U.S. Environmental Protection Agency

Second ToxCast Data Summit

September 29–30, 2014 | U.S. EPA Building Auditorium C111 | 109 T.W. Alexander Drive
Research Triangle Park, North Carolina

Website

Poster Abstracts

Models

1 Application of ToxCast and ToxRefDB to Develop a Quantitative Model of Systemic Toxicity

Lisa Truong[1], Gladys Ouedraogo[2], Sophie Loisel-Joubert[2], Matt Martin[1]
[1]US EPA- ORD – NCCT , RTP, NC 27711
[2]L'Oréal, Paris, France

EPA's ToxCast program profiles the bioactivity of chemicals in a diverse set of ~700 high throughput screening (HTS) assays. In collaboration with L'Oreal, a quantitative model of systemic toxicity was developed using lowest effect levels (LEL) from ToxRefDB for 633 chemicals with HTS data, chemical fingerprints, and a subset with reverse toxicokinetic (RtK) data. Floor and ceiling performance baselines (95% Confidence Intervals) were estimated to be 4.6 and 2.8 orders of magnitude uncertainty (OMU), respectively based on historical LEL distributions and reproducibility across study type and species. An initial read-across model was developed using chemical fingerprints to identify structurally similar neighbor LEL values resulting in 3.9 OMU, a 1/5th reduction in model uncertainty based on our performance baselines. HTS data was then incorporated into the model using 74 groups of assays based on biology (ie: response data, gene families), technology annotation (assay mechanisms, signal directions), and assay confounders (oxidative stress, cytotoxicity). For each assay grouping, a mean activity value was computed and adjusted for confounders. Incorporating HTS data with read-across resulted in a 3.7 OMU, a total reduction in model uncertainty of 2/5th. RtK steady-state concentrations were then incorporated to adjust *in vitro* concentration (uM) to *in vivo* dose (mg/kg/day). Although RtK values were only available for a subset of the total chemical set (211), including RtK further lowered the overall model uncertainty to 3.4 OMU, roughly 3/5th of the total model uncertainty we expect to be able to reduce. Herein, we have identified a model that incorporates HTS (dynamics), read-across (chemistry) and RtK (kinetics) to predict systemic LEL harnessing and incorporating the power of both new and existing data. This abstract does not necessarily represent EPA policy.

2 Is It Even Rational to Think We Should Be Able to Predict Animal Liver Toxicity from *In vitro* ToxCast Assays?

Neal Cariello, Constantine Kreatsoulas, David Cooper
GlaxoSmithKline

There is a tendency to view the liver as a bag of hepatocytes, but nothing could be further from the truth. ~40% of the liver cells are not hepatocytes and these cells play major roles in modulating liver toxicity by affecting liver immunological responses, cytokine production, cell death, liver regeneration and bile acid homeostasis.

In addition, hepatocytes metabolize many compounds and the chemical changes can have profound effects on liver toxicity. Few ToxCast assays use cells with significant metabolizing ability. Considering the above, it may be naive to believe we have ToxCast *in vitro* systems of sufficient complexity to predict *in vivo* liver toxicity.

Chemicals in ToxRefDB affecting the liver were selected, and importantly, neoplastic lesions were excluded. Liver pathologists classified the chemicals as hepatotoxins or non-hepatotoxins based on the ToxRefDB field effect_desc. This field has over 100 unique entries, such as apoptosis, atrophy, hypertrophy, etc.

AC50s from six *in vitro* ToxCast assays were used in the following machine learning algorithms: recursive portioning, adaptive boosting, random forest, general linear model, support vector machine, neural networks.

Very importantly, the ToxCast AC50 *in vitro* data is extremely sparse, often with ~95% missing data. Various methods of data imputation, binning and dichotomization were used in an effort to model the sparse data.

Our analysis indicates that no combination of the 6 ToxCast assays and the 6 machine learning algorithms produced a model much better than chance alone.

However, a chemoinformatics approach may fare better. A classical QSAR experiment was conducted using the same curated and annotated ToxRefDB data set. Several hundred models were built using a variety of topological, physicochemical and/or path based descriptors paired with a variety of categorical classifier techniques. Training, test and validation sets were varied to approximate best and worst case extrapolation or interpolation. Judged by their kappa statistic, most models were poorly predictive but a few had modest predictive ability.

3 Comparing OASIS Estrogen/Androgen Receptor Binding QSAR Predictions to Results from ToxCast II Estrogen/Androgen Receptor Binding Assays

Barun Bhhatarai, Dan Wilson, Paul Price, Sue Marty, Amanda Maguire and Edward Carney
Toxicology Environmental Research and Consulting, The Dow Chemical Company

One goal of the US-EPA ToxCast program is to use a combination of in silico and high-throughput (HT) assay-based signatures to screen untested chemicals as a means of prioritizing *in vivo* testing. In ToxCast II, >1800 compounds were tested in HT endocrine screening assays. We compared the three respective NovaScreen HT assays for estrogen - ER (human, bovine and mouse) and androgen - AR (human, chimp and rat) receptor binding to the predictions of OASIS, a three-dimensional quantitative structure-activity relationship (QSAR) model developed to assess the binding affinity to the mammalian nuclear receptors. Analysis of the ER - QSAR model predictions indicated that the in domain chemicals in the three assay platforms' results have low Sensitivity (< 56%) but high Specificity (95%). The analysis of the AR - QSAR model predictions indicated that the in domain chemicals had very high Sensitivity (92-100%) and acceptable Specificity (70-81%). When HT results were restricted to compounds within the respective domain of the ER and AR - QSAR model and showing consistent agreement of ER and AR binding at AC50 < 1 µM for the three binding assays, the ER and AR - QSAR model accurately predicted binding for the parent compounds 100% of the time. These results suggests the OASIS ER/AR QSAR models can be used to screen potential ER/AR binding but highlight the need to better understand the concordance between HT assay platforms as well as the sensitivity of the models.

4 Using High-Content Imaging Data from ToxCast to Analyze Toxicological Tipping Points

Imran Shah/EPA, Woodrow Setzer/EPA, John Jack/NCSU, Keith Houck/EPA, Richard Judson/EPA, Thomas Knudsen/EPA, Jie Liu/ORISE, Matt Martin/EPA, David Reif/NCSU, Ann M. Richard/EPA, Russell S. Thomas/EPA, Kevin Crofton/EPA, David J. Dix/EPA, Robert Kavlock/EPA

Translating results obtained from high-throughput screening to risk assessment is vital for reducing dependence on animal testing. We studied the effects of 976 chemicals (ToxCast Phase I and II) in HepG2 cells using high-content imaging (HCI) to measure dose and time-dependent perturbations in p53, JNK, oxidative stress, cytoskeleton, mitochondria, and cell cycle. A novel computational model was developed to describe the dynamic response of the system as cell-state trajectories based on multidimensional HCI datastreams. Cell-state trajectories produced by 10 concentrations (0.4 to 200 µM) of 976 chemicals showed resilience of the HepG2 system in many cases, however, we also found "tipping points" in system recovery. Further analysis of trajectories identified dose-dependent transitions, or critical points, in system recovery for 340/976 chemicals. The critical concentration was generally 5-times lower than the concentration that produced cell loss. We believe that HCI can be used to reconstruct cell state trajectories, and provide insight into adaptation and resilience for *in vitro* systems. With additional research, cellular tipping points could be used to define an *in vitro* point of departure (PoD) for risk-based prioritization of environmental chemicals.

This work does not reflect US EPA policy.

5 High-Throughput Physiologically Based Toxicokinetic Models for ToxCast Chemicals

Robert Pearce, John Wambaugh, Cory Strope, Woodrow Setzer

Physiologically based toxicokinetic (PBTK) models aid in predicting exposure doses needed to create tissue concentrations equivalent to those identified as bioactive by ToxCast. We have implemented four empirical and physiologically-based toxicokinetic (TK) models within a new R software package, vLiverPBPK. For the thousands of chemicals without *in vivo* TK data, all four TK models were designed to be parameterized with high-throughput (HT) *in vitro* TK experiments and structure-based physico-chemical property predictions. Two general types of predictions – steady-state serum concentration resulting from repeated exposures for use in reverse toxicokinetic (RTK) studies, and prediction of TK time course metrics such as Cmax and time-integrated plasma concentration (Area Under the Curve or AUC) for evaluating model prediction by comparison to *in vivo* data. In predicting the concentrations of a chemical over time, the HTTK models primarily use *in vitro* data for both the fraction of chemical unbound to plasma and the hepatic clearance, as well as structure-derived physicochemical properties for the calculation of partition coefficients and ratios of blood flows and tissue volumes to body weight for the models with multiple compartments. We have performed simulation studies using the more sophisticated HTPBTK model to evaluate key assumptions in the simpler three-compartment, steady-state model used in previous RTK studies and have found that although the majority of chemicals reach steady state within seven weeks, some never reach steady state within a typical human lifespan. We were also able to predict average steady state concentrations resulting from discrete dosing with predictions based on the infusion dosing assumption used in previous RTK studies; many of the chemicals that quickly reached lower steady state concentrations reached maximum concentrations of more than double the average steady state concentration. The package can currently make predictions for 369 chemicals, including 75 pharmaceuticals and 294 ToxCast chemicals. We will use the package to distribute additional data as it becomes available.

