effort to address this dilemma by providing a high-capacity alternative is underway, facilitated by integration of the fields of computational toxicology and high-throughput *in vitro* testing [4, 5]. The ultimate goals of this approach are the means to screen and prioritize thousands of chemicals, predict the potential for human health effects, and derive safe exposure levels for the myriad of chemicals to which we are exposed. This approach relies on a shift in toxicology research away from "black-box" testing on whole animals and toward an understanding of the direct interactions of chemicals with a broad spectrum of potential toxicity targets comprising specific molecular entities and cellular phenotypes. This bioactivity profiling of chemicals generated through the use of high-throughput approaches produces characteristic chemical response profiles, or signatures, which may describe the potential for toxicity of that chemical [6].

Computational analysis and modeling of the results are required to provide insight into complex datasets and support the development of predictive toxicity algorithms that ultimately may serve as the foundation of an *in vitro* toxicity testing approach replacing most or all animal testings. The groundwork required for a computational toxicology approach is the generation of datasets comprising the quantitative effects of chemicals on biological targets. Two types of data are required. The first are the test results from *in vitro* and/or in silico assays that can be run in high-throughput mode and provide bioactivity profiles for hundreds to thousands of chemicals. The second is a dataset that details the effects of these chemicals on whole organisms, ideally the species of interest. These data are used to anchor and build predictive models that can then be applied to chemicals that lack *in vivo* testing. Generation of the *in vitro* dataset has become feasible and widely available as high-throughput *in vitro* screening technology, developed in support of the drug discovery community. The selection and use of these assays for computational toxicology will be discussed further in Section 4. Obtaining the latter dataset of *in vivo* effects necessary to build the computational models presents unique challenges. Although thousands of chemicals have been tested using *in vivo* approaches, only a limited amount of this information has been readily available. Much of it lies in formats not readily conducive to computational analysis, for example, paper records, in the data stores of private corporations, or protected by confidentiality clauses [7], and generation of extensive new *in vivo* data to support the approach is cost prohibitive. The access and collation of these data into a relational database useful for computational toxicology will be discussed in Section 5.

Beyond the technical aspects of generating the data, assembling the collection of required datasets to support computational approaches is a challenging task in itself. Robust, efficient, and accurate knowledge discovery from large datasets require a

robust data infrastructure. There are a number of critical steps in the process beginning with designing an underlying architecture to manage the data. Appropriate data must be selected and preprocessed into common formats usable to computer programs (e.g., standardized field names for the types of attributes being measured, standardized chemical names and links to other data sources). The use of standardized ontologies can be particularly useful in the sharing of information across organizations [8]. Because of the complexities of achieving this on a large scale, these approaches are perhaps best conducted by large organizations with access to computational scientists in addition to experts in chemistry, toxicology, statistics, and high-throughput screening (HTS). Examples of integration of these diverse areas of expertise include the U.S. Environmental Protection Agency's (EPA) ToxCast program [4] and the Tox21 collaboration between the EPA, the National Toxicology Program, the National Institutes of Health Center for Translational Therapeutics—NCTT (formally the NIH Chemical Genomics Center [NCGC]), and the U.S. Food and Drug Administration [9, 10]. In addition, a number of large pharmaceutical companies have internal programs in this area relying on their own, extensive in-house expertise [11, 12].

As described, the ultimate goal is to use high-throughput *in vitro* assays to rapidly and inexpensively profile the bioactivity of chemicals of unknown toxicity and make predictions about their potential for causing various adverse endpoints [4]. Achieving a robust, predictive toxicology testing program is a long-range goal that will need to proceed through a number of systematic stages including proof-of-concept, extension of chemical and bioassay diversity, refinement, and ultimately, supplementation or replacement of existing methods. The initial stage involves multiple steps including (1) selecting an appropriate chemical test set for which *in vivo* data are available; (2) selecting high-throughput biological assays for screening the chemicals; (3) generating the screening data on the chemicals; (4) collating the *in vivo* anchoring data for the chemicals; and (5) building up predictive models. Such models can then be validated through testing of additional chemicals with known toxicity endpoints to determine the robustness of the models. It is likely that the development of the test systems, as well as the computational models, will be an iterative process. New biological assays and statistical approaches are evaluated for potential inclusion in the program, whereas assays and models not producing useful results are dropped.

The success of this stage of the process would be models judged useful for prioritizing chemicals for the potential to cause specific toxicity endpoints. This prioritization will be valuable in the short term by allowing focused and limited *in vivo* use of testing resources on chemicals most likely to be of concern. The results of targeted testing of designated chemicals for specific endpoints should ensure a reduced use of test animals as only limited endpoints would need to be evaluated. This targeted testing will also provide an additional validation method for the testing program, that is, do the adverse endpoints predicted by the models occur to a significant extent in the tested chemicals? Ultimately, refinement of the testing and modeling approaches should allow high-confidence prediction of the likelihood for toxicity, thereby avoiding animal testing altogether for many chemicals. The remainder of this chapter will focus more specifically on providing background on the steps undertaken in developing the initial stages of the ToxCast testing program at

EPA, as well as examples of applications of the program in prioritizing environmental chemicals for multiple toxicity endpoints.

1.2 CHEMICAL LANDSCAPE

A major driver of the development and use of HTS methods in toxicology is the scope of the chemical problem, that is, tens of thousands of chemicals to which individuals are potentially exposed, the majority of which have never been tested in any significant way [2] . What chemicals are of interest and the kind of data that is likely to be available depends on the use of the chemical, which in turn is related to the regulations to which the chemicals are subjected. To understand the world of chemicals that are of concern for potential toxicity and candidates for testing, it is useful to discuss a set of chemical inventories, some of which are overlapping.

1.2.1 Pesticide Active Ingredients

These are typically the active compounds in pesticide formulations, which are designed to be toxic against select types of organisms. A related category of compounds falling under this general label are antimicrobials, which are also designed to be toxic to certain organisms, in this case-targeting fungi or bacteria. These groups of chemicals are further divided into food-use and nonfood-use actives for the purpose of regulation. EPA sets tolerance levels for pesticides that may be used in specific foods, for particular reasons, and at particular exposure levels. Thus, EPA regulates the maximum amount of pesticide residue permitted to remain on a food approved for pesticide application. FDA, in contrast, has the authority to monitor and enforce levels of food-use pesticides and ensure that they comply with EPA regulations. FDA has additional authority regarding the use of antimicrobials in food packaging [13]. Food-use pesticide actives have the highest data requirements and, for these, a company will typically generate data from 2-year chronic/cancer bioassays in rats and mice, developmental toxicity studies in rats and rabbits, multigenerational reproductive toxicity studies in rats, and other specialized *in vivo* studies [14]. These are similar to the complete set of preclinical studies that are required for human pharmaceuticals. Because of this large data requirement, these chemicals are ideal for use in building up toxicity prediction models, since one will have near-complete *in vitro* and *in vivo* datasets. It is not surprising that pesticide actives have some of the same features and chemical properties as pharmaceutical products, given that they are often designed to interact with a specific molecular target.

1.2.2 Pesticidal Inerts

These are all of the ingredients in a pesticide product or formulation other than the active ingredients. Although they are labeled as “inert”, there is no requirement that they be nontoxic. These can range from solvents (e.g., benzene) to animal attractants, such as peanut butter or rancid milk. As with the actives, inerts are classified as

food-use and nonfood-use. Regulatory data requirements are, in general, limited, thus resulting in the availability of little *in vivo* data [15].

1.2.3 Industrial Chemicals

This is an extremely broad class of chemicals including solvents, detergents, plastic monomers and polymers, fuels, synthesis intermediates, and dyes. As such, they are typically not designed to be bioactive, although many do have bioactivity, sometimes through interaction with enzymes and receptors, or by chemically reacting with biomolecules or via physical interactions (e.g., by disrupting cell membranes). Many of these compounds are manufactured in very large quantities, posing greater potential risks. Such chemicals typically have less stringent regulatory oversight and toxicity testing requirements but are subject to reporting rules under the Toxics Substances Control Act (TSCA). Under TSCA, different reporting requirements and regulatory scrutiny are applied depending on production volume levels (MPV—medium production volume chemicals, >25 K tons/year; HPV—high-production volume chemicals, >1 M tons/year). On average, these industrial compounds have lower molecular weight than pesticidal actives or pharmaceuticals, and include many more volatile and semivolatile compounds.

1.2.4 Pharmaceuticals

These are the active ingredients in drugs and, hence, are designed to have specific bioactivity. It is well known that many drugs have toxic side effects, often through unexpected off-target interactions, and that this is a major economic concern for the pharmaceutical industry driving up the costs of drug development. In addition, there is increasing concern for toxicity, not only for patients directly taking the drug, but also for ecological species exposed to these compounds in waste water [16]. Despite large amounts of toxicity data submitted to the FDA during the drug-approval process, including clinical data on humans if the drug reaches clinical trials, as well as additional preclinical toxicity data generated within the pharmaceutical industry, little of these data see the light of day due to confidentiality concerns. As a result, public availability of toxicity data on pharmaceuticals is generally limited to what is available in the open literature.

1.2.5 Food Additives/Ingredients

This category includes both natural and synthetic small molecules that are intentionally added to food, often to enhance nutritional value (e.g., vitamins), to act as preservatives, such as in food packaging, or to enhance color or texture. FDA regulates allowed tolerances for such chemicals and has the authority to require a battery of *in vitro* (primarily genotoxicity) and *in vivo* toxicity testing to support such reviews within the Center for Food Safety and Nutrition (CFSAN) [17]. Such data can be made publicly available, hence providing a potentially rich source of additional *in vivo* data for computational toxicology modeling.

1.2.6 Water Contaminants

EPA regulates chemicals in surface and drinking water, and the relevant chemicals include any of the above categories that enter the water system, as well as metabolites or degradation products. One example of the latter is disinfection byproducts that can result from reactions of chlorine with organic molecules in a drinking water system to produce polychlorinated organic compounds. The regulatory authority in this instance is reactive. First, a chemical has to be detected in water at sufficient levels to cause some concern, and then sufficient scientific justification must be provided to warrant regulatory action. As a result, toxicity data is generally lacking for many of these chemicals, similar to the situation for industrial chemicals.

