ON THE USE AND FUTURE OF ALTERNATIVE TESTING
STRATEGIES IN REGULATORY TOXICOLOGY

by

Robert J. Borotkanics

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctorate of Public Health

Baltimore, Maryland
September 2014

© 2014 Robert Borotkanics
All Rights Reserved
Abstract

Regulatory toxicology has emerged as a necessary discipline to ensure that chemicals introduced into commerce are safe. Contemporary history has taught us the hard way that while many chemicals are not harmful to human health, there too are many chemicals that were found after introduction into commerce to cause harm. This is a serious public health concern. Regulatory toxicologists inform decisions in part based on alternative test methods in addition to traditional in vivo studies. This is in part due to legal constraints placed on regulators, and it has not been articulated in research what these constraints and data are, or how these considerations inform regulatory decisions. The findings of this body of research reveal that alternative test methods primarily inform decisions of the US Environmental Protection Agency’s New Chemicals Program, which administers the Toxic Substances Control Act. This research also finds that there exist data gaps that need to be filled so that the New Chemicals Program can be able to make more informed decisions. Various studies have blossomed in recent years on the potential of alternative test methods to inform decisions, which includes the work of the ToxCast program. ToxCast has been productive in research, but has yet to have a method validated and implemented by regulators. This study finds that there are ToxCast parameters that may be predictive of hepatocarcinogenesis, but further research is warranted. Finally, the US National Academies of Sciences has promoted the use of alternative test methods, specifically putting forward the notion that the perturbation of network motifs may be a means to fill information gaps. The literature on motifs to date is limited to basic research, but recurring motifs across species that were identified in this body of research can be applicable within toxicity testing moving forward.
Acknowledgements

Marsha Wills-Karp; CC Talbot; Allison Brooker, Agnes Brooker, Elizabeth Kasmeyer, Jennifer Watkins, Annette Boelman, Brian Olynek and the entire Roger’s Pass crew

In Memoriam

Iris, 1936-2007

Thesis readers

1. Harold Lehmann, PhD, MD, Johns Hopkins School of Medicine (chair)
2. Michael Trush, PhD, Johns Hopkins Bloomberg School of Public Health
3. Paul Locke, DrPH, JD, MPH, Johns Hopkins Bloomberg School of Public Health (advisor)
4. Mary Fox, PhD, MPH Johns Hopkins Bloomberg School of Public Health
5. Louis Scarano, PhD, US Environmental Protection Agency
**Table of Contents**

Abstract ................................................................................................................................. ii

Acknowledgments ................................................................................................................ iii

List of Tables ........................................................................................................................... v

List of Figures .......................................................................................................................... vii

Chapter I: Introduction ......................................................................................................... 1

Chapter II: A review of the literature with respect to the research outlined in this dissertation ......................................................... 10

Chapter III: EPA’s methodology to inform TSCA PMN decision-making: an analysis and critique ...................................................... 37

Chapter IV: Network motifs that recur across species: gene regulatory and protein-protein interaction networks ... 68

Chapter V: *In vitro* perturbation of the transcription factor e2f is a strong predictor of *in vivo* hepatocarcinogenesis resulting from chemical exposure ................................................. 97

Chapter VI: Conclusion ......................................................................................................... 131

Curriculum vitae ..................................................................................................................... 144
List of Tables

Chapter II
Table 1: Organisms across studies ......................................................... 34
Table 2: Statistical approaches preferred across science ...................... 35
Table 3: Statistical approaches recommended by validation organizations ................................................................. 36

Chapter III
Table 1: Chemicals meeting inclusion criteria ........................................ 60
Table 2: Test data provided ................................................................. 61
Table 3: EcoTox data generated using EcoSAR ............................... 62
Table 4: Decisions and basis for decisions ............................................. 63
Table 5: Human health concerns identified in EPA SAT reports ................................. 64

Chapter IV
Table 1: Studies meeting inclusion criteria ............................................. 89
Table 2: Two-node motifs .................................................................. 90
Table 3: Three-node motifs (1 of 2) .................................................... 91
Table 4: Three-node motifs (2 of 2) .................................................... 92
Table 5: Four-node motifs ................................................................ 93
Table 6: Five-node motifs ................................................................ 94

Chapter V
Table 1: Control group chemicals ........................................................ 120
Table 2: Case group chemicals ............................................................ 121
Table 3: Summary of chemicals by classification ..........................122

Table 4: Student’s t-tests: physiochemical properties and transcription factors/nuclear receptors .................................123

Table 5: AUC for physiochemical properties & Attagene assay .......124

Table 4: e2f cutoff values ................................................................125
List of Figures

Chapter 1
Figure 1: Future options of toxicity testing ........................................9

Chapter III
Figure 1: Sources for data abstraction .................................................65
Figure 2: Data abstraction steps ............................................................66
Figure 3: Inclusion criteria .................................................................67

Chapter IV
Figure 1: Inclusion criteria .................................................................95
Figure 2: CAR-PXR DOR .................................................................96

Chapter V
Figure 1: Study inclusion criteria .......................................................126
Figure 2: Data from Cellumen cell count .............................................127
Figure 3: Dose-Response of selected nuclear receptors and transcription factors .................................................................128
Figure 4: ROC curves, physiochemical properties ...............................129
Figure 5: ROC curves of weighted e2f and Nrf2 .................................130
Chapter I: Introduction

Problem statement

Regulatory toxicology has emerged as a necessary discipline to ensure that the chemicals introduced into commerce are safe. Contemporary history has taught us the hard way that while many chemicals are not harmful to the health of humans or the environment, there too are many chemicals that were found after introduction into commerce to cause harm. For instance, perfluoroalkylated substances are a group of fluorinated compounds with wide application, ranging from use in household products to outdoor clothing. Perfluorooctane sulfonate (PFOS) and perfluorooctanoic acid (PFOA) belong to this chemical class and have been in production since the 1940’s. It has been learned that increased exposure to PFOS is associated with bladder cancer, and PFOA is associated with kidney and testicular cancers (1, 2).

There exist in the United States (US) market today many chemicals that are used in mass quantity. For instance, the US Environmental Protection Agency (EPA) is required under the Toxic Substances Control Act (TSCA) to maintain a list of chemicals used in US commerce. There are 83,294 listed (3). We lack a full appreciation of the potential risks these chemicals may pose to human and environmental health. The US Centers for Disease Control and Prevention recently measured over 200 chemicals in human blood and urine and concluded that for most of these chemicals, there is a lack of sufficient information to determine whether these chemical exposures are of concern to human health (4). This lack of risk information is an important public health issue, because
Chemicals are ubiquitous in the US with nearly every person experiencing some level of exposure.

Regulatory toxicology is specifically designed to deal with these issues. This discipline is unique in that the decisions originating from the science are guided by a number of laws. For example, the Toxic Substances Control Act (TSCA) is a US law designed to regulate chemicals that could pose an unreasonable risk of harm to health or the environment that are not covered by other laws (5). Chapter III elaborates on select nuances of TSCA. Laws like TSCA have been enacted to offset imbalances and reduce risks to human health and the environment while still allowing chemicals into commerce. TSCA further sets a decision-making framework for the regulatory body and gives authorities - in this case the US Environmental Protection Agency (EPA) - to control a chemical’s entry into commerce and under what conditions entry may occur.

A necessary interrelationship exists between chemical safety laws and the science of toxicology. The decisions resulting from administering these chemical safety laws depend extensively on the science of toxicology, of which an integral component is toxicity testing. Tests are carried out on chemicals, ideally before they enter commerce, to increase the certainty that when used, they will not cause harm to health or the environment. These tests have traditionally been in vivo, or animal-based tests. However, alternative testing methods also exist and are starting to take prominence for a number of reasons. The term ‘alternative test methods’ even has legal meaning in the US. By definition, such a method “is a new or revised test method and reduces the
number of animals required, refines procedures to lessen or eliminate pain or distress to
animals or enhances animal well-being; or replaces animals with non-animal systems or
one animal species with a phylogenetically lower animal species, such as replacing a
mammal with an invertebrate (6).”

The creation and adoption of alternative test methods are important for ethical, cost and
efficiency reasons. First, the ethical treatment of animals is culturally important in
western society and increasingly, people are scrutinizing the use of animals in research.
The treatise entitled, “The Principles of Humane Experimental Technique” put forward
the tripartite notion of “replacement, refinement and reduction” to serve as the guiding
posts for the conduct of ethical research (7). In the US, multiple laws have been enacted
to promote animal welfare and adoption of alternative methods, including the Animal
Welfare Act (8) i, the Health Services Extension Act of 1985 (9) and Public Law 106-
545, which enshrined into law the Interagency Coordinating Committee on the Validation
of Alternative Methods (ICCVAM). The primary purpose of ICCVAM is to evaluate
alternative test methods that reduce, refine or replace animal tests (6)ii. Second, the
potential cost savings are a more pragmatic justification for creation and adoption of
alternative methods. It is generally accepted that alternative methods can potentially
result in significant cost savings over traditional, in vivo methods (10). Finally,
alternative methods can be carried out more quickly and the same test can be carried out
on a multitude of chemicals simultaneously (11).

---

i This act has undergone 7 revisions since its inception in 1966. The most recent update occurred in 2008.
ii In the European Union, the Lisbon treaty instills animal welfare as a common value of the community
(Lisbon, article 13). The European community has since 1986 implemented a series of directives
specifically designed to promote animal welfare and reduce animal use in experiments, including
The US National Academies of Sciences (NAS) also supports the use of alternative test methods. In 2007, the NAS published “Toxicity Testing in the 21st Century: a Vision and Strategy (12).” In this report a vision for toxicity testing was posited, in which the options for toxicity testing strategies in the future would rely less on in vivo methods and more extensively on in vitro and computational, also called in silico, methods. These in vitro and in silico methods would ideally be geared to identify those critical pathways that when sufficiently perturbed lead to adverse health outcomes (figure 1). The NAS further stated that, “perturbations of cell-signaling motifs . . . are obligatory changes related to chemical exposure that might eventually result in disease.”iii

The research arm of the EPA has been moving forward on this vision and initiated the ToxCast program. The ToxCast is a pilot program that conducts a variety of in vitro tests. Initiated in 2007 (13), ToxCast is an EPA-run program that generates large volumes of data on chemicals (11). The in vitro assays range in scale from enzyme challenge assays to cell-level assays. The information is supplemented by data libraries, detailing the varied properties of these very same chemicals. The ToxCast program completed their Phase I, “Proof of Concept” phase in 2009 analyzing 309 unique chemicals.iv

---

iii Four US agencies entered into a memorandum of understanding entitled, “High Throughput Screening, Toxicity Pathway Profiling” to collaborate on efforts related to the NAS report.

iv ToxCast Phase II data was released by the EPA in January 2014.
To date, ToxCast has analyzed many aspects of the program’s data. As noted in chapter V, these studies include but are not limited to analyses of pathway perturbations, models to potentially be predictive of developmental toxicity, a reproductive toxicity model, genotoxicity screening and endocrine disruption screening. The ToxCast battery of assays also includes data on perturbations of hepatocyte function. These studies have contributed to making important findings regarding the association and correlation of the activity of xenobiotic metabolizing enzymes and nuclear receptors with respect to liver carcinogenicity and chemical exposure. However, these perturbations have yet to be assessed as to their ability to discriminate and therefore predict the presence or absence of liver lesions of \textit{in vivo}, chronic toxicity tests. This important research question is addressed in chapter V.

The TSCA and ToxCast programs, via the NAS vision, have potentially important implications for one other. Regulators like EPA’s TSCA program continue to make regulatory decisions and rely on the science of toxicology - and the state-of-the-art thereof - to do so. If there comes to maturity scientifically sound alternative testing methods that the TSCA program can apply and use to inform decisions, particularly if the science can be applied more efficiently, then the TSCA program could potentially have an interest in adopting such methods. The application of ToxCast data and methods has the potential to demonstrate their relevance to regulatory bodies like TSCA.

Finally, a particular nuance to the NAS vision makes a curious if not provocative statement: “perturbations of cell-signaling motifs . . . are obligatory changes related to
chemical exposure that might eventually result in disease.” Not only can perturbations have potentially important application in the future of toxicity testing, but specifically perturbations of motifs. As chapter IV notes, motifs have only recently come to our scientific understanding. The notion of a motif comes from graph theory and its application in network analysis. Briefly a cellular network may be represented as a graph, which abstractly comprises nodes (vertices) and edges (arcs). Cellular molecules may be represented as nodes. Nodes may be connected by an edge, representing the interactions between the molecules. One may further break down a graph into subgraphs. These subgraphs are made up of a finite number of nodes and these nodes’ interactions with one another characterize a given network at the local level. Subgraphs that are statistically overrepresented in a network are referred to as network motifs. ToxCast has not characterized network motifs or perturbations thereof. It is, however, important to attain a better understanding of motifs at the most fundamental level so that this information can potentially be applied to the research of toxicity testing in the future.

Research questions and hypotheses

The specific aims of this body of research are:

Aim 1: Articulate the regulatory process with which the TSCA New Chemicals Program must legally comply:

1. Identify the in vivo and alternative testing methods currently applied in the decision-making process and examine how these methods are used to inform decisions.
2. Evaluate the New Chemicals Program premanufacturing notification review process.

3. Assess the potential applicability of the ToxCast system of assays in the assessment of hepatocarcinogenic hazard risk.

Hypothesis A: There exist toxicity testing gaps in the New Chemicals Program review process.

Hypothesis B: ToxCast phase I data can discriminate an apical endpoint, hepatocarcinogenesis, that is of interest to regulatory toxicologists.

Aim 2: Review and synthesize the current knowledge of network motifs.

1. Identify the network motifs that recur across species.
2. Outline and explain potential research agenda on motifs and their potential applicability to toxicity testing in the future.

Hypothesis C: There exist recurring motifs across species and biological pathways.

Aim 3: Integrate research findings and make recommendations to strengthen regulatory toxicology.
References


8. Title 7 - Agriculture; Chapter 54-Transportation, sale and handling of certain animals. Washington, DC; 2008.


### Figure 1: Future options of toxicity testing (12)

<table>
<thead>
<tr>
<th></th>
<th>Option I In Vivo</th>
<th>Option II Tiered In Vivo</th>
<th>Option III In Vitro/In Vivo</th>
<th>Option IV In vitro</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal biology</td>
<td>Animal biology</td>
<td>Primarily human biology</td>
<td>Primarily human biology</td>
<td></td>
</tr>
<tr>
<td>High doses</td>
<td>High doses</td>
<td>Broad range of doses</td>
<td>Broad range of doses</td>
<td></td>
</tr>
<tr>
<td>Low throughput</td>
<td>Improved throughput</td>
<td>High and medium throughput</td>
<td>High throughput</td>
<td></td>
</tr>
<tr>
<td>Expensive</td>
<td>Less expensive</td>
<td>Less expensive</td>
<td>Less expensive</td>
<td>Less expensive</td>
</tr>
<tr>
<td>Time consuming</td>
<td>Less time consuming</td>
<td>Less time consuming</td>
<td>Less time consuming</td>
<td>Less time consuming</td>
</tr>
<tr>
<td>Use of relatively large numbers of animals</td>
<td>Use of fewer animals</td>
<td>Use of substantially fewer animals</td>
<td>Use of virtually no animals</td>
<td></td>
</tr>
<tr>
<td>Based on apical endpoints</td>
<td>Based on apical endpoints</td>
<td>Based on perturbations of critical cellular responses</td>
<td>Based on perturbations of critical cellular responses</td>
<td></td>
</tr>
<tr>
<td>Some screening using computational and in vitro approaches; more flexibility than current methods</td>
<td>Screening using computational approaches possible; limited animal studies that focus on mechanism and metabolism</td>
<td>Screening using computational approaches</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Chapter II: A review of the literature with respect to the research outlined in this dissertation

Abstract

Toxicity testing is central to the process of regulatory decision making and ensuring the public is protected from the harm caused by chemicals. Regulatory toxicology has historically relied upon in vivo models to inform decisions. However, alternative testing methods are emerging as a preferred approach to toxicity testing. In fact, the U.S. National Academies of Sciences (NAS) in 2007, via its report entitled, “Toxicity Testing in the 21st Century,” posited a vision for toxicity testing that uses in vitro tests and in silico models more extensively and relies less on in vivo tests. The NAS, in particular, put forward the notion that the goals of toxicity testing should be to identify those critical pathways that when sufficiently perturbed lead to adverse health outcomes. The NAS gave special emphasis to network, or signaling motifs and their perturbation. But what are motifs and what is the state of science thereof?

It is important to validate alternative test methods before using such approaches to inform regulatory decisions. Test method validation is the process by which the reliability and relevance of a test, method, or approach are established for a particular purpose. Quantitative validation assessments are also important, particularly in the determination of a test’s or model’s accuracy.
This review of the literature review summarizes the body of research carried out to date on motifs and summarizes the state of science with respect to this area of interest. This review of the literature also summarizes the statistical methods used by major, alternative testing validation organizations and the broader scientific community in the assessment of accuracy.

Introduction

This body of research is made of up two components that while interrelated, necessitated distinct literature searches and reviews of the literature. The first review of the literature is a study of those network motifs that could potentially be of importance to toxicology, in particular, toxicity testing. The second review of the literature is on the quantitative methods used to assess the accuracy of alternative test methods.

Review of the literature with respect to network motifs:

Background

The U.S. National Academies of Sciences (NAS) in 2007, via its report entitled, “Toxicity Testing in the 21st Century,” posited a vision for toxicity testing, asserting that the goals of toxicity testing moving forward should be to identify those critical pathways that when sufficiently perturbed lead to adverse health outcomes. The NAS further stated that, “perturbations of cell-signaling motifs . . . are obligatory changes related to chemical exposure that might eventually result in disease (1).” The fundamental question is: what is a motif?
Mathematically, a cellular network may be represented as a graph (2), which abstractly comprises nodes (vertices) and edges (arcs). Cellular molecules may be represented as nodes. Nodes may be connected by an edge, representing the interactions between the molecules. If the nodes and edges are countable, then the graph is said to be finite. One may further break down a graph into subgraphs. These subgraphs are made up of a finite number of nodes and these nodes’ interactions with one another characterize a given network at the local level. Two or more graphs or subgraphs are considered isomorphic if they have the same node number and their adjacency is preserved (i.e., corresponding nodes are connected by edges the same way in both subgraphs). Subgraphs that are statistically overrepresented in a network are referred to as network motifs\(^v\) (3, 4). The theoretical foundations of network motifs are well established; however, it is not known what motifs are associated with toxicity, or which ones recur across species/biological pathways.

The aim of this review of the literature is therefore to identify the network motifs, if any, that are associated with toxicity.

**Methods:**

**Search overview**

A search of the peer reviewed literature was carried out, wherein our strategy was to maximize the chances of identifying all available articles identifying network motifs experimentally, communicating their findings in English.

\(^v\) Herein, called ‘motifs’ for brevity.
Search strategies and data sources

Peer reviewed research articles were identified using Google Scholar, PubMed and Scopus. Articles were found by the use of each search engine via keyword search, using the following search terms: ‘functional motif,’ ‘graph,’ ‘motif,’ ‘network motif,’ or ‘subgraph.’ Terms were used in their singular and plural forms. Supplementary searches were carried out using the aforementioned terms in combination with the following terms: ‘biological network,’ ‘cellular biology,’ ‘computational biology,’ ‘molecular biology,’ ‘network,’ ‘network topology,’ or ‘systems biology,’ again using both singular and plural forms.

Study eligibility and selection

Articles were included if they reported on research that quantitatively assessed the presence of motifs at the intracellular level, using experimental methods of non-plant organisms. In circumstances where studies were carried out using cell-lines, only those studies using non-carcinogenic cell lines were accepted for inclusion. Studies identifying intercellular motifs were excluded. Summaries and reviews were noted and used for informational purposes, but excluded from the analysis, so as to avoid the redundant reporting of findings.
Results

The literature search across both Google Scholar and PubMed resulted in the identification of 141 peer reviewed articles. No unique articles were found using Scopus. These articles were classified along five axes:

1) Does the article specifically identify the motif and provide a network topology;
2) Is the motif described intracellular in character;
3) If the motif is intracellular in character, is the motif identified in non-diseased cells;
4) Does the article use quantified, experimental methods to identify the motif;

A total of 46 articles specifically identified a motif and provided a network topology. A network topology is an abstract representation of a network. There were 42 articles that described intracellular motifs. Of these 42 articles, 36 articles described motifs of non-diseased cells. The remaining studies identified motifs in immortalized cell lines or cancer cell lines. One half of the articles that identified motifs, or 23 articles, detailed the quantified, experimental methods used to identify the motifs. Out of the entire literature search, 29 articles described experimentally the dynamic behavior of the motifs identified. Due to combinatorial differences, a total of 17 articles identified motifs along with topology AND described motifs at the intracellular level AND described intracellular motifs of non-diseased cells AND described the quantified, experimental
methods used to identify the motifs (table 1, chapter IV). All of these 17 articles were published between the years of 2002 to 2012.

While the notion of motifs for use in toxicology and toxicity testing has been described at the theoretical level (1), no articles meeting the literature search criteria were published in a toxicology journal nor were articles identified that had a specific, toxicology theme. All articles could at best be described as basic science discoveries; in which, the intracellular networks of specific cells within specific organisms had their respective motifs identified and described (table 1, chapter IV).

The motifs were identified across a range of species (table 1). In only two studies were specific cell lines identified. The remaining studies summarized intracellular motifs across the broad range of network types, including gene regulatory networks and protein-protein interaction networks, which include a multitude of cell lines or tissue types.

The majority of studies included in this review applied the isomorphism and randomization methods first developed by Milo et al. (3), and also by Kashtan et al. (5), with minor variations (n=13). There exist a number of distinct, motif isomorphism and randomization algorithms (6). Further, the choice of isomorphism algorithm is known to effect study results (7). This was reaffirmed by analysis of the studies included in this review. For instance, Konagurthu and Lesk (8) repeated the studies of Ma’ayan et al. (9) and Shen-Orr et al. (4), but used a more conservative isomorphism algorithm. This research team identified both the feed forward loop and 3-cycle motifs, consistent with
the other two studies. However, Konagurthu and Lesk did not identify the same extent of three node motifs found in Ma’ayan or the four node motifs of both Ma’ayan et al. (9) and Shen-Orr et al (4). Further, Kashtan et al. (5) applied a sampling based algorithm to a version of the data used in Shen-Orr et al. (4), identifying feed forward loops and dense overlapping regions, but like Konagurthu and Lesk (8), did not make a single input module motif finding.

