TOXICOLOGICAL SCIENCES **127(1)**, 1–9 (2012)
doi:10.1093/toxsci/kfs095
Advance Access publication March 2, 2012

Predictive Modeling of Chemical Hazard by Integrating Numerical Descriptors of Chemical Structures and Short-term Toxicity Assay Data

Ivan Rusyn,*[,1] Alexander Sedykh,† Yen Low,*[,†] Kathryn Z. Guyton,‡ and Alexander Tropsha†

*Department of Environmental Sciences and Engineering and †Division of Chemical Biology and Medicinal Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599; and ‡National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency, Washington, District of Columbia 20460

[1]To whom correspondence should be addressed. Fax: (919) 843-2596. E-mail: iir@unc.edu.

Received January 16, 2012; accepted February 20, 2012

Quantitative structure-activity relationship (QSAR) models are widely used for *in silico* prediction of *in vivo* toxicity of drug candidates or environmental chemicals, adding value to candidate selection in drug development or in a search for less hazardous and more sustainable alternatives for chemicals in commerce. The development of traditional QSAR models is enabled by numerical descriptors representing the inherent chemical properties that can be easily defined for any number of molecules; however, traditional QSAR models often have limited predictive power due to the lack of data and complexity of *in vivo* endpoints. Although it has been indeed difficult to obtain experimentally derived toxicity data on a large number of chemicals in the past, the results of quantitative *in vitro* screening of thousands of environmental chemicals in hundreds of experimental systems are now available and continue to accumulate. In addition, publicly accessible toxicogenomics data collected on hundreds of chemicals provide another dimension of molecular information that is potentially useful for predictive toxicity modeling. These new characteristics of molecular bioactivity arising from short-term biological assays, i.e., *in vitro* screening and/or *in vivo* toxicogenomics data can now be exploited in combination with chemical structural information to generate hybrid QSAR–like quantitative models to predict human toxicity and carcinogenicity. Using several case studies, we illustrate the benefits of a hybrid modeling approach, namely improvements in the accuracy of models, enhanced interpretation of the most predictive features, and expanded applicability domain for wider chemical space coverage.

Key Words: **QSAR; toxicity screening; hybrid modeling.**

Computational toxicology is a rapidly growing field that combines methodologies from computer science, bio and cheminformatics, chemistry and molecular biology (reviewed by Kavlock *et al.*, 2008; Nigsch *et al.*, 2009; Rusyn and Daston, 2010). Due to advances in biological screening technologies, multiple streams of novel toxicological data, ranging from short-term *in vitro* assays to various *in vivo* endpoints, are available for hundreds of chemicals (Martin *et al.*, 2009; Shukla *et al.*, 2010). The Tox21 consortium of the U.S. Environmental Protection Agency (EPA), National Toxicology Program (NTP), National Institutes of Health Chemical Genomics Center (NCGC), and U.S. Food and Drug Administration (FDA) is generating extensive quantitative *in vitro* data by screening hundreds to thousands of environmental chemicals in hundreds of experimental systems with the goal of re-establishing the field of predictive chemical toxicology under the paradigm of *in vitro-in vivo* extrapolation (Collins *et al.*, 2008). Many chemicals have been screened for toxicity phenotypes in cells from multiple individuals (Choy *et al.*, 2008; Lock *et al.*, 2012; O'Shea *et al.*, 2011). In addition, toxicogenomics data collected for hundreds of chemicals provide another dimension of experimental knowledge that is potentially useful for predictive chemical toxicity modeling (Fielden *et al.*, 2007; Uehara *et al.*, 2011). Innovative frameworks are required to integrate these rich and diverse new data for systematic investigation of the determinants of endpoint toxicity, including underlying chemical, biological, and genetic factors.

The explosive accumulation of biomolecular screening data that may help explain and predict toxicity mechanisms has led to the development of novel computational tools and databases (Barros and Martin, 2008; Blomme *et al.*, 2009; Fielden *et al.*, 2007; Waters and Fostel, 2004). The ultimate goal of computational modeling is fast and accurate estimation of environmental hazards and human health risks with minimal to no dependence on animal testing (National Research Council, 2007).

Cheminformatics approaches, such as quantitative structure-activity relationship (QSAR) modeling, have been traditionally used to rationalize biological screening data and employ resulting models, or predictors, as an initial virtual screen for efficacy and/or safety of candidate chemicals. The availability of predictive multidimensional *in vitro* and/or *in vivo* molecular data on a particular compound greatly facilitates decision making regarding its potential health hazard and

Disclaimer: The views expressed are those of the authors and do not necessarily reflect policy or endorsement of the U.S. Environmental Protection Agency or other funding agencies.

mechanisms thereof (Roth *et al.*, 2011). However, new regulations in Europe and initiatives in the United States (National Research Council, 2007) are applying pressure on the scientific and risk assessment communities to develop improved methods for evaluating thousands of chemicals (Rusyn and Daston, 2010; Schwarzman and Wilson, 2009).

Adverse outcomes *in vivo* depend both on the chemical's structure and the underlying toxicity mechanisms. In tune with the proliferation of transdisciplinary computational biology approaches to unravel chemical toxicity mechanisms, this review highlights several novel integrative strategies for prediction of *in vivo* chemical toxicity by concordant exploitation of both a chemical's structure and its short-term biological effects. Several recent studies demonstrate that statistically significant and externally predictive hybrid models can be developed. Hybrid modeling also affords a possibility of mechanistic interpretation both in terms of underlying chemical features and mechanisms of toxicity. Herein, we describe a general computational framework for modeling chemical toxicity using cheminformatics approaches, summarize recent hybrid modeling methodologies for *in vitro-in vivo* extrapolation paradigms, and comment on the outlook for the future use of these tools in computational toxicology.