Because there are so many chemicals to which humans and ecological species are potentially exposed, it is necessary to prioritize among them when setting up a large-scale screening program such as ToxCast or Tox21. The potential for exposure is one critical aspect of this prioritization, and these and further chemical use-categories are important indicators of the potential for exposure. For instance, any chemical that is directly in food or water (e.g., food additives or pesticides that leave residues on crops or chemicals found in drinking water) would have extra weight in a prioritization scheme. More detailed "use-categories" are also available to help refine estimates of potential exposure routes. For instance, if a chemical is found in products to which children are exposed (e.g., baby bottles, clothing), that chemical would have a heightened priority for screening. There is no general mapping of chemicals to use-categories that is publicly available, but the ExpoCast project, affiliated with the ToxCast project within EPA, is currently developing such a mapping based on merging data from many different sources [18].

The lack of data availability on chemicals, whether it is use-category, exposure potential, or toxicity data, is one of the major drivers of EPA's HTS computational toxicology program [4]. However, the success of this effort also relies upon the ability to collate as much available data as possible and systematize and format these data into computable forms to enable modeling efforts to proceed. To provide a central resource to support this effort, a large-scale database is being created to gather all publicly available data on chemicals in the environment through the Aggregated Computational Toxicology Resource (ACToR) effort [19]. Thus far, varying amounts and types of data have been compiled on several hundred thousand chemicals collected from over 1,000 different sources, consisting of data types that, for example, include information on hazard (i.e., *in vitro* and *in vivo* toxicity data), exposure, use, and production.

The above discussion focuses on the chemical landscape of concern for testing from a regulatory and use or exposure perspective, but an equally important consideration for our long-range purposes is providing adequate coverage of the chemical feature and property landscape spanned by the various use-category inventories of chemicals. Given the intimate relationship between the chemical structure and its biological activity, building a computational toxicology approach capable of predicting toxicity from HTS bioactivity profiles must provide for sufficient coverage of biological pathways and toxicity mechanisms across the chemical landscape of interest.

This means that a chemical testing library must also provide sufficient coverage of the diverse chemical property and features space capable of adequately probing this biological mechanism diversity.

1.3 THE CHEMICAL LIBRARIES

To generate the *in vitro* dataset required for the computational toxicology approach, a chemical library was assembled, with initial and later testing candidates largely drawn from the chemical inventories described above. Meeting the initial objectives of providing proof-of-concept of the HTS computational toxicology approach required a strong anchoring to *in vivo* animal toxicity studies. Hence, selection of the initial testing set for ToxCast, which we refer to as the Phase I chemical library, was primarily driven by the availability of detailed, *in vivo* toxicity data. The existence of high-quality regulatory guideline studies required for chemical safety evaluation of pesticide active ingredients by EPA motivated the selection of these compounds to fulfill these data requirement needs. Thus, the Phase I library consisted of 309 unique chemical substances, with more than 90% pesticides and the rest a mixture of *in vivo* data-rich industrial chemicals such as bisphenol A (BPA) and perfluorooctanoic acid (PFOA).

In vitro HTS testing procedures additionally have a number of practical requirements that affect the types of chemicals that can be tested using current technologies. Obvious concerns are the solubility of the chemical in aqueous buffer, which is the medium used to conduct HTS testing, as well as dimethyl sulfoxide (DMSO), which is the near universal solvent used to solubilize test chemicals for testing. Additionally, volatility is a concern, since the chemicals are run in batch mode and attention cannot be paid to special handling requirements for volatile or semivolatile chemicals. A few physical–chemical property filters, primarily molecular weight (MW) and octanol/water partition coefficient (logP), were used to choose the Phase I chemicals, but the structures of pesticides are such that most met the criteria for inclusion and were soluble in DMSO. The ToxCast Phase I chemical solutions that underwent the initial round of HTS testing were also post-analyzed by analytical quality control (QC) methods that are amenable to high-throughput application (primarily liquid chromatography–mass spectrometry [LC/MS] with gas chromatography–mass spectrometry [GC/MS] follow-up for compounds not suitable for LC/MS analysis). Identity and purity were confirmed for over 80% of the Phase I library, with the majority of the remaining compounds deemed unsuitable for analysis because they were metal containing or of low MW. One class of pesticides, consisting of 14 sulfurons, was found to significantly dissociate in DMSO over time, motivating the removal of these compounds from further ToxCast testing.

The ToxCast Phase I chemical library, despite its relatively small size, contained a significant amount of chemical and functional diversity, spanning over 40 chemical functional classes (e.g., pyrazoles, sulfonamides, organochlorines, pyrethroids, carbamates, organophosphates) and 24 known pesticidal functional classes (e.g., phenylurea herbicides, organophosphate insecticides, dinitroaniline herbicides). The

implication is that although the particular compounds included in this Phase I test set may not be representative of the larger chemical universe of potential interest such as antimicrobials, food-additives, drugs, and industrial compounds, the constituent features of these chemicals are potentially capable of representing a much broader set of chemicals from a wide range of use-categories.

Clearly, however, in order to meet the larger objectives of the ToxCast program for modeling *in vivo* toxicity, it is necessary to test larger chemical inventories that include greater representation of the various use-categories of high interest, as well as the more varied chemical and biological interactions that must be probed and characterized in order to build general models for predicting toxicity. Following the testing of the Phase I library, a much larger chemical collection was assembled based on these considerations for the dual purposes of expanding the ToxCast test library and constructing the EPA contribution to the Tox21 library. Nominations for this library were broadly drawn from the previously described inventories and initially exceeded 9,000 compounds. Given the much larger structural diversity of the chemicals nominated, a greater number of compounds were excluded from consideration on the basis of calculated physical–chemical properties, such as MW, vapor pressure, boiling point, solubility, and logP. Finally, practical considerations pertaining to physical samples, such as cost, availability, actual solubility in DMSO, and confirmed volatility, determined whether or not a compound was included in the final EPA Tox21 inventory, consisting of more than 3,700 unique chemical substances.

The ToxCast Phase II chemical library, currently undergoing testing, consists of 776 unique chemical substances, including nine Phase I compounds used as testing replicates, drawn from the expanded EPA Tox21 chemical inventory, spanning a much broader range of use-cases and chemical structures than in Phase I. For the selection of Phase II compounds, significant weight was given to those substances with extensive *in vivo* data available, as well as to toxicity reference substances with well-defined activities and mechanisms of action. Pursuant to this goal, approximately 30% of the Phase II compounds have *in vivo* data available from the National Toxicology Program or were generated to meet EPA or FDA's regulatory requirements for pesticide or food additives. However, due to the relative paucity of data for many of the use-categories described previously, many of the chemicals in this expanded collection had relatively little or no such data available. In addition, higher weight was given to chemicals on high-interest EPA inventories (such as listed above), as well as to chemicals that appeared on multiple inventories or use-categories. The Phase II inventory also benefitted from an unprecedented collaboration between EPA and the pharmaceutical industry, whereby 135 "failed drugs" were donated by six pharmaceutical companies (Pfizer, Merck, GlaxoSmithKline, Sanofi, Roche, and Astellas), along with preclinical and, in some cases, human clinical data reporting adverse effects. The value of these data in extending findings made on chemicals tested only in laboratory animals to those tested in humans may be significant.

The ToxCast Phase I and Phase II inventories total 1,060 unique compounds. These compounds are being run in the full suite of more than 500 ToxCast assays. Both of these chemical inventories are fully contained within the EPA Tox21 chemical inventory, which in turn is a subset of the complete Tox21 collection, totaling

approximately 8,200 unique chemical structures. In addition to the failed pharmaceuticals, the Tox21 library contains an extensive collection of human pharmaceuticals [20]. Although the Tox21 inventory is much larger and spans much greater chemical diversity, this library will only be tested in HTS assays being run at the NCTT and, thus, will have more limited bioactivity profiling data available. On the other hand, the smaller ToxCast Phase I and II chemical inventories will be run in the full suite of ToxCast assays, as well as in the additional Tox21 assays, providing a rich chemical and biological context for the interpretation of these data. Details of the chemical libraries can be accessed at *http://www.epa.gov/ncct/toxcast/chemicals.html.*

An expanded analytical quality control process to ensure that the tested chemicals are indeed what they are intended to be is accompanying the full Tox21 effort. Careful review and curation of chemical identifiers, including names and Chemical Abstracts Service Registry Numbers (CASRN), as well as reported purity were extracted from Certificates of Analysis at the time of procurement. Further review and chemical structure annotation of the full Tox21 inventory and component ToxCast inventories were carried out within EPA's Distributed Structure-Searchable Toxicity (DSSTox) project (see http://www.epa.gov/ncct/dsstox/ for access to downloadable structure files). Following solubilization in DMSO, the chemical identity, purity, and, concentration are determined by appropriate analytical techniques, including LC/MS and follow-up GC/MS. This analysis will be repeated over the course of the use of the library to assess compound stability during testing. While complex and costly, such efforts ensure that biological activity measured in an assay is associated with the appropriate chemical structure and, conversely, those negative results are associated with a chemical structure only if that chemical was indeed present.