6 Predicting Hepatotoxicity Using ToxCast *In vitro* Bioactivity and Chemical Structure

Jie Liu†‡, Kamel Mansouri†, Richard Judson†, Matthew T. Martin†, Minjun Chen§, Huixiao Hong§, Xiaowei Xu‡§, Russell Thomas†, Imran Shah†*

†National Center for Computational Toxicology, US EPA, RTP, NC, USA.

‡Department of Information Science, University of Arkansas at Little Rock, Little Rock, AR, USA.

§Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U.S. FDA, Jefferson, AR, USA

The U.S. EPA ToxCast™ program is screening thousands of environmental chemicals for bioactivity using hundreds of high-throughput *in vitro* assays to build predictive models of toxicity. We represented chemicals based on bioactivity and chemical structure descriptors then used supervised machine learning to predict their hepatotoxic effects. We systematically analyzed the relationship between predictive performance, the number and type of descriptors, and different classification algorithms. A set of 677 chemicals were represented by 711 *in vitro* bioactivity descriptors (from ToxCast assays), 4,379 chemical structure descriptors (from QikProp, OpenBabel, PADEL, and PubChem), and three hepatotoxicity categories (from animal studies). Hepatotoxicants were defined by rat liver histopathology observed after chronic chemical testing and grouped in hypertrophy (161), injury (101) and proliferative lesions (99). Classifiers were built using seven machine learning algorithms: linear discriminant analysis (LDA), naïve bayes (NB), ensemble learning (ENSMB), support vector machines (SVCL, SVCR), classification and regression trees (CART) and k-nearest neighbors (KNN). Classifiers of hepatotoxicity were built using chemical structure, bioactivity, and a hybrid representation. Predictive performance was evaluated using 10-fold cross-validation testing and in-loop, filter-based, feature subset selection. The best balanced predictive accuracy of classifiers constructed for hypertrophy, injury and proliferative lesions was 0.83 (CART/60 hybrid descriptors), 0.81 (CART/60 chemical structure descriptors), and 0.80 (CART/55 hybrid descriptors), respectively. ToxCast provides the largest and richest data set for mining linkages between the *in vitro* bioactivity of environmental chemicals and their adverse histopathological outcomes. Our findings

demonstrate the utility of high-throughput assays for characterizing rodent hepatotoxicants, the benefit of using hybrid representations that integrate bioactivity and chemical structure, and the need for objective evaluation of classification performance.

This abstract does not represent EPA policy.

7 Predicting Point-of-Departure Values from the ToxCast Data

Nisha S. Sipes, Richard Judson, John Wambaugh, Matt Martin, Rusty Thomas
US EPA/ORD/NCCT, Research Triangle Park, NC

There are less than two-thousand health assessments available for the tens of thousands of chemicals in commerce today. Traditional toxicity testing takes time, money, and resources leading in part to this large discrepancy. Faster and more efficient ways of understanding adverse chemical-biological interactions needed. We are evaluating ways to use the ToxCast high-throughput and high-content screening data to predict point-of-departure values, i.e., the dose-response point that marks the beginning of a low-dose extrapolation. First, we gathered lowest observed adverse effect levels (LOAEL) and highest dose tested (HDT) from ToxRefDB for the 239 ToxCast Phase I chemicals with human hepatic metabolic clearance and plasma protein binding data (Wetmore et al Toxicol Sci 2012). Next we converted calculated four ToxCast chemical specific values, such as lowest AC50, cytotoxicity, activity concentration at baseline (ACB), and activity concentration at the cutoff (ACC) values to oral equivalents using reverse toxicokinetic approaches. We compared the current LOAEL and HDT values with the calculated oral equivalents to evaluate the sensitivity of the ToxCast data with respect to the lowest doses where an effect is seen (LOAEL) or an estimate of the maximum tolerated doses (HDT). Initial analyses revealed that the four calculated oral equivalent doses from the ToxCast data (AC50, cytotoxicity, ACC, ACB) are generally more sensitive than the LOAEL or HDT values. The calculated cytotoxicity oral equivalents best fit a linear model with the slope closest to 1 (with respect to LOAEL) (r2=0.022). This finding indicates that cytotoxicity may provide a more precise estimate (over AC50, ACC, or ACB) of the dose at which we will first see an effect *in vivo*. Further work will explore other chemical specific values, limiting the analyses to specific assays or pathways, and endpoints. Applying bench mark dose modeling to the *in vivo* dose-response data will provide normalized bench mark doses. This work has the potential to provide rapid and efficient models that decrease uncertainties in predictions of points of departure for data-sparse chemicals. This abstract does not necessarily reflect US EPA policy.

8 Developing Predictions of *In vivo* Developmental Toxicity of ToxCast Chemicals Using Mouse Embryonic Stem Cells

S. Hunter, M. Rosen, M. Hoopes, H. Nichols, S. Jeffay, K. Chandler[1],
Integrated Systems Toxicology Division, National Health and Environmental Effects Research Laboratory, ORD, US EPA, RTP, NC
[1]Current Address: Office of Policy, Planning and Evaluation, National Institute of Environmental Health Sciences, NIH, RTP, NC
Contact: Hunter.sid@epa.gov

Developing predictive models of developmental toxicity is a focus of the virtual embryo project. Our experimental model uses a modified mouse embryonic stem cell (mESC) assay to assess chemical-induced cytotoxicity and altered differentiation. Two time points and differentiation outcomes are currently evaluated: cardiomyogenesis (MYH6/7) and gastrulation (GSC). ToxCast Phase I (TCP-I) compounds were evaluated using both differentiation time points, whereas, ToxCast Phase II (TCP-II) compounds were evaluated using the GSC biomarker alone. NCCT provided chemical stock solutions with 20µM being the highest concentration evaluated due to DMSO solvent toxicity. Forty-six percent (118/257) of TCP-I chemicals affected stem cells at the cardiomyogenic stage. Twenty-eight percent (302/1078) of TCP-I and II compounds affected stem cells at gastrulation. ToxPi tools were used to rank chemical potency. A subset of TCP-I chemicals, were identified as teratogen in rats or rabbits (ToxRefDB); 47% (80/170) of teratogenic chemicals produced effects in mESCs. Of the 118 TCP-1 chemicals that produced effects in mESCs, 80 (68%) are teratogens. Using cytotoxicity burst data, compounds were further classified as specific mESC toxicants

if they produced effects at concentrations < cytotoxicity burst. By combining gastrulation-stage cytotoxicity and cardiomyocyte differentiation data, 87% (26/30) of the most potent chemicals classified as specific mESC toxicants were teratogens. This association decreased to 73% as the potency of selective mESC toxicants decreased. Chandler et al., (PLoS One, 2011) developed a putative redox disrupting compound (pRDC) ToxPi based on analysis of TCP-I chemicals that affect cardiomyogenesis. Using this predictive model to evaluate TCP-II compounds, 71% (71/100) of the highest potency pRDCs and 2% (6/323) with no pRDC activity affected mESC at gastrulation. We have used mESCs as a model system to evaluate effects of ToxCast compounds and have shown an association between specific mESC toxicants and teratogens *in vivo*. Subsequent analysis will evaluate the relationship between constituents of the pRDC ToxPi and chemicals that are teratogenic *in vivo*. This abstract does not represent US EPA Policy.

9 Using Weighted Entropy to Rank Chemicals in ToxCast Phase II Data

Keith R. Shockley, National Institute for Environmental Health Sciences

Quantitative high throughput screening (qHTS) experiments simultaneously produce concentration-response profiles for thousands of chemicals. A metric that can effectively describe qHTS profiles is useful to rank chemicals for prioritization. Rankings are usually based on parameter estimates derived from fitting the data to a pre-specified model, such as the logistic Hill Equation model. For example, chemicals are often ranked based on AC50 parameter estimates calculated from the Hill Equation model, where an AC50 value represents the concentration at which the response is 50% of maximal induction. However, the estimated AC50 can be unreliable when response data does not adequately fit the assumed model structure. The weighted entropy score (WES) is an alternative metric describing the average activity level that can be calculated from the observed responses and the assay detection limits. Here, we use WES to rank chemicals tested in ToxCast Phase II efforts and uncover correlated response patterns. Area under receiver operator curves for data simulated from the four-parameter Hill equation showed that profile rankings based on WES outperform rankings based on AC50 estimates for a wide range of conditions and typical levels of qHTS measurement error (sigma = 5% and sigma =10%).

10 Stratifying Environmental Chemicals in the ToxCast I Library for Cancer Cell Proliferative Signature

Scott J. Sauer - Department of Surgery - Duke University Medical Center, Akshay V. Save - Department of Surgery - Duke University, Kevin P. Williams - Department of Pharmaceutical Sciences - BRITE - North Carolina Central University, Imran Shah - National Center for Computational Toxicology - Environmental Protection Agency, Gayathri R. Devi - Department of Surgery and Duke Cancer Institute - Duke University

The impact of environmental agents on cancer cell behavior through changes in cell proliferation and survival by modulation of signaling pathways is not well understood. Epidemiological studies have shown a correlation between pesticide or insecticide exposure and increased incidence of cancer. The purpose of this study was to develop a multi-step strategy for rapid, reproducible screening of compounds from the EPA ToxCast I library using both high-content and high-throughput cell based and cancer-specific proliferative/signaling assays.

We initially downselected the ~300 compound library to 31 high-priority compounds based on positive assay results from the ToxCast I database for endpoints of interest (epidermal growth factor receptor [EGFR], oxidative stress, AKT1, NFKB activity) with EC50 values ≤ 10 µM.