Discussion

The U.S. National Academies of Sciences (NAS) in 2007, via its report entitled, “Toxicity Testing in the 21st Century,” posited a vision for toxicity testing, asserting that the goals of toxicity testing moving forward should be to identify those critical pathways that when sufficiently perturbed lead to adverse health outcomes. The NAS further stated that, “perturbations of cell-signaling motifs . . . are obligatory changes related to chemical exposure that might eventually result in disease (1).” The notion of motifs is a recent phenomenon, a little over a decade old. This literature review demonstrates the scientific community only now is starting to systematically describe these important patterns of sub-cellular, network connectivity. This review of the literature shows that motifs associated with specific cells lines are not well described; nor are the motifs that recur across species well described. The notion put forward by the NAS is that perturbations of motifs resulting from chemical exposure could potentially be predictors of toxicity. In order to characterize these perturbations, it is first necessary to describe what the motifs are. Second, motifs need to be characterized across network types, cell lines, tissue types and species. The answers to these questions are fundamentally necessary before one can
proceed to dynamics of perturbations thereof. These fundamental gaps in our scientific understanding need to be fulfilled not only in toxicology, but in basic science in general.

Review of the literature with respect to quantitative methods used to assess alternative test method accuracy

Background
Many types of tests are used to assess whether or not a chemical causes harm. *In vivo*-based tests are animal-based studies used to assess any of a number of apical endpoints and may be carried out on any of a variety of species. For instance, the rat acute toxicity test, rabbit dermal toxicity and rabbit eye irritation assays are examples of the *in vivo* assays more commonly used to assess acute (i.e., short-term) hazard by the US Environmental Protection Agency’s (EPA) New Chemicals Program (chapter III; table 2). These *in vivo* methods generally follow approved methods and protocols, which regulators commonly apply. The Organisation for Economic Cooperation and Development (OECD) is an international body made up of 34 participating countries, including the United States. The OECD as an organization develops and publishes guidelines for toxicity testing that member countries, as signatories, agree to. For instance, test guideline OECD-423 is the standardized guideline to assess acute, oral toxicity (10). The EPA’s test guidelines are harmonized to conform to OECD test guidelines (11).
In vitro tests have also emerged as having value in predicting whether a chemical is likely to be a hazard to human health. Regulatory bodies such as the EPA commonly use in vitro assays, including the Ames assay and similar E. coli or chromosome aberration test assays (chapter III; table 2). Assays like these are also used internationally by regulators and also standardized by the OECD. For instance OECD-471 and OECD-473 are the standard, international test guidelines for the Ames assay and chromosome aberration test, respectively (12, 13).

Finally, there exist computationally-based approaches, sometimes referred to as in silico methods, that take information on a chemical or a physiological parameter to estimate the probability of whether or not a chemical causes harm. The US EPA has validated in silico methods to estimate the physiochemical properties and environmental fate of chemicals (14). Also used by the EPA is a platform called EcoSAR to estimate the probable aquatic toxicity (15). Both in silico methods have been validated by the EPA’s Scientific Advisory Board and are regularly used to inform regulatory decisions, which chapter III discusses in greater detail. In silico methods are less standardized internationally than its sibling, in vivo and in vitro methods. For instance, Canada uses a number of software platforms to screen chemicals (16). The North American Free Trade Agreement member countries have developed a guidance document on the use of quantitative structure activity relationship (QSAR) models to predict toxicity (17). The OECD has developed a QSAR Toolbox to make QSAR more accessible and to promote its use across the chemical industry, governments and other stakeholders (18). The European Union permits the use of in silico methods under its Registration, Evaluation,
Authorisation and Restriction of Chemicals (REACH) law so long as the models conform to specific criteria as laid out in Annex XI of this law (19). In general, models must have scientific validity established, the chemical being evaluated must fall within the applicability domain of the model, the results are adequate for the intended purposes and adequate and reliable documentation on the applied model is available and provided. The REACH framework sets similar standards for establishment and use of in vitro methods.

In vitro and in silico methods have multiple, potential benefits; however, it is also important for these alternative test methods to undergo validation before used to inform regulatory decisions, conceptually known as regulatory acceptance. The EPA’s in silico models have undergone validation, and REACH sets standards for adoption and use, which include validation. Many organizations are involved in validation, including the European Union Reference Laboratory for Alternatives to Animal Testing (EURL-ECVAM) and also the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM). The OECD is also involved in such activities. These organizations have laid out many guidance documents to facilitate validation, as will be described in greater detail later in this chapter. An important component of validation is the quantitative methods used to assess accuracy of any given candidate in vitro test or in silico model.

Test method validation is the process by which the reliability and relevance of a test, method, or approach are established for a particular purpose (20, 21, 22). Reliability is the extent to which a test method produces consistent results (20), which is often
considered across laboratories or over multiple experiments. Relevance is the extent to which a test method correctly predicts or measures the biological effect of interest (20). The closer the linkage between the effect measured and the toxicological effect of interest, the easier it will be to establish the relevance of the assay (20). An important component of a test method’s relevance is its accuracy. The accuracy of a test is the extent to which it agrees with accepted or reference values. In order to be acceptable to regulatory agencies, test methods must be reliable and relevant enough so that their results can be used in regulatory toxicity decision-making.

These validation processes, regardless of organization carrying out the study, rely in part on quantitative tools to analyze the accuracy of a proposed test method. But what are those statistical methods used in the quantification of test accuracy? Further, what are the quantitative methods applied to assess whether or not a candidate test is at least as good as or better than an existing test method? The aim of this literature review is to identify and summarize statistical methods applied in these two particular circumstances.

**Methods**

**Search overview**

The most prominent governmental organizations involved in validation of alternative test methods are EURL-ECVAM and ICCVAM. The OECD also plays a role in validation and promulgation of test guidelines. Therefore, a review of these organizations’ test validation guidelines was first carried out. A broader search of the scientific literature was then conducted, using Google Scholar and PubMed. The intent of this broader
search was to encapsulate the quantitative approaches used in the assessment of accuracy by the scientific community.

Data sources and search strategies
The three major international bodies involved in validation of alternative test methods or adoption thereof – EURL-ECVAM, ICCVAM and OECD – have established guidelines or similar documents that spell out the manner in which alternative test methods would be acceptable as replacements to existing tests. Therefore, a targeted literature review was first carried out on the guidance documents published by these organizations. These organizations make their reports publicly available via electronic means. Documents were identified by a literature search of their respective electronic libraries.

A broader review of the literature was then carried out, the purpose of which was to capture and encapsulate statistical methods used by the scientific community. This literature search relied upon both Google Scholar and PubMed to identify relevant articles, using the following combination of search terms: ‘Statistics’ or ‘Statistical’ or ‘Biostatistics’ or ‘Quantitative’ AND ‘Tests’ or ‘Screening Programs’ or ‘Screening Assay’ or ‘Diagnostic Assay’ or ‘Model’ AND ‘Accuracy’ or ‘Discrimination’ or ‘Sensitivity’ or ‘Specificity’ or ‘Classification’ or ‘Prediction’.

Study eligibility and selection
Peer reviewed journals, reports and books published in English were selected for consideration based on the following two inclusion criteria:
1. Quantitatively assessed how well a test performs against a known standard;
2. Quantitatively compares a candidate test against an existing test.

Guidance documents originating from EURL-ECVAM, ICCVAM and the OECD that articulated statistical methods with respect to quantification of accuracy were included in the review.

Results from the broader literature search that met inclusion criteria resulted in a unique circumstance. Initial rounds of literature searches using the stated search terms resulted in large quantities of documents that made it prohibitive to review each and every document. For instance, use of the search terms, ‘Statistics AND Test AND Accuracy’ returned 2.45 million results in Google Scholar; in PubMed, 4,470. Every search using the stated combination of search terms resulted in similar numbers of results. Therefore a convenience sample of 15 peer reviewed journals, reports and books was selected for review. Peer reviewed articles were selected using a stratification strategy where statistical methods used in the quantification of accuracy were selected from five unique, scientific disciplines. The rationale for this approach was to identify statistical methods preferred across the scientific community.

Results
EURL-ECVAM promulgates a number of guidance documents on the validation of alternative methods. The overall framework of EURL-ECVAM’s validation efforts is
guided by the laws that established the organization (23). These laws also provide a broad, regulatory framework for which EURL-ECVAM must operate within. This organization uses as its guide a number of peer reviewed scientific papers and similar reports that provide a conceptual framework of EURL-ECVAM’s validation activities (20, 24, 25). Many of these documents articulate important terms and associated definitions, and also outline the procedures and other important logistical aspects of validation. In one of these documents, Balls et al. summarizes two types of information required from validation studies (24). The first type of information is variability data both within and between participating laboratories. Analysis of variance (ANOVA) is specifically mentioned as an important tool in the assessment of variability. The second type of information is the direct comparison between the results of the alternative test and the associated in vivo test, of which this framework document only briefly highlights methodological considerations.

EURL-ECVAM established a task force on the use of biostatistical methods and the validation process (26). Their report summarizes the areas of validation for which statistical analyses are important. A section of the task force report addresses variability and strategies to reduce test variability both within and across laboratories. The report encourages the systematic alteration of test methods to identify key areas of variability. The task force recommends quantification of variability via ANOVA methods (table 2). The report also briefly touches upon outlier analysis and censoring of data. Dose-response analysis is identified in the report as the means to identify informative endpoints. The report does not speak to analysis of an in vitro test or in silico models.
against reference chemicals, but rather recommends the use of principal components analysis in the selection of test chemicals. With regard to tests of accuracy and discriminant analysis, which is the focus of this review of the literature, the task force recommends the use of contingency tables for the analysis of categorical data. For continuous data, the task force recommends the use of cluster analysis techniques or Fisher’s discriminant analysis. Cluster analysis is a generic term used to identify a wide range of quantitative methods that reveal natural groups or clusters and can be applied to both categorical and continuous data (27). Fisher’s discriminant analysis, or similar discriminant analysis tools, can take continuous predictors to discern two or more categories, classes, objects or events (28).

ICCVAM is the US federal government’s entity charged with the evaluation of alternative test methods (29). The ICCVAM does not develop alternative testing methods, but rather has developed and maintains a structured approach to review and evaluate alternative test proposals. An applicant can submit a proposed alternative test to ICCVAM, which ICCVAM then evaluates and reports on. The ICCVAM takes these evaluations and provides recommendations to US regulatory agencies. It is then up to the US regulators to adopt any alternative test method that has been evaluated by ICCVAM.

This validation organization requires specific statistics to accompany proposal submissions. For instance, ICCVAM requires analysis of within laboratory and inter-laboratory variability (30). Otherwise, ICCVAM is non-specific with regard to statistical methods to be applied in assessing the accuracy and discriminatory abilities of a given
test. Nor does the ICCVAM discuss methods for comparison of tests. The ICCVAM states only that statistical methods must be described, and the statistical methods used for data evaluation should be justified. The ad hoc Committee created prior to the passage of the ICCVAM Act of 2000 did mention that the operational characteristics of a test are desirable, including measures of sensitivity and specificity or the derivation of correlation coefficients (22). These are noted in table 2 of this chapter. Other statistical tools are briefly mentioned in this report; however, such approaches are discussed without any reference to any particular aspect of the validation process. ICCVAM does not discuss statistical approaches to summarize the accuracy of a given test, nor do they discuss methods to compare the performance of two or more tests.

The OECD has published and maintains hundreds of test guideline documents. These guidelines can be followed to determine the physiochemical properties of chemicals, their environmental fate, effect on organisms in the environment and also to assess hazards to human health. A guideline on the conduct of occupational/epidemiological studies exists. There is, however, only one report on the validation of new test methods for hazard assessment. This document is entitled, “Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment,” which was published in 2005 (20).

The main areas of validation, according to the OECD are that the reliability and relevance of the proposed test need to be fully assessed. The OECD guideline contains limited information about specific, statistical methods that can be applied to the exercise of
validation; however, the guidance document does note methods that are important to apply in the conduct of validation (table 2).

The OECD describes validation as a multi-phase process, where inter-laboratory testing is carried out using a blinded testing approach. Further, the candidate assay or method is tested against approved negative and positive controls. The OECD places great emphasis on variability. The guidance document calls for assessment of variability in a laboratory over time and also assessment of variability across laboratories, going so far as to recommend the ASTM E691 as the approach for assessment of interlaboratory variability. This standard was previously known as ASTM E691-92(62), but is now identified as ASTM E691-13, “Standard practice for conducting an interlaboratory study to determine the precision of a test method (31).”

The OECD does not address tests to apply to continuous data. OECD also does not discuss statistical approaches to summarize the accuracy of a given test, nor do they discuss methods to compare the performance of two or more tests.

A broader search of the literature revealed the preferred statistical approaches of the scientific community. For both quantitative assessment of how well a test performs against a known standard and quantitative comparison of a candidate test against an existing test, the preferred approach was receiver operating characteristic (ROC) analysis. As reflected in table 3, ROC analysis has been used in clinical medicine, ecology and weather, epidemiology and statistics and also in machine learning. ROC has been
applied less extensively in agriculture and forestry. It is of note that clinical medicine uses ROC analysis extensively and in particular, in radiology and nuclear medicine (32). The results reflected in table 3 likely underestimate the extent of ROC analysis in the field of psychology. Many psychology-based journals described ROC in their abstracts; however, due to journal access limitations, thorough assessment of specific journals was limited in this review of the literature.

**Discussion**

The most revealing finding from this review of the literature is the discrepancy between validation organizations and other scientific disciplines with regard to the statistical tools emphasized in assessing test or model accuracy. Validation organizations discussed variability most frequently. Methods to assess accuracy, discrimination and prediction are only lightly touched upon by these organizations. ROC analysis, for instance, is not mentioned by any of the validation organizations, representing a major gap in the foundational literature of the validation organizations.

ROC curves visually depict how well a diagnostic test, whose results are ordinal, distinguishes between these two states, where the sensitivity is plotted on the y axis and 1-specificity is plotted on the x axis. A benefit of ROC analysis is that a test’s performance can be considered across a range of cutpoint values (33). A ROC curve is based on the sensitivity and specificity of a given test and not the positive predictive value. Therefore, ROC does not exhibit the sensitivities to prevalence that positive predictive value does (34). The performance of a test can further be quantified via
calculation of the area under the curve (AUC). The AUC represents the test’s discriminatory power. A random test would have an AUC of 0.5. A perfect test would have an AUC of 1. The AUCs of two or more tests can be compared to one another and used to assess which test is more accurate. AUC comparisons are typically carried out using an adaptation of the Chi Square statistic developed by DeLong (33). This method is used by the statistical software package STATA. ROC analyses can be used to assess both the performance of diagnostic tests and also models (35). A limitation of ROC analysis is that it is difficult to assess the accuracy of combinations of tests (33). For example, the performance of parallel or sequential series of tests cannot be evaluated using this method.

With regard to the literature search carried out, the choice of search terms did influence search results. Terms like ‘screening assay,’ ‘screening test’ and ‘diagnostic assay’ resulted in many peer reviewed publications originating from clinical medicine. Likewise, use of the term ‘accuracy’ resulted in more search results related to contingency tables and ROC analyses; whereas, the search term ‘discrimination,’ produced more results related to discriminant analysis methods. A range of diverse search terms should be applied in reviews of the literature that are similar to this one.

The review of the literature revealed that ROC analysis is the preferred method in assessing the accuracy of a test or model. Of the scientific disciplines included in this analysis, only agriculture and forestry did not demonstrate a use of ROC.
The peer reviewed literature on statistical tools used to quantify accuracy is vast. Initial rounds of search engine results were in the millions and would have been prohibitive and impractical to carry out a detailed review of all results. The analysis in this review study is therefore a convenience sample. Therefore, the results of this review of the literature may be biased non-differentially.

It is reasonable, however, to conclude that ROC analysis is likely the more commonly accepted approach to quantify accuracy. For instance, Swets summarized ROC’s use in weather forecasting, machine learning, psychology, clinical medicine, engineering and polygraph lie detection (36). Swets noted the application of ROC in medical imaging, in particular. The classic example of ROC in medical imaging is the work by Hanley and McNeil (37). It would therefore be reasonable to apply the ROC methods to the analysis of alternative test methods within the discipline of toxicology as its application is generally accepted across the scientific community.
References


Table 1: Organisms across studies

<table>
<thead>
<tr>
<th>Organism/Cell line studied</th>
<th># of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. cerevisiae</td>
<td>7</td>
</tr>
<tr>
<td>E. coli</td>
<td>5</td>
</tr>
<tr>
<td>D. melanogaster</td>
<td>2</td>
</tr>
<tr>
<td>E. lupus</td>
<td>2</td>
</tr>
<tr>
<td>H. sapiens</td>
<td>2</td>
</tr>
<tr>
<td>Mammalian hippocampal CA1 neuron</td>
<td>2</td>
</tr>
<tr>
<td>B. subtilis</td>
<td>1</td>
</tr>
<tr>
<td>E. perischoechinoidea</td>
<td>1</td>
</tr>
<tr>
<td>M. musculus</td>
<td>1</td>
</tr>
<tr>
<td>Neonatal R. norvegicus myocytes</td>
<td>1</td>
</tr>
<tr>
<td>R. norvegicus</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2: Statistical approaches recommended by validation organizations

<table>
<thead>
<tr>
<th>Validation step &amp; associated statistical approach</th>
<th>OECD</th>
<th>ECVAM</th>
<th>ICCVAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification of informative endpoints</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dose-response analysis</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Assessment of reproducibility:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variability within laboratory</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Variability between laboratories</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Assess protocol components most responsible for</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>affecting reproducibility, contributing to variability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANOVA</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Selection of a set of test chemicals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Principal components analysis</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Determine how well results agree with reference data (i.e., test chemicals)</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analyses of accuracy and discrimination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ordinal data:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contingency tables</td>
<td>✓</td>
<td>✓</td>
<td>(implied)</td>
</tr>
<tr>
<td>Sensitivity and specificity of test results</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Interval data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster analysis</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Fisher’s discriminant analysis</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Miscellaneous statistics or summary measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient-of-variation</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measures of central tendency</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOEC/LOEC</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Statistical approaches preferred across science disciplines

<table>
<thead>
<tr>
<th>Discipline &amp; Study</th>
<th>Comparison against known standard</th>
<th>Compare 2 or more tests</th>
<th>Data type</th>
<th>Test method applied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture &amp; Forestry</td>
<td>Freese 1960 (38)</td>
<td>✓</td>
<td>Categorical; continuous</td>
<td>Chi square</td>
</tr>
<tr>
<td></td>
<td>Voet 1994 (39)</td>
<td>✓</td>
<td>Continuous</td>
<td>Randomization/permutation of the mean square error prediction &amp; applied t-test</td>
</tr>
<tr>
<td>Clinical medicine</td>
<td>Alonzo and Pepe 1999 (40)</td>
<td>✓</td>
<td>Categorical</td>
<td>Contingency table, applied the methods of discrepant analysis &amp; composite reference standard</td>
</tr>
<tr>
<td></td>
<td>Dudoit et al. 2002 (41)</td>
<td>✓</td>
<td>Categorical; continuous</td>
<td>Receiver operating characteristic (ROC)</td>
</tr>
<tr>
<td></td>
<td>Fahey et al. 1995 (42)</td>
<td>✓</td>
<td>Categorical; continuous</td>
<td>Receiver operating characteristic (ROC)</td>
</tr>
<tr>
<td></td>
<td>Obuchowski (43)</td>
<td>✓</td>
<td>Categorical; continuous</td>
<td>ROC</td>
</tr>
<tr>
<td></td>
<td>Zou et al. 2007 (44)</td>
<td>✓</td>
<td>Categorical; continuous</td>
<td>ROC</td>
</tr>
<tr>
<td>Ecology &amp; Weather</td>
<td>Fielding and Bell 1997 (45)</td>
<td>✓</td>
<td>Categorical; continuous</td>
<td>ROC</td>
</tr>
<tr>
<td></td>
<td>Willmont et al. 1985 (46)</td>
<td>✓</td>
<td>Continuous</td>
<td>Root mean square error with bootstrap; index of agreement</td>
</tr>
<tr>
<td>Epidemiology &amp; Statistics</td>
<td>Jesus and Delicisimo 2004 (47)</td>
<td>✓</td>
<td>Continuous</td>
<td>ROC</td>
</tr>
<tr>
<td></td>
<td>Katz and Foxman 1993 (35)</td>
<td>✓</td>
<td>Continuous</td>
<td>ROC</td>
</tr>
<tr>
<td></td>
<td>Swets 1988 (36)</td>
<td>✓</td>
<td>Categorical; continuous</td>
<td>ROC</td>
</tr>
<tr>
<td>Machine learning</td>
<td>Bradley 1997 (48)</td>
<td>✓</td>
<td>Categorical; continuous</td>
<td>ROC</td>
</tr>
<tr>
<td></td>
<td>Fawcett 2006 (49)</td>
<td>✓</td>
<td>Categorical; continuous</td>
<td>ROC</td>
</tr>
<tr>
<td>Psychology</td>
<td>Williams et al. 2005 (50)</td>
<td>✓</td>
<td>Categorical</td>
<td>Contingency table</td>
</tr>
</tbody>
</table>
Chapter III: EPA’s methodology to inform TSCA PMN decision-making: an analysis and critique

A version of this chapter will be submitted for review to the Journal of Toxicology and Environmental Health, date TBD.

Abstract

Background

Regulatory toxicology has emerged as a necessary discipline to ensure that the chemicals introduced into commerce are safe. A number of laws guide regulatory toxicology, including the Toxic Substances Control Act (TSCA). One aspect of TSCA is the New Chemicals Program, administered by US Environmental Protection Agency (EPA). This program evaluates notifications of new chemicals or significant new uses of some existing chemicals. Submitters are not required to include human health effects data or environmental effects data or to undertake any testing. These restrictions place severe constraints on the EPA with respect to the basic data that may be used to inform decisions, and the EPA has only 90 days to make decisions.

We set to inquire as to the methodology used by the TSCA New Chemicals Program, including the data generated or used, to inform regulatory decisions; specifically, decisions resulting in regulation of new chemicals via consent order.

Results

Consent order regulatory decisions are often made with little to no use of new in vivo tests and rely extensively on existing test data, using analog analyses, or generation of data via use of in silico methods. There were multiple human (3 or more) and
environmental health concerns associated with decisions to regulate. Submissions did include testing data about one half of the time. Tests submitted were not aligned to the health concerns identified by the EPA. It appears that better alignment of tests to health concerns may be warranted.

**Discussion**

The legally required structure of EPA’s New Chemicals Program necessitates the use of a methodology and form of data generation that can be applied quickly. The EPA is statutorily limited in requiring testing. A tiered testing strategy may be a testing approach to consider for TSCA new chemicals evaluation, as such an approach would better align health concerns to tests carried out. Adoption of such a model would require assistance from Congress, as the EPA is at present statutorily constrained.

**Introduction**

Regulatory toxicology has emerged as a necessary discipline to ensure chemicals introduced into commerce are safe. Contemporary history has taught us the hard way that while many chemicals are not harmful to the health of humans and the environment, there too are many chemicals that were found after their introduction into commerce to cause harm. For instance, perfluoroalkylated substances are a group of fluorinated compounds with wide application, ranging from use in household products to outdoor clothing. Perfluorooctane sulfonate (PFOS) and perfluorooctanoic acid (PFOA) belong to this chemical class and have been in production since the 1940’s. It has been
discovered over time that increased exposure to PFOS is associated with bladder cancer, and PFOA is associated with kidney and testicular cancers (1; 2).