1.4 THE BIOLOGICAL ASSAYS

Selection of *in vitro* assays for toxicity testing would be relatively straightforward if the molecular targets underlying mechanisms of toxicity were known. Advances in HTS technologies to support the drug discovery industry have provided the tools to develop assays for large numbers of biological targets, ranging from receptors to enzymes to ion channels and more. If a protein has a defined function, it is safe to say that an *in vitro* assay can be built to measure effects of chemicals on that function. Techniques such as surface plasmon resonance or LC–MS–MS exist that measure chemical–protein interactions even when the function is unknown [21]. Beyond assays focusing on specific molecular targets, many assays are available to probe phenotypic changes induced in cells by chemical exposure including effects on organelles and cellular structures such as mitochondria, nuclei, cytoskeleton, and cell membrane. Again, with advances in automated fluorescent microscopy screening platforms and associated imaging algorithms, the ability to measure altered cellular phenotypes is almost unlimited. However, assays targeting specific proteins or cellular phenotypes suffer from our lack of detailed knowledge with respect to mechanisms of toxicity that would guide high-confidence assay selection. Exceptions to this, while clear, are relatively few and include molecular targets such as the

potassium ion channel hERG [22], acetylcholinesterase [23], cytochrome P450s [24], drug transporters [25], nuclear receptors including the androgen, estrogen, and aryl hydrocarbon receptor (AhR) [26], as well as the 5-HT2b G-protein-coupled receptor (GPCR) [27]. In addition, cellular phenotypic assays for genotoxicity, oxidative stress, mitochondria energy homeostasis, calcium release from intracellular stores, and necrotic and apoptotic cell death can be used to determine toxicity, although with less specificity with respect to molecular target. Acceptance of these as valid toxicity targets usually resulted from many years of research, sometimes combined with serendipitous findings. Continuing with this model to complete our understanding of toxicology would be a long, expensive, and arduous route.

As an alternative approach, a broadly based interrogation of important families of biological targets and cellular phenotypes can be conducted efficiently using high-throughput *in vitro* screens, probing them with large chemical libraries with known animal and human health effects. The reference *in vivo* toxicity data for these chemicals are needed to correlate the *in vitro* findings with *in vivo* endpoints. The tools of computational toxicology can then be applied to analyze, interpret, and model the results, ultimately generating predictive signatures of toxicity compatible with cost-efficient, high-throughput assays conducive to screening unknown chemicals.

Defined toxicity targets are usually members of large protein families such as enzymes (e.g., acetylcholinesterase), receptors (e.g., estrogen receptor), and ion channels (e.g., voltage-gated sodium channels). These protein families make up the majority of what is called the "druggable genome", molecular targets thought to provide an opportunity for therapeutic intervention and of high interest to the pharmaceutical industry [28]. As a result, hundreds of HTS assays have been developed to support this drug discovery research. Since the vast majority of these potential drug targets have been selected based on believed critical roles in various pathological processes, extension of this thinking suggests that such targets could also be involved in toxicity when inappropriately perturbed by xenobiotic chemicals. This served as the impetus for developing a diverse suite of HTS assays to use for profiling the biological activity of chemical libraries by several groups including ourselves through the ToxCast program [4, 11, 12].

In vitro HTS assays facilitate the rapid, parallel generation of large numbers of individual assay data points through the use of miniaturized assay formats, automated liquid dispensers, and high-speed plate readers. The miniaturized assay formats are usually in multi-well plates with densities of 96, 384, or 1536 wells per plate in a single, standardized plate footprint, and use total assay volumes ranging from 200 μL down to 5 nL. The assay components can be highly varied and depend to a large degree on the biological target being measured. For instance, an assay measuring kinase activity would have a purified kinase, required cofactors, required substrates, appropriate buffer, and chemical to be tested. In addition, a means of measuring the assay endpoint, here the phosphorylation of the substrate, is required. This could be a radioactive or fluorescent technique, a means to detect the loss of ATP or the increase in ADP, or a separation of the phosphopeptide from the nonphosphorylated one by means of mobility shift microfluidics assay technology. Cellular phenotypic assays use *in vitro* cultured cells and automated, fluorescence microscopy to image

chemical perturbations of cellular structures, organelles, and functions followed by specific imaging algorithms to quantitate results [29]. Examples of this are assays using fluorescently tagged antibodies to actin microfilaments to monitor chemical affects on the stabilization or destabilization of the cytoskeleton [30].

The diversity of techniques used to quantitate HTS results underscores a critical point of understanding of HTS assays; all assays are susceptible to artifacts, and different assay formats are susceptible to different types of artifacts [31]. Assay artifacts are defined as test chemical-induced events that interfere with the ability to measure an accurate assay result such as chemical-induced fluorescent quenching, precipitation of the biological target by chemical aggregation, and inherent chemical fluorescence among others. Thus, an underlying caveat of any HTS assay is that all results must be interpreted with caution. In addition to artifacts induced by specific test chemicals, there are also experimental errors and normal assay variabilities that can affect the results. While HTS assays should be validated according to industry standards (http://spotlite.nih.gov/assay/index.php/) inherent in testing large numbers of chemicals is that some results will not be accurate. Inaccurate results can generate both false-positive and false-negative findings. Each has its own issues.

For toxicity testing, it is strongly desirable to avoid false-negative results that could miss important activities potentially resulting in endangering the health of exposed populations. Too many false-positives, however, can invalidate the utility of the screening by requiring extensive measures to follow-up on the large number of active chemicals. Unfortunately, decreasing the false-negative rate is usually at the expense of increasing the false-positive rate; thus, finding the right balance with a robust HTS assay is of high importance. Two methods of utility in providing high-confidence results for HTS toxicity testing are to use a concentration–response format for testing all chemicals and to have multiple assays using different assay technologies for important targets. Concentration–response testing allows testing concentrations high enough to detect the activity of weakly active chemicals, while minimizing concern for high concentration-induced artifacts such as cytotoxicity, which can mimic inhibition of functional activity in a cell-based assay. In addition, knowledge about the types of response expected for specific biological targets can help discriminate between chemicals affecting the target from those active by artifact. Receptor binding assays, for example, should follow the law of mass action and resulting concentration–response curves should display sigmoidal behavior with a slope near one on a semi-log plot [32]. Results with slopes of 10, for example, should flag the response as potentially suspect. Orthogonal assays are particularly useful as, for example, the use of a radioligand receptor binding assay and a cellular transactivation assay for the estrogen receptor. Chemicals active in both would have a high degree of confidence of being truly active at the receptor site and likely active *in vivo,* assuming the chemical reaches its receptor target. The efficiency of HTS supports both of these approaches by providing inexpensive screening methods with sufficient capacity to screen both large numbers of chemicals and at multiple concentrations [33].

However, given the sheer numbers of possibilities, testing of all potential toxicity targets is not feasible even with HTS technologies. Selection of the assays for testing

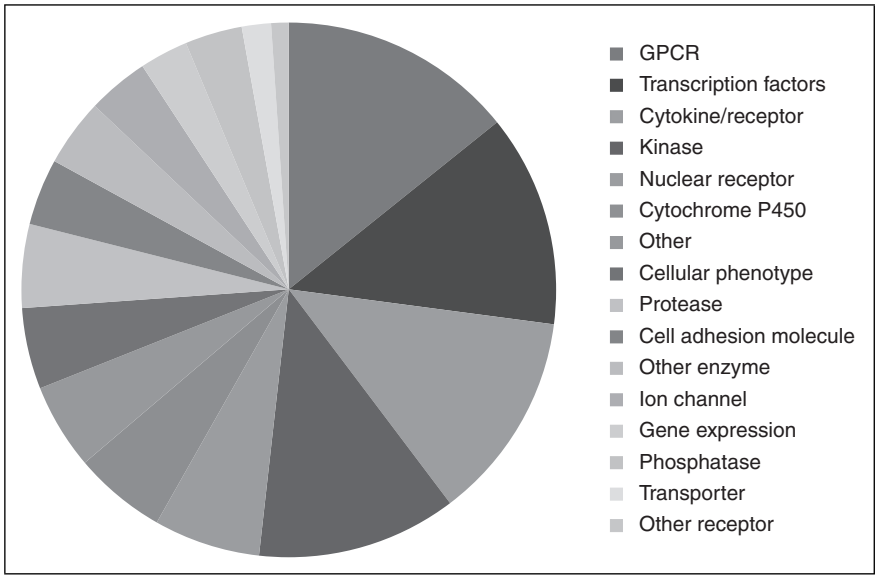


FIGURE 1.1 Distribution of assays categories in the ToxCast Phase I testing battery. (*See insert for color representation of this figure.*)

within the ToxCast program followed a strategy of selecting targets with known links to toxicity, for which assays were available, combined with widely sampling potential targets from the large protein superfamilies including GPCRs, kinases, phosphatases, nuclear receptors, chromatin-modifying enzymes, CYP P450s, ion channels, and transporters [34]. A list of the families and numbers of assays targeting specific molecular targets is shown in Figure 1.1. Sampling of these families may provide a window into potential chemical activity, even when the specific target of a chemical-induced toxicity is not included. This occurs through testing in a concentration–response format, which may allow the detection of chemical promiscuity at higher concentrations. This can be helpful when a specific target of toxicity is not included in the assay suite. Due to conservation of protein structure within families, it is somewhat more likely that a chemical will affect other closely related family members, but with different affinities. These may serve as assay surrogates for the actual target and may still be useful in developing signatures of toxicity.

The use of cellular assays provides a means to include large numbers of potential targets concurrently in a more physiologically relevant format. Such assays usually rely on coordinated signaling networks to carry out the downstream function being measured, for example, cell proliferation. There are many nodes in the pathways regulating cell proliferation that are potentially susceptible for chemical perturbation. These include growth factor receptors on the plasma membrane, kinase second messengers transmitting the growth signal to the nucleus, transcriptional regulatory and protein synthesis machinery, mitotic spindle apparatus, cytoskeletal components, and associated regulatory enzymes. It is because of this complexity that cell proliferation

has been described as one of the most sensitive endpoints for *in vitro* toxicology [35]. Endpoints may be more narrowly defined, such as mitochondrial function or DNA damage endpoints, but these also have many potential upstream targets. Thus, cellular assays, in general, lack the ability to clearly identify molecular mechanisms of action. They do, however, put the molecular targets in a more physiologically relevant context than generally are found with cell-free, biochemical assays. A valuable strategy is a combination of a biochemical assay, in which chemical specificity can be defined, as well as a cellular assay, in which functional efficacy can be demonstrated. The assays used in the ToxCast program provide both a broader coverage of toxicity targets as well as opportunities to define cellular efficacy for chemicals active in the biochemical screens.