CHEMINFORMATICS-BASED PREDICTORS IN TOXICOLOGY

Chemical structure–based predictors generally fall into two types: QSAR and expert systems (Valerio, 2009). QSAR are statistical models linking molecular structures (represented by chemical descriptors) to an activity such as an adverse health outcome (e.g., toxicity). QSAR embodies the principle of similarity, assuming that structurally similar chemicals may also have closely aligned activities. For example, chemicals with ≥ 0.85 similarity (based on the Tanimoto coefficient) to known actives were 30 times more likely to be confirmed as active than those picked randomly (Martin *et al.*, 2002). Expert systems, on the other hand, are models based on rules determined by the scientific consensus of the experts. For instance, the Ashby-Tennant structural alerts for carcinogenicity (Ashby and Tennant, 1994) have been incorporated into many software tools (Marchant *et al.*, 2008). There are a number of public and commercial stand-alone or web-based modeling systems that have been developed for prediction of a large number of toxicity-relevant endpoints (Tables 1 and 2). Several recent publications provide an excellent overview of the computational tools employed in toxicology (Nigsch *et al.*, 2009; Valerio, 2009).

Although QSAR modeling techniques are under continuous development, most predictors are not considered to be accurate enough for estimating complex biological phenotypes (Rusyn and Daston, 2010). Low quality of data, overextrapolation, and poor definition of the phenotypes to be predicted have been identified as factors limiting the accuracy of prediction of absorption, distribution, metabolism, excretion, and toxicity endpoints by QSAR (Penzotti *et al.*, 2004; Stouch *et al.*, 2003). In addition, the inherent limitations of QSAR lie in the general complexity of factors that impact the ultimate adverse health effect of a chemical, including pharmacokinetics, temporality, or the fact that multiple mechanisms and interconnected molecular signaling pathways may lead to the same toxicity phenotype. Thus, it is not surprising that the performance of QSAR models is inversely correlated to the complexity of the modeled endpoints (Hou and Wang, 2008; Penzotti *et al.*, 2004), higher accuracy being expected for predicting *in vitro* results, and lower accuracy observed for more complex *in vivo* endpoints, such as carcinogenicity (Benigni and Bossa, 2008). Given these limitations, it is unlikely that significant gains in prediction accuracy would be achieved by implementing alternative machine learning techniques or developing new chemical descriptors.

EXPLORING OMICS AND *IN VITRO* DATA FOR PREDICTIVE TOXICITY MODELING

Alternative methods have been proposed to improve predictive accuracy and take into account novel data streams that may help in overcoming some of the inherent limitations detailed above. Indeed, mechanistic toxicology research has taken advantage of technology developments in biomedical sciences. Toxicogenomics, proteomics, and metabolomics provide experimental approaches for viewing the complete biological system that is modulated by a chemical (Ekins *et al.*, 2005). These complex multidimensional data are now routinely used in drug and chemical safety evaluation, providing valuable mechanistic understanding of the molecular changes associated with the disease or treatment (Cui and Paules, 2010). The utility of these data in predictive toxicology has also been explored. A number of studies reported on the development of models that use omics data (most of these used transcriptional profiling) to predict chronic toxicity phenotypes (e.g., carcinogenic potential) with acute or subchronic study–derived information (Fielden *et al.*, 2007; Uehara *et al.*, 2011) or to classify chemicals with respect to their potential mode of toxicity (Fielden *et al.*, 2011; Uehara *et al.*, 2010; Waters *et al.*, 2010).

Recent advances in automated quantitative high-throughput screening (qHTS) have generated extensive biological data that can be modeled using statistical or machine learning techniques (Shukla *et al.*, 2010). The Tox21 program (Collins *et al.*, 2008), a partnership between EPA, NTP, NCGC, and FDA, is leading the field in use of a broad spectrum of *in vitro* assays, many in qHTS format, to screen thousands of environmental chemicals for their potential to disturb biological pathways that may result in human disease (Xia *et al.*, 2008). Such data on toxicologically relevant *in vitro* endpoints can be utilized as hazard-based triggers to inform

TABLE 1
Examples of Commercial Toxicity Predictors

Prediction tool	Categories of endpoints[a]	Features
ADMET Predictor www.simulations-plus.com	Irritation and adverse health effects Carcinogenicity and genotoxicity Acute and developmental toxicity Endocrine disruption and ecotoxicity	QSAR
ACD/Tox Suite www.acdlabs.com	Irritation and adverse health effects Genotoxicity Acute toxicity Endocrine disruption and ecotoxicity	Confidence intervals and probability of predictions
DEREK, DEREK Nexus www.lhasalimited.org	Irritation and adverse health effects Carcinogenicity and genotoxicity Developmental toxicity	Expert system
TOPKAT www.accelrys.com	Irritation Carcinogenicity and genotoxicity Acute, chronic, and developmental toxicity Ecotoxicity	QSAR
CASE www.multicase.com	Irritation and adverse health effects Carcinogenicity and genotoxicity Acute and developmental toxicity Endocrine disruption and ecotoxicity	Fragment-based QSAR
Leadscope Model Applier www.leadscope.com	Adverse health effects Carcinogenicity Reproductive and developmental toxicity	QSAR
HazardExpertPro, ToxAlert www.compudrug.com	Adverse health effects Carcinogenicity and genotoxicity Developmental toxicity	Expert system

[a]Irritation—skin, eye, or lung sensitization, allergies; adverse health effects—organ-specific toxicity; ecotoxicity—aquatic toxicity and related environmental endpoints.

prioritization for additional testing (Reif *et al.*, 2010), to predict *in vivo* toxicity (Martin *et al.*, 2010), or to generate testable hypotheses concerning the underlying mechanisms of toxicity (Xia *et al.*, 2009).

Statistical models employing biological data such as gene signatures or qHTS data as independent variables are in principle similar to QSAR models because both employ similar computational tools and focus on predicting similar toxicity phenotypes (Table 3). Importantly, the biological data–based models have been shown to be both predictive and interpretable (Coen, 2010; Van Hummelen and Sasaki, 2010; Wetmore and Merrick, 2004). Still, pure biological data–based predictive modeling approaches are not intended to explain chemical-induced factors but focus on the general biological processes related to toxicity. Furthermore, such models are inherently insensitive to explicitly defined chemical features of the tested compounds, and new biological data must be generated in order to predict the toxicity of novel compounds. In the case of biology-based approaches, additional factors such as experimental variability, interpretability, and data acquisition costs also need to be considered.