Regulatory toxicology is guided by a number of laws. For example, the Toxic Substances Control Act (TSCA) is a US law designed to regulate chemicals that could pose an unreasonable risk of harm to health or the environment that are not covered by other laws (3). Laws like TSCA have been enacted to offset imbalances and reduce risks to human health and the environment while still allowing chemicals into commerce. TSCA further sets a decision-making framework for the regulatory body and gives authorities - in this case the US Environmental Protection Agency (EPA) – to control a chemical’s entry into commerce and under what conditions production, including importation, may occur.

A necessary interrelationship exists between chemical safety laws and the science of toxicology. The decisions resulting from administering chemical safety laws like TSCA depend on the science of toxicology, of which an integral component is toxicity testing. Tests are carried out on chemicals, ideally before they enter commerce, to increase the certainty that when used, they will not cause harm to the health of humans or the environment. These tests have traditionally been in vivo, or animal-based tests. However, alternative testing methods also exist. For instance, in silico methods encompass a wide range of computationally-based approaches. Estimation of physiochemical properties, structure activity relationship (SAR) analysis and quantitative structure activity relationships (QSAR) analysis are all considered to be specific types of
in silico methods. Libraries of previously carried out in vivo and in vitro toxicity tests also exist, both in the public, peer reviewed literature and also in confidential form. These electronic libraries can be used to scientifically assess the toxicity of structurally similar chemicals. The TSCA New Chemicals Program has been applying these methods to inform regulatory decisions for over the past 30 years.

TSCA was passed into law in 1976 and implemented in 1979. What precipitated the law was the finding by Congress that there are some chemicals whose manufacture, processing, use in commerce, or disposal may present an unreasonable risk of injury to health or the environment. Policy was set out to develop data on chemicals and to regulate those chemicals posing unreasonable risks. The law further gives authority to the EPA to take action on those chemicals and mixtures identified as imminent hazards. TSCA was intended to be carried out in such a manner that it would not create unnecessary economic barriers to technological innovation. Therefore, the EPA was required under the statute to consider not only human health and environmental impacts, but also economic and social impacts of regulation as well.

The EPA implemented TSCA by creating two programs; the existing chemicals program and the New Chemicals Program. The existing chemicals program constitutes all chemicals in commerce within three years (35 months to be exact) of EPA enacting TSCA. The existing chemicals program also includes chemical substances that have fulfilled the evaluation requirements of the New Chemicals Program. There are many components of the existing chemicals program, including testing, reporting and record-
keeping of chemicals. The existing chemicals program is outside the scope of this study and so is not discussed further within this paper.

We focus this study on the EPA’s authority under TSCA to regulate new chemicals, also called a chemical substance, of which the law and associated regulations apply unique definitions. A chemical substance, under TSCA, is “any organic or inorganic substance of a particular molecular identity, including any combination of such substances occurring in whole or in part as a result of a chemical reaction or occurring in nature, and any chemical element or uncombined radical (4).” This definition does not include pesticides, tobacco, nuclear material, firearms, food, food additives, drugs, or cosmetics. A new chemical substance “is any chemical substance that is not currently listed on the Inventory.” The inventory is “the list of chemical substances manufactured or processed in the United States that EPA compiled and keeps current under section 8(b) of the Act.”

When TSCA was enacted, chemical manufacturers, distributors and importers were required to report to EPA all chemicals already in commerce as of the date TSCA was enacted. These chemicals are regulated as previously discussed under the existing chemicals program. The New Chemicals Program, in contrast, established a regulatory regime to evaluate and establish regulations for chemicals to be introduced into commerce after TSCA’s enactment. This program also sets regulation for significant new uses of some existing chemicals.

Before a new chemical or a significant new use of an existing chemical may enter commerce, a petitioner must submit a premanufacture notification (PMN) to the EPA (4).
The petitioner, “insofar as known to the person submitting the notice or insofar as reasonably ascertainable (3)” may provide a limited set of data that the EPA then uses as a foundation to evaluate whether or not the candidate chemical, also called a However, any data generated prior to the PMN petition must be submitted. PMN substance, may pose an unreasonable risk to health or the environment. Petitioners are not required to submit human health effects or environmental effects data (3). PMN submitters are also not required to undertake any testing. Further, if the PMN submitter volunteers a test result, they are not required to comply with any testing standards. The required data are limited to the chemical’s identity (name and CAS) and molecular structure (4). These restrictions place severe constraints on the EPA with respect to the basic data that may be used to inform decisions. As a matter of fact, the EPA states that for the most part it is unable to reach decisions based on the submitted data alone (5). Further, the EPA has only 90 days to carry out this decision process, making in vivo-based toxicity tests impractical. The EPA has, as a result, developed a series of methods and strategies to inform decisions. These approaches can be applied quickly to screen chemicals for potential, unreasonable risk (6). The results from these analyses are then used to inform regulatory decisions.

If, for example, a petitioner does not submit physical properties data, information which is widely known to be very inexpensive and quick for submitters to produce (7), then the EPA must make estimates. This can include physiochemical data as basic as a chemical’s melting point, vapor pressure or octanol-water partition coefficient (5). This
is in sharp contrast to EU reporting requirements, which have always called for submission of physical properties data (8; 9).

We therefore set to inquire as to the methodology used by the TSCA New Chemicals Program, including the data generated, to inform regulatory decisions. We specifically characterized the decisions that resulted in regulation of new chemicals via consent order. A consent order is a legal tool used by the EPA. It is a legally binding agreement that the PMN submitter enters into with the EPA. These orders specify regulatory parameters, for which the PMN submitter agrees to as part of the agreement to manufacture, distribute, import or dispose their chemical substance.

We find that the TSCA New Chemicals Program decision-making process largely utilizes a combination of analog analyses and data generated using in silico methods to inform their decisions to regulate via consent order. These methodologies of data generation and analysis resulted in decisions to regulate via consent order most commonly where the chemical substance may pose a risk to both human and environmental health.

Methods & Data

Our analysis of TSCA decisions is limited to chemicals that underwent the regulatory regime of EPA’s New Chemicals Program. This study focused on those chemicals that underwent PMN review and were permitted to enter commerce and were still in commerce as of January 2014, regulated by consent order. This excludes chemicals that qualify for low volume exemptions, low release exemptions, test market exemptions and
those chemicals used for research and development purposes. Also excluded are those chemicals that are not listed on the EPA 8(b) inventory, because their identities and supporting information are registered as confidential business information (CBI) by the PMN submitter, per section 14 of TSCA.

The initial list of chemicals meeting inclusion criteria were identified by the TSCA Chemical Substance Inventory (figure 1). The EPA is required under section 8(b) of TSCA to make public all non-CBI chemicals that are actively used in US commerce. The data set was accessed February 4, 2014 and so includes all chemicals in commerce as of January 3, 2014. The EPA has 30 days to update the inventory based on notices of commencement of manufacture or import received from petitioners, post PMN review. The inventory specifies the generic name of the chemical, PMN number and whether or not the chemical is subject to a consent order.

For this study, we were interested in the methods and data that informed and resulted in the regulatory decision. This information is not available in the 8(b) inventory. We therefore undertook to collect the documents supporting the consent order decisions and abstracted these documents for analysis. Briefly, the EPA is required under TSCA to communicate their PMN regulatory decisions publicly via the Federal Register. The Federal Register is a daily journal, established under the Federal Register Act of 1935 to communicate to the public the federal government’s activities (10). The Federal Register is available in electronic form from 1995 to the present; however, the electronic version can include decisions prior to this date. We carried out a search of all PMN-based
decisions of TSCA 8(b) chemicals resulting in a consent order as recorded by the electronic form of the Federal Register and recorded this data for inclusion into the study.

We were specifically interested in the basis of the PMN consent order decisions and in particular the methods and data used to inform the regulatory decisions. The basis for the decision to regulate is communicated in three documents: the Federal Register notice, the Structure Activity Team (SAT) report and the consent order itself (figure 1). The SAT report is a technical, EPA document that summarizes EPA’s PMN health and environmental health analyses. The consent order is the legally binding agreement between the EPA and the PMN submitter. The EPA electronically posts the SAT reports and consent orders, and they are available electronically via Regulations.gov. The Regulations.gov exchange - mandated by Executive Order of President Clinton (11), and implemented under President Bush - is the central federal repository of the US Federal Government where one may find and comment on proposed regulations, along with supporting documents. This central repository has been in existence since 2003, but contains information prior to this date. We collected this information, abstracted it and recorded it for inclusion into this study (figure 1). In those instances where the PMN submitter also provided test data, we abstracted and recorded this information for inclusion into the study also. These data were provided in the PMN submissions, available via Regulations.gov, and are often detailed in the SAT reports, or the consent order.

**Review of EPA New Chemical Review Methodologies**
This analysis focused on EPA’s approaches for new chemical review. The EPA uses a number of approaches to review and evaluate PMN submissions, briefly summarized here (6; 12). The New Chemicals Program can take basic information provided by the petitioner and estimate physiochemical properties of a given chemical (13). The tools used to estimate physiochemical properties have been found to produce accurate estimates (7). This capability is in addition to the established, manual protocol for identifying measured physiochemical properties and analogs of PMN substances (5). This physiochemical information can then be used to conduct SAR analyses. Analyses by SAR relate a chemical’s structure or substructure to the presence or absence of a property or activity of interest (14; 15). SARs are qualitative relationships, often in the form of structural alerts that incorporate molecular substructures or fragments related to the presence or absence of activity (16). EPA uses SAR to predict the toxicity to aquatic organisms (17). The Science Advisory Board (SAB) is the legally mandated, independent scientific advisor to the EPA, and as such is a critical point of peer review of EPA’s in silico models (18). The SAB has reviewed the New Chemicals Program in silico methods (19). These in silico methods can therefore be said to be validated on the basis of their passage through review and usage over time.

The EPA also relies upon libraries of publicly available information on chemicals that have previously undergone in vivo or in vitro tests. Included with this information is human epidemiological/occupational health data in limited instances. The EPA further retains confidential libraries of chemicals that have undergone testing, but which are not available to the public, because the tested chemicals are listed as CBI. What EPA does is
systematically identify an analog or analogs of a given PMN substance, as summarized previously. Available test results data, *in vivo* and *in vitro*, of these analogs are then retrieved from the libraries. Information is then reviewed and evaluated in a group setting of experts to infer the health and environmental effects that could potentially arise from exposure to the PMN substance. For this study, we refer to this approach as ‘analog analysis.’

The PMN decisions follow a standardized process where a determination of human and/or environmental health risk is made. Either of these determinations can be made based on any number of concerns, or toxicity axes (e.g., risk of neurotoxicity, risk of aquatic toxicity, risk of carcinogenicity, etc). The trail of logic is derived from the documents previously mentioned (figure 1). We reviewed these documents and abstracted the decisions using the protocol summarized in figure 2.

**Results**

As of January 2014, there were 67,253 chemicals listed in the TSCA 8(b) inventory. A total of 266 of the 8(b) chemicals are subject to regulation by consent order. Of the 266 chemicals, 33 chemicals met inclusion criteria for this study (figure 3 and table 1). The PMN substances that met inclusion criteria were decided upon by the EPA from 1999 to 2012.

**Information provided to EPA:**
The PMN petitioners are not required to generate and provide data on their chemical’s physiochemical properties to the EPA. Nor are the PMN petitioners required to generate tests and results on health and environmental effects. However, such information does accompany PMN submissions. Of the 33 submissions analyzed, 23 included physiochemical properties data. Specific analyses of these data could not be carried out, because all PMN submitters redacted this information in accordance with CBI claims. Just over one half, or 17, of the PMN submissions included test data on human and/or environmental effects; 16 submissions did not provide any test data or results. Of the 17 submissions that included human or environmental effects test results, only 5 provided original study reports. The PMN submissions that included original test reports were submitted by organizations originating from Germany, Japan or Switzerland. All other test results were either summaries provided by the PMN submitter (Number = 9) or summaries developed by the EPA (Number = 3).

Submissions included tests on actual PMN substances or analog substances. A total of 142 test results were reported on or summarized across 30 distinct test types (table 2). The maximum number of test results provided with a submission was 15; the minimum, 1. No discernible pattern existed as to the tests chosen by PMN submitters. Only 39 percent of the 142 test results stated their analyses complied with any recognized test guidelines or standards. Specific test protocols could not be compared across PMN submissions, because so few PMN submissions available to the public included test reports.
Information generated by EPA:

The PMN submitters are required to submit the molecular formula, molecular diagram and molecular weight of their PMN substances to the EPA. They are also required to identify the synthesis of the PMN substance, impurities present, estimated production volume and categorize one or more potential end uses. All of this information, including company name, can be claimed confidential. The EPA takes this information and generates estimates of the PMN substance’s physiochemical properties. This information is generated via a combination of analytical tools and manual searches (5). These data can be used, in part, to estimate the likely human and environmental health impacts of the PMN substance. This information is also used to estimate environmental fate. Solubility, boiling point, Henry’s Law Constant and vapor pressure estimates were summarized across the EPA’s SAT reports evaluated in this study. The EPA replaced predicted values with actual values in the instances where the PMN submitters provided actual data. Actual measurements are redacted from the PMN submissions on account of CBI claims by the PMN submitters. Other physiochemical properties estimates are generally not provided within the SAT reports and so could not be analyzed.

For some PMN substances, the EPA generated estimates of aquatic toxicity values (table 3). These SAR analyses were carried out using EcoSAR. The EPA replaced predicted values with actual values in the instances where the PMN submitters provided actual test results. The predicted values are generally presented within the SAT reports. Actual measurements are redacted for CBI reasons as are the names of analog chemicals or other physiochemical estimates used to arrive at the EcoSAR predictions.
Finally, the EPA compiled known human health effects information on each PMN substance using the analog analysis approach described in the methods section. Analog analyses were carried out for every chemical included in this study. The EPA summarized the results of the analog analyses in the SAT reports, consent orders and Federal Register notices. Analyses were based on analogs of the PMN substances where \textit{in vivo} information was available. The summaries described the specific health concerns; however, the identity of the PMN substance was redacted. Also redacted was the identity of the analog substance. Every analysis referenced \textit{in vivo} studies of analog substances/classes or to reports on chemical classes thereof.

\textbf{Decisions:}

For this analysis, 33 consent orders issued by the EPA were analyzed (table 4). A total of three consent orders were issued on the basis that the substance may present an unreasonable risk to human health. Five consent orders were issued on the basis that the substance may present an unreasonable risk to the environment. In 25 instances, or 76 percent of the time, the EPA issued consent orders on the basis that the substance may present an unreasonable risk both to human health and the environment. Of the 33 consent orders issued, a total of 31 Significant New Use Rules (SNUR) was also issued. There are two types of SNURs that the EPA may issue. In the instance relevant to this study, the EPA may issue a SNUR to extend the consent order regulations to other potential manufacturers, distributors, importers or disposers of the subject chemical substance.
Information used to inform decisions:

The majority of decisions were based on analog analyses in the determination of human health risk; for environmental health risk, SAR (table 4).

There were nine instances where ecological toxicity data was provided by the petitioner. These data replaced the EcoSAR estimates and served to inform the decisions to regulate based on unreasonable risk to the environment. The EPA published aquatic toxicity values in 15 of the 33 regulated PMN substances (table 3). A total of eight of 15 chemicals were also determined to be persistent or to bioaccumulate by the EPA. Additionally, 8 PMN substances were identified as persistent or to bioaccumulate. There were no EcoSAR data published in these 8 instances. All 8 instances were either perfluorooctane sulfonates or perfluorooctanoic acids. The remaining 10 PMN substances that were regulated in part due to health concerns to the environment and did not have EcoSAR data summarized and were also not identified as persistent or to bioaccumulate. In these 10 cases, it was unclear how EPA made an unreasonable environmental risk determination.

In the 12 instances of determining unreasonable health risk, submitted test data primarily informed the decision. In all other decisions of human health risk, analog analyses were the primary information source in making the regulatory decision. There were a median number of four human health concerns per PMN submission. The maximum number of health concerns for an individual PMN substance was 12; the minimum, three. Risk of
pulmonary toxicity recurred most often, followed by reproductive and developmental toxicity (table 5).

**Discussion**

The legally required structure of EPA’s New Chemicals Program, specifically its PMN process, necessitates the use of methodologies and data generation that can be applied quickly. The EPA has established such approaches, consistent with their need to make decisions on new chemicals and significant new uses of existing chemicals within 90 days. The majority of the regulatory decisions evaluated in this study were based on analog analyses in the determination of human health risk; for environmental risk, SAR. The majority of consent order decisions reviewed in this analysis concluded that the PMN substance may pose an unreasonable health and environmental health risk. There were no fewer than three health concerns associated with the regulated, PMN substance. This is the first study to evaluate the methods and strategies used to inform TSCA’s New Chemicals Program decisions to regulate via consent order.

It is revealing that the 8(b) inventory lists only 266 chemicals that are subject to consent order. However, in EPA’s Summary of Accomplishments dated September 30, 2010, a total of 1,492 chemicals are stated to be subject to consent order (20). This means that a mere 18 percent or less of all the chemicals of concern, by consent order, that are also actively used in commerce is known to the public via the 8(b) inventory. Over 80 percent of regulated chemicals under consent order are not made known to the public via the 8(b) inventory. This means that the majority of chemicals identified as a human or
environmental health concern are not readily known to the public under 8(b) inventory. Because these chemicals are on the CBI portion of the 8(b) inventory, the public has no way to know what chemicals are under a TSCA New Chemicals consent order.

A weakness of this study is that only 33 consent order decisions were available to be evaluated. The reason for this is twofold. First, most PMN submitters classify their submissions as confidential business information, which could include any of the following pieces of information: the substance itself, the company name, where it is made, what the production volume is, what the intended uses are, etc. As a result, the majority of the chemicals regulated by consent order are not known publicly via the 8(b) inventory. Second, the documents supporting consent order decisions has become available electronically only in the past decade. These two issues could bring into question the representativeness of this study’s findings. Personal conversation with an official at the EPA indicated that analog and SAR analyses are commonly used to evaluate PMN substances (personal communication, Dr. Louis Scarano, 28 February 2014). The consent orders evaluated in this study were issued from 1999 to 2012 and so results likely represent current patterns of analyses and decision making.

The results of this study further demonstrate the interdependence between law and science. Decision makers must render decisions and regulate based on the available scientific and analytical capabilities at the time. Further, regulatory bodies must render decisions within specific, legal constraints and timelines. The TSCA New Chemicals
Program is particularly challenged, because it must work with minimal data. PMN submitters, under the PMN submission process are not legally required to generate and submit health and/or environmental effects data. Only half of the submissions evaluated in this study provided test data and further in the rare instances where they did conduct *in vivo* studies and provide results, half of the study results did not reference a study protocol. Under TSCA, the EPA is statutorily limited in requiring testing, much less enforcement of standards on test data submitted. Further, the tests carried out by PMN submitters are generally not consistent with the toxicity concerns the EPA identifies with the PMN substances. For instance, the three most cited human health concerns across the PMN substances evaluated in this study were pulmonary toxicity, reproductive toxicity and developmental toxicity. In contrast, the PMN submissions carried out only one pulmonary toxicity study. There were only three reproductive toxicity studies and only five developmental toxicity studies carried out. Within the data analyzed here, the health concerns identified by the EPA and tests carried out by the PMN submitters are not aligned. It appears that better alignment of tests to areas of health concern may be warranted.

A tiered testing strategy may be a testing approach to consider for TSCA new chemicals evaluation. Such an approach would better align health concerns with tests carried out. Conceptually, a tiered testing strategy begins by assembling all relevant and existing data on the chemical, or similar analog, and reviewing the literature (21). An initial screening or battery of simple tests would be carried out. Results of this initial review of literature and battery of tests would inform important toxicological axes to evaluate further. More
targeted, or selective, tests would then be carried out on the axes of high priority (22). The manner in which a tiered testing strategy may be implemented can vary; however, the general approach is designed such that each test in the sequence of tests is selected so as to complement the preceding tests, where tests chosen in each succeeding level are determined by the results in the previous level of testing (23). In effect, the choice of tests are designed to generate more information only in those areas where such information is needed, resulting in more informed regulatory decisions.

This study revealed that of the PMN submissions regulated, potential human health concerns differed across PMN substances (table 5). It is impractical to require in vivo tests across all the potential toxicity axes. Alternatively, applying a tiered testing approach is practical. It is practical for PMN submitters to review the existing scientific literature on chemical analogs, using sound methodology and further to complete an in vitro assay battery or in silico screening. The PMN submitter would review and screen their candidate, PMN substance prior to submission to EPA, consistent with good research practices and testing standards. They would then submit this information as a component of their PMN submission to the EPA. The submission would be required to identify possible toxicological axes of concern. A time-defined process could then proceed where EPA has the opportunity to review the submission and carry out a similar analysis on the chemical using appropriate methodologies. The EPA has already developed a foundation that identifies axes of concern, including analog analyses and in silico methods. This can be expanded on further by validation of alternative methods, like in silico or in vitro. The EPA would make their conclusions and compare their
conclusions to that of the PMN submitter. If more targeted tests are deemed warranted, then the PMN submitter would then be required to conduct these selective tests over a time defined period. EPA would be required to review these final results in a defined time frame and prescribe a regulatory framework appropriate to the chemical prior to the chemical entering commerce.

While the details of such an evaluation methodology would certainly need to be worked out, such an approach would better align testing to toxicity axes of concern and in parallel, still allow the quick progression of chemicals through regulatory review in a systematic and consistent manner. This tiered approach would also promote efficiencies by requiring only tests needed to inform decisions where there could be a specific safety concern. This publicly documented approach would further increase the transparency of PMN submitters, reducing their litigation risk and further, increase the comfort of consumers regarding the safety of products built on chemicals developed. Adoption of such a model would require assistance from Congress, as the EPA is at present statutorily constrained.

This study further demonstrates that laws can and do promote innovation and efficiencies. When TSCA was enacted, no validated, in silico or formalized analog analysis strategies existed to evaluate the toxicity of chemicals. TSCA forced the creation of alternative methods, in this case, the development of a combination of in silico methods and libraries of chemical toxicity tests to be used in analog analyses.
Laws that promote the refinement, reduction or replacement of *in vivo* methods are generally recognized to be more cost effective and efficient than equivalent *in vivo* methods. TSCA is not the only law to promote such innovation. An ad hoc group of experts on behalf of the European Commission in 2003 reviewed the state of the art of safety assessment by animal and non-animal tests for 11 different human health effects of concern in the frame of the seventh amendment of the Cosmetics Directive 76/768/EEC, aiming at establishing time tables for the implementation of marketing and testing bans including deadlines for the phasing out of the various animal tests, which has been carried out (24). Such a strategy is consistent with the National Academies vision of toxicity testing in the 21st century to use alternative methods more extensively, with much reduced need for whole animal testing (25).