Choosing the appropriate cell type for HTS assays supporting predictive toxicology approaches is important to the success of the approach. Many factors need to be considered and these vary, depending on the goal of the assay. In measuring the ability of a chemical to perturb a specific molecular target such as a kinase or a nuclear receptor, it may be appropriate to use standard, immortalized cell lines that provide robust and highly reproducible results. However, in determining the effects of chemicals on complex signaling pathways, the use of such cells may be of little value if these pathways have been altered during the immortalization process and adaptation to growth under standard cell culture conditions. In this case, the use of primary cells may have distinct advantages and provide more physiologically relevant information useful to predicting *in vivo* effects [36, 37]. However, the use of primary cells also has its limitations in terms of limited passage numbers, batch-to-batch variation, difficulty to engineer with respect to introducing reporter genes, and lack of large signal-to-background ratios for the endpoints being measured.

To effectively use data from HTS assays for computational toxicology approaches, it is very useful to acquire complete testing datasets, meaning testing all chemicals against all assays in the testing set and to define standard data handling and analysis procedures. The ToxCast project used defined chemical libraries, described earlier, in testing against suites of *in vitro* assays in a concentration–response format. All chemicals were run in all assays as minimizing missing values in the data matrix greatly enhances the value of the dataset for subsequent analysis. Screening results were used to generate AC50 values, the concentration at which an assay is activated or inhibited by 50% when compared to the control values, for each chemical–assay pair. The AC50 is somewhat arbitrary in that it often has no direct toxicological interpretation. However, it does provide a means of comparing chemicals within an assay, serves as a flag for activity for a chemical in a given assay, and provides information as to its general potency range.

The concentration–response curves for ToxCast are modeled by the four-parameter Hill equation [38] implemented in the R programming language [39]. Heuristics are employed to accommodate aspects of assay results that cause implementation of the Hill equation to fail. The reasons underlying the curve-fitting failures may apply to all assays or be specific to a given assay or platform. For example, results that show no concentration-dependent increase in activity but rather maximal activity at all concentrations tested must be flagged with an AC50 less than the lowest

concentration tested. Assays susceptible to cytotoxicity at high concentrations often need the responses obtained at these cytotoxic concentrations removed from the curve-fitting routine. Since it would be very difficult to generate a universally accepted best method for doing curve-fitting to a wide variety of biological assays, it is important to provide transparency to the process used, as well as access to the underlying unprocessed data in order for others to apply their own techniques.

The combination of chemicals and assay AC50 values defines the basic data matrix required for the computational toxicology input. Value can be added to the matrix through additional metadata. One very useful type of metadata is the mapping of the assays to specific gene ontologies, which are tied to biological pathways annotated by databases such as GO or KEGG [40]. This bioinformatics approach links chemical effects to biological pathways that can provide an additional connection to toxicity endpoints. The annotation is relatively straightforward for most biochemical assays targeting single proteins. However, the ability to do this properly with cellular assays is more challenging, since often a specific molecular target of the assay is not known. In some cases, specific biological pathways could be used to annotate the assay endpoint. This approach will be illustrated in Sections 6 and 7.

Development of a complete, well-annotated data matrix consisting of curated chemical structures and their activity against well-characterized biological targets is the core component of a computational toxicology approach. Such a dataset could be the final product for a predictive toxicology effort, if the biological assays were all highly validated surrogates for *in vivo* toxicology. However, as previously discussed, few such validated targets exist. Therefore, one needs to identify which assays or groups of assays are linked to toxicity endpoints and can serve as signatures for *in vivo* toxicity. We thus focused much of our early ToxCast screening on chemicals with rich *in vivo* toxicity information to use as an anchor for our *in vitro* results. The development of the *in vivo* database to support this effort will be described next.

1.5 *IN VIVO* TOXICITY DATABASE

ToxRefDB (Toxicity Reference Database) aims to capture traditional animal toxicity studies across a variety of study types and endpoints, including short-term and long-term systemic toxicity, cancer, reproductive toxicity, and developmental toxicity [7]. The ToxRefDB project initially focused on capturing previously unpublished high-quality regulatory guideline studies required for chemical safety evaluation by the EPA. The study submissions were reviewed by the EPA's Office of Pesticide Programs (OPP) and results consolidated into Data Evaluation Record (DER), which is the primary data source for ToxRefDB. Study results from this DER as well as from other high-quality, publically available studies have been manually curated into ToxRefDB's relational database model. The relational database for ToxRefDB ensures data integrity by forcing specific vocabulary to be used across all major ToxRefDB fields. The ToxRefDB relational format follows the following logic: A chemical can have many studies performed, each study can have multiple treatment groups (male and female, low-, mid-, and high-dose), and each treatment group can

TABLE 1.1 Study and Chemical Counts from the ToxRefDB Website

Summary Statistics	
Study count	1,978
Chemical count	474
Combined chronic/cancer rat	324
Combined chronic/cancer mouse	324
Multigenerational reproductive rat	352
Prenatal developmental rat	365
Prenatal developmental rabbit	331
Subchronic rodent	302

observe many effects. ToxRefDB has subsequently been integrated into the ACToR system, primarily through generic chemical linkages (i.e., CASRN) and is available as a searchable database (http://actor.epa.gov/toxrefdb). ToxRefDB was designed to capture detailed study design, dosing, and treatment-related effect information. In addition to the relational design of the database, controlled and standardized vocabularies were used for the vast majority of fields to ensure the uniformity of the manually curated and entered legacy toxicity information. The current publically released version of ToxRefDB has study and chemical effect information on 474 chemicals, primarily pesticides due to their consistent and large data coverage of chronic, cancer, reproductive, and developmental studies. The "Basic Info" page on the ToxRefDB website contains summary information about the database and the associated manuscripts. Importantly, the manuscripts release supplemental files with aggregated and detailed endpoints across the full ToxRefDB chemical library. These "flattened" endpoint files (i.e., flat tabular listings) have been directly incorporated into the ToxCastDB system for predictive modeling exercises. The "Basic Info" page also provides information on the current database and chemical coverage counts for each study type (Table 1.1).

The "home" page of ToxRefDB, similar to that of all ACToR system databases, allows the user to search by generic chemical. As an example, the key word "azole" was used to search all 474 chemicals in ToxRefDB, by both their assigned chemical name and all synonyms, and resulted in the return of 46 chemicals (Fig. 1.2). The red boxes indicate whether or not a study is available in ToxRefDB for the particular study type. A "Generic Chemical Page" is displayed, as shown in the ACToR website. However, when accessing the ToxRefDB portion of ACToR, only chemicals with traditional toxicity data captured in ToxRefDB can be viewed. Under the "Toxicology Data" heading, all ToxRefDB data are displayed in a three-tiered structure. The first tier contains the study design information, including data quality, species and strain, dose administration, study type, and citation information. The second tier contains treatment group and dosing information, while the third tier indicates the treatment-related effects observed at the various dose levels. The study information is available for viewing, but, due to the amount of detailed information stored within each tier, the system does not currently allow for detailed filtering of the data. However, a full

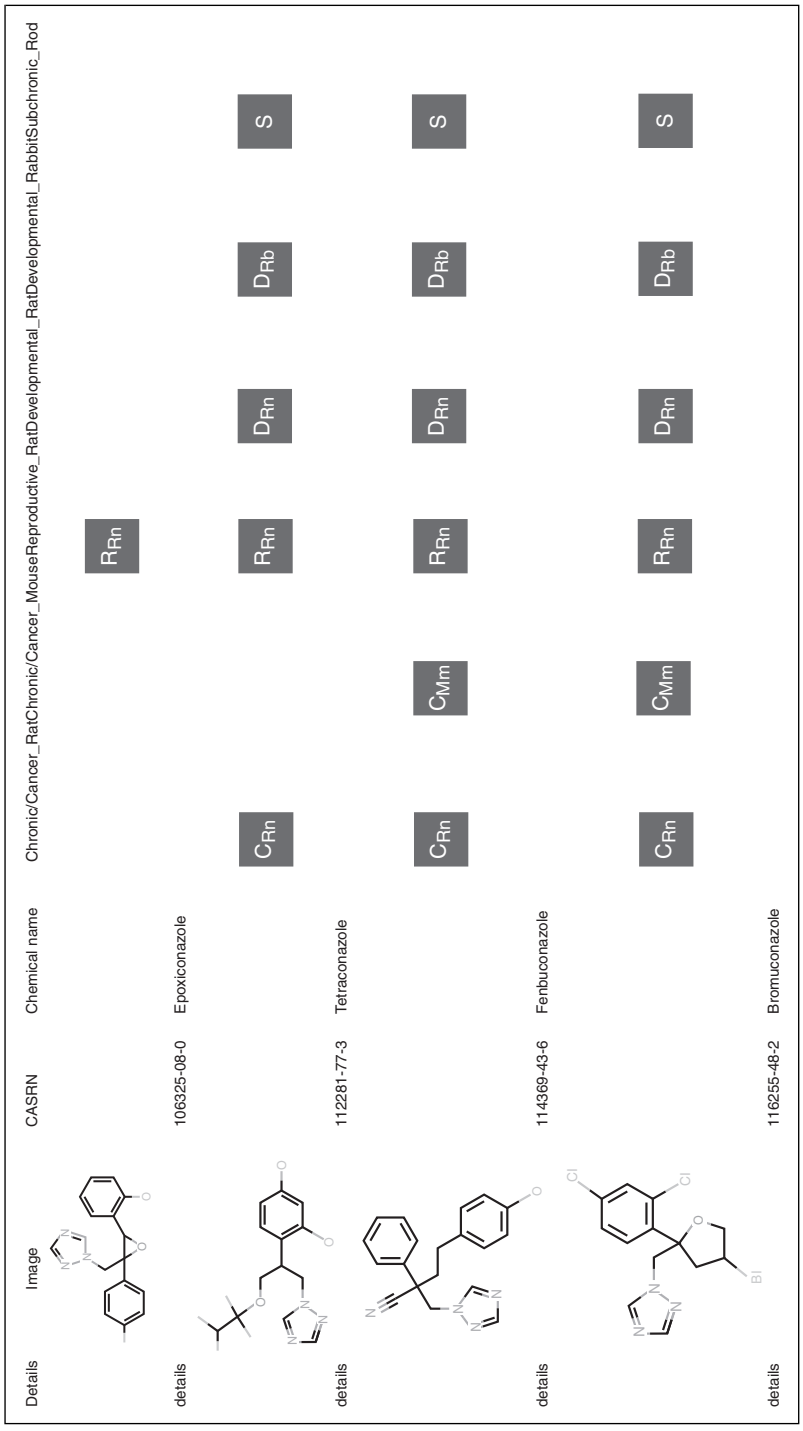


FIGURE 1.2 Screen shot from the ToxRefDB website for the high-level view of data available for a set of chemicals across the various available study types.