HYBRID MODELING APPROACHES

To properly realize the joint benefits of bioinformatics- and cheminformatics-based approaches, several strategies can be envisioned (Fig. 1). The simplest approach is to utilize a ''consensus'' of QSAR and biological models that were derived independently to predict the same endpoint (Fig. 1A). Consensus modeling is an approach to developing an overall prediction by combining multiple classifiers, and it is widely used in traditional QSAR (reviewed in Dearden, 2003; Tong *et al.*, 2006). Proponents of the consensus approach expound that combining multiple models that otherwise individually encode for different relationships would result in a more robust prediction (Tong *et al.*, 2006). On the other hand, opponents question if the marginal predictivity gains are worth the added complexity of consensus modeling (Hewitt *et al.*, 2007). Success of consensus prediction depends on the relative performance, applicability domain, and the number of included individual models (Penzotti *et al.*, 2004). Although there are no published examples of consensus between QSAR and biological data–based models, this approach is likely to yield models of predictive performance in between that of contributing models

TABLE 2
Examples of Toxicity Predictors in Public Domain

Prediction tool	Categories of endpoints	Features
T.E.S.T. (EPA) www.epa.gov/nrmrl/std/cppb/qsar	Carcinogenicity and genotoxicity Acute and developmental toxicity Ecotoxicity	Consensus and batch prediction modes by QSAR
OncoLogic (EPA) http://www.epa.gov/oppt/sf/pubs/oncologic.htm	Carcinogenicity	Expert system
OpenTox www.opentox.org	Irritation Carcinogenicity and genotoxicity	Expert system (ToxTree); QSAR (Lazar); ontology of toxic endpoints
OECD QSAR Toolbox www.qsartoolbox.org	Irritation Carcinogenicity and genotoxicity Ecotoxicity	Prediction by ''read across'' analysis or by QSAR
OCHEM www.ochem.eu	Genotoxicity Ecotoxicity	Online chemical database and QSAR modeling environment
ChemBench chembench.mml.unc.edu	Genotoxicity Ecotoxicity	Web-based platform for QSAR modeling or prediction

as a consequence of statistical averaging. For example, in the simplest instance, predictions from a QSAR model and a biological model would be averaged into a final consensus score. Further improvements in consensus prediction may lie in adjusting relative contributions of the individual models.

There are several examples of how the modeling routine may use a ''hierarchical'' approach (Fig. 1B). First, it was suggested by several groups that a hierarchy of chemical descriptors of increased complexity may be used to improve a model's accuracy. For instance, Basak *et al.* (2003) developed models of cytotoxicity of halocarbons by utilizing a hierarchy of different types of computed descriptors of inherent chemical properties. In this method, model building begins with descriptors, which can be computed most easily, and additional descriptors that may demand more computational resources are added only if the easily calculable ones do not give satisfactory results. A similar approach was incorporated into hierarchical QSAR (HiT QSAR) software (Kuz'min *et al.*, 2008). Both studies showed that the complexity of chemical descriptors has an impact on the accuracy of model predictions.

Second, a hierarchy of computational methods was used, whereby compounds are classified into subgroups with different levels of response using liner discriminant analysis followed by recursive partitioning for each subgroup (Manga *et al.*, 2003). This study developed a model of drug biotransformation using physicochemical and structural descriptors to predict the percent of unmetabolized drug excreted after iv dose. The resultant hierarchical model for biotransformation was a three-level decision tree that incorporated various classification techniques and a series of arbitrary cutoffs.

Third, a hierarchical workflow was proposed to explore chemical structure/*in vitro*/*in vivo* relationships (Zhu *et al.*, 2009). Under this approach, *in vitro*/*in vivo* correlation patterns for all compounds in the modeling set could be ascertained, and compounds may be clustered into several subsets (e.g., toxic both *in vitro* and *in vivo*; nontoxic in both cases; toxic *in vitro* but nontoxic *in vivo*) based on the discovered relationships. The modeling set compounds were partitioned into two or more subclasses, and a classification QSAR model was developed using chemical descriptors only. Then, subclass-specific QSAR models were developed. Thus, for any external compound, the classification model is used first to make assignment to one of the subclasses, and then a subclass-specific model is used to make a quantitative prediction of a compound's toxicity.

An alternative strategy is a "hybrid" approach (Fig. 1C), in which biology-derived features and chemical structural properties are pooled into a joint descriptor matrix, which is then used for modeling. Although, in principle, such joint descriptors may have limitations (i.e., data quality, cost of data acquisition, etc.), recent studies suggest that hybrid descriptors do afford improvement to the accuracy of prediction of *in vivo* toxicity. Several recent publications (Low *et al.*, 2011; Sedykh *et al.*, 2011; Zhu *et al.*, 2008) provide illustrative examples of hybrid modeling.

For example, Zhu *et al.* (2008) have introduced a concept of chemical-biological descriptors where conventional chemical descriptors are augmented by binary qHTS results ("active" response is encoded as "1," "inactive" as "0") from a variety of assays to create a single combined array of hybrid descriptors. Using chemical descriptors only, QSAR modeling resulted in 62.3% prediction accuracy for rodent carcinogenicity applied to the data set of over 300 chemicals for which rodent 2-year cancer bioassay data were available. Importantly, the prediction accuracy of the model was significantly improved (to 72.7%) when chemical descriptors were augmented by qHTS cytotoxicity data on six rodent and human cell lines, which were regarded as biological descriptors.