In conclusion, this first study to evaluate the methodology used to inform TSCA’s New Chemicals Program decisions to regulate via consent order found the majority of decisions are made by a combination of analog analyses and *in silico* methods. The majority of decisions to regulate by consent order were instances where chemicals were reasoned to pose a risk to both human and environmental health.
References


<table>
<thead>
<tr>
<th>Number</th>
<th>Premanufacturing Number</th>
<th>Generic Name (as given in Federal Register)</th>
<th>CAS #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P-00-0346</td>
<td>Dialkyl hydroxybenzenealkanoic acid ester</td>
<td>CBI</td>
</tr>
<tr>
<td>2</td>
<td>P-00-0368</td>
<td>Benzylsulfonamide, alkylphenylsubstitutedphenyl substituted carbonyl.                                                                ----------------------------------------------------------------------------------------------------------</td>
<td>CBI</td>
</tr>
<tr>
<td>3</td>
<td>P-01-0470</td>
<td>Ethoxylated alkylphenol sulfate, ammonium salt</td>
<td>CBI</td>
</tr>
<tr>
<td>4</td>
<td>P-01-0595</td>
<td>Tertiary amine salt of glycol succinate</td>
<td>CBI</td>
</tr>
<tr>
<td>5</td>
<td>P-01-0918</td>
<td>Isocyanate compound, modified with methoxyisilane</td>
<td>CBI</td>
</tr>
<tr>
<td>6</td>
<td>P-02-0382</td>
<td>Alkylbenzene sulfonate</td>
<td>CBI</td>
</tr>
<tr>
<td>7</td>
<td>P-04-0640</td>
<td>Diisocyanate terminated polycarboxamide</td>
<td>CBI</td>
</tr>
<tr>
<td>8</td>
<td>P-04-0834</td>
<td>HDI biuret, hydroxyethyl methacrylate prepolymer</td>
<td>CBI</td>
</tr>
<tr>
<td>9</td>
<td>P-06-0153</td>
<td>Iso-tridecanol</td>
<td>CBI</td>
</tr>
<tr>
<td>10</td>
<td>P-06-0816</td>
<td>Modified reaction products of alkyl alcohol, halogenated alkane, substituted epoxyde, and amino compound</td>
<td>CBI</td>
</tr>
<tr>
<td>11</td>
<td>P-07-0087</td>
<td>Partially fluorinated condensation polymer</td>
<td>CBI</td>
</tr>
<tr>
<td>12</td>
<td>P-07-0328</td>
<td>Phenol, 4-{(alkyl)[4-(phenylsubstituted)phenyl substituted]methyl]-2,6-bis(1,1-dimethylthyl)</td>
<td>CBI</td>
</tr>
<tr>
<td>13</td>
<td>P-07-0537</td>
<td>Alkane, bis(cyanoalkyl)amino</td>
<td>CBI</td>
</tr>
<tr>
<td>14</td>
<td>P-09-0048</td>
<td>Alky carbonyl polyester acetate reaction products with mixed metal oxides</td>
<td>CBI</td>
</tr>
<tr>
<td>15</td>
<td>P-10-0060</td>
<td>Partially fluorinated alcohol substituted glycols</td>
<td>CBI</td>
</tr>
<tr>
<td>16</td>
<td>P-10-0367</td>
<td>Carbon black derived from the pyrolysis of rubber tire shreds</td>
<td>CBI</td>
</tr>
<tr>
<td>17</td>
<td>P-10-0470</td>
<td>Dimethyl siloxynapfluoro methyl siloxynap(oxyalkeynediyl) methyl silox copolymer</td>
<td>CBI</td>
</tr>
<tr>
<td>18</td>
<td>P-10-0471</td>
<td>Alkyl acrylatepolyfluoro methacrylatepoly(oxyalkeynediy)-methacrylates</td>
<td>CBI</td>
</tr>
<tr>
<td>19</td>
<td>P-10-0472</td>
<td>Alkyl acrylatepolyfluoro methacrylatepoly(oxyalkeynediyl)-methacrylates</td>
<td>CBI</td>
</tr>
<tr>
<td>20</td>
<td>P-10-0485</td>
<td>Alkethyl methacrylates, polymer with substituted carbonmonocycle, hydroxyethyl acrylamide and fluorinatedalkyl acrylate</td>
<td>CBI</td>
</tr>
<tr>
<td>21</td>
<td>P-10-0546</td>
<td>Modified lithium iron phosphates</td>
<td>CBI</td>
</tr>
<tr>
<td>22</td>
<td>P-11-0048</td>
<td>Diethylene glycol, polymer with disocyanatoamine, polyethylene glycol monomethyl ether- and fluorinatedalkanol -blocked</td>
<td>CBI</td>
</tr>
<tr>
<td>23</td>
<td>P-11-0063</td>
<td>Perfluoroalkyl acrylate copolymer</td>
<td>CBI</td>
</tr>
<tr>
<td>24</td>
<td>P-11-0203</td>
<td>Perfluoroalkylethylmethacrylate copolymer with dialkylaminoethylmethacrylate</td>
<td>CBI</td>
</tr>
<tr>
<td>25</td>
<td>P-11-0247</td>
<td>Perfluoroalkylethyl methacrylate copolymer with hydroxymethyl acrylamide, vinyl chloride and long chain fatty alkyl acrylate</td>
<td>CBI</td>
</tr>
<tr>
<td>26</td>
<td>P-11-0264</td>
<td>Brominated polyphenyl ether</td>
<td>CBI</td>
</tr>
<tr>
<td>27</td>
<td>P-11-0384</td>
<td>Fluorinated alkylsulfonamidol urethane polymer</td>
<td>CBI</td>
</tr>
<tr>
<td>28</td>
<td>P-11-0557</td>
<td>2-Propenoic acid, 2-methyl-, 2-hydroxyethyl ester, telomers with C18-26-alkyl acrylate, 1-dodecanethiol, N-(hydroxymethyl)-2-methyl-2-propenamide, polyfluoroocetyl methacrylate and vinylidene chloride, 2,2'-(1,2-diazenediybis(1-methylethylidene)bis[4,5-dihydro-1Himidazole] hydrochloride (1:2)-initiated</td>
<td>CBI</td>
</tr>
<tr>
<td>29</td>
<td>P-11-0561</td>
<td>Tetrafluoroethylene chlorotrifluoroethylene copolymer</td>
<td>CBI</td>
</tr>
<tr>
<td>30</td>
<td>P-11-0646</td>
<td>Perfluoroalkylethyl methacrylate copolymer</td>
<td>CBI</td>
</tr>
<tr>
<td>31</td>
<td>P-11-0653</td>
<td>Perfluoroalkylethyl methacrylate copolymer</td>
<td>CBI</td>
</tr>
<tr>
<td>32</td>
<td>P-12-0031</td>
<td>Modified fluorinated acrylates</td>
<td>CBI</td>
</tr>
<tr>
<td>33</td>
<td>P-99-0848</td>
<td>Silicic acid (H6SiO2O7), magnesium strontium salt(1:1:2), dysprosium and europium-doped.</td>
<td>CBI</td>
</tr>
</tbody>
</table>
Table 2: Test data provided

<table>
<thead>
<tr>
<th>Test</th>
<th>No.</th>
<th>Standard (no. tested using std/no. tested not using std)</th>
<th>Test</th>
<th>No.</th>
<th>Standard (no. tested using std/no. tested not using std)</th>
<th>Test</th>
<th>No.</th>
<th>Standard (no. tested using std/no. tested not using std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ames</td>
<td>13</td>
<td>OECD 471; Japanese MOL (6/7)</td>
<td>Mouse Local Lymph Node Assay</td>
<td>4</td>
<td>None Stated (0/4)</td>
<td>Algal Inhibition (72hr)</td>
<td>7</td>
<td>OECD 201; EPA-6009-78-018 (3/4)</td>
</tr>
<tr>
<td>E. Coli</td>
<td>13</td>
<td>OECD 471/2; Japanese MOL (6/7)</td>
<td>Rat LD&lt;sub&gt;50&lt;/sub&gt;</td>
<td>4</td>
<td>OPPTS 870.1100 (1/3)</td>
<td>Daphnia Magna Acute Toxicity</td>
<td>7</td>
<td>OECD 202 (2/5)</td>
</tr>
<tr>
<td>Rat Acute Oral Toxicity</td>
<td>11</td>
<td>OECD 423; EEC Guideline 92/69 (3/8)</td>
<td>Rat Reproduction</td>
<td>3</td>
<td>OPPTS 870.3100 (1/2)</td>
<td>Fish Acute Toxicity (96hr)</td>
<td>7</td>
<td>OECD 203 (3/4)</td>
</tr>
<tr>
<td>Rabbit Acute Dermal</td>
<td>10</td>
<td>OECD 404 (4/6)</td>
<td>Pharmacokinetic Monkey</td>
<td>2</td>
<td>None Stated (0/2)</td>
<td>Biodegradation (microorganisms)</td>
<td>5</td>
<td>OECD 301b (2/3)</td>
</tr>
<tr>
<td>Rabbit Acute Eye Irritation</td>
<td>9</td>
<td>OECD 405 (4/5)</td>
<td>14d Rat Repeat Dose Oral Toxicity</td>
<td>2</td>
<td>None Stated (0/2)</td>
<td>Bioaccumulation (carp)</td>
<td>3</td>
<td>OECD 210 (1/2)</td>
</tr>
<tr>
<td>28d Rat Oral Toxicity (systemic)</td>
<td>6</td>
<td>OECD 407 (3/3)</td>
<td>10d Rat Oral Gavage</td>
<td>1</td>
<td>None Stated (0/1)</td>
<td>Avian Reproduction</td>
<td>2</td>
<td>None Stated (0/2)</td>
</tr>
<tr>
<td>Chromosomal Aberration (in vitro mammalian cells)</td>
<td>6</td>
<td>OECD 473; Japanese MOL (3/3)</td>
<td>14d Rat Repeat Dose Rats &amp; Mice</td>
<td>1</td>
<td>None Stated (0/1)</td>
<td>Activated Sludge Respiration inhibition</td>
<td>1</td>
<td>OECD 209 (1/0)</td>
</tr>
<tr>
<td>Guinea Pig Skin Sensitization</td>
<td>5</td>
<td>OECD 406 (2/3)</td>
<td>90d Rat Neurotoxicity</td>
<td>1</td>
<td>OECD 424 (1/0)</td>
<td>Daphnia Magna Reproduction</td>
<td>1</td>
<td>OECD 211 (1/0)</td>
</tr>
<tr>
<td>90d Rat Oral Toxicity</td>
<td>5</td>
<td>None Stated (0/5)</td>
<td>Human Occupational/ Epidemiology</td>
<td>1</td>
<td>None Stated (0/1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rat Acute Dermal</td>
<td>5</td>
<td>OECD 402 (3/2)</td>
<td>Mouse Micronucleus Test</td>
<td>1</td>
<td>OECD 474 (1/0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rat Developmental Toxicity</td>
<td>5</td>
<td>OECD 414 &amp; 422 (4/1)</td>
<td>Rat Inhalation</td>
<td>1</td>
<td>None Stated (0/1)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: EcoTox data generated using EcoSAR

<table>
<thead>
<tr>
<th>Test Organism</th>
<th>Test Duration (hr)</th>
<th>Endpoint</th>
<th>Median (mg/L) across regulated substances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish</td>
<td>96</td>
<td>LC₅₀</td>
<td>20</td>
</tr>
<tr>
<td>Daphnid</td>
<td>48</td>
<td>LC₅₀</td>
<td>4</td>
</tr>
<tr>
<td>Green Algal</td>
<td>96</td>
<td>EC₅₀</td>
<td>6.95</td>
</tr>
<tr>
<td>Daphnid</td>
<td>N/A</td>
<td>Chronic value</td>
<td>3.1</td>
</tr>
<tr>
<td>Algal</td>
<td>N/A</td>
<td>Chronic value</td>
<td>1.1</td>
</tr>
</tbody>
</table>
Table 4: Decisions and basis for decisions

<table>
<thead>
<tr>
<th>Basis for Consent Order Decision (unreasonable risk of injury to health or the environment)</th>
<th>Number</th>
<th>&gt;1 Human Health Concern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human health</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Environmental health</td>
<td>5</td>
<td>N/A</td>
</tr>
<tr>
<td>Both human &amp; environmental health</td>
<td>25</td>
<td>22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of information/analysis that informed the decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human health risk</td>
</tr>
<tr>
<td>Analog search</td>
</tr>
<tr>
<td>In vivo or in vitro test data submitted by petitioner</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Environmental health risk</td>
</tr>
<tr>
<td>Structure Activity Relationship (aquatic toxicity)</td>
</tr>
</tbody>
</table>
Table 5: Human health concerns identified in EPA SAT reports

<table>
<thead>
<tr>
<th>Health concern</th>
<th>No.</th>
<th>Health concern</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulmonary toxicity</td>
<td>18</td>
<td>Respiratory sensitization</td>
<td>4</td>
</tr>
<tr>
<td>Developmental toxicity</td>
<td>17</td>
<td>Kidney toxicity</td>
<td>4</td>
</tr>
<tr>
<td>Reproductive toxicity</td>
<td>17</td>
<td>Neurotoxicity</td>
<td>2</td>
</tr>
<tr>
<td>Immuno toxicity</td>
<td>13</td>
<td>Bone marrow toxicity</td>
<td>1</td>
</tr>
<tr>
<td>Liver toxicity</td>
<td>12</td>
<td>Spleen toxicity</td>
<td>1</td>
</tr>
<tr>
<td>Blood toxicity</td>
<td>10</td>
<td>Thymus toxicity</td>
<td>1</td>
</tr>
<tr>
<td>Dermal sensitization</td>
<td>8</td>
<td>Thyroid toxicity</td>
<td>1</td>
</tr>
<tr>
<td>Eye irritation</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: Sources for data abstraction

1) Federal Register notice of regulation by consent order
   • Decision of potential risk to human and/or environmental health and/or produced in large quantities
   • Basis for decision stated

2) Structure Activity Team Report
   • Identification/estimation of physiochemical properties
   • Test data submitted with PMN (in vitro, in silico & in vivo); also available in PMN itself
   • Analyses based on structural analogs
   • Structure Activity Relationship Analyses
   • Other analyses

3) Consent Order
   • Statement of human and/or environmental health concerns and supporting information
Figure 2: Data abstraction steps

1) Basis for decision
   • Determination of human and/or environmental health risk
   • Health and/or environmental health risk concern
   • Basis for concern

2) Method(s) used to inform decision
   • Physiochemical properties estimates/measures
   • SAR (chemical or analog)
   • Analog search of existing studies
   • Submitted test data (chemical or analog)
     • Test by type (i.e., acute oral, dermal, etc)
Figure 3: Inclusion criteria

TSCA 8(b) Inventory (67,253)

TSCA 8(b) chemicals regulated by Consent Order (266)

Decisions documented in electronic, Federal Register (87)

Analyses supporting the decisions were published electronically (33)
Chapter IV: Network motifs that recur across species: gene regulatory and protein-protein interaction networks

This chapter was published in the Archives of Toxicology, May 2014. No additional permissions required.

Abstract

Background

Cellular molecules interact in complex ways, giving rise to a cell’s functional outcomes. Conscientious efforts have been made in recent years to better characterize these patterns of interactions. It has been learned that many of these interactions can be represented abstractly as a network and within a network there in many instances are network motifs. Network motifs are subgraphs that are statistically overrepresented within networks. To date, specific network motifs have been experimentally identified across various species and also within specific, intracellular networks; however, motifs that recur across species and major network types have not been systematically characterized. We reason that recurring network motifs could potentially have important implications and applications for toxicology and in particular, toxicity testing. Therefore, the goal of this study was to determine the set of intracellular, network motifs found to recur across species of both gene regulatory and protein-protein interaction networks.

Results

We report the recurrence of 13 intracellular, network motifs across species. Ten recurring motifs were found across both protein-protein interaction networks and gene regulatory networks. The significant pair motif was found to recur only in gene
regulatory networks. The diamond and one way cycle reversible step motifs were found to recur only in protein-protein interaction networks.

Conclusion

This study is the first formal review of recurring, intracellular network motifs across species. Within toxicology, combining our understanding of recurring motifs with mechanism and mode of action knowledge could result in more robust and efficient toxicity testing models. We are sure our results will support research in applying network motifs to toxicity testing.

Introduction

Cellular molecules interact in complex ways, giving rise to a cell’s functional outcomes. These interactions span an array of functions ranging from glycolysis to transcription to apoptosis. Scientists and researchers have produced volumes of information on how these interactions occur, often describing experimentally the interaction between two or more molecules and the associated mechanism of action. While the interactions between individual molecules and in some instances larger pathways have been characterized to varying degrees, this understanding is variable and often quite incomplete. Systems biology has in recent years made a conscientious effort to better characterize these broader patterns of interactions.
These complex molecular interactions can be represented abstractly as a network. Biological networks exhibit organizing principles and further this organization may be represented by the network’s topology (1). This organization is distinct from random networks and is very often scale free, which means that their degree distribution approximates a power law, $P(k) \sim k^{-\gamma}$ (2, 3).

Mathematically, a cellular network may be represented as a graph (4), which abstractly comprises nodes (vertices) and edges (arcs). Cellular molecules may be represented as nodes. Nodes may be connected by an edge – directed or undirected, representing the interactions between the molecules. If the nodes and edges are countable, then the graph is said to be finite. One may further break down a graph into subgraphs. These subgraphs are made up of a finite number of nodes and these nodes’ interactions with one another characterize a given network at the local level. Two or more graphs or subgraphs are considered isomorphic if they have the same node number and their adjacency is preserved (i.e., corresponding nodes are connected by edges the same way in both subgraphs). Subgraphs that are statistically overrepresented in a network are referred to as network motifs (5, 6). Milo et al. (5) and Shen-Orr et al. (6) were the first to use graph theory principles and the experimental method to identify motifs within the transcriptional networks of simple organisms. Since their discovery, much research has been carried out identifying and characterizing motifs at the intracellular level across many species.

vi Herein, called ‘motifs’ for brevity.
However, motifs that recur across species have not been systematically characterized to date. We reason that motifs that recur across species could potentially have important implications and applications. Recurring motifs could improve upon our existing knowledge of toxicology and in particular, could have practical applications for toxicity testing. The U.S. National Academies of Sciences (NAS) in 2007, via its report entitled, “Toxicity Testing in the 21st Century,” posited a vision for toxicity testing, asserting that the goals of toxicity testing moving forward should be to identify those critical pathways that when sufficiently perturbed lead to adverse health outcomes (7). An important foundation and component of these critical pathways, we reason, is motifs. In particular, it is important from the perspective of toxicology to identify and characterize those motifs that recur across species as such knowledge could enable more robust toxicity tests and better assist in informing cross-species extrapolation. This, we reason, sets a stronger foundation for alternative testing methods, like in vitro and in silico, consistent with the NAS report.

The goal of this study therefore is to determine the set of motifs that have been identified at the intracellular level and are found to recur across species and major types of biological networks (i.e., protein-protein interaction and gene regulation). To date, a finite number of motifs have been identified. Motif identification has been carried out extensively in simple organisms; for instance, the transcription networks in *E. coli* and *S. cerevisiae* (5, 6, 8-15). Select motifs have been summarized to varying degrees in the literature (16). However, these motifs have not been characterized via a formal review or
in a systematic manner. Herein, we report the intracellular motifs that recur across species. Further, we reaffirm that select motifs are more or less common within broad networks and across species, with specific exceptions. We finally discuss the potential implications and applicability of these findings across disciplines, including medicine and toxicology.

**Methods**

We carried out a search of the peer reviewed literature, wherein our strategy was to maximize the chances of identifying all available articles, communicating their findings in English. Peer reviewed research articles were identified using Google Scholar, PubMed and Scopus. Articles were found by the use of each search engine via keyword search, using the following search terms: ‘functional motif,’ ‘graph,’ ‘motif,’ ‘network motif,’ or ‘subgraph.’ Terms were used in their singular and plural forms. Supplementary searches were carried out using the aforementioned terms in combination with the following terms: ‘biological network,’ ‘cellular biology,’ ‘computational biology,’ ‘molecular biology,’ ‘network,’ ‘network topology,’ or ‘systems biology,’ again using both singular and plural forms.

Articles were included if they reported on research that quantitatively assessed the presence of motifs at the intracellular level, using experimental methods of non-plant organisms. In circumstances where studies were carried out using cell-lines, only those studies using non-carcinogenic cell lines were accepted for inclusion. Studies identifying intercellular motifs were excluded. Summaries and reviews were noted and used for
informational purposes, but excluded from the analysis, so as to avoid to the redundant reporting of findings.

Pertinent motif data from each study – motif structure, organism, data source, biological network and methods – were documented. A network motif is a recurring pattern within the network that recurs more often than at random (17). These comparisons are made by evaluating the subject network to reasonably equivalent, randomized networks. These can be described by a number of statistical tests. These can be described by a Z score, which is simply the difference in occurrence of a pattern in a selected network compared to a randomized network, divided by the standard deviation of occurrence in the randomized network. If multiple, randomized networks are used, then the mean frequency of the randomized networks is applied (18). The result of this calculation can alternatively be described as a p value. Statistically significant motifs from within each study were recorded, where a motif exhibited a p < 0.05, z > 2 or a normalized Z score > 0.5 compared to a randomized network or networks. Motif significance profiles are a vector of Z-scores of a set of motifs. These allow for comparison of networks of different sizes (18), as was the case in some studies evaluated. Similarly, if a motif exhibited a positive significance profile (SP), it too was accepted for inclusion. The statistically significant motifs from a given study were then compared to the statistically significant motifs identified across all other studies where statistically significant motifs were also identified. A motif was considered to recur if it was identified in at least two independent studies. The result of this analysis was a set of recurring, statistically significant motifs. These recurring motifs were also associated with multiple organisms.
and intracellular, molecular networks. The primary author carried out the analysis, of which the secondary author carried out a sequential review.

**Results**

The literature search resulted in the identification of 141 articles, of which 17 met inclusion criteria (Figure 1 & Table 1). Based on these studies, we report 13 recurring motifs (Tables 2-6). Each edge is directional, but is non-specific with regard to type (i.e., induction, inhibition or binding), unless otherwise stated, consistent with the protocol laid out in the Methods section. We organized motif sub-types according to the direction of the higher order motifs’ edges. Of the 13 recurring motifs, there were 2 subtypes also found to be statistically significant across two or more studies.

Most motifs were found to recur across an evolutionary range of organisms. For instance, the feed-forward loop motif was found to recur in organisms ranging from *S. cerevisiae* to *H. sapiens* and 10 other organisms (Table 3). Only two motifs did not recur across an evolutionary range of organisms: the mixed feedback loop and diamond motifs. Most motifs were found to recur in both the broad categories of gene regulation and protein-protein interaction (PPI). There were, however, exceptions. The one way cycle reversible step and diamond motifs were found only in PPI networks. The significant pair motif was found solely in gene regulatory networks.

Two studies included in this analysis evaluated motifs where gene regulation and protein-protein interaction were considered jointly as a single network (14, 15). Both studies
identified the same, 3-node motifs in S. cerevisiae, including the following motifs: feed forward loop, co-regulator, co-regulated interacting, and protein clique.

Sub-networks, or pathways, of gene regulatory and PPI networks were not evaluated, on account of the limited number of studies that evaluated any given pathway.