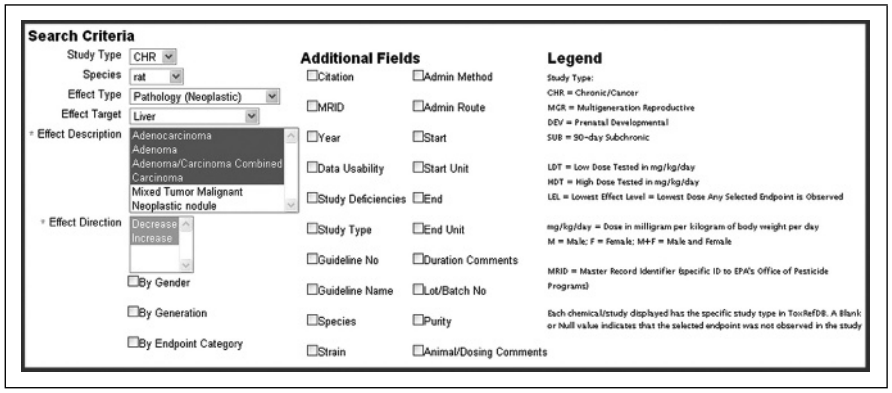
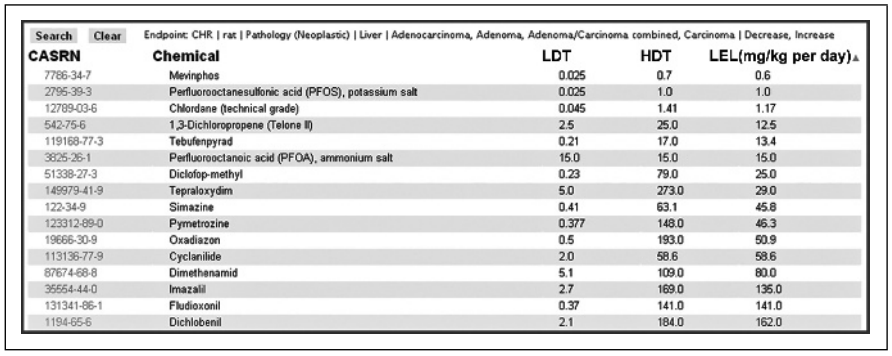


FIGURE 1.3 Screen shot from the ToxRefDB website of the endpoint search page with the search criteria and additional field information to be included.

download of the ToxRefDB data is available for each chemical as a csv file, enabling further analysis and viewing options.

The primary search tool currently available within the ToxRefDB system is located in the "Search by Endpoint" tab. The page allows the user to select from the standardized effect vocabulary, the exact search criteria of interest as well as the additional field information to be displayed (Fig. 1.3). The results of searching, for example, "Chronic/Cancer Rat Liver Neoplastic Pathology" returns the lowest effect level (LEL) in mg/kg/day dose, which represents the lowest dose at which a treatment-related change in the selected effect or effects was observed (Fig. 1.4). Each row from the returned search represents a unique study in ToxRefDB, with the low and high dose tested (LDT and HDT) provided for reference. If multiple effects are selected, a single LEL is returned, which aggregates all selected effects with a primary goal of providing the field of predictive toxicology a tool for rapidly defining


Search Clear Endpoint: CHR | rat | Pathology (Neoplastic) | Liver | Adenocarcinoma, Adenoma, Adenoma/Carcinoma combined, Carcinoma | Decrease, Increase

CASRN	Chemical	LDT	HDT	LEL(mg/kg per day)
7786-34-7	Mevinphos	0.025	0.7	0.6
2795-39-3	Perfluorooctanesulfonic acid (PFOS), potassium salt	0.025	1.0	1.0
12789-03-6	Chlordane (technical grade)	0.045	1.41	1.17
542-75-6	1,3-Dichloropropene (Telone II)	2.5	25.0	12.5
119168-77-3	Tebufenpyrad	0.21	17.0	13.4
3825-26-1	Perfluorooctanoic acid (PFOA), ammonium salt	15.0	15.0	15.0
51338-27-3	Diclofop-methyl	0.23	79.0	25.0
149979-41-9	Tepraloxydim	5.0	273.0	29.0
122-34-9	Simazine	0.41	63.1	45.8
123312-89-0	Pymetrozine	0.377	148.0	46.3
19666-30-9	Oxadiazon	0.5	193.0	50.9
113136-77-9	Cyclanilide	2.0	58.6	58.6
87674-68-8	Dimethenamid	5.1	109.0	80.0
35554-44-0	Imazalil	2.7	169.0	135.0
131341-86-1	Fludioxonil	0.37	141.0	141.0
1194-65-6	Dichlobenil	2.1	184.0	162.0

FIGURE 1.4 Screen shot from the ToxRefDB website of the endpoint search page with the results of the search displayed.

endpoints across a large chemical library. The endpoint search tool can also be used for researchers interested in delineating a set of reference chemicals with positive and negative outcomes for a particular effect or endpoint.

ToxRefDB has been applied to multiple problem types, including retrospective and prospective questions. Retrospectively, ToxRefDB has been used to assess the impact of specific traditional toxicity endpoints and parameters on the safety regulation of chemicals. For example, traditional toxicity testing for reproductive toxicity potential has relied heavily on a two-generation reproductive toxicity study in rats. However, the importance of the second generation has come into question [41]. An extended one-generation protocol has been proposed that would only produce a second generation when triggered, would require far fewer animals, and would derive more toxicological and kinetic information from each animal used. To assess the impact of the second generation on risk assessment, ToxRefDB was used as a data source to systematically evaluate the question, relying on the highly standardized vocabulary and relational format of ToxRefDB. Based on the data in ToxRefDB, the analysis indicates that the second generation does not greatly impact the interpretation of the reproductive study from a risk assessment perspective. The two-generation retrospective analysis demonstrated the ability of ToxRefDB to provide a systematic review of traditional toxicity studies. Additional retrospective analyses are underway, including the analysis of the relative impact and importance of running both rat and rabbit in prenatal developmental toxicity studies.

ToxRefDB also stores no-observed and lowest-observed adverse effect levels (NOAEL and LOAEL) for studies reviewed by EPA and used in the chemical registration process. The Threshold of Toxicological Concern (TTC) is an approach that uses NOAEL/LOAEL distributions and chemical structure characteristics to establish safe exposure levels for chemicals with limited to no toxicity information [42]. ToxRefDB is currently being applied to TTC approaches in numerous venues, including assessing the applicability of the standard TTC to antimicrobial pesticide products and the refinement of TTC approaches for specific chemical classes. In the example of the antimicrobial TTC study, all available toxicity study information on antimicrobials is being collected and entered into ToxRefDB. Antimicrobial pesticides typically have less available toxicity data when compared to conventional pesticides, and this underscores the need for alternative safety assessment approaches. With the full food-use antimicrobial traditional toxicology dataset available in a standardized and relational format, detailed analysis of the NOAEL/LOAEL distributions across study type, endpoint categories, and structural classes can be obtained and compared to other TTC analyses. If found to be similar, then all or a portion of antimicrobials could be evaluated using a TTC approach.

1.6 PREDICTIVE MODELS

An important use of the HTS data we produce is to develop predictive signatures of particular types of toxicity. A signature is a pattern of *in vitro* assay hits that is predictive of a particular toxic endpoint. The basic approach used is as follows.

First, we select a set of chemicals for which we have both *in vitro* data and *in vivo* phenotype information, based on data from ToxRefDB. The phenotype or endpoint can either be a quantitative value (e.g., the LEL) or a yes/no (either causes or does not cause the phenotype, at any dose). We then look for statistical correlations between the chemicals active in one or more *in vitro* assays and the endpoint. Concretely, we ask if chemicals that are active in a particular assay are significantly more likely than chance to be positive for the phenotype. A variety of standard statistical methods can be used, including multivariate machine learning techniques. Whenever doing these types of analyses, one needs to be particularly careful to not overfit a statistical model, which would result in the creation of a model with little or no forward predictive value. One method that one almost always uses is called cross-validation, in which the dataset is divided into training and test sets. A model is built on the training set and evaluated on the test set. If the performance in the latter is not significantly better than chance, the model is rejected. After the statistical model is built, one typically looks at the biological meaning of the signature, for example, is there a support in the literature for a linkage between the target that is probed by the assay and the endpoint being evaluated? Finally, one may then modify the model to enhance the biological content while retaining the statistical predictive power. The ultimate test of a model is always a forward validation against an entirely new dataset.

Here we briefly describe several signatures we have developed using the ToxCast and ToxRefDB data from Phase I of the ToxCast project. The chemicals in this set are largely pesticidal active ingredients, so they have a wealth of high-quality *in vivo* animal toxicology data. Judson et al. [6] show significant correlations between several pathways and preneoplastic and neoplastic liver lesions in rats. The target genes included the peroxisome proliferator-activated receptors (PPARs)—PPARα and PPARγ. Activation of these receptors is a well-documented cause of liver tumors in rodents [43], so that this is an external validation that our screening and statistical approach can recover correct biological links. Kleinstreuer et al. [44] have demonstrated associations between particular pathways leading to disrupted vasculogenesis and involving inflammatory chemokine signaling, the vascular endothelial growth factor pathway and the plasminogen-activating system. Activity in these pathways can lead to limb malformations during embryonic development, as demonstrated in prenatal developmental studies in rats and rabbits. This analysis is based on the data from ToxCast and ToxRefDB. Martin et al. [45] have used these same approaches to develop signatures for predicting reproductive toxicity. They built composite endpoints of male and female effects, including fertility and reproductive fitness. The signature included assays related to endocrine disruption (estrogen and androgen receptor), PPAR activity, liver metabolism as evidenced through activity in a pregnane X receptor (PXR) assay and in CYP450 assays, and generalized activity against GPCRs. This model produced a balanced accuracy for sensitivity and specificity of >0.70 in both a cross-validation test set and an independent forward validation set [45]. Finally, Sipes et al. [46] have developed a model of cleft palate and urogenital defects in rat and rabbit. The assays statistically associated with these endpoints include the retinoic acid receptor (RAR), interleukins 1A and 8, and the transforming growth factor β (TGF-β).