Sedykh *et al.* (2011) have employed concentration-response qHTS data reported by Xia *et al.* (2008) by transforming them

TABLE 3
Examples of Toxicity Data–Based Predictive Models

Predicted endpoints	Input variables	Publication
Reproductive toxicity	500 *in vitro* assays on 256 compounds from ToxCast Phase I	Martin *et al.*, 2011
Hepatotoxicity	Liver gene expression from rats, rat hepatocytes, and human hepatocytes treated with two compounds for 24 h	Roth *et al.*, 2011
Hepatobiliary injury	Blood gene expression and urine metabolomics of rats treated with 16 compounds for 1, 3, 14 days	Ellinger-Ziegelbauer *et al.*, 2011
Hepatotoxicity	Liver and blood gene expression of rats treated with eight compounds at three doses over 6, 24, and 48 h	Huang *et al.*, 2010
Acetaminophen-, phenobarbital- and methapyrilene-induced hepatotoxicity	Gene expression, proteomics, and metabolomics of mice treated for short period of time	Coen *et al.*, 2004; Craig *et al.*, 2006; Waterman *et al.*, 2010
Hepatotoxicity mechanisms	Liver gene expression of rats treated with 150 compounds at three doses from 1 to 28 days	Open TG-GATEs database (Uehara *et al.*, 2008)
Hepatotoxicity	Liver gene expression of rats (rat hepatocytes) treated with 111 (86) compounds for 24 h	ToxExpress database (Barros and Martin, 2008)
Hepatotoxicity, hepatotumorigenesis, nephrotoxicity, etc.	Liver and kidney gene expression of rats treated with 344 compounds for 1–7 days	Iconix database (Blomme *et al.*, 2009; Fielden *et al.*, 2007; Wang *et al.*, 2008)
Hepatotoxicity	Liver gene expression of rats treated with 15 compounds for 6, 24, and 72 h	Zidek *et al.*, 2007
Nongenotoxic carcinogenicity	Liver gene expression of rats treated with 52 compounds for 24 h	Nie *et al.*, 2006
Hepatocarcinogenicity	Liver gene expression of rats treated with seven congeneric compounds for 1–28 days	Nakayama *et al.*, 2006
Valproic acid–induced hepatotoxicity	Liver gene expression and proteomics and urine metabolomics of rats treated for 6, 12, and 24 h	Schnackenberg *et al.*, 2006
Toxicity related to pancreas, liver, kidney, testes, and bladder	Blood and urine metabolomics of rats treated with 147 compounds for 1–8 days	COMET database (Lindon *et al.*, 2005)
Bromobenzene-induced hepatotoxicity	Gene expression and proteomics of rats treated for 6 and 24 h	Heijne *et al.*, 2004
Mechanisms of hepatotoxicity	Gene expression of rat primary hepatocytes treated with 15 compounds for 24 h	Waring *et al.*, 2001

into quantitative biological descriptors of chemicals. The *in vitro* data, especially concentration-response qHTS profiles, were shown to improve the results of QSAR modeling of *in vivo* end points (i.e., rat LD_{50}) as compared with conventional QSAR models that used only chemical structure descriptors. Furthermore, the biological qHTS descriptors data also enhanced the model's coverage (i.e., the number of compounds within the applicability domain of the model), which is essential for applying models to large and diverse chemical libraries of environmental concern.

Toxicogenomic data provide another example of high-dimensional biological information that may be used for hybrid modeling. A comparative analysis of QSAR- and toxicogenomics data–based models was recently reported (Liu *et al.*, 2011). The authors used gene expression profiles of liver tissue obtained from rats treated with 62 chemicals at different time points (1, 3, and 5 days) to predict rat liver carcinogenicity and concluded that the toxicogenomics data–based models outperformed QSAR. Low *et al.* (2011) reported a similar outcome when gene expression data–based models (24-h rat liver toxicogenomics profiles of 127 compounds) were compared with conventional QSAR in modeling 28-day hepatotoxicity in the rat. However, the latter study also attempted to combine toxicogenomics data and chemical descriptors for a hybrid approach. Although hybrid models did not afford prediction accuracy higher than that of toxicogenomics data–based models, they identified both chemical features and transcripts predictive of the phenotype, which provided additional insight regarding the mechanistic basis of subchronic liver injury.

CONCLUSIONS AND FUTURE DIRECTIONS

Accurate and high-throughput predictive methods are needed to support efficient decision making regarding the efficacy and/or safety of candidate compounds and in tiered screening and assessment schemes. Chemical structure–based