Methodologically, studies used wide ranging, overlapping and distinct data sets. We evaluated the effect the choice of data set would have on results, comparing the motifs found in the gene regulatory network of S. cerevisiae. Six studies evaluated motifs based on this condition, of which all six experimentally identified the feed forward loop. Two studies, Joshi et al. (19) and Mazurie et al. (12) identified the significant pair motif. Four other studies individually identified the two-way interaction, co-regulator, 3-cycle and bifan motifs. These four motifs were found to recur only in the larger analysis, which included all organisms and/or networks/pathways studied. This finding indicates that the choice of data set in motif identification analyses is important. It is based on this finding that we reason that it is crucial to continue to conduct motif identification analyses across data sets and spanning a range of organisms, so as to minimize the risk of spurious results.

The majority of studies included in this review applied the isomorphism and randomization methods first developed by Milo et al. (5) and Kashtan et al. (20), with minor variations (n=13). There exist a number of distinct, motif isomorphism and randomization algorithms (21). Further, the choice of isomorphism algorithm is known
to effect results (18). This was reaffirmed by analysis of the studies included in this review. For instance, Konagurthu and Lesk (9) repeated the studies of Ma’ayan *et al.* (22) and Shen-Orr *et al.* (6), but used a more conservative isomorphism algorithm. This research team identified both the feed forward loop and 3-cycle motifs, consistent with the other two studies. However, Konagurthu and Lesk did not identify the same extent of three node motifs found in Ma’ayan or the four node motifs of both Ma’ayan *et al.*, and Shen-Orr *et al.* Further, Kashtan *et al.* (20) applied a sampling based algorithm to a version of the data used in Shen-Orr *et al.* (6) identifying feed forward loops and dense overlapping regions, but like Konagurthu and Lesk, did not make a single input module motif finding.

Only seven studies provided explicit motif definitions. Therefore, in many instances, many terms and definitions in tables 2-6 had to be inferred based on the larger study description and associated diagrams. The unique term to identify a motif type in this analysis was generally that term used by the research team that made the initial finding. In limited instances, more refined or concise terms were applied for sake of clarity. For instance, Shen-Orr *et al.* (6) were the first to use ‘feed forward loop’ and so is applied (table 3). We applied the same rule to the definitions of each motif. We did abstract each definition, so that it could be applied to describe such motifs regardless of network type in those instances where motifs were found across both network types, with the goal of affording the motif adequate, generalized description. For instance, Shen-Orr *et al.* (6) first defined the FFL as, “*defined by a transcription factor X that regulates a second transcription factor Y, such that both X and Y jointly regulate an operon Z.*” As FFLs are
known to reside in gene regulatory and PPI networks, we abstracted this definition to, “A regulates B, such that both A and B jointly regulate an operon C.” We hope this approach to terms and definitions minimizes controversy.

Finally, a limited set of motifs could not be curated. This is because these motifs were found in only single studies or lacked complete information.

**Discussion**

These 13 recurring motifs suggest that they are more than just statistical anomalies; they potentially represent critical patterns of biological activity essential for cellular homeostasis. This study constitutes the first formal review of recurring molecular motifs across species and pathways at the intracellular level. The review was based on primary research that drew conclusions originating from distinct research teams, unique information sources, using overlapping methods.

Other studies generally support this review’s findings. Alon (16) reported that feed forward loops recur in the transcription networks that respond to environmental stimuli of *E. coli, H. sapiens* and *S. cerevisiae*. The same summary reported on the recurrence of dense overlapping regions in the transcription networks in *E. coli* and *S. cerevisiae*. Eom *et al.* (23) carried out a taxonomic analysis of motif triads in energy metabolism pathways across 43 species and found many motifs were largely conserved within broad taxonomic categories.
While these findings reinforce this study’s strength, there too are study limitations. For instance, it could be that motifs simplify a multi-step processes into over-simplified interactions and so could under or over-state the existence of a motif non-differentially. This could be due to the fact that the knowledge of a given network or pathway may not be complete or the methods used to integrate network data from multiple sources are unclear or not specified to the right level of detail. Second, potentially important motifs that do not recur would not be detected based on the general methods used to inform this review. Third, this study applied conservative criteria in accepting a motif as recurring. Further, most studies did not specify the type of relationship (edge) between nodes (i.e., induction, inhibition or binding). Only those studies published in English were included. The results of this study, therefore, most likely under-represent the full catalogue of recurring motifs and subtypes across species and networks.

These limitations are largely overcome by the overall strengths of this review. This review includes analyses from 17 studies. These studies experimentally investigated the existence of motifs within sub-cellular networks ranging from simple organisms, like S. cerevisiae, to H. sapiens. Further, these motifs were found consistently regardless of data source or research method and across distinct research teams.

It is important to note the absence of three subgraphs from the recurring set: 1) auto-regulation; 2) cascade and; 3) single input module (SIM). Auto-regulation is where a molecule directly induces or inhibits the activity or production of itself. Auto-regulatory subgraphs are known to occur with some transcription factors, where the feedback edge is
often inhibitory; e.g., Nrf2\textsuperscript{vii} exhibits this pattern. The cascade subgraph is where molecule A regulates the activity of B; B regulates the activity of C, as demonstrated in the MAPK\textsuperscript{viii} cascade. The SIM is “defined by a set of operons that are controlled by a single transcription factor (6).” Aryl hydrocarbon receptor, for example, is known to induce the transcription of a multitude of mRNA.

Each of these subgraphs was identified experimentally in single studies; however, none were experimentally identified in two or more studies using the inclusion criteria and in subgraphs of the same node number. The absence of auto-regulation may be an artifact. Many studies excluded auto-regulatory feedback loops from their network graphs before carrying out analyses. Further, most studies carried out analyses only at subgraphs of node size three or greater. Such study designs would preclude identification of auto-regulatory loops in most circumstances. The cascade subgraphs may too be underrepresented, but for a different reason. Joshi \textit{et al.} (19) identified a three node cascade subgraphs in the post transcriptional control of mRNA of \textit{S. cerevisiae}. Kashani, \textit{et al.} (24) identified a four node form of the same subgraph in the metabolic network of \textit{E. coli}. We reason that underrepresentation of the cascade subgraph exist, in part because most studies evaluated pathways associated with the early steps of gene regulation, and few studies researched the complete gene regulation pathway. The regulatory properties of transcription factors and associated molecules typically do not exhibit cascade-like topology. Further, signaling pathways, which are generally regarded to exhibit cascade-like topology, have not been as extensively researched as have

\textsuperscript{vii} NF-E2 related factor 2
\textsuperscript{viii} Mitogen activated protein kinase
transcription networks. Of the studies included in this systematic review, only seven evaluated PPIs.

Finally, the SIM subgraph was experimentally identified in only a single study (6). We reason, based on our own experiences that this subgraph is likely a recurring motif. We evaluated potential causes of underrepresentation, but analysis of the studies used in this review did not reveal one single cause. Choice of isomorphism algorithm does impact results. Further, selected pathways, data sets and definitions may also be contributors. Clear articulation of these important methodological factors is critical in motif identification studies.

Four studies did research subgraphs of node sizes five or greater (6, 20, 22, 24). The five node, dense overlapping region (DORs) motif was found across three studies, as noted in the results section. Other five node or larger sized motifs were not found to recur. The reason for this may in part be due to the fact that the computational complexity associated with detecting larger motifs is expensive, and it has only been in recent years that the ability to conduct such analyses at this level has become available.

Motifs that may exist within specific biological pathways, particularly in H. sapiens compared to other species, represent areas where further research is needed. Related, analytical tools now exist where specific motif subtypes may be identified. For instance, researchers may now use color techniques to discern motifs of the same node size with specific types of nodes and edges (e.g., FANMOD or similar tools). This is an important
analytical capability for motif identification in more complex pathways like signaling or pathways involved in adaptation to environmental stimuli, as these pathways exhibit complex patterns of induction, inhibition and binding.

We reason that these recurring motifs could have important implications. For instance, recurring motifs could bring increased knowledge in toxicology and in particular, toxicity testing. The U.S. National Academies of Sciences (NAS) in 2007, via its report entitled, “Toxicity Testing in the 21st Century,” posited a vision for toxicity testing, asserting that the goals of toxicity testing moving forward should be to identify those critical pathways that when sufficiently perturbed lead to adverse health outcomes. The NAS further stated that, “perturbations of cell-signaling motifs . . . are obligatory changes related to chemical exposure that might eventually result in disease (7).” This new paradigm is a sharp contrast to the reductionist tradition of toxicology. However, the rich literature on mechanism and mode of action could be a potentially important source to build upon and apply to our existing knowledge of recurring motifs. For instance, two nuclear receptors, constitutive androstane receptor (CAR) and pregnane X receptor (PXR) are known to regulate the transcription of a common number of phase I and phase II metabolizing enzymes (25). Nuclear receptors are a super-family of transcription factors and are commonly associated with apical outcomes of interests to toxicologists involved in regulatory decision making (26). As we see in figure 2, the simplified relationship between CAR and PXR resembles a bipartite graph and a potential Bifan or DOR motif. The DOR motif is poorly understood and simplified dynamic models based on structure alone yield seemingly contradictory results (16, 27). However, in the instance of CAR
and PXR, it is known that PXR is more promiscuous and able to recognize a wider range of xenobiotics than CAR; whereas, CAR is expressed at higher basal levels (28). Further, nuclear receptors are known to undergo post-transcriptional regulation (29). This toxicological information has the potential to better inform dynamic models of recurring motifs and in turn be applied to analyze the effects of chemical perturbations on key components of regulatory pathways; i.e., recurring motifs. This combined approach has the potential to result in validated models, better informing more rapid commercial and regulatory toxicology decisions. Practically, this methodological approach has the potential to reduce reliance on whole animal testing, which is a cost burden to both regulators and industry (30).

This first formal review of network motifs finds the recurrence of 13 motifs across species. These 13 recurring motifs are more than just statistical anomalies; they potentially represent critical patterns of biological activity essential for cellular homeostasis and are an important, scientific step forward in further understanding the wiring of sub-cellular pathways. While further research is required, these findings also have potentially important and practical implications for medical science and toxicology. We encourage increased applied research to analyze how this knowledge can be applied to medical science and chemical safety.
References


Table 1: Studies meeting inclusion criteria<sup>ix</sup>

<table>
<thead>
<tr>
<th>Study</th>
<th>Organism/Cell Line</th>
<th>Data Source</th>
<th>Network; Pathway (where applicable)</th>
<th>Motif Identification Method (isomorphism; randomization)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eom et al. 2006 (23)</td>
<td>S. cerevisiae</td>
<td>Alon (provided)</td>
<td>Gene regulation; transcription</td>
<td>Milo et al. 2002 (5); edge swapping</td>
</tr>
<tr>
<td>Joshi et al. 2011 (19)</td>
<td>S. cerevisiae</td>
<td>1) Halbeisen &amp; Gerber 2009 (31); 2) Harbison et al. 2004 (32)</td>
<td>Gene regulation; post transcriptional and translational control of mRNA</td>
<td>FANMOD (Wernicke &amp; Rasche 2006 (33)); unstated</td>
</tr>
<tr>
<td>Kashani et al. 2009 (24)</td>
<td>E. coli</td>
<td>KEGG</td>
<td>PPI; metabolism</td>
<td>Demonstrates algorithm based on revolving door ordering and NAUTY; Milo et al. 2002 (5)</td>
</tr>
<tr>
<td>Kashan et al. 2004 (20)</td>
<td>E. coli</td>
<td>Shen-Orr et al. 2002 (data v. 1.1) (6)</td>
<td>Gene regulation; transcription</td>
<td>Demonstrates an algorithm for motif identification based on sampling; Switching, matching, go with winners</td>
</tr>
<tr>
<td>Kim et al. 2011 (34)</td>
<td>D. melanogaster</td>
<td>Nikitin et al. 2003 (35)</td>
<td>Protein-protein interaction</td>
<td>Milo et al. 2002 (5); edge swapping</td>
</tr>
<tr>
<td></td>
<td>E. lupus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H. sapiens</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M. musculus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R. norvegicus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S. cerevisiae</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Konagurthu &amp; Lesk 2008 (9)</td>
<td>E. coli</td>
<td>Shen-Orr et al. 2002 (data v 1.0) (6)</td>
<td>Gene regulation; transcription</td>
<td>Ulmann 1976 (36); Shen-Orr et al. 2002 for 3 node randomization (6); For N node randomization: Konagurthu &amp; Lesk 2008 (9).</td>
</tr>
<tr>
<td></td>
<td>Mammalian hippocampal CA1 neuron</td>
<td>Ma'ayan et al. 2005 (22)</td>
<td>PPI: signal propagation resulting from ligand occupancy.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S. cerevisiae</td>
<td>Luscombe et al. 2004 (37)</td>
<td>Gene regulation; transcription</td>
<td></td>
</tr>
<tr>
<td>Ma'ayan et al. 2005 (22)</td>
<td>Mammalian hippocampal CA1 neuron</td>
<td>Literature search</td>
<td>PPI: signal propagation resulting from ligand occupancy.</td>
<td>Kashan et al. 2004 (20); full enumeration (3-4 node), Random sampling (5-6 node); unstated</td>
</tr>
<tr>
<td>Martinez et al. 2008 (38)</td>
<td>C. elegans</td>
<td>miRNA-Yeast 1-hybrid method</td>
<td>Gene regulation: transcription of miRNAs &amp; post transcriptional control of transcription Factors by miRNAs</td>
<td>Kashan et al. 2004 (20); edge switching, node replacement, complete randomization</td>
</tr>
<tr>
<td>Milo et al. 2002 (5)</td>
<td>E. coli</td>
<td>Shen-Orr et al. 2002 (6)</td>
<td>Gene regulation: transcription</td>
<td>Developed method; randomization by 2 algorithms: 1) Markov, based on Kannan et al. 1997 (41) &amp; Maslov &amp; Snippen 2002 (40); 2) modification of Newman et al. 2001 (42)</td>
</tr>
<tr>
<td></td>
<td>S. cerevisiae</td>
<td>Yeast Proteome Database</td>
<td>Gene regulation: transcription</td>
<td></td>
</tr>
<tr>
<td>Milo et al. 2004 (13)</td>
<td>B. Subtilis</td>
<td>Ishii et al. 2001 (43)</td>
<td>Gene regulation: transcription</td>
<td>Mfinder 1.1 (Milo et al. 2002 (5) &amp; Kashtan et al. 2004 (20)); edge switching</td>
</tr>
<tr>
<td></td>
<td>D. melanogaster</td>
<td>GeneNet</td>
<td>Gene regulation: developmental transcription networks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E. coli</td>
<td>Shen-Orr et al. 2002 (6)</td>
<td>Gene regulation: transcription</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E. perischoechinoidea</td>
<td>Davidson et al. 2002 (44)</td>
<td>Gene regulation: endomesoderm development transcription networks</td>
<td></td>
</tr>
</tbody>
</table>

<sup>ix</sup> Table summarizes only those organisms and pathways the studies reported statistical results on.
<table>
<thead>
<tr>
<th>Study</th>
<th>Organism/Cell Line</th>
<th>Data Source</th>
<th>Network; Pathway (where applicable)</th>
<th>Motif Identification Method (isomorphism; randomization)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ryall et al. 2012 (45)</td>
<td>H. sapiens</td>
<td>Signal Transduction Knowledge Environment</td>
<td>PPI: signal transduction</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S. cerevisae</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neonatal R. norvegicus myocytes</td>
<td>Literature search</td>
<td>PPI: cardiac hypertrophy</td>
<td>NetMatch; RandomNetworks</td>
</tr>
<tr>
<td>Shalgi et al. 2007 (46)</td>
<td>E. lupus</td>
<td>1) TargetScan; 2) PicTar</td>
<td>Gene regulation: transcription of miRNAs &amp; post transcriptional control of transcription factors by miRNAs</td>
<td>Shin-Orr et al. 2002 (6); edge swapping</td>
</tr>
<tr>
<td></td>
<td>H. sapiens</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M. musculus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R. norvegicus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shellman et al. 2013 (47)</td>
<td>H. salinarum</td>
<td>Gonzalez et al. 2008 (48)</td>
<td>PPI: metabolism</td>
<td>FANMOD (Wernicke &amp; Rasche 2006 (33)); unstated</td>
</tr>
<tr>
<td></td>
<td>H. sapiens</td>
<td>Duarte et al. 2007 (49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M. acetivorans</td>
<td>Kumar et al. 2011 (50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M. barkeri</td>
<td>Feist et al. 2006 (51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M. musculus</td>
<td>Sigurdsson et al. 2010 (52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S. cerevisae</td>
<td>Duarte et al. 2004 (53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shen-Orr et al. 2002 (6)</td>
<td>E. coli</td>
<td>1) Regulon Database (v3.2); 2) Literature search</td>
<td>Gene regulation: transcription</td>
<td>Connectivity matrix for non-dense overlapping regions and standard average-linkage algorithm for dense overlapping regions; Newman et al. 2001 (42) &amp; Kannan et al. 1997 (41)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yeger-Logem et al. 2004 (14)</td>
<td>S. cerevisae</td>
<td>1) Biomolecular Interaction Network Database; 2) Database of Interacting Proteins; 3) Lee et al. 2002 (10); 4) Munich Information Center for Protein Sequences database; 5) Ren et al. 2000 (54); 6) SCPD Promoter Database; 7) Simon et al. 2001 (55); 8) Yeast Proteome Database</td>
<td>Combined gene regulation &amp; PPI</td>
<td>Extension of Shen-Orr et al. 2002 (6), considering extended degree of a node and the edge profile of 2 nodes; unstated</td>
</tr>
<tr>
<td>Zhang et al. 2005 (15)</td>
<td>S. cerevisae</td>
<td>1) Gavin et al. 2002 (56); 2) Ho et al. 2002 (57); 3) Hughes et al. 2000 (58); 4) Lee et al. 2002 (10); 5) Munich Information Center for Protein Sequences database; 6) Protein sequence homology; relationships from a genome-wide BLAST search; 7) Tong et al. 2004 (59)</td>
<td>Combined gene regulation &amp; PPI</td>
<td>Milo et al. 2002 (5); Park &amp; Newman 2003 (60)</td>
</tr>
</tbody>
</table>
Table 2: Two-node motifs

<table>
<thead>
<tr>
<th>Motif</th>
<th>Two-way interaction</th>
<th>Mixed feedback loop</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>A regulates the activity of or the production of B. B or the product of B regulates the activity of or the production of A.</td>
<td>A regulates the production of B. The product of B inhibits the activity of A.</td>
</tr>
<tr>
<td><strong>Network: Pathways</strong></td>
<td>GR: transcription of miRNAs &amp; post transcriptional control of transcription factors by miRNAs; PPI</td>
<td>GR: transcription of miRNAs &amp; post transcriptional control of TFs by miRNAs; Combined transcription-protein-protein interaction</td>
</tr>
<tr>
<td><strong>Organisms</strong></td>
<td>D. melanogaster; E. lupus; H. sapiens; M. musculus; R. norvegicus; S. cerevisiae</td>
<td>C. elegans; S. cerevisae</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td>Kim et al., 2011; Mazurie et al., 2005</td>
<td>Martinez et al., 2008; Yeger-Lotem et al., 2004</td>
</tr>
</tbody>
</table>

Legend:
- Gene regulation: GR
- Protein-protein interaction: PPI

---

Induces: A
Inhibits: B
Table 3: Three-node motifs (1 of 2)

<table>
<thead>
<tr>
<th>Motif</th>
<th>Feed Forward Loop</th>
<th>Co-regulator</th>
<th>3-Cycle</th>
<th>One way cycle, reversible step</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>A regulates B, such that both A and B jointly regulate an operon C.</td>
<td>Two regulators, A &amp; B that interact with one another, both jointly regulating the activity or production of the target, C.</td>
<td>A regulates the activity or production of B, B regulates the activity or production of C and C regulates the activity or production of A.</td>
<td>A regulates the activity of B and C, B regulates the activity of C, C regulates the activity of A.</td>
</tr>
<tr>
<td><strong>Pathways</strong></td>
<td>GR: post transcriptional control of transcription factors by miRNAs, transcription; PPI: cardiac hypertrophy, metabolism, signal transduction</td>
<td>GR: post transcriptional &amp; translational control of mRNA, transcription; PPI: signal propagation resulting from ligand occupancy</td>
<td>GR: transcription, transcription in response to stress; PPI: signal propagation resulting from ligand occupancy</td>
<td>PPI: metabolism</td>
</tr>
<tr>
<td><strong>Organisms</strong></td>
<td>B. subtilis; D. melanoogaster; E. coli; E. lupus; E. persicochelinoides; H. salinarum; H. sapiens; Mammalian hippocampal CA1 neuron; M. acetivorans; M. barkeri; M. musculus; R. norvegicus; S. cerevisiae</td>
<td>D. melanoogaster; E. lupus; H. sapiens; Mammalian hippocampal CA1 neuron; M. acetivorans; M. barkeri; M. musculus; R. norvegicus; S. cerevisiae</td>
<td>E. coli; Mammalian hippocampal CA1 neuron; S. cerevisiae</td>
<td>H. salinarum; H. Sapiens; M. musculus; R. norvegicus; S. cerevisiae</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td>Eom et al., 2012; Joshi et al., 2011; Kashtan et al., 2004; Kim et al., 2011; Konagurthu &amp; Lesk 2008; Mazure et al., 2005; Milo et al., 2002; Milo et al., 2004; Ryall et al., 2012; Shalgi et al., 2007; Shellman et al., 2013; Shen-Orr et al., 2002; Yeger-Lotem et al., 2004; Zhang et al., 2005</td>
<td>Joshi et al., 2011; Kim et al., 2011; Ma’ayan et al., 2005; Shalgi et al., 2007; Yeger-Lotem et al., 2004; Zhang et al., 2005</td>
<td>Kim et al., 2011; Konagurthu &amp; Lesk 2008; Ma’ayan et al., 2005</td>
<td>Kim et al., 2011; Shellman et al., 2013</td>
</tr>
</tbody>
</table>

A \rightarrow B \rightarrow C

Binds: A

Induces: B

References:
Eom et al., 2012; Joshi et al., 2011; Kashtan et al., 2004; Kim et al., 2011; Konagurthu & Lesk 2008; Mazure et al., 2005; Milo et al., 2002; Milo et al., 2004; Ryall et al., 2012; Shalgi et al., 2007; Shellman et al., 2013; Shen-Orr et al., 2002; Yeger-Lotem et al., 2004; Zhang et al., 2005
Table 4: Three-node motifs (2 of 2)