Using HTS data for predictive modeling may have opened the door for developing an accurate predictive model of reproductive toxicity. Reproductive toxicity is an aggregated multi-modal and multi-effect outcome. No single assay has the ability to broadly identify reproductive toxicants and, to date, traditional structure-based and other methods, have not been able to produce an externally validated predictive model of reproductive toxicity. Computational modeling of HTS data allows one to explore the complex relationships between *in vivo* observations and networks of *in vitro* activity. One of the more simplistic computational modeling approaches is the development of a classification model, which aims to accurately classify or predict an outcome based on a training set with known outcomes. The training set for modeling reproductive toxicity was the set of chemicals in the ToxCast library with high-quality reproductive toxicity data [45]. The initial inputs into the model were the hundreds of ToxCast assays that were collectively mapped to genes and the aggregate activity across the assays per gene provided the quantitative inputs into the model. The assay–gene combinations were further filtered based on a feature selection process that evaluated the statistical association to the training set data. The filtered gene set was then weighted in a multivariate model using linear discriminate analysis (LDA) and fivefold cross-validation. Many other approaches and methods could have been deployed, but our observation has been that the use of complex machine learning algorithms has a tendency to over-fit the data lowering the output model's ability to be externally predictive. The resulting internal model performance statistics were greater than or equal to 75% balanced accuracy, and there was no significant difference between the training and test set accuracies. The final combined model produced a balanced accuracy of 80%.

Among the chemicals selected for external validation, the model provided accurate predictions for 16 of the 21 chemicals. The five chemicals with inaccurate predictions provide valuable insight into potential limitations or gaps of the model. Interestingly, the five chemicals had a common phenotypic profile with respect to reproductive toxicity causing reduced early offspring survival, particularly the litter size decreases with little to no accompanying effects on reproductive performance or reproductive tract pathology. The reproductive LOAEL for all the five chemicals was set at the high dose tested based on the early offspring survival effects, and the parental and offspring LOAEL were set at the lower dose levels. Based on the inclusive definition used for defining a positive for reproductive toxicity for model development, all the five were considered positive but lack evidence of specific fertility-related or developmentally sensitive reproductive outcomes. Nonetheless, a gap in model predictivity was identified and could potentially be filled using additional assay technologies, physical–chemical properties, or structural descriptors.

The model development process identified biologically plausible features and pathways from over 500 assays mapped to less than 100 genes or gene sets and spanning many reproductively relevant MOA. PPARα activity was clearly associated with reproductive toxicity, with all 10 PPARα agonists in the training set causing reproductive toxicity. Although a mechanistic link between PPAR activity and fertility or other reproductive impairments remains unclear [47], the role of PPAR in steroid

metabolism and its activity in reproductive tissues infers that it is a plausible target for the disruption of endocrine signaling and altered gametogenesis. Androgen and estrogen α activities were also associated with reproductive toxicity. The ToxCast receptor profiling identified, most if not all, the known antiandrogenic and estrogenic chemicals in the current dataset, but the causal relationship between reproductive toxicity and steroid receptor activity, absolute and relative potency and efficacy needs to be explored further. CYP enzyme inhibition, as compared to gene induction, was significantly associated with reproductive toxicity. Alterations in steroid metabolism through CYP induction have previously been associated with reproductive impairment [48]. However, the nonspecific inhibition of CYPs may be a surrogate for the capacity of a chemical to disturb steroid metabolism including inhibition of key CYPs involved in steroidogenesis (e.g., CYP19 and CYP17). Related to CYP activity, PXR interestingly displayed a negative correlation/association with reproductive toxicity. In general, PXR lowered the false positive rate of the model by lowering the model score of chemicals with nonspecific and low-potency nuclear receptor activity. Robust PXR activity is an indication of potent xenobiotic sensing and potentially rapid metabolism. A major component of the model not directly related to nuclear receptor biology and xenobiotic/steroid metabolism was GPCR binding. Numerous GPCR binding assays were significantly associated with reproductive toxicity. Those chosen to represent the GPCR family were selected for statistical and not for biological reasons as there is limited literature information on their role in reproduction in contrast to their well-characterized role in the nervous system function. Platforms measuring epidermal growth factor receptor, TGFβ1 and NF-κB activity were also associated with reproductive toxicity. All the three gene products have been shown to modulate the relative sensitivity of developmental toxicants, especially AhR signaling [49, 50], and may be indicative of altered xenobiotic metabolism, cellular proliferation, cell–cell signaling, or potential epigenetic effects [51, 52]. Overall, the key targets in the model identify plausible MOA leading to reproductive toxicity and covering antiandrogenic, estrogenic, cholesterol/steroid metabolism, limited coverage of disruption of steroidogenesis, and altered xenobiotic metabolism MOA.

With the availability of an externally validated classification model predicting reproductive toxicity, the bottleneck of uncharacterized chemicals can be evaluated either through improvement in the overall statutory authority to request multigenerational reproductive studies or in the ability to quantitatively identify reproductive toxicants. If the statutory authority to request these studies were improved, then the current model in concert with other models, alternative methods, and institutional knowledge could identify with fairly good accuracy and efficiency for all chemicals for which a multigenerational reproductive study should be requested. If the latter were improved to the point of accurate adverse dose predictions, then the model could drastically decrease the need for multigenerational studies and be used in the assessment of the majority of environmental chemicals. To do this, improvements in HTS assay reproducibility, metabolic capacity, mode-of-action coverage, reverse toxicokinetics, and overall model accuracy would need to be made. Placing the classification

model into a system modeling context will begin to address this next generation of research questions. For now, the predictive model of reproductive toxicity can go a long way in improving reproductive chemical testing efficiency and decision making.

1.7 CHEMICAL PRIORITIZATION

Alternative testing methods (*in vitro* and in silico) will require extensive refinement and validation before they ultimately replace standardized animal testing. Their more immediate utility may lie in their use as tools for chemical prioritization of the extensive chemical inventories of exposure concern. Prioritization is the key linkage between rapidly deployed, computational models and downstream applications. These downstream applications may address the question "In what order do we test chemicals?" or "Which sets of chemicals deserve targeted testing first?" Both of these questions involve prioritization in terms of ranking, but the second question also involves communication of reasons underlying a ranking. Thus, the ideal prioritization approach yields both an explicit, prioritized order of chemicals as well as a transparent look at the evidence used.

There are several prioritization and decision-analysis approaches being developed to support chemical prioritization needs [53]. One example is the ToxPi (Toxicological Prioritization Index) framework [54]. This approach can be tailored to diverse sets of chemicals, evidence (data), and prioritization tasks, because it is based on relative—rather than absolute—rankings. It satisfies both the rank and evidence aspects above, as ToxPi provides a visual, weight-of-evidence index that can be used to rank and compare chemicals. Its initial application was to aid in the prioritization of chemicals for putative endocrine activity, in support of the Endocrine Disruptor Screening Program (EDSP). Composite activity scores across sets of endocrine relevant data from the ToxCast assay battery for each chemical were calculated to rank all 309 ToxCast Phase I chemicals from the highest to lowest priority. Inclusion of well-studied reference chemicals from the domain of the prioritization is particularly valuable. In the EDSP example, the Phase I chemicals contained BPA, methoxychlor and its active metabolite 2,2-bis(*p*- hydroxyphenyl)-1,1,1-trichloroethane (HPTE), all well-studied chemicals with estrogenic activity. Such chemicals serve to put results for unknown chemicals in a better toxicological perspective. In the face of practical temporal and economic limitations, this estimate of potential endocrine activity provides a formal rationale for prioritizing resources toward further testing. Alternatively, the ToxPi profiles could be used for chemical "read-across", analogous to QSAR structural alert models. The read-across can be implemented in terms of overall ToxPi bins/clusters of chemicals having similar profiles, or subsets of slices can be interpreted as *in vitro* "alerts" to support targeted testing decisions.

Future prioritization approaches will have an increasing demand for transparency and interactivity. This demand is driven by a more informed public, as well as the thoughtful work of many non-governmental organizations (NGOs) and government entities. Transparency in both data and models is facilitated by web-accessible

databases and software tools. These interactive prioritization interfaces allow formal input into decisions by stakeholders, regulators, and the public.

1.8 TARGETED TESTING

The objective of targeted testing of chemicals is to reduce the use of test animals by limiting the endpoints that need to be evaluated. The way we identify the endpoints for targeted testing on a given chemical is not straightforward. For new chemicals and for thousands of existing chemicals in commerce, there is little or no information about biological effects. In such cases structural similarity between a new chemical and a well-studied chemical may be used to infer plausible similarity of biological effects. While this approach works for some endpoints such as genotoxicity, it is currently difficult to accurately classify chemicals by hazard based on structural information alone. The availability of high-throughput bioactivity profiling makes it feasible to rapidly produce a rich overview of molecular and cellular effects of thousands of chemicals. Bioactivity profiles can be used to guide targeted tests in two main ways. First, bioactivity profiles can be phenotypically anchored to adverse effects from animal testing to discover predictive signatures to classify chemicals by hazard. Second, the molecular and cellular activities in the signatures can be used to relate early events in the pathways to adverse outcomes. Knowledge about such pathways, which are also known as adverse outcome pathways (AOPs) [55], can be combined with signatures to improve confidence in their predictions. Predictive signatures can serve as a practical tool for identifying endpoints for targeted testing and Section 4 describes their development from bioactivity profiles. Briefly, a signature is developed using computational tools that mine hundreds of thousands of associations between bioactivity patterns and an endpoint. We select signatures that are statistically significant and that accurately classify known toxicants. This approach has been used to develop and evaluate signatures for cancer, reproductive toxicity, and developmental toxicity endpoints using ToxCast bioactivity profiles, and animal toxicology outcomes from ToxRefDB. In principle, the bioactivity profiles of new chemicals can be matched with signatures to select endpoints for targeted testing. For instance, if a new chemical activates inflammatory chemokine signaling, the vascular endothelial growth factor pathway and the plasminogen-activating system according to the signature proposed by Kleinstreuer et al. [44], it may be a developmental toxicant. Similarly, if a chemical is a potent activator of multiple nuclear receptors, it could be a hepatocarcinogen [43] and may warrant testing in a 2-year bioassay.