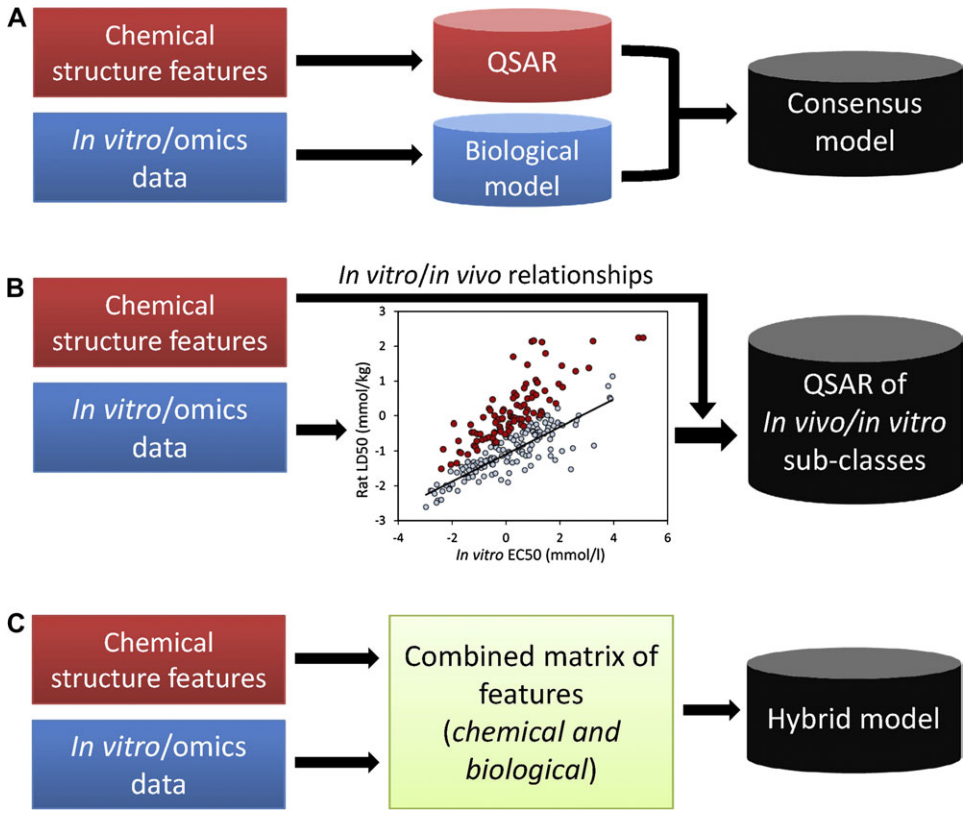

FIG. 1. Strategies for utilizing biological and chemical data in predictive modeling of *in vivo* toxicity.

predictive methods have been widely applied in the screening and ranking of thousands of chemicals for bioactivity and have demonstrated the promise of *in silico* approaches for achieving these goals. However, predictive methods based on chemical structure alone have limitations, especially for accurately projecting complex *in vivo* outcomes. Integration of chemical features and biological screening and/or toxicogenomic data provides important advantages (i.e., improved prediction accuracy, greater chemical space coverage, and interpretability of predictive features) over traditional cheminformatic methods such as QSAR modeling. As shown in Figure 1, novel strategies for integrating chemical structural information with bioactivity data include consensus (Tong *et al.*, 2006), hybrid (Sedykh *et al.*, 2011; Zhu *et al.*, 2008), and hierarchical approaches (Zhu *et al.*, 2009).

Data limitations are currently the major obstacle to advancing these transdisciplinary integration approaches. In particular, the database of toxicity studies is limited to a small number of chemicals. These chemicals are both too few in number and too limited in structural diversity for reliable QSAR analysis. At present, there are only several sufficiently large omics data sets (e.g., Open Toxicogenomics Project Genomics Assisted Toxicity Evaluation System [http://toxico.nibio.go.jp/], Chemical Effects in Biological Systems Database [http://www.niehs.nih.gov/research/resources/databases/cebs/], ToxExpress [http://www.genelogic.com/knowledge-suites/toxexpress-program]) with hundreds of compounds of largely disparate chemotypes selected for phenotypic diversity. As such, most omics data sets are poorly suited for machine learning by QSAR. This deficit supports the more general and recognized need for hazard characterization of a greater number of more varied chemicals, including a larger proportion of the tens of thousands of yet untested chemicals in commerce and the environment. Other outstanding data needs concern the classification of chemicals according to a wider array of hazard traits and susceptibility factors (Guyton *et al.*, 2009). There are ongoing efforts to address these significant data limitations by characterizing multiple *in vitro* and *in vivo* toxicological phenotypes (Martin *et al.*, 2009; Padilla *et al.*, 2012; Shukla *et al.*, 2010), including in cells from genetically diverse individuals (Choy *et al.*, 2008; Lock *et al.*, 2012; O'Shea *et al.*, 2011). The large-scale screening efforts of Tox21 (Huang *et al.*, 2011) and other public-private partnerships (Cavero, 2011) hold particular promise for vastly expanding the database of chemicals and endpoints for which experimental data are available.

Additional types of hybrid/hierarchical modeling approaches can be envisioned to address the dependency of hybrid approaches on the availability of experimental data, a current limitation for the wide use of these models in predictive

toxicology. In principle, QSAR models could be developed to predict the results of short-term toxicity assays (once enough data for a sufficiently large chemical library is available) because this task is inherently less challenging than modeling complex *in vivo* endpoints. This represents an intriguing possibility of applying QSAR methods to build predictive models of each of many individual molecular endpoints from which the resulting "predicted *in vitro*" data can then be used as inputs into models of *in vivo* toxicity (Martin *et al*., 2011; Sipes *et al*., 2011). The application of this strategy could potentially enable a predictive modeling workflow that does not require new experimental data and employs only compound descriptors that can be computed from chemical structure.

Future computational methods should aim to optimize the use of both chemical- and biological-based data domains to achieve the most accurate predictions possible, because each one individually provides limited and complementary insights regarding toxicity. To this end, studies can be designed with both approaches in mind, so as to provide sufficient diversity from both chemical and biological data domains. The goal should be to generate data matrices with broad and dense coverage of chemical structure and bioactivities for hybrid data analysis, i.e., combining chemical and biological data for machine learning. Additional improvements can be achieved by using mechanistically relevant short-term toxicity assays. The resulting integrative approaches have the potential to become a powerful tool for elucidating both relevant biological interactions and structural motifs that together better represent the underlying complex mechanisms by which toxic effects of chemicals develop. Systematic investigation of genetic and other determinants of chemical toxicity can also be envisioned. These approaches can, in turn, support applications in the design of new products and chemical processes as well as in the evaluation of in-use chemicals and environmental contaminants, based on comprehensive and integrative characterization by both chemical structural features and the results of multiple and diverse short-term biological assays and/or omics studies.

FUNDING

This work was supported, in part, by grants from the National Institutes of Health (NIH) (R01 ES015241) and U.S. Environmental Protection Agency (EPA) (RD83382501).

ACKNOWLEDGMENTS

The research described in this article has not been subjected to each funding agency's peer review and policy review and therefore does not necessarily reflect their views and no official endorsement should be inferred. The authors declare no competing financial interests.

REFERENCES

Ashby, J., and Tennant, R. W. (1994). Prediction of rodent carcinogenicity for 44 chemicals: Results. *Mutagenesis* **9,** 7–15.