These compounds were categorized using a high-throughput high-content analysis using a breast cancer cell model for identification of compounds that increased, decreased or had no effect on nuclear count/cell number. These compounds were also analyzed in parallel with a high-throughput MTT proliferation assay as a secondary screen. Nine representative compounds with diverse applications that either increased cell number, had no effect or were cytotoxic were evaluated by a multiparametric high-content analysis. We employed Hoechst, YOYO-1 and TMRE dyes to determine the impact on cell number, health and mitochondrial function, respectively. These 9 chemicals were then characterized for their effect on cancer specific assays, such as colony forming and anchorage-independent growth, followed by western

immunoblot analysis of their impact on the proliferative signaling markers extracellular regulated kinase (ERK) and EGFR. This multi-step stratification of environmental compounds in the ToxCast I library using imaging, quantitative cell morphological parameters along with cancer-phenotypic and signaling assays allowed for a mechanism-based identification of a subset of compounds that have the potential to increase cancer cell proliferative signaling.

Funding support by: DCI Cancer and Environment Initiative P3917733, Department of Surgery Bolognesi award, ACS RSG-08-290-01-CCE (GRD) and NCI training grant T32CA009111 (SJS).

11 Predicting Acute Toxicity Using *In vitro* ToxCast™ HTS Mitochondrial Inhibition Assays

Barun Bhhatarai, Dan Wilson, Paul Price, Mike Bartels, Shubhra Chaudhuri and Ed Carney
Toxicology Environmental Research and Consulting
The Dow Chemical Company

Mitochondrial inhibition is a mechanism known to drive acute toxicity for certain chemicals. High Throughput Screening (HTS) assays have been developed to test if a chemical's toxicity operates by this mechanism. We hypothesized that, for chemicals that cause mitochondrial inhibition in HTS assays, acute toxicity is conserved across invertebrate, aquatic and mammalian species, suggesting that 1) in the absence of pre-systemic metabolism or limited absorption, *in vitro* mechanistic data could predict responses in multiple species, 2) under conditions of similar bioavailability, concordance of dose response between species would be high, and 3) predictions of oral toxicity from HTS assays routes would often be confounded by chemical-specific differences in uptake and metabolism. To validate our hypothesis, we determined whether, 1) *in vitro* data for mitochondrial inhibition generated by the US-EPA ToxCast™ program was correlated with various measures of acute toxicity, 2) read-across could be used between rat, daphnia and fish acute toxicity for those substances that cause mitochondrial inhibition in *in vitro* systems, and 3) incorporation of predictions of gastro-intestinal absorption and first pass metabolism improved correlations between acute toxicity observed by the oral and intravenous routes. We observed that, 1) plots of mitochondrial inhibition versus acute toxicity showed a well-delineated cusp (upper boundary) for daphnia, fish and rat intravenous data but were scattered with no delineation of a cusp for rat oral data, 2) simulations of fractional absorption and systemic bioavailability (including metabolism for three Cyp enzymes-2C, 2D, 3A) and hydrolysis improved trending of the rat oral data towards definition of a cusp, 3) plots of acute toxicity showed strong concordance between fish, daphnia and rat intravenous data but not rat oral data consistent with compounds being ~ 25-fold on average more toxic intravenously than orally, 4) mitochondrial inhibition at low concentrations predicted high acute toxicity in fish and daphnia but not in the rat because of often limited oral bioavailability. Predictive models for non-Cyp phase 1 and 2 metabolism are not available and would likely provide more insight. We believe that these findings demonstrate that HTS assays of mitochondrial toxicity can be used to predict high acute systemic toxicity in multiple species, and with proper in silico data on absorption and first pass metabolism, can predict acute oral toxicity in mammals. Such predictions could reduce the need for nonclinical regulatory safety testing for acute toxicity.

12 Applying an Active Machine Learning Process to Build Predictive Models of *In vivo* Toxicity from ToxCast Screening Data

Richard J. Brennan[1], Joshua D. Kangas[2], Friedemann Schmidt[3], Richard Khan-Malek[4], and Douglas A. Keller[4]
[1]Sanofi, Waltham, MA, USA
[2]Quantitative Medicine, Gainesville, FL, USA
[3]Sanofi, Frankfurt, Germany
[4]Sanofi, Bridgewater, NJ, USA

We have rigorously mined all available *in vivo* toxicity data for the ToxCast Phase II chemical library to identify high-confidence training sets for pathological findings of concern to preclinical drug discovery programs. An iterative machine learning process called Active Machine Learning was applied, using CoRE™ technology from Quantitative Medicine, to identify assay sets from the ToxCast panel of assays with predictive power for 7 *in vivo* endpoints of varying specificity: liver injury, hepatocelluar necrosis (with and

without serum ALT increase, or serum ALT + bilirubin increase), bile duct hyperplasia, renal tubular injury, and necrosis (any systemic finding). Active Machine Learning is an approach to directing experimentation that seeks to iteratively identify the most informative experiments prior to execution. This controlled learning process was applied to identify informative subsets of *in vitro* results (compound/assay combinations), and predictive models built at each iteration using the limited dataset exposed. Model performance was evaluated by validation against a non-incorporated portion of the *in vivo* results at each iteration, by measuring sensitivity, specificity, and area under the curve of the ROC plot. Model building continued until the models appeared to stabilize based on measured changes in model outputs. The best model for each endpoint was then tested against a separate set of forward-validation compounds with known, but blinded (to the model) activity. Performance of the resulting best models, and the ability of the CoRE™ Active Machine Learning process to efficiently identify the most informative training sets will be discussed.

13 Gas Phase Probe Molecules for Assessing *In vitro* Metabolism to Infer an *In vivo* Response

Michelle M. Angrish[1], Michael C. Madden[2], and Joachim D. Pleil[3]
[1]ORISE Research Fellow, US EPA, Research Triangle Park, NC
[2]Environmental Public Health Division, NHERL/ORD, US EPA, Chapel Hill, NC
[3]Human Exposure and Atmospheric Sciences Division, NERL/ORD, US EPA, Research Triangle Park, NC

Efficient and accurate *in vitro* high-throughput screening (HTS) methods use cellular and molecular based adverse outcome pathways (AOPs) as central elements for exposure assessment and chemical prioritization. However, not all AOPs are based on human or animal systems biology, but rather supported by *in vitro* to *in vivo* extrapolation and other computational modeling. The challenge is to develop unambiguous quantitative links between *in vitro* responses and corresponding *in vivo* effects. The use of gas phase probe molecules (PrMs) supported by relevant human exposure studies and pharmacokinetic (PK) parameters may address this gap. Furthermore, existing HTS assays that require liquid handling robotics would be complemented by quantitative ultrasensitive gas phase PrM assays. We previously determined the kinetic parameters for methyl-tertiary butyl ether (MTBE) metabolism to tertiary butyl alcohol (TBA) via CYP2A6 and sevoflurane metabolism to hexafluoroisopropanol via CYP2E1 pathways in the liver from human empirical data. In this study, we constructed a one-compartment PK model based on differential equations to estimate MTBE and SEV probe pathways for establishing steady state *in vitro* liver function. Because the MTBE and SEV metabolic pathways are well characterized from *in vivo* data, we can use them as PrM to explore the effects of chemicals of interest on their respective CYP pathways. We found that the PrM concept may provide a quantitative real time measurement from air of an *in vitro* response with a well-defined and corresponding *in vivo* effect. PrM methodology could be easily applied to a broad range of *in vitro* cell models and would provide a novel approach to assess chemicals of concern including endocrine disruptors, air toxics, and particulate matter.

14 Computational Embryology and Computational Modeling of Embryonic Limb Development Using ToxCast™High-Throughput Screening Data for Predictive Toxicology

*BK Ahir[1], NS Sipes[1], NC Baker[2], MCK Leung[1], RS DeWoskin[3], RM Spencer[2], RS Judson[1], MT Martin[1] and TB Knudsen[1]**
US-EPA/ORD/[1]NCCT/[2]Lockheed Martin/[3]NCEA, RTP, NC, USA

Skeletal defects are one of the major adverse outcomes observed across many prenatal developmental toxicity studies. We mined the ToxCast database (ToxCastDB) for *in vitro* chemical-bioactivity profiles that significantly correlated with skeletal defects in the Toxicity Reference Database (ToxRefDB). Using high-throughput screening (HTS) data from >1060 chemicals tested in 860 assays, 734 chemicals had *in vivo* developmental toxicity data in ToxRefDB and 44 (4-5%) produced fetal limb defects. In addition, we identified 112 ToxCast chemicals not in ToxRefDB that according to published articles produced limb defects. Significant univariate associations (e.g., assay-endpoint) were used to filter HTS assays based on statistical correlation ($p<0.05$) with distinct *in vivo* developmental limb defects. Retinoic acid receptors (RARs) assays emerged as the top target and were activated by 17 ToxCast chemicals. Retinoic acid (RA)

signaling plays an essential role in limb patterning as a ligand for nuclear RA receptors. We incorporated RA signaling into a multicellular agent-based model (ABM) of early limb development in CompuCell 3D. The ABM stimulates complex cellular interactions (adhesion, apoptosis, chemotaxis, migration, mitosis, secretion) through formation of apical epidermal ridge, zone of polarizing activity, and expansion of mesenchyme driven by morphogenetic signals (BMPs, FGFs, SHH, RA). To evaluate the model we selected two known prototypes: Dieldrin (AC50 = 4.8 µM on RARγ, and 0.974 µM on RARα) and Aldrin (AC50 = 45.5 µM on RARγ, and inactive on RARα) based on ToxCast data. Simulating the impact of RAR disruption via down-regulating GREM1 expression by these compounds on other pathways (subsequently impact on other signaling e.g. SHH, FGFs) in the limb model can provide insight into the spatial-temporal dynamics of altered limb development as a tool for predictive toxicology. [This abstract does not necessarily reflect EPA policy.]