To use a predictive signature to propose an endpoint for targeted testing, it should be objectively evaluated using known chemicals first. The predictive accuracy of the signature is a measure of confidence but does not guarantee future performance. Future performance depends on the reproducibility and biological breadth of the assays (Section 4), the diversity of the chemical landscape (Section 2), and a sufficient number of chemicals to build statistically rigorous signatures of adverse effects. Another approach for improving confidence in empirically derived signatures is to highlight their role in pathophysiological processes that lead to toxicity.

A major challenge for using bioactivity profiling for targeted testing will be to provide confidence in their ability to extrapolate from *in vitro* to chronic human health outcomes. It is extremely difficult at this time to elucidate the detailed sequence of cause and effect relationships, so translating *in vitro* assays and *in vivo* outcomes will always involve a level of uncertainty. In addition to this uncertainty, a significant difference between the two approaches further confounds correlation due to the general lack of xenobiotic metabolism with *in vitro* approaches. This metabolism could either activate a chemical to a toxic metabolite or inactivate a toxic chemical to a less toxic one, thus leading to incongruent *in vitro* and *in vivo* results. However, it is important to note that extrapolating adverse effects from test species to humans is also fraught with considerable uncertainty. For instance, a number of PPAR activators are rodent hepatocarcinogens, but it is difficult to evaluate the relevance of this outcome in humans given the widespread use of pharmaceuticals targeting the same receptor family [56].

The effectiveness of targeted testing will improve as diverse evidence about the events that lead to human toxicity is organized, such as AOPs. An AOP describes the initiation of molecular events by chemicals, followed by a complex sequence or network of key molecular, cellular, and tissue level changes that culminate in an adverse effect [55]. In the context of an AOP, we can assume that bioactivity assays directly or indirectly measure changes in early molecular events, and that the endpoints are the toxic outcomes. For instance, PPARα activation may be considered a molecular initiating event that begins a cascade of subsequent changes. Persistent stimulation with PPARα activators may lead to proliferative lesions that can progress to neoplasms. However, the AOP in this case includes a series of intermediate effects, among others sustained cell proliferation, hepatic hyperplasia, and preneoplastic lesions.

This knowledge can be used to assess the events preceding the apical endpoint such as cell cycle progression. This is valuable because apical outcomes may only be measured in animals, whereas the incipient events such as cell cycle changes may be observed in cell culture models. Targeted testing approaches are feasible now by using predictive signatures derived from high-throughput bioactivity assays. We believe that confidence in these signatures will improve over time as more chemicals undergo HTS via programs such as ToxCast. In addition, growing knowledge about AOPs may enable additional tiers of *in vitro* tests, thereby further reducing the need for animal testing. Sophisticated system models [57, 58] that can accurately estimate the dynamic changes in AOPs in humans at environmentally relevant exposures may one day help us design a battery of tiered tests that eliminate the need of animals to evaluate chemical safety.

1.9 CONCLUSION

Changing the nature of toxicity testing in the interest of better identifying and characterizing the potential for risk to the health of human and other populations is a central need in the environmental toxicology field. The use of HTS technologies

and computational toxicology approaches is currently in the initial proof-of-concept phase. Much progress has been achieved to date including the demonstration of the feasibility of testing large chemical libraries in diverse HTS assays, collating *in vivo* toxicity information on myriads of chemicals in a relational database, and building predictive models and prioritization schemes for a number of important toxicological endpoints. However, much work remains to be performed. Many important toxicity endpoints have not yet been modeled. Many chemical classes have not yet been tested and remain as major obstacles to HTS screening. Lack of significant xenobiotic metabolic activity in *in vitro* assays remains yet another challenge. Finally, models will require extensive refinement and validation before they will serve regulatory purposes. However, the success of these early steps provides hope that continued research efforts in this arena will eventually lead us to our goal of an efficient, robust, *in vitro* predictive toxicity screening program that will serve the needs of the public by providing the capacity to routinely screen existing inventories as well as new chemical entities for the potential for harm to the health of human and other populations.

DISCLAIMERS

The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the US Environmental Protection Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

ACKNOWLEDGMENTS

The authors thank Robert Kavlock and David Dix for their invaluable contributions toward developing the computational toxicology approach described here. In addition, we thank Andrew Kligerman and Kevin Crofton for their comments and review of this chapter.

REFERENCES

1. Hartung, T. (2009). Toxicology for the twenty-first century. *Nature*, *460*, 208–212.
2. Judson, R.A., Dix, D.J., Houck, K., Martin, M., Kavlock, R., Dellarco, V., Henry, T., Holderman, T., Sayre, P., Tan, S., Carpenter, T., Smith, E. (2009). The toxicity data landscape for environmental chemicals. *Environmental Health Perspectives*, *117*, 685–695.
3. Wataru, Y., Peterson, R.E., Tohyama, C. (2010). Molecular targets that link dioxin exposure to toxicity phenotypes. *Journal of Steroid Biochemistry and Molecular Biology*, *127*, 96–101.
4. Dix, D.J., Houck, K.A., Martin, M.T., Richard, A.M., Setzer, R.W., Kavlock, R.J. (2007). The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicological Sciences*, *95*, 5–12.

5. Kavlock, R., Chandler, K., Houck, K., Hunter, S., Judson, R., Kleinstreuer, N., Knudsen, T., Martin, M., Padilla, S., Reif, D., Richard, A., Rotroff, D., Sipes, N., Dix, D. (2012). Update on EPA's ToxCast program: Providing high throughput decision support tools for chemical risk management. *Chemical Research in Toxicology*, *25*, 1287–1302.
6. 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., Dix, D.J. (2010). *In vitro* screening of environmental chemicals for targeted testing prioritization: The ToxCast project. *Environmental Health Perspectives*, *118*, 485–492.
7. Martin, M.T., Judson, R.S., Reif, D.M., Kavlock, R.J., Dix, D.J. (2008). Profiling chemicals based on chronic toxicity results from the U.S. EPA ToxRef database. *Environmental Health Perspectives*, *117*, 392–399.
8. Visser, U., Abeyruwan, S., Vempati, U., Smith, R.P., Lemmon, V., Schürer, S.C. (2011). BioAssay Ontology (BAO): A semantic description of bioassays and high-throughput screening results. *BMC Bioinformatics*, *12*, 257.
9. Collins, F.S., Gray, G.M., Bucher, J.R. (2008). Toxicology: Transforming environmental health protection. *Science*, *319*, 906–907.
10. Kavlock, R.J., Austin, C.P., Tice, R.R. (2009). Toxicity testing in the 21st Century: Implications for human health risk assessment. *Risk Analysis*, *29*, 485–487.
11. Fliri, A.F., Loging, W.T., Thadeio, P.F., Volkmann, R.A. (2005). Biological spectra analysis: Linking biological activity profiles to molecular structure. *Proceedings of the National Academy of Sciences USA*, *102*, 261–266.
12. Bender, A., Scheiber, J., Glick, M., Davies, J.W., Azzaoui, K., Hamon, J., Urban, L., Whitebread, S., Jenkins, J.L. (2007). Analysis of pharmacology data and the prediction of adverse drug reactions and off-target effects from chemical structure. *ChemMedChem*, *2*, 861–873.
13. U.S. Food and Drug Administration. (1999). Guidance for industry: Antimicrobial food additives. Available at http://www.fda.gov/Food/GuidanceComplianceRegulatory Information/GuidanceDocuments/FoodIngredientsandPackaging/ucm077256.htm. Accessed May 4, 2012.
14. U.S. Environmental Protection Agency. (2012). Data requirements for pesticide registration. Available at http://www.epa.gov/pesticides/regulating/data_requirements.htm. Accessed May 4, 2012.
15. Cox, C., Surgan, M. (2006). Unidentified inert ingredients in pesticides: Implications for human and environmental health. *Environmental Health Perspectives*, *114*, 1803–1806.
16. Kolpin, D.W., Furlong, E.T., Meyer, M.T., Thurman, E.M., Zaugg, S.D., Barber, L.B., Buxton, H.T. (2002). Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999–2000: A national reconnaissance. *Environmental Science & Technology*, *36*, 1202–1211.
17. U.S. Food and Drug Administration. CFSAN: Scope of responsibility. Available at http://www.fda.gov/AboutFDA/CentersOffices/OfficeofFoods/CFSAN/WhatWeDo/default.htm. Accessed May 22, 2012.
18. Cohen-Hubal, E.A., Richard, A., Aylward, L., Edwards, S., Gallagher, J., Goldsmith, M.R., Isukapalli, S., Tornero-Velez, R., Weber, E., Kavlock, R. (2010). Advancing exposure characterization for chemical evaluation and risk assessment. *Journal of Toxicology and Environmental Health Part B: Critical Reviews*, *13*, 299–313.