<table>
<thead>
<tr>
<th>Motif</th>
<th>Co-regulated Interacting</th>
<th>Protein Clique</th>
<th>Significant Pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td>B and C are regulated by a common regulator; A, B and C and their products interact with each other.</td>
<td>A, B and C join, forming a complex.</td>
<td>A and B regulate the production of C.</td>
</tr>
<tr>
<td>Pathways</td>
<td>GR: transcription; PPI: signal propagation resulting from ligand occupancy; signal transduction</td>
<td>GR: transcription; PPI: metabolism, signal propagation resulting from ligand occupancy; signal transduction</td>
<td>GR: post transcriptional &amp; translational control of miRNA, transcription</td>
</tr>
<tr>
<td>Organisms</td>
<td>H. sapiens; Mammalian hippocampal CA1 neuron; M. musculus; R. norvegicus; S. cerevisiae</td>
<td>E. coli; H. sapiens; M. bakeri; Mammalian hippocampal CA1 neuron; M. musculus; R. norvegicus; S. cerevisiae</td>
<td>S. cerevisiae</td>
</tr>
<tr>
<td>References</td>
<td>Kim et al., 2011; Ma’ayan et al., 2005; Milo et al., 2004; Yeger-Lotem et al., 2004; Zhang et al., 2005</td>
<td>Kim et al., 2011; Ma’ayan et al., 2005; Milo et al., 2004; Shellman et al., 2013; Yeger-Lotem et al., 2004; Zhang et al., 2005</td>
<td>Joshi et al., 2011; Mazutie et al., 2005</td>
</tr>
</tbody>
</table>

![Motif Diagram](image-url)
Table 5: Four-node motifs

<table>
<thead>
<tr>
<th>Motif</th>
<th>Bifan</th>
<th>Diamond</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>Two upstream components, A &amp; B, independently or co-regulate the activity or production of two downstream components, C &amp; D.</td>
<td>A regulates the activity of B &amp; C. C &amp; D regulate the activity of D.</td>
</tr>
<tr>
<td><strong>Pathways</strong></td>
<td>GR: transcription; PPI: cardiac hypertrophy, signal propagation resulting from ligand occupancy</td>
<td>PPI: cardiac hypertrophy, signal propagation resulting from ligand occupancy</td>
</tr>
<tr>
<td><strong>Organisms</strong></td>
<td>E. coli; mammalian hippocampal CA1 neuron; Neonatal R. norvegicus myocytes; S. cerevisiae</td>
<td>Mammalian hippocampal CA1 neuron; Neonatal R. norvegicus myocytes</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td>Kashtan et al., 2004; Ma’ayan et al., 2005; Milo et al., 2002; Ryall et al., 2012; Yeger-Lotem et al., 2004</td>
<td>Ma'ayan2005; Ryall2012</td>
</tr>
</tbody>
</table>

![Bifan diagram](image1)

![Diamond diagram](image2)
Table 6: Five node motif

<table>
<thead>
<tr>
<th>Motif</th>
<th>Dense Overlapping Region 1</th>
<th>Dense Overlapping Region 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>Is an overlapping layer of regulators, A &amp; B, independently or co-regulating the production or activity of operons, C, D &amp; E.</td>
<td>Is an overlapping layer of regulators, A, B &amp; C, independently or co-regulating the production or activity of operons, D &amp; E.</td>
</tr>
<tr>
<td><strong>Pathways</strong></td>
<td>GR: transcription; PPI: signal propagation resulting from ligand occupancy</td>
<td>GR: transcription; PPI: signal propagation resulting from ligand occupancy</td>
</tr>
<tr>
<td><strong>Organisms</strong></td>
<td>E. coli; Mammalian hippocampal CA1 neuron</td>
<td>E. coli; Mammalian hippocampal CA1 neuron</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td>Kashtan et al., 2004; Ma’ayan et al., 2005; Shen-Orr et al., 2002</td>
<td>Kashtan et al., 2004; Ma’ayan et al., 2005</td>
</tr>
</tbody>
</table>

![Diagram](https://via.placeholder.com/150)
Figure 1: Inclusion criteria

- Literature search (141)
  - Motif specified (46)
    - Intracellular (42)
      - Non-carcinogenic cells (36)
        - Quantified, experimental method (23)
          - Total publications for inclusion (17)
Figure 2: CAR-PXR DOR

```
CAR          PXR
  ↓          ↓
CYP 3A4     CYP 2C9
  ↓          ↓
CYP 2C19    GST A2
```
Chapter V: *In vitro* perturbation of the transcription factor e2f is a strong predictor of *in vivo* hepatocarcinogenesis resulting from chemical exposure

This paper will be submitted for review, journal to be determined.

Abstract

Background

Interest has grown in recent years in the potential for *in vitro* toxicity testing methods to replace *in vivo* methods. In fact, the US National Academies of Sciences in their report entitled, “Toxicity Testing in the 21st Century,” promoted such strategies. The research arm of the EPA has embraced this challenge and in 2007 initiated the ToxCast program. This program has contributed to making important findings regarding the association and correlation of the activity of xenobiotic metabolizing enzymes and nuclear receptors with respect to liver carcinogenicity and chemical exposure. However, these perturbations have yet to be assessed for their ability to discriminate and therefore predict the presence or absence of liver lesions of *in vivo*, chronic toxicity tests. We therefore set out to contribute to the literature and answer the basic question: could ToxCast phase I data be used to discriminate and therefore predict chemicals that either do or do not produce liver lesions of *in vivo*, chronic toxicity tests.

Results

Chemicals known to produce preneoplastic and neoplastic lesions in the livers of *M. musculus* and *R. norvegicus* were found to be more lipophilic, have higher molecular weights, exhibited higher polarizability and solvent accessible surface area than chemicals that did not produce lesions. Lesion-causing chemicals also had a higher,
predicted ability to bind to human serum albumin. The activity of select nuclear
receptors exhibited modest increases in activity with Nrf2 exhibiting statistically
significant results. The transcription factor e2f also exhibited a statistically significant
increase in activity. Receiver operating characteristic analysis revealed that the
physiochemical properties previously mentioned are modest discriminators as is Nrf2.
Weight adjusted e2f produced an AUC of 0.931, suggesting it is a stronger discriminator
and therefore predictor of chemicals that do cause liver lesions versus those that do not.

Discussion
This study demonstrates a targeted use of high throughput screening data towards a
potential risk assessment application like hazard identification. The finding made with
respect to the e2f transcription factor is intriguing. E2f is part of a critical corridor
involved in cell cycle control and apoptosis. Perturbations of e2f could be a potentially
important discriminator of hepatocarcinogenic risk and possible candidate as a
carcinogenesis biomarker of promotion.

Introduction
In 2007, the US National Academies of Sciences (NAS) published “Toxicity Testing in
the 21st Century: a Vision and Strategy (1).” A vision was put forward for toxicity
testing, where testing strategies in the future would rely less on in vivo methods and more
extensively on in vitro or in silico methods. The NAS posited that these in vitro and in
Methods would ideally identify those critical pathways that when sufficiently perturbed lead to adverse health outcomes (2).

The creation and adoption of *in silico* and *in vitro* methods are important for ethical, cost and efficiency reasons. The ethical treatment of animals is culturally important in western society and increasingly, people are scrutinizing the use of animals in research. Second, the potential cost savings are a more pragmatic justification. It is generally accepted that alternative methods can potentially result in significant cost savings over traditional, *in vivo* methods (3). Finally, alternative methods can be carried out more quickly and the same test can be carried out on a multitude of chemicals simultaneously (4). Such methods could bring many potential benefits for groups like regulators and manufacturers.

The research arm of the EPA has taken on the NAS challenge and has been moving forward to bring to reality the NAS vision. The EPA initiated ToxCast in 2007 (5, 6, 7). ToxCast is an EPA pilot program that continues to carry out a variety of tests on thousands of chemicals. The ToxCast program generates large volumes of data on chemicals using a number of *in vitro* assay batteries. The *in vitro* assays range in scope from enzyme challenge assays to cell-level assays. This information is supplemented by data that detail the varied properties of these thousands of chemicals. The ToxCast program completed their phase I, “Proof of Concept,” phase in 2009, the data of which is available to the public.

---

ToxCast has entered into phase II, of which the program made this data publically available early in 2014. Due to the recent release of this data, it is excluded from this study.
The ToxCast program has analyzed many aspects of the program’s data. These studies include but are not limited to analyses of pathway perturbations (4), a predictive model of developmental toxicity (8), a reproductive toxicity model (9), genotoxicity screening (10) and endocrine disruption screening (11).

We were interested in chemical exposure and associated liver carcinogenicity. The liver, which is made up predominately of hepatocytes, is a primary area to study the toxicity arising from chemical exposure. Its specialized processes and physical features - first pass effect - result in higher exposure levels of toxicants than other tissues (12). Hepatocytes have protective systems - a high concentration of phase I and phase II enzymes - whose reactions often define the difference between cell injury or detoxification and restoration of homeostasis. These protective systems also include safety controls and shut down systems (13). We know these as cell cycle control mechanisms, including apoptosis. It is in part for these reasons that the liver is the most common target organ affected in repeat dose toxicity studies (14). Hepatotoxicity is also a leading reason why the US Food and Drug Administration withdraws drugs from the market (15). If hepatocyte functions deviate substantially from safety limits – i.e., are sufficiently perturbed - then the hepatocyte can shut down or cease to replicate. Failure of the protective systems could result in necrosis or an ongoing system whose contributions are no longer safe; e.g., propagation of carcinogenesis. The multi-stage nature of carcinogenesis in particular is important to note. Upon exposure, a cell can be initiated and undergo a somatic cell mutation. Promotion then occurs through selective
clonal expansion of initiated cells. The cancer process progresses with the maturation of neoplastic phenotypes, including angiogenesis, invasiveness and metastasis.

The ToxCast battery of assays includes data originating from hepatocyte-based assays. These include broad measures of cell level outcomes, xenobiotic metabolizing enzyme activity, nuclear receptor activity and transcription factor activity (10, 16, 17, 18). These studies have contributed to making important findings regarding the association and correlation of the activity of xenobiotic metabolizing enzymes and nuclear receptors with respect to liver carcinogenicity and chemical exposure. However, these perturbations have yet to be assessed as to their ability to discriminate and therefore predict the presence or absence of liver lesions of *in vivo*, chronic toxicity tests.

We therefore set out to contribute to the literature by studying this question. In this study, we assessed whether or not ToxCast phase I data can be used to discriminate and therefore accurately predict chemicals that do or do not produce liver lesions of *in vivo*, chronic toxicity tests. Herein, we find that select physiochemical properties and the Nrf2 assay data are weak discriminators; whereas, the e2f assay appears to be a stronger discriminator.

**Methods**

**Selection of chemicals**

We began by establishing a list of chemicals that have undergone *in vivo*, chronic toxicity testing and are known to either produce or not produce liver lesions in both *M. musculus*
and R. norvegicus. We relied upon the ToxRef database (ToxRefDB) for this task. The ToxRefDB is a curated summary of animal toxicity tests developed by the EPA’s National Center for Computational Toxicology (19, 20). The ToxRefDB is a curated, comparative database of apical endpoints across species. The data was collected predominately as a result of regulatory, animal testing carried out under the Federal Insecticide, Fungicide, and Rodenticide Act registration and re-registration processes. These animal studies were carried out using standard protocols. The source protocols originate from what are known as the 870 series. The 870 series is a harmonized series of protocols that EPA uses to fulfill their statutory mandates. Toxicity studies were carried out on M. musculus and R. norvegicus, evaluating numerous apical endpoints. The in vivo metrics were assessed based on chronic toxicity studies, where animals were exposed over a period of 18 to 24 months. The EPA curated the ToxRefDB using a two step validation process to ensure accuracy of data.

Chemicals selected for entry in to the case sample were defined as chemicals that produced preneoplastic lesions or neoplastic lesions in the liver at 1 mg/kg/day, over a two year period, in both M. musculus and R. norvegicus. Chemicals that were selected for entry into the control sample were defined as those chemicals that did not produce preneoplastic or neoplastic liver lesions at any dose of the same animal species. All chemicals must have also been included and analyzed in the ToxCast phase I data set (figure 1).

Chemical physiochemical data
The ToxCast phase I repository includes data on each chemical’s physiochemical properties. The physiochemical properties for each chemical meeting ToxRefDB inclusion criteria were collected and recorded. These physiochemical properties data were generated by use of both the Estimation Program Interface Suite (EPI Suite) modeling system and also QikProp (21, 22). The EPI Suite was developed by the EPA and Syracuse Research Corporation. This software estimates a number of physiochemical properties and also environmental fate parameters. The estimators are based primarily on a fragment constant approach (23, 24). The components and suite has been validated by an independent, scientific review board (25). QikProp is an estimation platform owned by Schrödinger. This platform has undergone benchmarking and performed well on common physiochemical parameters (26).

Assays

ToxCast also includes data from a number of *in vitro*, cell free and whole cell assays (4). Assays that measured the activity of molecules associated with cell cycle control and xenobiotic metabolism of the hepatocyte were included for study consideration (figure 1). Herein are summaries of the two assays included into this study.

The Celluman assay measures multiple indicators of cellular toxicity (27, 28). This platform measures multiple endpoints across both HepG2 and primary rat hepatocyte cells. The assay is run with endpoints measured at 1, 24 and 72 hours. For this study, only endpoints measured at 24 hours were used, measuring cell counts in both HepG2 cells and primary rat hepatocyte cells. Raw data were provided by ToxCast where the
data for each chemical were sampled four times per dose, using the following serial
dilutions (μM): 0.39, 0.78, 1.56, 3.12, 6.25, 12.5, 25, 50, 100, 200.

The Attagene assay quantifies the activity of select nuclear receptors and transcription
factors (1; 2). This is a multiple, gene reporter assay that includes both CIS and TRANS
acting elements. Activity is measured as fold change over control, where the control
compound is DMSO. Homogenous transcripts are distinguished by distinct DNA
fragments, quantitatively evaluated via capillary electrophoresis. The assay uses HepG2
cells and measures the activity over 24 hours. The ToxCast data provides only average
value per dose and oversamples at low concentrations (μM): 0.129, 0.391, 1.186, 3.594,
10.89, 33, 100. We evaluated only those transcription factors and nuclear receptors
involved in xenobiotic metabolism and cell cycle control pathways, including apoptosis.
The CIS-based factors selected were: Ahr, e2f, FoxO1, NF-kB, Nrf2, and p53. The
TRANS-based factors were: CAR and PXR.

Statistical Analyses

Data from both of the Cellumen and Attagene assays were tested for normality using the
Shapiro-Wilk method (30). With a 95% (p≤0.05) confidence interval, outliers in the data
were identified and removed prior to further statistical analyses. Removal of outliers was
necessary, as Receiver Operating Characteristic (ROC) analysis, which was applied in
this study, is sensitive to extreme values. Review of Cellumen outliers found that no
individual chemical was an outlier, but rather only single samples within the
quadruplicate sampling. Descriptive analyses were carried out along with development
of dose-response curves. As the assays used either serial dilution or oversampling at small doses approaches, the concentrations were log transformed to provide a more clear representation of the dose-response patterns. Two-way, student’s t-tests were carried out using unequal variances to identify those physiochemical properties and assays that showed statistically significant differences between the case and control groups. Physiochemical properties and assays found to differ at the level of statistical significance then underwent ROC curve analyses.

ROC analyses were used to assess the ability of select variables to accurately predict in vivo liver lesions (31, 32, 33). ROC analysis assesses the predictive ability of tools and models where the outcome variable is binary and the predictor variable is continuous. By use of this method, one can evaluate the performance of a test at numerous combinations of true positive (sensitivity) and false positive (1-specificity) rate cut points (34). This is plotted as a ROC curve. We further calculated the area under the curve (AUC) for each ROC curve and statistically compared competing AUCs to identify which physiochemical parameter or assay (i.e., test) is a better discriminator and therefore predictor of the apical endpoint of interest (35). Included was calculation of the Youden’s index, which is simply the cutoff value at which sensitivity and specificity are maximized (36). ROC analyses were carried out using un-weighted and weighted assay values. Weighted analyses were carried out using the standard inverse of the variance approach.
Tests for normality, descriptive analyses, dose-response analyses and t-tests were carried out using SAS 9.1. The ROC analyses were carried out using STATA 10.

**Results**

There were a total of 94 chemicals that met our defined inclusion criteria (figure 1). There were 33 chemicals that fit the criteria for the case study group of producing either liver preneoplastic lesions or neoplastic lesions in the liver at 1 mg/kg/day, x2yr. The remaining 61 chemicals did not produce any lesions at any dose. Tables 1, 2 and 3 summarize the chemicals by both case and control groups and also the chemical classification used by the ToxCast program. The totals for each classification group are higher than sample sizes, because some chemicals fall into more than one chemical class.

The chemicals of the case group had higher molecular weights than controls and were also more lipophilic than controls. The case group chemicals were also found to have higher polarizability and solvent accessible surface area than controls. Finally, the case group chemicals had a higher, predicted ability to bind to human serum albumin. All of these physiochemical properties were found to be statistically significant (table 4).

We then turned to the ToxCast *in vitro* assays and started with analysis of cell counts of rat hepatocytes. This was carried out using ToxCast’s Cellumen data. We determined the median cell count at each dose for both the case group and the control group. We chose median values as the measure of central tendency, because the data remained weakly skewed even after removal of outliers (see methods section). Median values are
less sensitive to outliers than mean values and so is a preferred measure of central
tendency. We then calculated the difference in cell count of the cases versus controls by
dose and plotted (figure 2). The results indicate little difference in cell count between the
case group and the control group.

The activity of the nuclear receptors and transcription factors selected for this study did
exhibit differences between the case group and control group. These differences were
generally modest. Only Nrf2 and e2f exhibited statistically significant differences (table 4).

Dose-response analyses of the nuclear receptors and transcription factors selected for
inclusion into this study was carried out next (figure 3). These analyses were carried out
using ToxCast’s Attagene data. The nuclear receptors – Ahr, Nrf2, CAR, PXR and Nf-kβ
– exhibited increases in activity as expressed by fold change. The extent of these
increases in the case group varied in reference to the control group. Ahr, CAR and Nf-kβ
exhibited modest fold change increases in the case group over the controls. In contrast,
the fold change of Nrf2 and PXR in the case group exhibited greater fold changes over
the control group. The three transcription factors included in this study – FoxO1, e2f and
p53 – exhibited varying dose-response patterns. The fold change of the FoxO1
transcription factor decreased with increasing dose. Neither e2f nor p53 exhibited a
consistent, dose-response pattern over the exposure ranges evaluated in this study.
The nuclear receptors and transcription factors used in this study interact with one another under physiological, real circumstances. Analysis of curated data from Ingenuity IPA (37) revealed that CAR and PXR inhibit FoxO1. Ahr induces Nrf2 (38). E2f affects the activity of p53 via the ARF-MDM2 route (39). Therefore, Spearman correlation analyses were carried out on these relationships, again using ToxCast’s Attagene data. The intent of the correlation analysis was to determine if the cells used in this study under the given exposure scenarios exhibited patterns consistent with current, biochemical knowledge, as previously referenced. Increases in median activity of PXR were negatively correlated with FoxO1 activity: -0.96. Increases in median activity of CAR were also negatively correlated with FoxO1 activity: -0.64. Ahr and Nrf2 were strongly correlated: 0.96. Finally, we conducted a Spearman correlation analysis of e2f and p53 and again found a positive correlation: 0.68. With the exception of the e2f-p53 relationship, all correlations were consistent with current, scientific knowledge, as previously referenced.

Receiver operating characteristic analyses (ROC) of the physiochemical properties that were found to exhibit statistically significant differences between study groups revealed that the predicted binding of the chemical to human serum albumin exhibited the largest area under the curve (table 5 and figure 4). All these predictors yielded AUCs and accompanying confidence intervals that modestly discriminated between the case and control groups. Comparisons of the AUCs across these physiochemical properties did not yield statistically significant differences, suggesting none of these parameters statistically performed better in discriminating between the two study groups.
We carried out the same ROC analysis on the nuclear receptors and transcription factors selected for inclusion into this study where the concentration applied in the assays was 100μM. Only two predictors yielded AUCs and accompanying confidence intervals that modestly discriminated between the case and control groups (table 5 and figure 5). These two predictors were the e2f transcription factor and the Nrf2 nuclear receptor. Comparison of these two predictors’ AUCs did not yield statistically significant differences. However, when we weight-adjusted the values using the inverse of the variance, marked differences in the e2f and Nrf2 ROC curves emerged, and the accompanying differences were statistically significant (table 5 and figure 5). The weighted AUC for e2f was 0.931, which suggests that this assay is a stronger discriminator of the case chemicals versus control chemicals.

**Discussion**

This study demonstrates a targeted use of high throughput screening data towards a potential risk assessment application like hazard identification. As previously stated, the purpose of this study was to assess whether ToxCast phase I data could be used to discriminate and therefore potentially predict chemicals that either do or do not produce liver lesions of *in vivo*, chronic toxicity tests. We find that select physiochemical properties and the Nrf2 assay data are weak discriminators; whereas, the e2f assay appears to be a stronger discriminator. In this respect, our data and results contributes to the literature and demonstrates a targeted use of high-throughput screening data towards potential risk assessment applications, in particular, hazard identification.
This study also yielded results consistent with previous studies. Shah et al. evaluated the activity of nuclear receptors with respect to hepatocarcinogens using ToxCast phase I data and found statistically significant differences in the activity of PXR and CAR (18). We noticed a difference in the activity of CAR and PXR between the case and control groups, but the differences between the case and control groups in this study did not exhibit a statistically significant difference. The difference in the results using the same data set can be explained by the fact that we applied strict criteria to this study. Chemicals were included into our case study group only if they produced preneoplastic or neoplastic liver lesions at 1mg/kg/day over 18-24 months. We also had a well defined control group with a specified cutoff dose. Shah et al. did not specify a cut-off dose as part of their inclusion criteria into either study group and also included non-proliferative lesions into their case study group.

Martin et al. (17) found Nrf2 activity to be higher across ToxCast assays compared to the activity of other assays. Our study also found Nrf2 activity to be higher in the case group than the control group and the difference was found to be statistically significant (table 4). Nrf2 also showed a clear dose-response pattern where activity increased with increasing dose (figure 3). Further, the strong correlation between Ahr activity and Nrf2 activity is consistent with current, scientific knowledge (40). Nrf2 is an indicator of oxidative stress and the presence of electrophiles (41, 42). The Nrf2 transcription factor is kept in its inactive state by Keap1. Upon activation by reactive oxygen species (ROS), Keap1, releases Nrf2 (43). The Nrf2 then moves into the nucleus and activates the
transcription of any of a number of genes associated with chemical defense and also genes associated with proteasomal degradation (44, 45). However, ROC analysis revealed Nrf2 to be only a weak predictor of liver preneoplastic and neoplastic lesions (table 5; figure 5). Nrf2 activity exhibited a modest variance across chemicals, as a result, when weight-adjusting the nuclear receptor’s activity, using within group inverse of the variance, did not produce results that differed from the un-weighted, ROC analyses. In this respect, the statistically significant increase in Nrf2 suggests that the cells under study were undergoing stress response as a result of chemical exposure. Further, if this alteration in activity is in response to electrophilic agents, it could be suggestive of potential DNA damage, as DNA exhibits a slight negative charge. If such is the case, then increases in Nrf2 activity could be an indicator of a chemical’s capacity to initiate cells.