Barros, S. A., and Martin, R. B. (2008). Predictive toxicogenomics in preclinical discovery. *Methods Mol. Biol.* **460,** 89–112.

Basak, S. C., Balasubramanian, K., Gute, B. D., Mills, D., Gorczynska, A., and Roszak, S. (2003). Prediction of cellular toxicity of halocarbons from computed chemodescriptors: A hierarchical QSAR approach. *J. Chem. Inf. Comput. Sci.* **43,** 1103–1109.

Benigni, R., and Bossa, C. (2008). Predictivity and reliability of QSAR models: The case of mutagens and carcinogens. *Toxicol. Mech. Methods* **18,** 137–147.

Blomme, E. A., Yang, Y., and Waring, J. F. (2009). Use of toxicogenomics to understand mechanisms of drug-induced hepatotoxicity during drug discovery and development. *Toxicol. Lett.* **186,** 22–31.

Cavero, I. (2011). Promises and partnership: FDA's Critical Path Initiative and its intersection with pharmacology: An ASPET 2011 annual meeting symposium. *Expert Opin. Drug Saf.* **10,** 827–837.

Choy, E., Yelensky, R., Bonakdar, S., Plenge, R. M., Saxena, R., De Jager, P. L., Shaw, S. Y., Wolfish, C. S., Slavik, J. M., *et al*. (2008). Genetic analysis of human traits in vitro: Drug response and gene expression in lymphoblastoid cell lines. *PLoS Genet.* **4,** e1000287.

Coen, M. (2010). A metabonomic approach for mechanistic exploration of pre-clinical toxicology. *Toxicology* **278,** 326–340.

Coen, M., Ruepp, S. U., Lindon, J. C., Nicholson, J. K., Pognan, F., Lenz, E. M., and Wilson, I. D. (2004). Integrated application of transcriptomics and metabonomics yields new insight into the toxicity due to paracetamol in the mouse. *J. Pharm. Biomed. Anal.* **35,** 93–105.

Collins, F. S., Gray, G. M., and Bucher, J. R. (2008). Toxicology. Transforming environmental health protection. *Science* **319,** 906–907.

Craig, A., Sidaway, J., Holmes, E., Orton, T., Jackson, D., Rowlinson, R., Nickson, J., Tonge, R., Wilson, I., and Nicholson, J. (2006). Systems toxicology: Integrated genomic, proteomic and metabonomic analysis of methapyrilene-induced hepatotoxicity in the rat. *J. Proteome Res.* **5,** 1586–1601.

Cui, Y., and Paules, R. S. (2010). Use of transcriptomics in understanding mechanisms of drug-induced toxicity. *Pharmacogenomics* **11,** 573–585.

Dearden, J. C. (2003). In silico prediction of drug toxicity. *J. Comput. Aided Mol. Des.* **17,** 119–127.

Ekins, S., Nikolsky, Y., and Nikolskaya, T. (2005). Techniques: Application of systems biology to absorption, distribution, metabolism, excretion and toxicity. *Trends Pharmacol. Sci.* **26,** 202–209.

Ellinger-Ziegelbauer, H., Adler, M., Amberg, A., Brandenburg, A., Callanan, J. J., Connor, S., Fountoulakis, M., Gmuender, H., Gruhler, A., Hewitt, P., *et al*. (2011). The enhanced value of combining conventional and "omics" analyses in early assessment of drug-induced hepatobiliary injury. *Toxicol. Appl. Pharmacol.* **252,** 97–111.

Fielden, M. R., Adai, A., Dunn, R. T., Olaharski, A., Searfoss, G., Sina, J., Aubrecht, J., Boitier, E., Nioi, P., Auerbach, S., *et al*. (2011). Development and evaluation of a genomic signature for the prediction and mechanistic assessment of nongenotoxic hepatocarcinogens in the rat. *Toxicol. Sci.* **124,** 54–74.

Fielden, M. R., Brennan, R., and Gollub, J. (2007). A gene expression biomarker provides early prediction and mechanistic assessment of hepatic tumor induction by nongenotoxic chemicals. *Toxicol. Sci.* **99,** 90–100.

Guyton, K. Z., Kyle, A. D., Aubrecht, J., Cogliano, V. J., Eastmond, D. A., Jackson, M., Keshava, N., Sandy, M. S., Sonowane, B., Zhang, L., *et al*. (2009). Improving prediction of chemical carcinogenicity by considering multiple mechanisms and applying toxicogenomic approaches. *Mutat. Res.* **681,** 230–240.

Heijne, W. H., Slitt, A. L., van Bladeren, P. J., Groten, J. P., Klaassen, C. D., Stierum, R. H., and Van Ommen, B. (2004). Bromobenzene-induced hepatotoxicity at the transcriptome level. *Toxicol. Sci.* **79,** 411–422.

Hewitt, M., Cronin, M. T., Madden, J. C., Rowe, P. H., Johnson, C., Obi, A., and Enoch, S. J. (2007). Consensus QSAR models: Do the benefits outweigh the complexity? *J. Chem. Inf. Model.* **47,** 1460–1468.

Hou, T., and Wang, J. (2008). Structure-ADME relationship: Still a long way to go? *Expert Opin. Drug Metab. Toxicol.* **4,** 759–770.

Huang, J., Shi, W., Zhang, J., Chou, J. W., Paules, R. S., Gerrish, K., Li, J., Luo, J., Wolfinger, R. D., Bao, W., *et al.* (2010). Genomic indicators in the blood predict drug-induced liver injury. *Pharmacogenomics J.* **10,** 267–277.

Huang, R., Xia, M., Cho, M. H., Sakamuru, S., Shinn, P., Houck, K. A., Dix, D. J., Judson, R. S., Witt, K. L., Kavlock, R. J., *et al.* (2011). Chemical genomics profiling of environmental chemical modulation of human nuclear receptors. *Environ. Health Perspect.* **119,** 1142–1148.