Adverse Outcome Pathways

15 Adverse Outcome Pathway for Embryonic Vascular Disruption: ToxCast HTS Predictive Model Qualified by a Validated Human Angiogenesis Assay

Knudsen TB, Kleinstreuer NC**, Heinonen T****
** U.S. EPA/ORD/NCCT, RTP, NC*
*** ILS/NICEATM, RTP, NC*
**** FICAM, University of Tampere, Tampere, Finland*

Project 1.6 in the OECD AOP development programme workplan for 2013 addresses molecular initiating events leading to disruption of blood vessel formation and remodeling, intended for use in the predictive toxicology of developmental hazards and translating ToxCast HTS data. The cardiovascular system is the first to function in the embryo. Shaped by genetic signals, physiological cues and mechanical forces in the micro-environment, it is sensitive to chemicals that disrupt receptor tyrosine kinases (e.g., VEGFR-2), G-protein coupled receptors (e.g., CCR2), and GPI-anchored receptors (e.g. uPAR). The AOP was anchored to the MGI mammalian phenotype ontology browser [http://www.informatics.jax.org/] and 25 functional molecular targets in ToxCast. It was used to rank 1060 chemicals by predicting anti-angiogenic activity. A subset of 36 ToxCast chemicals was tested in an *in vitro* human cell-based angiogenesis assay (FICAM) co-culturing human endothelial cells (HUVEC) and fibroblasts. Effects (10 pM to 10 mM) were assessed by neutral red uptake (24h cytotoxicity) and, for concentrations that produced less than 20% cytotoxicity, endothelial tubule formation (6-days). The five most potent anti-angiogenic compounds were (FICAM AC50 in uM; ToxCast rank # in 1060): Pyridaben (0.005; #75); Cladribine (0.6; #179); Triclocarban (1.4; #201); Triclosan (1.7; #8); and Octyl gallate (0.8; #28). The five with the least rank across ToxCast were: Decane (#957; no tubule effect); Diethanolamine (#925; 1788.2 uM); Methimazole (#889; no tubule effect); TCEP (#839; no tubule effect) and Pymetrozine (#837; no tubule effect). Model sensitivity (0.91), specificity (0.79), and balanced accuracy (0.85) were excellent. A subset of compounds were also run in a CompuCell3D computer model that simulates an emergent endothelial network (angiogenic sprouting, exploratory filopodia, growth, apoptosis, connectivity). To date the simulation predicted anti-angiogenic activity for: Pyridaben; Octyl gallate; Fluazinam (2.5 uM; #82); Disulfiram (23.2 uM; #47); Bisphenol A (47.9 uM; #193); and Oxytetracycline dihydrate (200 uM; #319). It predicted non anti-angiogenic activity for Imazamox (#770; no tubule effect) but not PFOS (#48; no tubule effect). A simulated concentration response for Octyl gallate recapitulated the FICAM assay, with visible disruption at 1 uM. Overall, *in vitro* and in silico results were concordant with the ToxCast predictive model and bolster weight-of-evidence for plausibility of key event relationships for the AOP framework. DISCLAIMER: this abstract does not necessarily reflect US EPA policy.

16 Identifying Adverse Outcome Pathways Through the Integration of High-Throughput *In vitro* Assays and Corresponding *In vivo* Data: A Case Study in Integration of ToxCast, ToxRefDB and a Medium-Throughput *In vivo* Zebrafish Screen

Noffisat O. Oki^, Lisa Truong*~, Matthew Martin~, Richard Judson~, and Stephen W. Edwards^*
** Oak Ridge Institute for Science and Education*
^ Integrated Systems Toxicology Division, National Health and Environmental Effects Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA
~ National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA

High-throughput screening results are a cost effective way of gathering toxicity information for a variety of chemicals, but the connection of these assays with commonly known adverse outcomes is often unknown. The Adverse Outcome Pathway (AOP) framework is a useful tool for making these biological connections, as it summarizes the key information needed across the different levels of biological organization. Computational approaches to speed the assembly of these AOPs will enhance our ability to interpret the results from HTS screens such as ToxCast. For this study, we utilize data from a medium-throughput screen using zebrafish to better connect ToxCast assays with rodent endpoints from ToxRefDB as this should provide additional information in interpreting ToxCast results.

We integrated information from the U.S. EPA ToxCast dataset and *in vivo* data from a zebrafish developmental study to determine which assays are associated with endpoints from both zebrafish and another the U.S. EPA ToxRefDB rat *in vivo* data. The experimental design of the zebrafish dataset gives an added value by allowing the use of these data as an intermediary for finding candidate associations between the ToxCast assays and the ToxRefDB phenotypes. We use Frequent Itemset Mining to make our associations between datasets by using the chemicals as our “market basket” or aggregating variable for the analysis. Rather than drawing conclusions about specific chemicals, the focus of this analysis is on connecting the high-throughput molecular/cellular effects with organism-level adverse outcomes as an initial step towards the development of AOPs.

By using the zebrafish data as an intermediary, we connected phenotypes between two organisms (zebrafish and rat) as well as the assays predictive for these shared associations. The results also suggest zebrafish assays that may be candidates for more rigorous follow-up testing of certain HTS results. These methods are one component of a workflow being designed to fill in missing information in the AOP framework for the purpose of translating HTS screening data to adverse human health and ecological outcomes.

Disclaimer: The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA

17 Using Putative AOPs from High Content Data to Inform High Throughput Chemical Screening

Shannon M. Bell, Oak Ridge Institute for Science and Education, Stephen W. Edwards,U.S. Environmental Protection Agency

Adverse outcome pathways (AOPs) are useful tools for connecting molecular changes in response to chemicals, which can be monitored via high throughput screening (HTS), to an adverse outcome of regulatory concern. Development of detailed AOPs from traditional experimental results is a slow, tedious process and is unrealistic when covering the breadth of perturbations in response to the >83,000 chemicals in commerce. Large toxicogenomic screening studies, such as the Toxicogenomics Project-Genomics Assisted Toxicity Evaluation system (TG-GATEs), offer an opportunity to link molecular changes in response to chemical exposure to adverse outcomes. High throughput chemical screening data, like that in ToxCast, helps to provide information on the potential target AOPs of environmental chemicals. In this work we

illustrate how putative AOPs developed using toxicogenomic data can be used with HTS data from ToxCast to distinguish putative chemical modes of action and identify additional targets for first and second-tier screening assays using the case study of a putative AOP for fatty liver.

This is an abstract or a proposed presentation and does not necessarily reflect EPA policy.

Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

18 Integration of ToxCast and ToxRefDB in the Development of a Mode of Action (MoA)-Based Classification Model for Hepatotoxicity

Elisabeth Joossens, Julien Burton, Tomislav Horvat, Brigitte Landesmann, Alfonso Lostia
Institute for Health and Consumer Protection, European Commission Joint Research Centre, 21027 Ispra (VA), Italy

SEURAT-1 is a public – private partnership between the European Commission and Cosmetics Europe to develop a long term research strategy for the development of new non-animal test systems in the field of repeated dose systemic toxicity for the innovative assessment of human safety. Partners are now undertaking a proof-of-concept exercise to demonstrate how a mode-of-action based approach can be used to predict aspects of repeated dose target organ toxicity.

This poster presents the chemical selection strategy followed and the initial steps of the experimental design as a part of a SEUART-1 case study which aims to develop a Mode of Action (MoA)-based classification model to distinguish between potential hepatotoxicants and non-hepatotoxicants. The classification model should be constructed such that it allows correct discrimination between hepatotoxic and non-hepatotoxic chemicals with a minimal probability of false negatives using a test set of 92 chemicals of which 75% are positive hepatotoxicants.

To identify a robust set of chemicals known to cause hepatotoxicity at the initial phase literature information from human data from pharmaceuticals was used. This data set has been expanded with additional non-pharmaceutical chemicals from ToxRefDB and/or TG-GATE. The additional chemicals are selected on the basis of their known *in vivo* liver activity also ensuring that they provide a good coverage of the chemical space in the light of Structure Activity Relationship principles. A second phase of the study concentrates on an analysis of the available *in vitro* toxicity data for the selected chemicals and an optimization of the experimental design using ToxCast data. By extracting information on the chemicals for ToxCast liver-relevant assays, including the number of active assays and level of activity (AC50) in each assay the data could be used in selection of the appropriate dose levels for optimizing the experimental design.