The lack of difference in the activity of p53 of the HepG2 cell-based assay is of note (figure 3 and table 4). P53 is understood to play an important role in apoptosis and cell cycle control. Changes in p53 activity are associated with cancer (39). However, the most intriguing finding was the statistically significant difference in the activity of e2f between the two study groups and the robust AUC value that this particular assay produced. The difference in activity, does not show a consistent, dose-response pattern (figure 3); however at 100 μM, the small, within group variances (control group: 0.027; case group: 0.018) associated with a full unit difference in fold change contributed to a statistically significant difference finding and also a ROC analysis result that is suggestive of being a strong predictor of preneoplastic and neoplastic liver lesions in M.
*musculus* and *R. norvegicus*. These changes in the activity of transcription factors involved in cell cycle control and apoptosis are suggestive of chemicals capable of causing the expansion of initiated cells.

We evaluated whether the finding with respect to e2f could be anomalous and concluded that it likely is not. The e2f transcription factor is involved in both cell cycle control and apoptosis (46). E2f is responsible for the transcription of a number of S phase transition genes. It is also responsible for the transcription of Apaf1, which is a critical component of the apoptosome heterodimer, whose formation is necessary for apoptosis to be carried out in hepatocytes. E2f is also on one pathway that inhibits p53, which is carried out via the ARF-MDM2 route. The retinoblastoma protein (pRb) actively represses e2f and maintains e2f in an inactive state. Upon full phosphorylation, pRb releases e2f, and it becomes active. Other proteins are also involved in the e2f regulatory machinery. The cell count assay, which used rat hepatocytes, revealed little to no difference in cell count between the case and control groups across the chemicals (figure 2), consistent with the finding that e2f activity is depressed when exposed to hepatocarcinogen-causing chemicals used in this study.

Differences in e2f expression and activity are associated with hepatocarcinogenesis (47, 48). However, these studies often find increases in expression or activity, particularly in proliferating cell lines or susceptible animal models. We inquired as to whether HepG2 cells exhibit different, basal expression and activity levels of e2f compared to other non-proliferative cell types. Harris *et al.* compared the basal, mRNA expression of primary
human hepatocytes and HepG2 cells (49). E2f was not identified as exhibiting basal differences. However, expression is not necessarily indicative of activity.

The suggestion that changes in e2f expression or activity are predictive of hepatocarcinogenesis does not necessarily mean this is the mechanism by which lesions occur. In this study, reduced activity of e2f served as a strong discriminator of the two study groups. E2f is in the central corridor of both cell proliferation and apoptosis, making e2f perturbation a relevant candidate as a carcinogenesis biomarker of promotion. We are confident that further research will result in a more definitive conclusion.

A limitation of this study is that the sample size of the case group was smaller than desired. However, the size of the control group was large, improving overall study power. It also would have been more desirable to have had analyses carried out on primary human hepatocytes, rather than HepG2 cells and rat hepatocytes. In this respect, the potential distinction between whether these agents are either initiating agents or promoting agents would be clearer. In contrast, the strength of this study is the stringent inclusion and exclusion criteria applied to the case and control groups along with a sound statistical approach.

The use of ROC analyses in this research is also an important contribution. This approach has been applied as part of an overall strategy to evaluate toxicity tests in Europe (50). Toxicity tests, like the ToxCast, in vitro assays, are ultimately assessed for their potential ability to correctly distinguish a diseased state (i.e., an apical outcome)
from a non-diseased state arising from a chemical exposure. Tests with continuous outcome measures, like those in the ToxCast battery of assays, require a cutoff value to be set, because the data are not binary. ROC analyses are uniquely suited to evaluate the performance of such data against known and accepted toxicity data.

ROC curves visually depict how well a diagnostic test, whose results are ordinal, distinguishes between two states. Sensitivity is plotted on the y axis and 1-specificity is plotted on the x axis. The area under the curve (AUC) represents the test’s discriminatory power. Further, the performance of the diagnostic test can be evaluated at any of a number of different cut points. For instance, table 6 is a cross section of the weight adjusted e2f results. In bold is the cut point where sensitivity and specificity are maximized, which is quantified using the Youden’s index. If one were to equally weigh the importance of sensitivity and specificity, the cutoff of 43.25 may be the point at which the threshold decision with respect to hazard risk could be made. Accepting values higher and to the right of the ROC curve are considered more lenient; values lower and to the left, more stringent (31). Such a determination would in principle be based on the policy of a regulatory body, for example.

In closing, this study demonstrates a targeted use of high throughput screening data towards a potential risk assessment application like hazard identification. We specifically assessed whether ToxCast phase I data could be used to discriminate chemicals that either do or do not produce liver lesions from in vivo, chronic toxicity tests. We found that select physiochemical properties and the Nrf2 assay data are weak discriminators;
whereas, the e2f assay appears to be a stronger discriminator. In this respect, our study adds to the literature. However, considering the aforementioned study limitations, we consider these results tentative. We are sure our results will support research in applying these findings to future studies and further, targeted use of high throughput data and its potential application to informing hazard risk.

These innovative findings demonstrate the ability to link high throughput ToxCast assays with in vivo toxicity data in looking at one endpoint (liver lesions) and selected assays. We have only scratched the surface as far as the potential of ToxCast data. Our method offers a useful model for continual research exploring the predictive value of ToxCast for other endpoints.
References


45. Kensler TW, Wakabayashi N, Biswal S. Cell survival responses to environmental stresses via the Keap1-Nrf2-ARE pathway. Annual Reviews of Pharmacology and Toxicology 2007: 47: 89-119.


Table 1: Control group chemicals

<table>
<thead>
<tr>
<th>Number</th>
<th>CASRN</th>
<th>Name</th>
<th>Attagene</th>
<th>Number</th>
<th>CASRN</th>
<th>Name</th>
<th>Attagene</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10043-35-3</td>
<td>Boric acid</td>
<td>✓</td>
<td>31</td>
<td>23135-22-0</td>
<td>Oxamyl</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>100784-20-1</td>
<td>Halosulfuron-methyl</td>
<td>✓</td>
<td>32</td>
<td>23422-53-9</td>
<td>Formetanate hydrochloride</td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>101-05-3</td>
<td>Anilazine</td>
<td>✓</td>
<td>33</td>
<td>24307-26-4</td>
<td>Mepiquat chloride</td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>103361-09-7</td>
<td>Flumioxazin</td>
<td>✓</td>
<td>34</td>
<td>25606-41-1</td>
<td>Propamocarb hydrochloride</td>
<td>✓</td>
</tr>
<tr>
<td>5</td>
<td>104088-48-8</td>
<td>Imazapic</td>
<td>✓</td>
<td>35</td>
<td>296-00-0</td>
<td>Parathion-methyl</td>
<td>✓</td>
</tr>
<tr>
<td>6</td>
<td>111812-58-9</td>
<td>(Z,E)-Fenpyroximate</td>
<td></td>
<td>36</td>
<td>296-04-4</td>
<td>Disulfoton</td>
<td>✓</td>
</tr>
<tr>
<td>7</td>
<td>1134-23-2</td>
<td>Cychoilate</td>
<td>✓</td>
<td>37</td>
<td>31218-83-4</td>
<td>Propetamphos</td>
<td>✓</td>
</tr>
<tr>
<td>8</td>
<td>115-29-7</td>
<td>Endosulfan</td>
<td></td>
<td>38</td>
<td>39515-41-8</td>
<td>Fenpropatrin</td>
<td>✓</td>
</tr>
<tr>
<td>9</td>
<td>118134-30-8</td>
<td>Spiroxamine</td>
<td>✓</td>
<td>39</td>
<td>41198-08-7</td>
<td>Profenofos</td>
<td>✓</td>
</tr>
<tr>
<td>10</td>
<td>120116-88-3</td>
<td>Cyazoamid</td>
<td></td>
<td>40</td>
<td>420-04-2</td>
<td>Cyanamide</td>
<td>✓</td>
</tr>
<tr>
<td>11</td>
<td>124465-18-7</td>
<td>Quinoxyfen</td>
<td></td>
<td>41</td>
<td>42599-80-8</td>
<td>Isazofos</td>
<td>✓</td>
</tr>
<tr>
<td>12</td>
<td>126833-17-8</td>
<td>Fenhexamid</td>
<td>✓</td>
<td>42</td>
<td>43222-48-6</td>
<td>Difenzoquat methisulfate</td>
<td>✓</td>
</tr>
<tr>
<td>13</td>
<td>127277-53-6</td>
<td>Nonhexadione-calcium</td>
<td>✓</td>
<td>43</td>
<td>54593-83-8</td>
<td>Chlorothoxylos</td>
<td>✓</td>
</tr>
<tr>
<td>14</td>
<td>13194-48-4</td>
<td>Ethoprop</td>
<td>✓</td>
<td>44</td>
<td>55-38-9</td>
<td>Fenithion</td>
<td>✓</td>
</tr>
<tr>
<td>15</td>
<td>134-62-3</td>
<td>Diethyltoluamide</td>
<td>✓</td>
<td>45</td>
<td>55283-68-6</td>
<td>Ethallfluralin</td>
<td>✓</td>
</tr>
<tr>
<td>16</td>
<td>138261-41-3</td>
<td>Imidacloprid</td>
<td>✓</td>
<td>46</td>
<td>56-72-4</td>
<td>Coumaphos</td>
<td>✓</td>
</tr>
<tr>
<td>17</td>
<td>145701-21-9</td>
<td>diclosulam</td>
<td>✓</td>
<td>47</td>
<td>66215-27-8</td>
<td>Cyromazine</td>
<td>✓</td>
</tr>
<tr>
<td>18</td>
<td>156052-68-5</td>
<td>Zoxamide</td>
<td></td>
<td>48</td>
<td>68359-37-5</td>
<td>Cyfluthrin</td>
<td>✓</td>
</tr>
<tr>
<td>19</td>
<td>1698-60-8</td>
<td>Chloridazon</td>
<td>✓</td>
<td>49</td>
<td>69377-81-7</td>
<td>Fluroxypr</td>
<td>✓</td>
</tr>
<tr>
<td>20</td>
<td>1702-17-6</td>
<td>Clopyralid</td>
<td>✓</td>
<td>50</td>
<td>74223-64-6</td>
<td>Metsulfuron-methyl</td>
<td>✓</td>
</tr>
<tr>
<td>21</td>
<td>173159-57-4</td>
<td>Foramsulfuron</td>
<td>✓</td>
<td>51</td>
<td>75-60-5</td>
<td>Cacodylic acid</td>
<td>✓</td>
</tr>
<tr>
<td>22</td>
<td>181274-15-7</td>
<td>Propoxy carbazole-sodium</td>
<td>✓</td>
<td>52</td>
<td>81335-37-7</td>
<td>Imazaquin</td>
<td>✓</td>
</tr>
<tr>
<td>23</td>
<td>19044-88-3</td>
<td>Oryzalin</td>
<td></td>
<td>53</td>
<td>83-79-4</td>
<td>Rotenone</td>
<td>✓</td>
</tr>
<tr>
<td>24</td>
<td>1912-24-9</td>
<td>atrazine</td>
<td>✓</td>
<td>54</td>
<td>84087-01-4</td>
<td>Quinclorac</td>
<td>✓</td>
</tr>
<tr>
<td>25</td>
<td>1918-00-9</td>
<td>dicamba</td>
<td></td>
<td>55</td>
<td>86-50-0</td>
<td>Azinphos-methyl</td>
<td>✓</td>
</tr>
<tr>
<td>26</td>
<td>199119-58-9</td>
<td>trifluousulfuron-sodium</td>
<td></td>
<td>56</td>
<td>94-74-6</td>
<td>MCPA</td>
<td>✓</td>
</tr>
<tr>
<td>27</td>
<td>21087-64-9</td>
<td>metribuzin</td>
<td>✓</td>
<td>57</td>
<td>95737-68-1</td>
<td>Pyriproxyfen</td>
<td>✓</td>
</tr>
<tr>
<td>28</td>
<td>2212-67-1</td>
<td>molinate</td>
<td>✓</td>
<td>58</td>
<td>96182-53-5</td>
<td>Tebufentin</td>
<td>✓</td>
</tr>
<tr>
<td>29</td>
<td>2222-92-6</td>
<td>fenamiphos</td>
<td>✓</td>
<td>59</td>
<td>96489-71-3</td>
<td>Pyridaben</td>
<td>✓</td>
</tr>
<tr>
<td>30</td>
<td>2278-23-3</td>
<td>bendiocarb</td>
<td>✓</td>
<td>60</td>
<td>97780-06-8</td>
<td>Ethamsulfuron methyl</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>61</td>
<td>98967-40-9</td>
<td>flumetsulam</td>
<td>✓</td>
</tr>
</tbody>
</table>
Table 2: Case group chemicals

<table>
<thead>
<tr>
<th>Number</th>
<th>CASRN</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>101-21-3</td>
<td>Chlorpropham</td>
</tr>
<tr>
<td>2</td>
<td>10453-86-8</td>
<td>Resmethrin</td>
</tr>
<tr>
<td>3</td>
<td>105512-06-9</td>
<td>Clodinafop-propargyl</td>
</tr>
<tr>
<td>4</td>
<td>113-48-4</td>
<td>MGK</td>
</tr>
<tr>
<td>5</td>
<td>117-81-7</td>
<td>Diethylhexyl phthalate</td>
</tr>
<tr>
<td>6</td>
<td>117337-19-6</td>
<td>Fluthiacet-methyl</td>
</tr>
<tr>
<td>7</td>
<td>121-75-6</td>
<td>Malathion</td>
</tr>
<tr>
<td>8</td>
<td>122-39-4</td>
<td>Diphenylamine</td>
</tr>
<tr>
<td>9</td>
<td>123312-89-0</td>
<td>Pyretrin</td>
</tr>
<tr>
<td>10</td>
<td>123343-16-8</td>
<td>Pyrinthioc-ac-sodium</td>
</tr>
<tr>
<td>11</td>
<td>129630-19-9</td>
<td>Pyraflumen-ethyly</td>
</tr>
<tr>
<td>12</td>
<td>131341-86-1</td>
<td>Fludioxonil</td>
</tr>
<tr>
<td>13</td>
<td>136-45-8</td>
<td>2.5-Pyrinedicarboxylic acid, dipropyl ester</td>
</tr>
<tr>
<td>14</td>
<td>141112-29-0</td>
<td>Isoxaflutole</td>
</tr>
<tr>
<td>15</td>
<td>148477-71-8</td>
<td>Spirodiclofen</td>
</tr>
<tr>
<td>16</td>
<td>149979-41-9</td>
<td>Tepraloxydin</td>
</tr>
<tr>
<td>17</td>
<td>161328-34-7</td>
<td>Fenamidone</td>
</tr>
<tr>
<td>18</td>
<td>19666-30-9</td>
<td>Oxadiazan</td>
</tr>
<tr>
<td>19</td>
<td>35554-44-0</td>
<td>Imazalil</td>
</tr>
<tr>
<td>20</td>
<td>50471-44-8</td>
<td>Vinclozolin</td>
</tr>
<tr>
<td>21</td>
<td>51-03-6</td>
<td>Piperonyl butoxide</td>
</tr>
<tr>
<td>22</td>
<td>51338-27-3</td>
<td>Diclofop-methyl</td>
</tr>
<tr>
<td>23</td>
<td>60168-88-9</td>
<td>Fenarimol</td>
</tr>
<tr>
<td>24</td>
<td>63-25-2</td>
<td>Carbaryl</td>
</tr>
<tr>
<td>25</td>
<td>68694-11-1</td>
<td>Triflumizole</td>
</tr>
<tr>
<td>26</td>
<td>69327-76-0</td>
<td>Buprofezin</td>
</tr>
<tr>
<td>27</td>
<td>72490-01-8</td>
<td>Fenoxy carb</td>
</tr>
<tr>
<td>28</td>
<td>74051-80-2</td>
<td>Sethoxydim</td>
</tr>
<tr>
<td>29</td>
<td>741-58-2</td>
<td>Bensoleide</td>
</tr>
<tr>
<td>30</td>
<td>76-87-9</td>
<td>Fentin</td>
</tr>
<tr>
<td>31</td>
<td>77501-63-4</td>
<td>Lactofen</td>
</tr>
<tr>
<td>32</td>
<td>79622-59-6</td>
<td>Fluazinam</td>
</tr>
<tr>
<td>33</td>
<td>97886-45-8</td>
<td>Dithiopyr</td>
</tr>
</tbody>
</table>
### Table 3: Summary of chemicals by classification (Provided by ToxCast)

<table>
<thead>
<tr>
<th>Chemical Class</th>
<th>Cases: No. (%)</th>
<th>Controls: No. (%)</th>
<th>Chemical Class</th>
<th>Cases: No. (%)</th>
<th>Controls: No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>aliphatic organothiophosphate</td>
<td>1 (0.03)</td>
<td>3 (0.05)</td>
<td>oxadiazolone</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>amide</td>
<td>0 (0.00)</td>
<td>6 (0.10)</td>
<td>oxazole</td>
<td>2 (0.06)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>anilide</td>
<td>0 (0.00)</td>
<td>3 (0.05)</td>
<td>oxime_carbamate</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>aromatic acid</td>
<td>1 (0.03)</td>
<td>3 (0.05)</td>
<td>phenoxy</td>
<td>2 (0.06)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>arsenical</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
<td>phenoxyacetic</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>aryloxyphenoxypropionic</td>
<td>2 (0.06)</td>
<td>0 (0.00)</td>
<td>phenoxypropionic</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>benzamide</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
<td>phenyl_organothiophosphate</td>
<td>0 (0.00)</td>
<td>3 (0.05)</td>
</tr>
<tr>
<td>benzoic acid</td>
<td>1 (0.03)</td>
<td>1 (0.02)</td>
<td>phosphoramidate</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>bridged_diphenyl</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
<td>phosphoramidothioate</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>Carbanilate</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
<td>phthalate</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>chlorotriazine</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
<td>picolinic_acid</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>conazole</td>
<td>3 (0.09)</td>
<td>1 (0.02)</td>
<td>pyrazole</td>
<td>1 (0.03)</td>
<td>2 (0.03)</td>
</tr>
<tr>
<td>cyclodiene</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
<td>pyrethroid</td>
<td>1 (0.03)</td>
<td>2 (0.03)</td>
</tr>
<tr>
<td>cyclohexene oxide</td>
<td>2 (0.06)</td>
<td>0 (0.00)</td>
<td>pyridazinone</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>cyclopropylisoxazole</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
<td>pyridine</td>
<td>2 (0.06)</td>
<td>2 (0.03)</td>
</tr>
<tr>
<td>dicarboximide</td>
<td>1 (0.03)</td>
<td>1 (0.02)</td>
<td>pyridylmethylamine</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>dichlorophenyl_dicarboximide</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
<td>pyrimidine</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>dinitroaniline</td>
<td>0 (0.00)</td>
<td>2 (0.03)</td>
<td>pyrimidinylsulfonylurea</td>
<td>0 (0.00)</td>
<td>2 (0.03)</td>
</tr>
<tr>
<td>diphenyl_ether</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
<td>pyrimidinylthiobenzoic_acid</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>formamidine</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
<td>pyrrole</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>heterocyclic_organothiophosphate</td>
<td>0 (0.00)</td>
<td>4 (0.07)</td>
<td>quaternary_ammonium</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>midazole</td>
<td>3 (0.09)</td>
<td>1 (0.02)</td>
<td>quinoline</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>midazolone</td>
<td>0 (0.00)</td>
<td>2 (0.03)</td>
<td>quinolinecarboxylic_acid</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>inorganic</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
<td>sulfonanilide</td>
<td>0 (0.00)</td>
<td>2 (0.03)</td>
</tr>
<tr>
<td>nicotinoid</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
<td>sulfonylurea</td>
<td>0 (0.00)</td>
<td>4 (0.07)</td>
</tr>
<tr>
<td>nitroguanidine</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
<td>tetronic_acid</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>nitrophenyl_ether</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
<td>thiocarbamate</td>
<td>0 (0.00)</td>
<td>2 (0.03)</td>
</tr>
<tr>
<td>organochlorine</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
<td>triazine</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>organophosphate</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
<td>triazinone</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>organophosphorus</td>
<td>2 (0.06)</td>
<td>12 (0.20)</td>
<td>triazinylsulfonylurea</td>
<td>0 (0.00)</td>
<td>4 (0.07)</td>
</tr>
<tr>
<td>organothiophosphate</td>
<td>1 (0.03)</td>
<td>10 (0.16)</td>
<td>triazole</td>
<td>0 (0.00)</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>organotin</td>
<td>1 (0.03)</td>
<td>0 (0.00)</td>
<td>triazolopyrimidine</td>
<td>0 (0.00)</td>
<td>2 (0.03)</td>
</tr>
</tbody>
</table>
Table 4: Student’s t-tests: physiochemical properties and transcription factors/nuclear receptors

<table>
<thead>
<tr>
<th>Physiochemical properties†</th>
<th>Controls</th>
<th>Cases</th>
<th>t value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logkhsa*</td>
<td>-0.15</td>
<td>0.22</td>
<td>3.58</td>
<td>0.001</td>
</tr>
<tr>
<td>LogP*</td>
<td>2.80</td>
<td>3.82</td>
<td>2.92</td>
<td>0.005</td>
</tr>
<tr>
<td>Molecular Weight*</td>
<td>295</td>
<td>334.17</td>
<td>2.37</td>
<td>0.020</td>
</tr>
<tr>
<td>Polarz*</td>
<td>27.96</td>
<td>32.65</td>
<td>3.15</td>
<td>0.002</td>
</tr>
<tr>
<td>SASA*</td>
<td>523.39</td>
<td>584.61</td>
<td>2.81</td>
<td>0.006</td>
</tr>
<tr>
<td>Water Solubility*</td>
<td>97632.77</td>
<td>2901.51</td>
<td>2.55</td>
<td>0.014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attagene (activity)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahr</td>
<td>1.417</td>
<td>1.213</td>
<td>0.63</td>
<td>0.532</td>
</tr>
<tr>
<td>CAR</td>
<td>1.118</td>
<td>1.165</td>
<td>0.99</td>
<td>0.325</td>
</tr>
<tr>
<td>e2f*</td>
<td>1.012</td>
<td>0.921</td>
<td>2.46</td>
<td>0.017</td>
</tr>
<tr>
<td>Fox01</td>
<td>0.950</td>
<td>0.902</td>
<td>1.36</td>
<td>0.178</td>
</tr>
<tr>
<td>Nf-κβ</td>
<td>1.140</td>
<td>1.250</td>
<td>1.68</td>
<td>0.010</td>
</tr>
<tr>
<td>Nrf2*</td>
<td>2.040</td>
<td>2.674</td>
<td>2.04</td>
<td>0.049</td>
</tr>
<tr>
<td>p53</td>
<td>1.081</td>
<td>1.063</td>
<td>0.40</td>
<td>0.689</td>
</tr>
<tr>
<td>PXR</td>
<td>1.705</td>
<td>1.930</td>
<td>1.00</td>
<td>0.325</td>
</tr>
</tbody>
</table>