Kavlock, R. J., Ankley, G., Blancato, J., Breen, M., Conolly, R., Dix, D., Houck, K., Hubal, E., Judson, R., Rabinowitz, J., *et al.* (2008). Computational toxicology—A state of the science mini review. *Toxicol. Sci.* **103,** 14–27.

Kuz'min, V. E., Artemenko, A. G., and Muratov, E. N. (2008). Hierarchical QSAR technology based on the Simplex representation of molecular structure. *J. Comput. Aided Mol. Des.* **22,** 403–421.

Lindon, J. C., Keun, H. C., Ebbels, T. M., Pearce, J. M., Holmes, E., and Nicholson, J. K. (2005). The Consortium for Metabonomic Toxicology (COMET): Aims, activities and achievements. *Pharmacogenomics* **6,** 691–699.

Liu, Z., Kelly, R., Fang, H., Ding, D., and Tong, W. (2011). Comparative analysis of predictive models for nongenotoxic hepatocarcinogenicity using both toxicogenomics and quantitative structure-activity relationships. *Chem. Res. Toxicol.* **24,** 1062–1070.

Lock, E. F., Abdo, N., Huang, R., Xia, M., O'Shea, S. H., Zhou, Y. H., Sedykh, A., Tropsha, A., Austin, C. P., Tice, R. R., *et al.* (2012). Quantitative high-throughput screening for chemical toxicity in a population-based in vitro model. *Toxicol. Sci.* **126,** 578–588.

Low, Y., Uehara, T., Minowa, Y., Yamada, H., Ohno, Y., Urushidani, T., Sedykh, A., Muratov, E. N., Kuz'min, V., Fourches, D., *et al.* (2011). Predicting drug-induced hepatotoxicity using QSAR and toxicogenomics approaches. *Chem. Res. Toxicol.* **24,** 1251–1262.

Manga, N., Duffy, J. C., Rowe, P. H., and Cronin, M. T. (2003). A hierarchical QSAR model for urinary excretion of drugs in humans as a predictive tool for biotransformation. *QSAR Comb. Sci.* **22,** 263–273.

Marchant, C. A., Briggs, K. A., and Long, A. (2008). In silico tools for sharing data and knowledge on toxicity and metabolism: Derek for windows, meteor, and vitic. *Toxicol. Mech. Methods* **18,** 177–187.

Martin, M. T., Dix, D. J., Judson, R. S., Kavlock, R. J., Reif, D. M., Richard, A. M., Rotroff, D. M., Romanov, S., Medvedev, A., Poltoratskaya, N., *et al.* (2010). Impact of environmental chemicals on key transcription regulators and correlation to toxicity end points within EPA's ToxCast program. *Chem. Res. Toxicol.* **23,** 578–590.

Martin, M. T., Judson, R. S., Reif, D. M., Kavlock, R. J., and Dix, D. J. (2009). Profiling chemicals based on chronic toxicity results from the U.S. EPA ToxRef Database. *Environ. Health Perspect.* **117,** 392–399.

Martin, M. T., Knudsen, T. B., Reif, D. M., Houck, K. A., Judson, R. S., Kavlock, R. J., and Dix, D. J. (2011). Predictive model of rat reproductive toxicity from ToxCast high throughput screening. *Biol. Reprod.* **85,** 327–339.

Martin, Y. C., Kofton, J. L., and Traphagen, L. M. (2002). Do structurally similar molecules have similar biological activity? *J. Med. Chem.* **45,** 4350–4358.

Nakayama, K., Kawano, Y., Kawakami, Y., Moriwaki, N., Sekijima, M., Otsuka, M., Yakabe, Y., Miyaura, H., Saito, K., Sumida, K., *et al.* (2006). Differences in gene expression profiles in the liver between carcinogenic and non-carcinogenic isomers of compounds given to rats in a 28-day repeat-dose toxicity study. *Toxicol. Appl. Pharmacol.* **217,** 299–307.

National Research Council (2007). *Toxicity Testing in the 21st Century: A Vision and a Strategy*. National Academies Press, Washington, DC.

Nie, A. Y., McMillian, M., Parker, J. B., Leone, A., Bryant, S., Yieh, L., Bittner, A., Nelson, J., Carmen, A., Wan, J., *et al.* (2006). Predictive toxicogenomics approaches reveal underlying molecular mechanisms of nongenotoxic carcinogenicity. *Mol. Carcinog.* **45,** 914–933.

Nigsch, F., Macaluso, N. J., Mitchell, J. B., and Zmuidinavicius, D. (2009). Computational toxicology: An overview of the sources of data and of modelling methods. *Expert Opin. Drug Metab. Toxicol.* **5,** 1–14.

O'Shea, S. H., Schwarz, J., Kosyk, O., Ross, P. K., Ha, M. J., Wright, F. A., and Rusyn, I. (2011). In vitro screening for population variability in chemical toxicity. *Toxicol. Sci.* **119,** 398–407.

Padilla, S., Corum, D., Padnos, B., Hunter, D. L., Beam, A., Houck, K. A., Sipes, N., Kleinstreuer, N., Knudson, T., Dix, D. J., *et al.* (2012). Zebrafish developmental screening of the ToxCast™ Phase I chemical library. *Reprod. Toxicol.* **33,** 174–187.

Penzotti, J. E., Landrum, G. A., and Putta, S. (2004). Building predictive ADMET models for early decisions in drug discovery. *Curr. Opin. Drug Discov. Devel.* **7,** 49–61.

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

Roth, A., Boess, F., Landes, C., Steiner, G., Freichel, C., Plancher, J. M., Raab, S., de Vera Mudry, C., Weiser, T., and Suter, L. (2011). Gene expression-based in vivo and in vitro prediction of liver toxicity allows compound selection at an early stage of drug development. *J. Biochem. Mol. Toxicol.* **25,** 183–194.

Rusyn, I., and Daston, G. P. (2010). Computational toxicology: Realizing the promise of the toxicity testing in the 21st century. *Environ. Health Perspect.* **118,** 1047–1050.