19 Mitochondrial Function in Mesc Differentiation: An Endpoint for ToxCast Predictive Modeling

Joyce E. Royland, Sarah H. Warren, Susan Jeffay, Maria R. Hoopes , Harriette P. Nichols, Sid Hunter
U.S. Environmental Protection Agency, Integrated Systems Toxicology Division, 109 TW Alexander Dr., B105-03, Research Triangle Park, NC 27711, USA

Differentiating embryonic stem cells (ESCs) undergo mitochondrial maturation leading to a switch from a system dependent upon glycolysis to a reliance on oxidative phosphorylation. This switch acts as a signal to the cell to decrease the rate of proliferation and initiate differentiation into downstream cell lineages, a critical step in embryogenesis and a potential target for disruption by xenobiotic compounds. The research characterizes the cellular responses to specific mechanisms of mitochondrial Ox-Phos perturbation. Since mitochondrial dysregulation is a component of the putative Redox Disrupting Chemical ToxPi (Chandler et al., PLoS ONE 2012), this research adds to our understanding of predictive models for altered mESC differentiation and ultimately *in vivo* teratogenesis. We investigated the sensitivity of the developmental metabolic shift to interruption of oxidative phosphorylation at various points of the mitochondrial electron transport chain in a differentiating mouse ESC model system. The impacts of exposures to chemical inhibitors of electron transport complexes were assessed using in-cell Western analysis of the differentiation biomarker, Goosecoid (GSC), a transcription factor highly expressed during embryonic

gastrulation. Additional fluorescent bioassays were used to monitor cell number, mitochondria number and cell death. The effect on cell proliferation by inhibitor exposure varied depending on the complex tested. We found inhibitors of complex I led to a 91% decrease in cell number, which was > complex V (69% decrease) ≈ complex IV (66% decrease) > complex III (25% decrease) ≈ complex II (21% decrease). The pattern of altered differentiation as measured by GSC was similar (complex I inhibitor, 97%; complexes IV and V, ~50%; and little effect in complexes II and III, 12 and <1%, respectively). In contrast, inhibitors of electron transport at complexes I, IV and V showed increased number of mitochondria/cell in the surviving cell population. Mitochondria per cell decreased following inhibition of complexes II and III (~56%). These data will be combined with other measures of mitochondrial function to help define a mitochondrial adverse outcome pathway, to expand the application of the mESC developmental model and to aid in developing in silico models for toxicity testing. (This abstract does not necessarily reflect USEPA policy).

20 Causal Inferences from ToxCast Data: Molecular Pathways and Cellular Processes for Cleft Palate

Nancy C. Baker[1], Nisha S. Sipes[2], Christopher M. Grulke[1], Richard S. Judson[2], Thomas B. Knudsen[2]
[1]Lockheed Martin, RTP, NC
[2]National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, RTP, NC

Sixty-two chemicals in the ToxCast high-throughput screening (HTS) dataset have been linked to cleft palate based on data from ToxRefDB (rat or rabbit prenatal developmental toxicity studies) or from literature reports. These compounds are structurally diverse and thus likely to perturb prenatal development in mechanistically diverse ways. Integration of the HTS *in vitro* profiling data with information from chemotype profiling and automated literature survey provides a generalizable approach for adverse outcome pathway (AOP) elucidation. We generated a heatmap by clustering the 62 chemicals using as attributes 287 ToxCast Gene Scores and 228 chemistry structural elements (chemotypes). Hierarchical relationships in the heatmap revealed several cohesive bioactivity-chemotype clusters. For example, the conazoles, retinoids, and phthalates formed clusters, as did the chemicals that hit GPCRs, angiogenic targets, and neuro-active targets. These cohesive clusters will enable focused literature mining for cellular and tissue effects linked to the ToxCast targets and enable extension and enhancement of AOPs for cleft palate. This abstract does not necessarily represent U.S. EPA policy.

21 Quantitative High-Throughput Screening and Confirmation Studies for Identification of Compounds that Activate the Aryl Hydrocarbon Receptor Pathway

J.A. Franzosa (U.S. EPA/ORD/NCCT), K. Mansouri (U.S. EPA/ORD/NCCT), A. A. Soshilov (U.C. Davis), J.H.Hsieh (NTP/NIEHS), S.S. Ferguson (NTP/NIEHS) , X. Chang (ILS/NICEATM), N. C. Kleinstreurer (ILS/NICEATM), J. Hamm (ILS/NICEATM), D.L. Villeneuve (U.S. EPA/ORD/NCCT), R.S. Judson (U.S. EPA/ORD/NCCT), R.S. Thomas (U.S. EPA/ORD/NCCT), R. Huang (NIH/NCATS), M. Xia (NIH/NCATS), S.W. Kullman (NCSU), M.S. Denison (U.C. Davis), K.A. Houck (U.S. EPA/ORD/NCCT)

The aryl hydrocarbon receptor (AhR) is a transcription factor that mediates adaptive responses to known environmental pollutants, such as aromatic hydrocarbons, through regulation of Phase I and II xenobiotic metabolizing enzymes as well as important growth and differentiation pathways. The AhR is activated by a wide range of structurally diverse chemicals that include naturally occurring, synthetic, and environmental chemicals. As part of the U.S. Tox21 collaboration, an approximately 8.5K chemical library was profiled in a quantitative high-throughput screening (qHTS) platform in a 1536-well plate format with an AhR luciferase reporter gene assay in human HepG2 cells at 15 concentrations run in triplicate. Results from the qHTS data identified 768 compounds as potential AhR ligands; of which ~300 demonstrated activity at less than 10 µM. Artificial neural networks were used to train supervised and unsupervised self-organizing maps to identify structural features associated with potency and efficacy of AhR activity. A subset of the Tox21 library (~1800 compounds) and 24 surface water samples collected from various locations across the U.S. were also screened through the EPA ToxCast Program in the Attagene Factorial assay, which evaluates the activity of transcription factors using reporter DNA constructs in HepG2 cells. The *in vitro* library data combined with in

silico (ADMET-Predictor) predictions of metabolic clearance rates and literature studies, were used to guide the selection of 50 compounds with a range of potencies, efficacies, structures and clearance rates to be tested in orthogonal *in vitro* and *in vivo* assays. Results of *in vitro*, DNA binding (GRAB bioassay) and ligand binding (HAP assay), and preliminary *in vivo* evaluations in teleost models, were used to confirm the qHTS data. Taken together, these data support increased confidence in the use of robotic qHTS data to predict AhR activity. In addition, many of the environmental water samples elicited AhR activity along with various other endpoints measured. These results show the application of qHTS assays for testing relevant environmental samples and linking potential exposures directly to molecular initiating events to facilitate the assessment and characterization of chemical induced toxicity. This abstract does not necessarily reflect US EPA policy.

22 A Method for the Assignment of Chemicals to Mode of Action (MoA) Categories Using *In vitro* Cellular Time-Response and Dose-Response Curves

Zhankun Xi[1], Swanand Khare[1], Tianhong Pan[1,2], Fadi Ibrahim[1], Biao Huang[1], Stephan Gabos[3,5], Weiping Zhang[4], Dorothy Huang[5], David Kinniburgh[5], Can Jin[6], David Santa Anna[6]
[1]Department of Chemical and Materials Engineering, University of Alberta
[2]School of Electrical and Information Engineering, Jiangsu University
[3]Department of Laboratory Medicine and Pathology, University of Alberta
[4]Alberta Health, Health
[5]Alberta Center for Toxicology, University of Calgary
[6]ACEA Biosciences Inc.

Measurements of bioactivity (AC50) or inhibition of cell proliferation (IC50) are often used in high throughput screening. However, the informational value of single measures may be quite limited. Therefore, as part of the Cytotoxicity Profiling Program, Alberta, Canada we have studied 90 compounds from the original ToxCast™ program using the same instrumentation, HepG2 cells, and multiple chemical concentrations. Our objective was to determine if dynamic information on cell growth characteristics could provide more insight into the MoA of chemicals compared to AC50 alone.

In our study, the cytotoxic responses of chemicals in HepG2 cells were dynamically recorded as multi-concentration time-dependent cellular response curves using the xCELLigence real-time cell analysis high-throughput (RTCA-HT) system. Some chemicals were labeled as unclassified if cellular responses did not show a response compared to the control. A model-based hierarchical approach, incorporating principal component analysis (PCA) and functional data analysis (FDA), was developed to extract informational feature vectors from the response curves in order to cluster the chemicals. Dendrograms were generated and cut at proper heights that resulted in four clusters. The clustering results show satisfactory discrimination of chemicals based on the similarity of their cytotoxic response patterns. The results were further validated by examining the commonality of their MoA category labels obtained from the literature, and by calculating the matching rate between the results and the labels. The matching rates showed that the 90 chemicals were properly clustered within major MoA categories. The MoA label matching rate was at least 60%.

This study demonstrated that dynamic time-response and dose-response data are superior to endpoint measurements in terms of richer data content, which could potentially be utilized for predicting chemical MoAs.

Assays

23 New Approaches for Cytotoxicity Profiling of ToxCast™ Chemicals Based on Cellular Growth Dynamics

Stephan Gabos[1,3], Fred Ackah[2], Weiping Zhang[2], Dorothy Y. Huang[3], David Kinniburgh[1,3], Can Jin[4], David Santa Anna[4], Keith A. Houck[5]

[1]Department of Laboratory Medicine and Pathology, University of Alberta

[2]Alberta Health

[3]Alberta Center for Toxicology, University of Calgary

[4]ACEA Biosciences Inc.

[5]National Center for Computational Toxicology, US Environmental Protection Agency

A fundamental goal of *in vitro* toxicity programs is to collect high throughput cytoxicity and other biological information in human cell lines, useful for chemical safety and health risks assessments. ToxCast™ includes data on cell proliferation collected primarily using the ACEA xCELLigence Real-Time Cell Analysis (RTCA) system. However, traditional measures such as compound bioactivity (AC50) and cytotoxicity (IC50) have limitations, that must be overcome in order to better understand the effect of chemical exposure on cell proliferation and putative mode of action. The Cytotoxicity Profiling Program, Alberta, Canada re-screened the ToxCast™pesticide library consisting of 303 compounds, using the same RTCA instrumentation, in a panel of five different cell lines, using multiple concentrations. Our objective was to determine if dynamic information on cell growth characteristics could provide more insight into the mode of action of chemicals, and possibly associated health and safety risks.