† Only statistically significant physiochemical properties shown for sake of brevity.
Logkhsa: predicted probability of chemical binding to human serum albumin
Polarz: polarizability
SASA: Solvent accessible surface area

* Statistically significant at p≤0.05.
Student’s t-test carried out using unequal variances.
Table 5: AUC for physiochemical properties & Attagene assay

<table>
<thead>
<tr>
<th>Physiochemical properties</th>
<th>Predictor</th>
<th>AUC (ϴ)</th>
<th>CI (95%)</th>
<th>SE ϴ</th>
<th>p-value</th>
<th>X² (df=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Logkhsa</td>
<td>0.720</td>
<td>0.614 – 0.827</td>
<td>0.054</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>LogP</td>
<td>0.658</td>
<td>0.512 – 0.769</td>
<td>0.058</td>
<td>2.19</td>
<td>0.139</td>
</tr>
<tr>
<td></td>
<td>MW</td>
<td>0.627</td>
<td>0.515 – 0.740</td>
<td>0.515</td>
<td>2.65</td>
<td>0.104</td>
</tr>
<tr>
<td></td>
<td>Polarz</td>
<td>0.689</td>
<td>0.577 – 0.800</td>
<td>0.057</td>
<td>0.260</td>
<td>0.610</td>
</tr>
<tr>
<td></td>
<td>SASA</td>
<td>0.664</td>
<td>0.544 – 0.783</td>
<td>0.061</td>
<td>0.364</td>
<td>0.550</td>
</tr>
<tr>
<td></td>
<td>WS*</td>
<td>0.680</td>
<td>0.566 – 0.795</td>
<td>0.058</td>
<td>0.372</td>
<td>0.542</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attagene Assay (unweighted)</th>
<th>Predictor</th>
<th>AUC (ϴ)</th>
<th>CI (95%)</th>
<th>SE ϴ</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>e2f*</td>
<td>0.673</td>
<td>0.534 – 0.812</td>
<td>0.071</td>
<td>0.723</td>
</tr>
<tr>
<td></td>
<td>Nrf2</td>
<td>0.652</td>
<td>0.513 – 0.790</td>
<td>0.071</td>
<td>--</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attagene Assay (weighted, using 1/var)</th>
<th>Predictor</th>
<th>AUC (ϴ)</th>
<th>CI (95%)</th>
<th>SE ϴ</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>e2f*†</td>
<td>0.931</td>
<td>0.873 – 0.990</td>
<td>0.0295</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Nrf2</td>
<td>0.640</td>
<td>0.491 – 0.789</td>
<td>0.076</td>
<td>--</td>
</tr>
</tbody>
</table>

Physiochemical properties were compared to Logkhsa
*1-WS and 1-e2f was calculate so as to standardize summary statistics across predictors.
Logkhsa: predicted probability of chemical binding to human serum albumunin
Polarz: polarizability
SASA: Solvent accessible surface area

† Statistically significant at p≤0.05.
Student’s t-test carried out using unequal variances.
Table 6: e2f cutoff values (cross-section)

<table>
<thead>
<tr>
<th>Cutpoint (Wt adj. e2f)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Youden's Index</th>
<th>1-Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;= 39.429</td>
<td>0.957</td>
<td>0.630</td>
<td>0.587</td>
<td>0.370</td>
</tr>
<tr>
<td>&gt;= 39.577</td>
<td>0.913</td>
<td>0.630</td>
<td>0.543</td>
<td>0.370</td>
</tr>
<tr>
<td>&gt;= 39.684</td>
<td>0.913</td>
<td>0.652</td>
<td>0.565</td>
<td>0.348</td>
</tr>
<tr>
<td>&gt;= 40.450</td>
<td>0.913</td>
<td>0.696</td>
<td>0.609</td>
<td>0.304</td>
</tr>
<tr>
<td>&gt;= 40.610</td>
<td>0.913</td>
<td>0.717</td>
<td>0.630</td>
<td>0.283</td>
</tr>
<tr>
<td>&gt;= 40.686</td>
<td>0.913</td>
<td>0.739</td>
<td>0.652</td>
<td>0.261</td>
</tr>
<tr>
<td>&gt;= 40.705</td>
<td>0.913</td>
<td>0.761</td>
<td>0.674</td>
<td>0.239</td>
</tr>
<tr>
<td>&gt;= 41.437</td>
<td>0.913</td>
<td>0.783</td>
<td>0.696</td>
<td>0.217</td>
</tr>
<tr>
<td>&gt;= 42.49</td>
<td>0.913</td>
<td>0.804</td>
<td>0.717</td>
<td>0.196</td>
</tr>
<tr>
<td>&gt;= 42.780</td>
<td>0.913</td>
<td>0.826</td>
<td>0.739</td>
<td>0.174</td>
</tr>
<tr>
<td><strong>&gt;= 43.257</strong></td>
<td><strong>0.913</strong></td>
<td><strong>0.848</strong></td>
<td><strong>0.761</strong></td>
<td><strong>0.152</strong></td>
</tr>
<tr>
<td>&gt;= 43.272</td>
<td>0.870</td>
<td>0.848</td>
<td>0.717</td>
<td>0.152</td>
</tr>
<tr>
<td>&gt;= 44.485</td>
<td>0.826</td>
<td>0.848</td>
<td>0.674</td>
<td>0.152</td>
</tr>
<tr>
<td>&gt;= 44.69</td>
<td>0.783</td>
<td>0.848</td>
<td>0.630</td>
<td>0.152</td>
</tr>
<tr>
<td>&gt;= 44.981</td>
<td>0.739</td>
<td>0.848</td>
<td>0.587</td>
<td>0.152</td>
</tr>
<tr>
<td>&gt;= 45.540</td>
<td>0.739</td>
<td>0.870</td>
<td>0.609</td>
<td>0.130</td>
</tr>
<tr>
<td>&gt;= 45.823</td>
<td>0.739</td>
<td>0.891</td>
<td>0.630</td>
<td>0.109</td>
</tr>
<tr>
<td>&gt;= 45.937</td>
<td>0.696</td>
<td>0.891</td>
<td>0.587</td>
<td>0.109</td>
</tr>
<tr>
<td>&gt;= 46.293</td>
<td>0.696</td>
<td>0.913</td>
<td>0.609</td>
<td>0.087</td>
</tr>
<tr>
<td>&gt;= 46.914</td>
<td>0.696</td>
<td>0.935</td>
<td>0.631</td>
<td>0.065</td>
</tr>
<tr>
<td>&gt;= 46.965</td>
<td>0.696</td>
<td>0.957</td>
<td>0.652</td>
<td>0.044</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Figure 1: Study inclusion criteria

ToxRefDB
Cases: Chemicals that produced liver preneoplastic lesions and neoplasitic lesions in the liver at 1 mg/kg/d, x2yr, both rats and mice.
- 33 chemicals

Controls: Chemicals that did not produce preneoplastic nor neoplastic liver lesions at any dose, both rats and mice.
- 61 chemicals

ToxCast Phase I data
1) Physiochemical properties: Epi Suite and QikProp
   - Cases: 33 chemicals
   - Controls: 61 chemicals

2) Cell level assay: Cellumen
   - Cases: 33 chemicals
   - Controls: 61 chemicals

3) Transcription factor assay: Attagene
   - Cases: 24 chemicals
   - Controls: 46 chemicals
Figure 2: Data from Cellumen cell count

Cell count difference

Difference in cell count, absolute

Dose, uM, log transformed
Figure 3: Dose-Response of selected nuclear receptors and transcription factors
Figure 4: ROC curves, physiochemical properties
Figure 5: ROC curves of weighted e2f and Nrf2
Chapter VI: Conclusion

Introduction

The specific aims of this body of research were to articulate the regulatory process that the TSCA, New Chemicals Program must legally comply with. One aspect of this aim was to assess the potential applicability of the ToxCast system of assays in the assessment of hazard; specifically, accurate prediction of hepatocarcinogenesis. This body of research also sought to review and synthesize the current knowledge of network motifs as they apply to toxicology.

Summary of results

Policy analysis of the EPA’s New Chemicals Program

Analysis of the 33 regulatory decisions that resulted in consent orders being issued by the EPA found that the New Chemicals Program premanufacturing notification (PMN) review process relied on a variety of data to inform regulatory decisions, with most decisions based on analog analyses or in silico methods. Just over one half, or 17, of the PMN submissions included test data on human and/or environmental effects; 16 submissions did not provide any test data or results. There were a total of 142 test results reported on or summarized across 30 distinct test types (table 2 of chapter III). The EPA generated data to help inform decisions, most of which originated from aquatic toxicity models (table 3 of chapter III). The EPA also compiled known human health effects information on each PMN substance using the analog analysis approach described in chapter III. In 25 instances, or 76 percent of the time, the EPA issued consent orders on
the basis that the substance may present an unreasonable risk to both human health and the environment. The majority of decisions were based on analog analyses in the determination of human health risk; for environmental health risk, SAR (table 4 of chapter III). In 12 instances of determining unreasonable health risk, submitted test data primarily informed decisions. In all other determinations of human health risk, analog analyses were the primary information source in making the regulatory decision. Risk of pulmonary toxicity was the most often cited health concern, followed by reproductive and developmental toxicity (table 5 of chapter III). The PMN submitters rarely provided test results on these endpoints, demonstrating that the lack of alignment between the human health concerns identified by the EPA and the data provided by PMN submitters.

Network motifs and their potential application to toxicity testing
A natural outgrowth of the review of the literature was a systematic characterization of motifs; specifically, identification of those motifs found to recur across species and major types of biological networks (i.e., protein-protein interaction and gene regulation). Chapter IV provides an overview network motifs and graph theory.

The literature search resulted in the identification of 13 recurring motifs (Tables 2-6 of chapter IV). Of the 13 recurring motifs, there were 2 subtypes also found to be statistically significant.

Most motifs were found to recur across an evolutionary range of organisms. Most motifs were also found to recur in both the broad categories of gene regulation and protein-
protein interaction. There were various methodological nuances to this study of recurring motifs. Studies meeting inclusion criteria of this study used wide ranging, overlapping and distinct data sets, the impact of which does impact study results and are described more thoroughly in chapter IV. The choice of isomorphism and randomization methods used to identify motifs also impacts study results, which again are described in greater detail in chapter IV.

**ToxCast and its potential application to toxicity testing**

ToxCast is an EPA pilot program that continues to carry out a variety of tests on thousands of chemicals, supplemented by data that detail each chemical’s physiochemical properties. The ToxCast program completed their Phase I, “Proof of Concept,” phase in 2009, the fruits of which has been many research findings and publications. However, to date, no ToxCast assay or model has assessed with regard to prediction of apical endpoints *in vivo*. Chapter V therefore set out to contribute to the literature and answer the basic question: could ToxCast phase I data be used to accurately discriminate chemicals that whose exposure results in the presence or absence of an apical endpoint? Specifically, could ToxCast phase I data be used to discriminate chemicals that either do or do not produce liver lesions of *in vivo*, chronic toxicity tests. Chapter V details the methods applied to answer this research question.

Receiver operating characteristic analyses (ROC) of the physiochemical properties that were found to exhibit statistically significant differences between study groups revealed that the predicted binding of the chemical to human serum albumin exhibited the largest
area under the curve (table 5 and figure 4 of chapter V). All these predictors yielded AUCs and accompanying confidence intervals that modestly discriminated between the case and control groups. Comparisons of the AUCs across these physiochemical properties did not yield statistically significant differences.

ROC analysis of nuclear receptors and transcription factors selected for inclusion into this study yielded only two predictors of which their AUCs and accompanying confidence intervals modestly discriminated between the case and control groups (table 3 and figure 5 of chapter V). These two predictors were the e2f transcription factor and the Nrf2 nuclear receptor. Weight-adjusted values using the inverse of the variance revealed marked differences in the e2f and Nrf2 ROC curves, and the accompanying differences were statistically significant (table 5 and figure 5 of chapter V). The weighted AUC for e2f was 0.931, which suggests that the assay for this transcription factor is a stronger discriminator of the chemicals known to produce preneoplastic and neoplastic lesions of the liver versus chemicals that do not.

**Policy recommendations**

This body of research started with an analysis of the data used to inform regulatory decisions. Chapter III encapsulates the decision process of the EPA’s New Chemicals Program, and the information ultimately relied upon to inform regulatory decisions. The results articulated in chapter III found that the EPA primarily relies on alternative test methods to inform regulatory decisions. The tests actually submitted by PMN submitters and the human health concerns articulated by the EPA are not aligned. EPA generates its
own information to inform decisions. Further, the EPA must make decisions in a short time frame, making use of in vivo tests impractical. The EPA is also limited under their statutory authority to require tests to be carried out by PMN submitters and the tests results provided by these petitioners are of variable quality and not always consistent with established standards.

The method of analysis used to characterize the New Chemicals Program decision making process is important and provides much information for the researcher and public health advocate alike.

For the researcher, these findings mean that attempts to develop new in vivo tests, or refinements thereof would not likely be adopted and used by the New Chemicals Program to inform regulatory decisions. In vivo tests take more time to carry out than the EPA has to render regulatory decisions. The short time span in which decisions must be made necessitates the use of in vitro and in silico methods. Such alternative test methods can be applied relatively quickly; in a matter of just a few days. With no potential changes in the TSCA law in the foreseeable future, the EPA will continue to need to use or generate its own data to inform decisions. Therefore, a researcher’s efforts would be more wisely spent on the development and validation of methods and models that accommodate these particular needs. The EPA, for instance, realized this early in the implementation of the New Chemicals Program. Since TSCA’s implementation in 1979, the EPA has sponsored the development, validated and implemented EpiSuite and EcoSAR (1, 2). Combined, these computational approaches estimate physiochemical
properties of chemicals, environmental fate parameters and hazard of toxicity to aquatic species. These in silico models have been applied to inform many regulatory decisions over the past 30 years, chapter III of which highlight more recent decisions that were communicated to the public. The EPA needs more tools like this, so that they can make more informed decisions.

To improve policy making, is important for a researcher’s efforts to be relevant and of practical use. Otherwise, research results are left to the dust of the bookshelf or the computer hard drive. As noted in chapter III, recent regulatory decisions were driven in large part by three health concerns identified by the EPA: risk of pulmonary toxicity (pulmonary overload), developmental toxicity and reproductive toxicity. These three health concerns are also health concern areas where PMN submitters provided little to no data. Developmental and reproductive toxicity tests in particular are burdensome not only for EPA but also for PMN submitters. Analysis of such endpoints requires years of research and millions of dollars to carry out (3). One may therefore reasonably conclude that this would be a fruitful area to research and develop and validate a battery of in vitro assays or in silico models and is therefore a reasonable policy recommendation.

The method of analysis used to assess the PMN decision process is also informative for public health professionals and law makers. The abstraction of regulatory decisions lays out the decision process under study, specifies the information used to inform decisions and articulates the decisions made. TSCA has been criticized by many and calls for this law’s reform have been numerous (4, 5). However, the law’s specific weaknesses to date
have not been well articulated, as a result law makers have not had rigorous information, grounded in research, for which to use to construct sound TSCA revisions. The method of analysis used in chapter III serves to inform law makers on how this specific aspect of TSCA was actually implemented and is presently carried out. Such a method of analysis has the potential to make more informed law and regulatory implementations in the future by identifying gaps, inconsistencies and inefficiencies for which only law makers can address. Such information is also useful for the public health professional as a tool to promote improved public health policy and also to further educate the public on how decisions are actually made with regard to chemical safety.

The results of this policy study also clearly show there is much room for improvement in EPA’s decision making process. Improvements can in part be realized by the development, validation and implementation of alternative test methods better suited to the constraints the EPA must work within. Other improvements can only be realized by changes to the existing TSCA law. The EPA fundamentally lacks the authority to require PMN submitters to provide test data. Paradoxically, the EPA has the authority to regulate, using the least burdensome standard if, “the information is insufficient to permit a reasoned evaluation of the health and environmental effects (6).”

Testing as it applies to PMN submissions is covered under sections 4, 5 and 8 of TSCA. Sections 5 and 8 limit the information required to be supplied to the EPA and specifically excludes PMN submitters from being required to provide new environmental or human
health effects data to the EPA. This exclusion needs to be stricken from the law. This is this body of research’s second policy recommendation.

The EPA can require testing of a new chemical if it is covered under a test regulation implemented as part of section 4. The paradox here is obvious. The EPA must formulate a new testing rule and yet it through the regulatory review process. This is a time consuming process, and chemical manufacturers have delayed test rules by tying up such proposed rules in court challenges (7).

On the basis of this work, a tiered testing strategy could be a testing approach to consider for the TSCA new chemicals evaluation. Such an approach would better align health concerns with tests carried out. The manner in which a tiered testing strategy could be implemented can vary; however, the general approach is designed so each test in the sequence of tests is selected to complement the preceding tests, where tests chosen in each succeeding level are determined by the results in the previous level of testing. In effect, the choice of test is designed to generate more information only in those areas where such information is needed, resulting in more informed regulatory decisions. While the details of such an evaluation methodology would certainly need to be worked out, such an approach would better align testing to toxicity axes of concern and in parallel, still allow the quick progression of chemicals through regulatory review in a systematic and consistent manner. This tiered approach would also promote efficiencies, requiring only tests needed to inform decisions where there could be a specific safety concern. This publicly documented approach would further increase the transparency of
PMN submissions and further, increase the comfort of consumers regarding the safety of products built on chemicals developed. Adoption of such a model would require assistance from Congress, as the EPA is at present statutorily constrained.

**Research recommendations**

The notion of network motifs is a recent phenomenon, and their potential application to toxicity testing is in its infancy. This study identified 13 recurring motifs across species. These 13 recurring motifs suggest that they are more than just statistical anomalies; they potentially represent critical patterns of biological activity essential for cellular homeostasis.

Motifs that may exist within specific biological pathways, particularly in *H. sapiens* compared to other species, represent areas where further research is needed. Related, analytical tools now exist where specific motif subtypes may be identified. For instance, researchers may now use color techniques to discern motifs of the same node size with specific types of nodes and edges (e.g., FANMOD or similar tools). This is an important analytical capability for motif identification in more complex pathways like signaling or pathways involved in adaptation to environmental stimuli, as these pathways exhibit complex patterns of induction, inhibition and binding.

Near term opportunities to apply research to regulatory toxicity testing may possibly emerge from the ToxCast program. Results from this body of research indicate perturbation of the transcription factor e2f is a stronger predictor of hepatocarcinogenesis.
than the other transcription factors evaluated in this study. Considering methodological issues, further study on the role of e2f as a potential biomarker requires further research. Study of this topic represents a practical and focused use of high throughput data to inform risk assessment, particularly hazard identification.

Tangentially, the complexity of ToxCast’s data and accompanying assays is both an asset to this program and also an issue that warrants further research. ToxCast is now in its eighth year and at a point where it needs to demonstrate its relevance to EPA. The EPA is ultimately a regulator and the US’s lead in the regulation of chemicals. ToxCast’s data and methods need to be assessed with regard to apical endpoints of interest to the regulatory side of EPA with greater vigor. Further, those models that demonstrate the ability to accurately predict apical endpoints need to undergo validation, consistent with generally accepted validation standards. It is by this process that the EPA regulators can adopt such tests and methods. Failure to do so means that ToxCast runs the serious risk of becoming an artifact.

**Conclusion**

This dissertation explored more closely alternative testing methods, including their present use in regulatory decision making, along with the potential applicability of ongoing research to regulatory toxicity testing in the near and intermediate future.

The review of the literature found that the three most prominent organizations, internationally, involved in validation of alternative test methods place great emphasis on
the statistical analysis of variability; while, assessment of accuracy is not addressed as thoroughly. A broader search of the literature revealed that the preferred statistical approach by the scientific community to quantitatively assess accuracy is receiver operating characteristic analysis (ROC).

Analysis of EPA’s regulatory decision making process with respect to new chemicals found that most decisions are based on a combination of analog analysis and in silico methods. The alignment between tests actually submitted by PMN submitters and the human health concerns articulated by the EPA is poor. Further, the EPA must make decisions in a short time frame, making use of in vivo tests impractical. The EPA is also limited in its ability to require tests to be carried out by PMN submitters and the tests results provided by these petitioners are of variable quality and not always consistent with established standards.

While the notion of network motifs has entered into the lexicon of toxicology, its specific application has not yet emerged. It was realized in this body of research that there exist a finite number of motifs that recur across species. These statistical anomalies represent potentially important regulatory pathways. Their importance requires further research as to their potential application in toxicology.

Data originating from the EPA’s ToxCast program have yielded many results; however, these efforts have not been validated and translated into regulatory toxicology practice. Therefore, exploration of whether ToxCast phase I data could be used to discriminate and
predict the presence or absence liver lesions of *in vivo*, chronic toxicity tests was carried out. It was found that chemicals known to produce liver preneoplastic and neoplastic lesions *in vivo* exhibited specific physiochemical properties and further that two transcription factors, e2f and Nrf2, were found to exhibit greater activity. E2f was found to be the stronger of the two predictors.

The policy recommendations of this body of research are three-fold. First, there is a need for researcher to better align with policy realities to which the research could potentially apply. Second, there exist fundamental flaws in the testing and information submission requirements of TSCA’s New Chemicals Program. PMN submitters should not be precluded from providing data on environmental or human health effects. Finally, a tiered testing strategy may be a more efficient approach to evaluate new chemicals.

Overall, this policy work addressed one of EPA’s most important functions: toxicity evaluation of new chemicals, a function shown to be severely hampered by the provisions of the authorizing statute. The NAS has proposed that new testing modalities have the potential to improve the efficiency of toxicity testing, but the value of these potential methods remains unknown. This work found that an element of ToxCast can accurately discern chemicals that do or do not produce in vivo liver lesions. This body of dissertation research contributes to the literature in these many respects and a step forward in the advancement of regulatory toxicity testing.
References


5. Locke PA, Myers DB. A replacement-first approach to toxicity testing is necessary to successfully reauthorize TSCA. ALTEX 2011: 28: 266-271.


Curriculum vitae

Robert Borotkanics

Education
- Doctorate of Public Health (DrPH), Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, 2014.
- Masters of Science in Informatics, Johns Hopkins School of Medicine, Baltimore, MD, USA, 2010
- Masters of Public Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, 2007
- Bachelor of Science in Cognitive Science, Defiance College, Defiance, Ohio, USA, 1996

Professional Experience
- Assistant, Johns Hopkins Bloomberg School of Public Health March 2011 – present (part time)

Awards
- Director’s Award for Merit, AHRQ, HHS, 2010
- Secretary’s Award for Distinguished Service, HHS, 2003
- Team Recognition Award, Booz Allen Hamilton, 2001
- High Five Award, Booz Allen Hamilton, 2001

Publications
- Borotkanics R, Lehmann H. Network motifs that recur across species, including gene regulatory and protein-protein interaction networks. Archives of Toxicology: May 204.

Posters
- Borotkanics RJ, Ying J, Samore M. Antibiotic Resistant Escherichia Coli, Food Exposures and Household Clustering. 2010. 48th Infectious Disease Society of America Annual Meeting, Vancouver, Canada.