19. Judson, R., Richard, A., Dix, D., Houck, K., Elloumi, F., Martin, M., Cathey, T., Transue, T.R., Spencer, R., Wolf, M. (2008). ACToR—Aggregated Computational Toxicology Resource. *Toxicology and Applied Pharmacology*, *233*, 7–13.
20. Huang, R., Southall, N., Wang, Y., Yasgar, A., Shinn, P., Jadhav, A., Nguyen, D.T., Austin, C.P. (2011). The NCGC pharmaceutical collection: A comprehensive resource of clinically approved drugs enabling repurposing and chemical genomics. *Science Translational Medicine*, *3*, 80ps16.
21. Boozer, C., Kim, G., Cong, S., Guan, H., Londergan, T. (2006). Looking towards label-free biomolecular interaction analysis in a high-throughput format: A review of new surface plasmon resonance technologies. *Current Opinion in Biotechnology*, *17*, 400–405.
22. Roy, M., Dumaine, R., Brown, A.M. (1996). HERG, a primary human ventricular target of the nonsedating antihistamine terfenadine. *Circulation*, *94*, 817–823.
23. Marrs, T.C. (1993). Organophosphate poisoning. *Pharmacology & Therapeutics*, *58*, 51–66.
24. Guengerich, F.P. (2001). Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chemical Research in Toxicology*, *14*, 611–650.
25. Szakács, G., Váradi, A., Özvegy-Laczka, C., Sarkadi, B. (2008). The role of ABC transporters in drug absorption, distribution, metabolism, excretion and toxicity (ADME–Tox). *Drug Discovery Today*, *13*, 379–393.
26. Barlow, S., Kavlock, R.J., Moore, J.A., Schantz, S.L., Sheehan, D.M., Shuey, D.L., Lary, J.M. (1999). Teratology Society Public Affairs Committee position paper: Developmental toxicity of endocrine disruptors to humans. *Teratology*, *60*, 365–375.
27. Hutcheson, J.D., Setolab, V., Roth, B.L., Merryman, W.D. (2011). Serotonin receptors and heart valve disease—it was meant 2B. *Pharmacology & Therapeutics*, *132*, 146–157.
28. Hopkins, A.L., Groom, C.R. (2002). The druggable genome. *Nature Reviews of Drug Discovery*, *1*, 727–730.
29. Taylor, D.L., Giuliano, K.A. (2005). Multiplexed high content screening assays create a systems cell biology approach to drug discovery. *Drug Discovery Today*, *2*, 149–154.
30. Cartera, C.A., Hamm, J.T. (2009). Multiplexed quantitative high content screening reveals that cigarette smoke condensate induces changes in cell structure and function through alterations in cell signaling pathways in human bronchial cells. *Toxicology*, *261*, 89–102.
31. Inglese, J., Johnson, R.L, Simeonov, A., Xia, M., Zheng, W., Austin, C.P., Auld, D.S. (2007). High-throughput screening assays for the identification of chemical probes. *Nature Chemical Biology*, *3*, 466–479.
32. Hulme, E.C., Trevethick, M.A. (2010). Ligand binding assays at equilibrium: Validation and interpretation. *British Journal of Pharmacology*, *161*, 1219–1237.
33. Inglese, J., Auld, D.S., Jadhav, A., Johnson, R.L., Simeonov, A., Yasgar, A., Zheng, W., Austin, C.P. (2006). Quantitative high-throughput screening: A titration-based approach that efficiently identifies biological activities in large chemical libraries. *Proceedings of the National Academy of Sciences USA*, *103*, 11473–11478.
34. Knudsen, T.B., Houck, K.A., Sipes, N.S., Singh, A.V., Judson, R.S., Martin, M.T., Weissman, A., Kleinstreuer, N.C., Mortensen, H.M., Reif, D.M., Rabinowitz, J.R., Setzer, R.W., Richard, A.M., Dix, D.J., Kavlock, R.J. (2011). Activity profiles of 309 ToxCast™ chemicals evaluated across 292 biochemical targets. *Toxicology*, *282*, 1–15.
35. O'Brien, P.J., Irwin, W., Diaz, D., Howard-Cofield, E., Krejsa, C.M., Slaughter, M.R., Gao, B., Kaludercic, N., Angeline, A., Bernardi, P., Brain, P., Hougham, C. (2006). High

concordance of drug-induced human hepatotoxicity with in vitro cytotoxicity measured in a novel cell-based model using high content screening. *Archives of Toxicology*, *80*, 580–604.

36. Houck, K.A., Dix, D.J., Judson, R.S., Kavlock, R.J., Yang, J., Berg, E.L. (2009). Profiling bioactivity of the ToxCast chemical library using BioMAP primary human cell systems. *Journal of Biomolecular Screening*, *14*, 1054–1066.
37. Berg, E.L., Yang, J., Melrose, J., Nguyen, D., Privat, S., Rosler, E., Kunkel, E.J., Ekins, S. (2010). Chemical target and pathway toxicity mechanisms defined in primary human cell systems. *Journal of Pharmacological and Toxicological Methods*, *61*, 3–15.
38. Hill, A.V. (1910). The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *Journal of Physiology*, *40*, 4–7.
39. Ihaka, R., Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, *5*, 299–314.
40. Bader, G.D., Cary, M.P., Sander, C. (2006). Pathguide: A pathway resource list. *Nucleic Acids Research*, *34*, D504–D506.
41. Cooper, R.L., Lamb, J.C., Barlow, S.M., Bentley, K., Brady, A.M., Doerrer, N.G., Eisenbrandt, D.L., Fenner-Crisp, P.A., Hines, R.N., Irvine, L.F., Kimmel, C.A., Koeter, H., Li, A.A., Makris, S.L., Sheets, L.P., Speijers, G., Whitby, K.E. (2006). A tiered approach to life stages testing for agricultural chemical safety assessment. *Critical Reviews in Toxicology*, *36*, 69–98.
42. Kroesa, R., Gallib, C., Munroc, I., Schilterd, B., Trane, L.-A., Walkerf, R., Würtzeng, G. (2000). Threshold of toxicological concern for chemical substances present in the diet: A practical tool for assessing the need for toxicity testing. *Food and Chemical Toxicology*, *38*, 255–312.
43. Klaunig, J.E., Babich, M.A., Baetcke, K.P., Cook, J.C., Corton, J.C., David, R.M., DeLuca, J.G., Lai, D.Y., McKee, R.H., Peters, J.M., Roberts, R.A., Fenner-Crisp, P.A. (2003). PPARalpha agonist-induced rodent tumors: Modes of action and human relevance. *Critical Reviews in Toxicology*, *33*, 655–780.
44. Kleinstreuer, N.C., Judson, R.S., Reif, D.M., Sipes, N.S., Singh, A.V., Chandler, K.J., Dewoskin, R., Dix, D.J., Kavlock, R.J., Knudsen, T.B. (2011). Environmental impact on vascular development predicted by high-throughput screening. *Environmental Health Perspectives*, *119*, 1596–1603.
45. Martin, M.T., Knudsen, T.B., Reif, D.M., Houck, K.A., Judson, R.S., Kavlock, R.J., Dix, D.J. (2011). Predictive model of rat reproductive toxicity from ToxCast high-throughput screening. *Biology of Reproduction*, *85*, 327–339.
46. Sipes, N.S., Martin, M.T., Reif, D.M., Kleinstreuer, N.C., Judson, R.S., Singh, A.V., Chandler, K.J., Dix, D.J., Kavlock, R.J., Knudsen, T.B. (2011). Predictive models of prenatal developmental toxicity from ToxCast high-throughput screening data. *Toxicological Sciences*, *124*, 109–127.
47. Peraza, M.A., Burdick, A.D., Marin, H.E., Gonzalez, F.J., Peters, J.M. (2006). The toxicology of ligands for peroxisome proliferator-activated receptors (PPAR). *Toxicological Sciences*, *90*, 269–295.
48. Goetz, A.K., Ren, H., Schmid, J.E., Blystone, C.R., Thillainadarajah, I., Best, D.S., Nichols, H.P., Strader, L.F., Wolf, D.C., Narotsky, M.G., Rockett, J.C., Dix, D.J. (2007). Disruption of testosterone homeostasis as a mode of action for the reproductive toxicity of triazole fungicides in the male rat. *Toxicological Sciences*, *95*, 227–239.

49. Abbott, B.D., Buckalew, A.R., DeVito, M.J., Ross, D., Bryant, P.L., Schmid, J.E. (2003). EGF and TGF-alpha expression influence the developmental toxicity of TCDD: Dose response and AhR phenotype in EGF, TGF-alpha, and EGF + TGF-alpha knockout mice. *Toxicological Sciences*, *71*, 84–95.
50. Tian, Y., Ke, S., Denison, M.S., Rabson, A.B., Gallo, M.A. (1999). Ah receptor and NF-kappa B interactions, a potential mechanism for dioxin toxicity. *Journal of Biological Chemistry*, *274*, 510–515.
51. Tian, Y.A. (2009). Ah receptor and NF-kappa B interplay on the stage of epigenome. *Biochemical Pharmacology*, *77*, 670–680.
52. Tian, Y.N., Rabson, A.B., Gallo, M. A. (2002). Ah receptor and NF-kappa B interactions: Mechanisms and physiological implications. *Chemico-Biological Interactions*, *141*, 97–115.
53. Gabbert, S., Weikard, H.P. (2010). A theory of chemicals testing and regulation. *Natural Resources Forum*, *34*, 155–164.
54. Reif, D.M., Martin, M.T., Tan, S.W., Houck, K.A., Judson, R.S., Richard, A.M., Knudsen, T.B., Dix, D.J., Kavlock, R.J. (2010). Endocrine profiling and prioritization of environmental chemicals using ToxCast data. *Environmental Health Perspectives*, *118*, 1714–1720.
55. Ankley, G.T., Bennett, R.S., Erickson, R.J., Hoff, D.J., Hornung, M.W., Johnson, R.D., Mount, D.R., Nichols, J.W., Russom, C.L., Schmieder, P.K., Serrrano, J.A., Tietge, J.E., Villeneuve, D.L. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry*, *29*, 730–741.
56. Michalik, L., Desvergne, B., Wahli, W. (2004). Peroxisome-proliferator-activated receptors and cancers: Complex stories. *Nature Reviews Cancer*, *4*, 61–70.
57. Shah, I., Wambaugh, J. (2010). Virtual tissues in toxicology. *Journal of Toxicology and Environmental Health Part B: Critical Reviews*, *13*, 314–328.
58. Wambaugh, J., Imran, S. (2010). Simulating microdosimetry in a virtual hepatic lobule. *PLoS Computational Biology*, *6*(4), e1000756.