Different approaches for data analysis of the effects of chemicals on cell population growth dynamics have been developed based on the time-response and dose-response curves. These include a model-based hierarchical method to extract informational feature vectors from the response curves, a scheme for separating cellular growth dynamics into its individual components - lag phase, exponential growth rate, and cumulative growth, as well as the NCI mean-graph method.

The results show that the dynamic analysis of cell growth curves increases the predictive power of chemical cytotoxicity screens. The effect of chemical exposure on cellular growth dynamics also appears to represent functionally distinct fingerprints characteristic to the mode of action of chemicals. These findings suggest that the resolution and quantification of all facets of cellular growth, as compared to single measurements (e.g., AC50, IC50), increases the informational and interpretational output of cytotoxicity screening using the RTCA system.

24 Toxicity Screening of the ToxCast Phase II Chemical Library Using a Zebrafish Developmental Assay

Mosher S[1], Houck K[1], Frady S[1], Judson R[1], Wambaugh J[1], Reif D[2], and Padilla S[1]

[1]ORD, US EPA, RTP, NC

[2]Department of Biological Sciences, NCSU, Raleigh, NC

As part of the chemical screening and prioritization research program of the US EPA, the ToxCast Phase II chemicals were assessed using a vertebrate screen for developmental toxicity. Zebrafish embryos (Danio rerio) were exposed in 96-well plates from late-blastula stage (6hr post fertilization, pf) through day 5pf (1-2 days post-hatch). All exposures were by immersion and renewed daily. The 700 chemicals included food additives, consumer use product ingredients, pesticides, failed pharmaceuticals, and “green” plasticizers (http://epa.gov/ncct/toxcast/chemicals.html). Intra- and inter-plate replicates were included for quality control. Developmental toxicity was initially assessed using a single nominal concentration of 80 µM: positives and a selection of negatives were confirmed by concentration-response determinations. On day 5pf, larvae were moved from exposure solution to a control solution without chemical, and on day 6pf were assessed for overt toxicity (i.e., death, non-hatching and dysmorphology; n=4 embryos per chemical). Dysmorphology was a combined score using both in-life observation and brightfield, high-content image analysis. Overt toxicity was noted with 46% of the chemicals tested, compared to 62% positive chemicals

when the ToxCast Phase I library, consisting of mostly pesticide active ingredients, was previously tested. As with the Phase I library, the octanol-water partition coefficient (logkow) of the Phase II library chemicals was positively correlated with overt toxicity: there were 18% positive chemicals with logkow <0; 41% positive chemicals with logkow of 0 to 4; and 67% positive chemicals with a logkow >4. All chemicals positive at the single concentration were further assessed for potency using a Dose-Response Study (8-point, semi-log concentration curve: n=3 embryos per concentration). These data demonstrate the utility of zebrafish in medium-throughput chemical testing programs for detection of adverse developmental outcomes. This abstract may not necessarily reflect official Agency policy.

25 Developmental Toxicity of Phase I and II ToxCast™ Chemicals to *Caenorhabditis elegans*

*Windy A Boyd[1], Marjolein V Smith[2], Caroll A Co[2], Jason R Pirone[2], **Julie R Rice[1]**, Jonathan H Freedman[1]*
[1]NIEHS, NIH, DHHS, Research Triangle Park, NC 27709
[2]Social & Scientific Systems, Inc., Durham, NC 27713

Alternative animal models using lower organisms share many advantages with in vitro assays, while also exhibiting complex whole organism responses to chemical exposures. Over 950 unique compounds in the ToxCast Phase I and II libraries were screened using a high-throughput *C. elegans* larval growth and development assay. Changes in the size of individual nematodes were measured using COPAS Biosort flow cytometry after 48-hour exposures to chemicals over seven concentrations (0.5-200 µM). Activity of chemicals at each concentration was classified using a previously determined effect size threshold, which indicates a biologically significant decrease in nematode larval development and growth. Using this threshold, LECs were defined as the lowest concentration at which the mean size of exposed nematodes was less than the effect size threshold and remained below this threshold for subsequent, higher concentrations; 63% (603/959) of Phase I and II compounds were active for at least the highest concentration tested, with a higher percentage of active compounds in the Phase I library (71% [207/292] versus 59% [396/676] in Phase II). The 35 most active compounds disrupted *C. elegans* development at the lowest concentration tested including a number of organotins, avermectin insecticides and anthelmintics, organophosphates, and organochlorines. Compound activities in the *C. elegans* growth assay were compared to those from two zebrafish embryonic development studies, as well as mammalian data available in the US EPA's ToxRefDB for many of the Phase I compounds. Concordance between *C. elegans* and one zebrafish study was higher (79%; n=292) than between rat and rabbit developmental outcomes (58%; n=200). Concordance for 959 Phase I and II compounds between *C. elegans* and a second zebrafish study was more modest at 59%. Using data from 200 chemicals tested in all four species, modest balanced accuracies of 45–53% were observed when using nematode or zebrafish data to predict rat or rabbit developmental effects.

26 Identification of Chemical Vascular Disruptors During Development Using an Integrative Predictive Toxicity Model and Zebrafish and *In vitro* Functional Angiogenesis Assays

Tamara Tal[1], Nicole Kleinstreuer[2], Tarja Toimela[3], Riina Sarkanen[3], Tuula Heinonen[3], Thomas Knudsen[4], and Stephanie Padilla[1]
[1]U.S. EPA/ORD/ISTD, RTP, NC
[2]ILS, inc./NICETUM, RTP, NC
[3]FICAM, University of Tampere, Tampere, Finland
[4]U.S. EPA/ORD/NCCT RTP, NC

Chemically-induced vascular toxicity during embryonic development can result in a wide range of adverse prenatal outcomes. Previously, we constructed an embryonic vascular disruption adverse outcome pathway (AOP) based on molecular initiating events corresponding to genetic models with phenotypic evidence of abnormal embryonic vascular development in the Mouse Genome Informatics Database. Here we used ToxCast high throughput screening data for 25 assays mapping to targets in the Vascular Disruption AOP to prioritize 1060 chemicals for their potential to disrupt vascular development. A subset of 37 predicted vascular disrupting chemicals (pVDCs) or non-pVDCs, including pesticides, flame retardants, and endocrine active compounds, were selected for targeted testing in zebrafish (D. rerio). To test computational

predictions, TG(flk1:GFP) zebrafish embryos were used to visualize and quantify blood vessel formation during development. Manual and automated methods of vessel quantification were developed, and the assay was evaluated with anti-angiogenic reference compounds PTK787 and AG1478, small molecule inhibitors of VEGFR2 and EGFR, respectively. The zebrafish assay was then used to test the effects of 37 chemicals in combination with a functional angiogenesis assay comprised of a human endothelial cell and fibroblast co-culture system. Chemical rankings were well correlated among the predictive signature and zebrafish and *in vitro* tubulogenesis assays. Taken together, the zebrafish assay meets a critical need for an *in vivo* platform that can assess predictions generated by computational models of developmental vascular toxicity. This abstract does not necessarily reflect EPA policy.

27 Establishment of Sensitive, Quantitative and Real-Time Cellular Assays for Assessment and Screening of Modulators of Endogenous Androgen Receptor Signaling Pathways

Can Jin, Yama Abassi, Xiao Xu, Xiaobo Wang
ACEA Biosciences, Inc.

Androgen receptor (AR) plays critical role mediating androgen function for the development and maintenance of male sexual characteristics. AR is also responsible for the carcinogenesis and metastasis of certain prostate cancer. The situation can be further complicated by AR mutations. One obstacle to develop drugs targeting AR is the lack of sensitive and specific *in vitro* assays reflecting endogenous AR signaling pathways. Here, we report the development of cell based assays for detection of AR modulators using real time impedance technology. Two androgen responsive human prostate cancer cell lines 22Rv1 and LNCaP, were used for the study. Stimulation of these cells with androgen agonists such as R1881 and dihydrotestosterone (DHT), lead to alterations in cell number and cell adhesion, which can be detected by gold microelectrodes embedded in the bottom of the well of specialized microelectronic plates. The time-dependant cellular kinetic response profiles were different in 22Rv1 and LNCaP cells, indicating distinctive endogenous androgen signaling pathways in these two cell lines. Both cell types exhibited EC50 values in the picomolar range, indicating high sensitivity to androgen receptor stimulation. The specificity of the assay for AR activity was established using “pure” AR antagonists, such as bicalutamide and nilutamide, and chemicals known with anti-androgen side effect, such as vinclozolin. More interestingly, when LNCaP cells were starved, the kinetic response profile to AR agonist was changed, reflecting altered native androgen response pathways in response to changes in growth condition. In addition, under this condition, nilutamide and vinclozolin displayed AR agonist rather than AR antagonist effects in the real time cellular assay, consistent with reported effect of the T877A mutation in LNCaP AR. The data suggests that the impedance based real time cellular assay system has the capacity to sensitively, selectively and quantitatively detect endogenous AR responses. The information can be useful to understand endogenous AR signaling pathways, and to develop new AR modulators for therapeutic applications.