Schnackenberg, L. K., Jones, R. C., Thyparambil, S., Taylor, J. T., Han, T., Tong, W., Hansen, D. K., Fuscoe, J. C., Edmondson, R. D., Beger, R. D., *et al.* (2006). An integrated study of acute effects of valproic acid in the liver using metabonomics, proteomics, and transcriptomics platforms. *OMICS* **10,** 1–14.

Schwarzman, M. R., and Wilson, M. P. (2009). New science for chemicals policy. *Science* **326,** 1065–1066.

Sedykh, A., Zhu, H., Tang, H., Zhang, L., Richard, A., Rusyn, I., and Tropsha, A. (2011). Use of in vitro HTS-derived concentration-response data as biological descriptors improves the accuracy of QSAR models of in vivo toxicity. *Environ. Health Perspect.* **119,** 364–370.

Shukla, S. J., Huang, R., Austin, C. P., and Xia, M. (2010). The future of toxicity testing: A focus on in vitro methods using a quantitative high-throughput screening platform. *Drug Discov. Today* **15,** 997–1007.

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., and Knudsen, T. B. (2011). Predictive models of prenatal developmental toxicity from ToxCast high-throughput screening data. *Toxicol. Sci.* **124,** 109–127.

Stouch, T. R., Kenyon, J. R., Johnson, S. R., Chen, X. Q., Doweyko, A., and Li, Y. (2003). In silico ADME/Tox: Why models fail. *J Comput. Aided Mol. Des.* **17,** 83–92.

Tong, W., Fang, H., Xie, Q., Hong, H., Shi, L., Perkins, R., Scherf, U., Goodsaid, F., and Frueh, F. W. (2006). Gaining confidence on molecular classification through consensus modeling and validation. *Toxicol. Mech. Methods* **16,** 59–68.

Uehara, T., Hirode, M., Ono, A., Kiyosawa, N., Omura, K., Shimizu, T., Mizukawa, Y., Miyagishima, T., Nagao, T., and Urushidani, T. (2008).

A toxicogenomics approach for early assessment of potential non-genotoxic hepatocarcinogenicity of chemicals in rats. *Toxicology* **250,** 15–26.

Uehara, T., Minowa, Y., Morikawa, Y., Kondo, C., Maruyama, T., Kato, I., Nakatsu, N., Igarashi, Y., Ono, A., Hayashi, H., *et al*. (2011). Prediction model of potential hepatocarcinogenicity of rat hepatocarcinogens using a large-scale toxicogenomics database. *Toxicol. Appl. Pharmacol.* **255,** 297–306.

Uehara, T., Ono, A., Maruyama, T., Kato, I., Yamada, H., Ohno, Y., and Urushidani, T. (2010). The Japanese toxicogenomics project: Application of toxicogenomics. *Mol. Nutr. Food Res.* **54,** 218–227.

Valerio, L. G. J. (2009). In silico toxicology for the pharmaceutical sciences. *Toxicol. Appl. Pharmacol.* **241,** 356–370.

Van Hummelen, P., and Sasaki, J. (2010). State-of-the-art genomics approaches in toxicology. *Mutat. Res.* **705,** 165–171.

Wang, E. J., Snyder, R. D., Fielden, M. R., Smith, R. J., and Gu, Y. Z. (2008). Validation of putative genomic biomarkers of nephrotoxicity in rats. *Toxicology* **246,** 91–100.

Waring, J. F., Ciurlionis, R., Jolly, R. A., Heindel, M., and Ulrich, R. G. (2001). Microarray analysis of hepatotoxins in vitro reveals a correlation between gene expression profiles and mechanisms of toxicity. *Toxicol. Lett.* **120,** 359–368.

Waterman, C. L., Currie, R. A., Cottrell, L. A., Dow, J., Wright, J., Waterfield, C. J., and Griffin, J. L. (2010). An integrated functional genomic study of acute phenobarbital exposure in the rat. *BMC Genomics* **11,** 9.

Waters, M. D., and Fostel, J. M. (2004). Toxicogenomics and systems toxicology: Aims and prospects. *Nat. Rev. Genet.* **5,** 936–948.

Waters, M. D., Jackson, M., and Lea, I. (2010). Characterizing and predicting carcinogenicity and mode of action using conventional and toxicogenomics methods. *Mutat. Res.* **705,** 184–200.

Wetmore, B. A., and Merrick, B. A. (2004). Toxicoproteomics: Proteomics applied to toxicology and pathology. *Toxicol. Pathol.* **32,** 619–642.

Xia, M., Huang, R., Guo, V., Southall, N., Cho, M. H., Inglese, J., Austin, C. P., and Nirenberg, M. (2009). Identification of compounds that potentiate CREB signaling as possible enhancers of long-term memory. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 2412–2417.

Xia, M., Huang, R., Witt, K. L., Southall, N., Fostel, J., Cho, M. H., Jadhav, A., Smith, C. S., Inglese, J., Portier, C. J., *et al*. (2008). Compound cytotoxicity profiling using quantitative high-throughput screening. *Environ. Health Perspect.* **116,** 284–291.

Zhu, H., Rusyn, I., Richard, A., and Tropsha, A. (2008). Use of cell viability assay data improves the prediction accuracy of conventional quantitative structure-activity relationship models of animal carcinogenicity. *Environ. Health Perspect.* **116,** 506–513.

Zhu, H., Ye, L., Richard, A., Golbraikh, A., Wright, F. A., Rusyn, I., and Tropsha, A. (2009). A novel two-step hierarchical quantitative structure-activity relationship modeling work flow for predicting acute toxicity of chemicals in rodents. *Environ. Health Perspect.* **117,** 1257–1264.

Zidek, N., Hellmann, J., Kramer, P. J., and Hewitt, P. G. (2007). Acute hepatotoxicity: A predictive model based on focused Illumina microarrays. *Toxicol. Sci.* **99,** 289–302.