28 Building Predictive Gene Signatures through Simultaneous Assessment of Transcription Factor Activation and Gene Expression

Natalia M. VanDuyn (USEPA/ORD/NHEERL/ISTD), Jill A. Franzosa (USEPA/ORD/NCCT), Keith Houck (USEPA/ORD/NCCT), William Ward (USEPA/ORD/NHEERL/ISTD), Brian N. Chorley (USEPA/ORD/NHEERL/ISTD), J. Christopher Corton (USEPA/ORD/NHEERL/ISTD)

Exposure to many drugs and environmentally-relevant chemicals can cause adverse outcomes. These adverse outcomes, such as cancer, have been linked to molecular initiating events (MIE) and downstream key events to define adverse outcome pathways (AOP). Identification of gene sets (signatures) that are predictive of either MIEs (e.g., transcription factor (TF) activation) or key events (e.g., cell proliferation) would be useful in predicting AOP modulation after chemical exposure. The goal of this project is to identify signature genes for TF activation via simultaneous assessment of TF activity and global gene expression in the same cell system. RNA isolated from HepG2 cells exposed in concentration-response to ~800 ToxCast or reference chemicals was used to assess TF activation at over 50 cis elements with Attagene FACTORIAL assays and expression of more than 47,000 RNA targets was measured for 100 of these samples using the Illumina HumanHT-12 v4 Expression BeadChip. To evaluate the effects of chemical exposure on gene

expression and TF activation, these data were fit to a four-parameter Hill function with the ToxCast program's data processing pipeline. Pearson's correlation coefficient and associated p-value were used to identify genes with a significant correlation between the change in expression and activation of one or more TF/cis-elements. These results were used to derive gene expression signatures for the aryl hydrocarbon receptor (AhR), thyroid hormone receptor alpha (TRα), and peroxisome proliferator-activated receptor gamma (PPARγ) activation. Utilizing the fold-change rank-based Running Fisher's algorithm, the signatures were compared to ~1500 biosets in an annotated human primary hepatocyte gene expression database to identify those that exhibited a significant positive correlation to the signature. Each signature identified an independent set of chemicals known to activate the TF. For example, the AhR signature identified biosets associated with known AhR-activating chemicals including TCDD, benzo[a]pyrene and quercetin (p-value < 1x10-33), thus validating the method. Future work will focus on expanding the analysis to other TFs, allowing a comprehensive, simultaneous assessment of the modulation of multiple human TFs by chemicals in large, publically available, genomic datasets. (This abstract does not represent EPA policy).

29 Towards a Live Cell High Throughput Microscopy Pathway of Toxicity Reporter Platform for Chemical Safety Assessment

Bob van de Water, Steven Wink, Steven Hiemstra, Suzanna Huppelschoten, Bram Herpers, Marije Niemeijer, Luc Bischoff

Division of Toxicology, Leiden Academic Centre for Drug Research, Leiden University, The Netherlands

Contact: b.water@lacdr.leidenuniv.nl

Adaptive cellular stress responses are paramount in the healthy control of cell and tissue homeostasis after cell injury during hypoxia, oxidative stress or unanticipated side-effect of medications and other chemical exposures. To increase our understanding of chemically-induced adaptive stress response pathway activation and its contribution to safety assessment a time-resolved, sensitive and multiplex readout of chemical-induced toxicological relevant cellular stress responses is essential. For this we develop a unique innovative platform containing a broad panel of distinct adaptive stress response fluorescent protein reporter HepG2 cell lines that represent both upstream as well as downstream components in the different toxicity pathways, including amongst others oxidative stress, DNA damage, inflammation and unfolded protein response signaling. In addition we have generated morphological reporters for the various cell organelles. All reporter cell lines are applied in automated high content live cell imaging settings followed by quantitative multi-parameter image analysis to elucidate the critical temporal dynamics of adaptive stress response pathway activation which can contribute to human chemical safety assessment [1]. To conserve the endogenous gene regulatory programs, we tag selected reporter target genes with GFP using BAC-transgenomics approaches. We have validated the functionality of individual BAC-GFP pathway in toxicity reporter cell lines to their respective specific model compounds in HepG2 2D as well as 3D spheroid culture in 384 well format. We have applied these reporters in chemical safety assessment in relation to drug-induced liver injury [2]. Moreover we have integrated these reporter cell lines in RNAi screening settings to define the signaling components that control the toxicant relevant adaptive stress response pathways [2]. We anticipate that ultimately a phenotypic adaptive stress response profiling platform will allow a high throughput and time-resolved classification of chemical-induced stress responses assisting in the safety assessment of chemicals and likely also contributing to the Tox21/ToxCast program objectives.

[1] Wink et al. Chem Res Toxicol. 2014 Mar 17;27(3):338-55.

[2] Fredriksson et al. Toxicol Sci. 2014 Jul;140(1):144-59.

This work is part of the MIP-DILI project supported by the Innovative Medicines Initiative (grant agreement n° 115336), and the FP7 SEURAT-1 DETECTIVE project (grant agreement 266838).

30 High Content Screening of ToxCast Compounds Using VALA Sciences' Complex Cell Culturing Systems

Kristin Connors
United States Environmental Protection Agency
National Center for Computational Toxicology

The US EPA's ToxCast research program gathers toxicity information for over 1000 chemicals utilizing high-throughput toxicity screening (HTS) assays with human gene and protein targets to inform chemical prioritization. Vala Sciences provides high content, multiplexed assays that utilize quantitative cell-based digital image analysis. We measured the ToxCast Phase I_v2 chemical library with 12 Vala Sciences assays in six-point concentration response from 0.013 uM to 100 uM. Assays examined chemical effects on embryonic stem cell differentiation, neuronal function, pancreatic beta cell differentiation, germ layer proliferation, adipogenesis, adipocyte lipolysis, hepatic steatosis, and junctional proteins critical to developmental angiogenic processes and tumor progression. Due to the diversity of assays and the large number of cellular imaging features from each assay, a systematic selection process was used to derive the assay readouts with the greatest signal and reproducibility. We then subjected the concentration response data to the ToxCast Data Analysis Pipeline performing concentration response modeling and activity calling. Early response targets in this data set provide toxicological and mechanistic insight to complement other ToxCast assays and lead to stronger predictive toxicity models. Collectively, these results add important data to ToxCastDB that will help identify and prioritize possible developmental toxicants, endocrine disruptors, liver toxicants, and carcinogens to inform target testing strategies.

31 ToxCast- Applications in Predicting Toxicity of Harmful and Potentially Harmful Tobacco Smoke Toxicants

Emma Bishop, Damien Breheny, Marianna Gaça, Debbie Dillon & Clive Meredith
British American Tobacco, Group R&D, Southampton, Hampshire, SO15 8TL UK

Tobacco smoke is a complex mixture of over 6,000 identified constituents. Over 150 constituents are known to have defined toxicological properties (toxicants) and prioritisation of these (and other) tobacco toxicants has received increased attention, particularly selective toxicant reduction to support tobacco harm reduction or future tobacco regulation. The US-FDA Center for Tobacco Products has established a list of 93 "harmful and potentially harmful constituents" (HPHCs) in tobacco products and tobacco smoke and include polycyclic aromatic hydrocarbons, heterocyclic amines, nitrosamines, carbonyls and metals, few of which are existing ToxCast chemicals. Understanding of the toxicities of individual smoke constituents and subsequently smoke mixtures using *in vitro* assays and computer models, such as those applied in ToxCast will support current generated in-house data and the subsequent prediction of the toxicity of complex mixtures such as tobacco smoke and e-cigarette vapour.

For example, we have developed in-house *in vitro* and computational methods to support initial toxicity prediction of tobacco smoke toxicants. Cadmium, a HPHC, has been identified on the basis of Margin of Exposure (MoE) calculations as a high priority for exposure reduction. A study was designed to inform the MOA using *in vitro* cultures of primary human bronchial epithelial cells treated with cadmium chloride over 28 days investigating endpoints including cytotoxicity, membrane integrity, MUC5AC production, cell morphology, cilia beat frequency, 8-oxo-dG adduct formation and secreted protein expression.

Generated *in vitro* data supported the plausibility of a number of key events in the postulated cancer and non-cancer MOAs for cadmium and its compounds and adds weight of evidence for its status as a high priority for exposure reduction in tobacco smoke. The integration of assessment techniques such as *in vitro* assays and computer models may give guidance for research and regulatory priorities for selective tobacco smoke toxicant reduction. A single chemical approach (93 HPHCs) will need to be integrated in a mixture (smoke/ vapour) or modified mixture assessment as a longer term strategy for toxicant management and regulatory assessment. Such research would be greatly accelerated with the inclusion of the HPHCs to the ToxCast list and the application of recently generated ToxCast data

32 Role of Abcg2 During Mouse Embryonic Stem Cell Differentiation

Mitchell B. Rosen (USEPA/ORD/NHEERL/ISTD), Kelly J. Chandler (NIEHS/Office of Policy, Planning, and Evaluation), Susan C. Jeffay (USEPA/ORD/NHEERL/ISTD), Harreitte P. Nichols (USEPA/ORD/NHEERL/ISTD), Maria R. Hoopes (USEPA/ORD/NHEERL/ISTD), and E. Sidney Hunter(USEPA/ORD/NHEERL/ISTD)

Abcg2 is a multidrug resistance ATP-binding cassette (ABC) transporter whose activity may be considered a hallmark of stem cell plasticity. The role of Abcg2 during early embryogenesis, however, is unclear. Studies done with mouse embryonic stem cells (mESCs) suggest that inhibition of Abcg2 by either chemical means or RNAi may reduce colony expansion of pluripotent cells and promote differentiation. Our group has also observed a correlation between xenobiotics that alter mESC differentiation and ToxCast Phase I chemicals that induce expression of Abcg2. On the other hand, knockout mice display a normal phenotype unless challenged by certain toxicants. Chemical inhibitors of Abcg2, such as K0143 and fumitremorgin C (FTC) were utilized to evaluate the role of Abcg2 in cultured J1 cells. MK571 and verapamil, inhibitors of Abcc1 and Abcb1, respectively, were utilized as well. Based on cellular accumulation of the Abcg2 substrate, pheophorbide A, maximum inhibition of Abcg2 was observed at approximately 1µm K0143 or FTC. Neither verapamil nor MK571 had a notable effect on Abcg2 function. While growth and differentiation of mESC could be affected by FTC, K0143, and verapamil at concentrations greater than 10 µm, lower concentrations did not influence cell proliferation or alter the expression of selected pluripotency/lineage markers on the appropriate days of culture (culture days 2, 4, 6, 9). These transcript markers included Pou5f1, Nanog, T, Gsc, Bmp4, Nes, Ncam1, Des, Ttr, Gata4, Abcg2, Myl4, and Myl7. Hence, Abcg2 does not appear to play a fundamental role in mESC differentiation based on observations in an adherent culture system. The possibility that Abcg2 plays a role in protecting embryonic cells from damage caused by xenobiotics, or that regulation of this efflux transporter by exogenous compounds may affect early differentiation is being considered. (This abstract does not reflect EPA policy.)

33 Agilent's GeneSpring Multi-Omics Analysis and Integration Solution Applied to a Comparative Transcriptomics Analysis of Estradiol and PPT Treatment Effects in MCF7 Cells as Part of the NIH Transformative Research Grant for Mapping the Human Toxome by Syst

Carolina B. Livi[1], Rick Fasani[1], Michael Rosenberg[1], Renxiang Chen[2], Heng-Hong Li[2], Albert J. Fornace[2], James D. Yager[3], Shelly Odwin-Dacosta[3], Kim Boekelheide[4], Marguerite Vantangoli[4], Melvin Andersen[5], Patrick McMullen[5], Salil Pendse[5], Alexandra Maertens[6], Tom Luechtefeld[6], Andre Kleensang[6], Mounir Bouhifd[6], Thomas Hartung[6]*

[1]Agilent Technologies, Inc.

[2]Georgetown University

[3]Johns Hopkins University

[4]Brown University

[5]Hamner Institute

[6]Center for Alternatives to Animal Testing, Johns Hopkins University

**carolina.livi@agilent.com*

Here we give one illustrative example on how-to integrate different platform data sources for gene expression data (microarrays, qPCR and RNA-seq) in a combined analysis to define a biomarker candidates list of reproducible gene expression changes in response to estrogen agonist treatment. The next step will include integration of metabolomics data. GeneSpring provides a platform for multi-omic data analysis allowing for comparisons and results interpretation across different platforms and technologies well suited to the modern challenges of system toxicology including pathway centered analysis as well as a comprehensive descriptive and inferential statistics and QC methods that allow for example classifier /biomarker development, identification of co-regulated genes, metabolites, and proteins in an intuitive and easy-to-use manner. A major challenge for identifying biomarkers of environmental effects is the perceived lack of reproducibility from metadata analysis of the literature. Often this analysis relies on the results reported in publications as part of author conclusions and not in the data measurement itself. As part of The Human Toxome Project we re-analyzed publicly available datasets of estrogen agonist exposed commonly

studied breast cancer cell line and combined this with our transcriptome profiling of MCF7 cells treated with vehicle control, estradiol or propyl pyrazole triol as part of a time course experiment using Agilent human gene expression microarrays and qPCR. In addition to these experimental variables, cell cultures were setup at two independent laboratories (at Johns Hopkins and Brown Universities). Microarray and RNA-seq gene expression profiling datasets deposited in GEO were downloaded and analyzed as per standard methods. Quality controls as described by manufacturer were used to evaluate data prior to further analysis. Differential expression analysis comparing estrogen agonist treated MCF7 cells to their respective controls were performed taking into consideration the unique experimental designs described in publications or data repository entries. While previous work comparing results reported in tables of differentially expressed genes suggested none were reproducible across a dozen published datasets, we found a short list of estrogen response biomarker candidates with reproducible gene expression changes in response to estrogen agonist treatment.

34 Annotating Biological Activities of Environmental Toxicants by Transcription Factor Profiling

Sergei S. Makarov[1], Matt Moeser[1], Liubov Medvedeva[1], Elena Martsen[1], Alexander Granick[1], Lydia Raines[1], Ming Zeng[1], Keith A. Houck[2], Alexander Medvedev[1]
[1]Attagene Inc, P.O. Box 12054, Research Triangle Park, North Carolina 27709
[2]US Environmental Protection Agency, 109 TW Alexander Dr., D343-03, Research

Rationale. A main objective of molecular toxicology is the identification of the underlying molecular mechanisms. Here, we describe a systems biology approach that enables annotation of broadly different bioactivities by profiling perturbations of signal transduction pathways that regulate transcription.

Experimental design. Using a multiplexed reporter system allowing quantitative assessments of activities of 48 transcription factors (TFs) (cis-FACTORIAL™), we examined responses of human cells to a series of prototypical inhibitors of cellular systems mediating proteasomal degradation, histone deacetylation, oxidative phosphorylation, and cytoskeleton assembly.

Results. Strikingly, inhibition of each system produced its own, distinct, “TF signature” regardless different chemical structures of inhibitors and their exact targets within the system. Using the common TF signature of mitochondrial dysfunction we interrogated the database containing TF signatures of over 1,800 of ToxCast chemicals and correctly identified novel mitochondrial inhibitors.

Conclusion. Our data indicate that inhibitors of certain cellular systems produce specific TF response patterns that can be used as “archetypal signatures” indicating impairment of of corresponding systems. Using these archetypal TF signatures one can identify inhibitors of these systems. These findings indicate the feasibility to develop a unified, mechanism-based, approach to annotation of diverse bioactivities by TF profiling, with numerous ramifications for drug discovery and toxicity testing.

The views expressed in this abstract are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.

Chemistry and Quantitative Structure–Activity Relationship

35 In Silico Modeling Workflows in Support of Exploratory Computational Toxicology

Michael-Rock Goldsmith, Daniel T. Chang, Alain Deschênes, Christopher Williams, Alain Ajamian*
Chemical Computing Group Inc, 1010 Sherbrooke St West, Suite 910, Montreal QC H3A 2R7 Canada
**Presenter Contact: rgoldsmith@chemcomp.com phone: 514-393-1055*

The ToxCast™dataset in the iCSS (interactive Chemical Safety for Sustainability) dashboard is ideal for in silico model development to support the rational molecular-based screening and prioritization of bioassay properties. Combined in silico and *in vitro* techniques have previously been used to explore nuclear receptor (NR)-based binding properties (e.g., ER, AR and TR) for existing endocrine panels. The present work demonstrates the development and treatment of NR family related iCSS data, and integration with structural information from DSSTox through the application of the Molecular Operating Environment (MOE)

for in silico "NR superfamily" model development, although these approaches extend to other classes within ToxCast™including but not limited to oxidoreductases, kinases, phosphatases, GPCRs and ion channels. The scope of the data development ranges from ligand/protein structure preparation for in silico data generation (such as ligand dataset optimization and computational molecular docking studies) to QSAR-derived descriptor-based filters and binary classification trees (for apical endpoint estimates and targeted protein-ligand structure-based classifiers) which include both qualitative and quantitative in silico toxicogenomic datasets and workflows. These models and underlying molecular toxicoinformatics approaches can be used independently, or in tandem with other techniques that have been developed, to provide rapid assay-chemical and target-toxicant relationship networks for interpreting adverse outcome pathways (AOPs) and to gain toxicological insight with a molecular and atomic level of resolution.

36 Identifying Structural Alerts Based on Zebrafish Developmental Morphological Toxicity

Jui-Hua Heish, Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709
Nicole Kleinstreuer, Integrated Laboratory Systems, Inc., Research Triangle Park, NC, 27560, USA
Jill Franzosa, U.S. Environmental Protection Agency, Research Triangle Park, NC
Alex Sedykh, MultiCASE Inc. Cleveland, OH 44122, USA
Ray Tice, Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709

Zebrafish constitute a powerful alternative animal model for chemical hazard evaluation. To provide an *in vivo* complement to high-throughput screening data from the ToxCast program, zebrafish developmental toxicity screens were conducted on the ToxCast Phase I (Padilla et al., 2012) and Phase II chemical libraries (Truong et al., 2014). The screens had several differences in experimental design and data analysis. Padilla et al. assigned Toxicity Scores based on descriptive data (lethality, hatching) and malformation assessments (binary and numeric). The Toxicity Scores were used to plot concentration response curves and estimate half-maximal activity concentrations. Truong et al. evaluated morphological endpoints by assigning binary incidence counts and computing lowest effect levels. This binomial count approach preserves independent morphological information and does not constrain dose-response behavior (e.g., response can be non-monotonic). In contrast, the Toxicity Score approach aggregates morphological data and assumes monotonic dose-response behavior. Utilizing the relative strengths of both approaches, we conducted a study to identify and analyze correlated morphological endpoints based on the Phase II data. We found endpoints that are closely related, for example, the jaw, snout, and pericardial and yolk sac edema, and grouped them into developmental process endpoints. Based on the new endpoints, we generated Severity Scores to plot concentration response curves. Based on the curve assumed to be monotonic, we calculated the point-of-departure (POD), which is the concentration where a response exceeds the baseline noise level. Quantitative structure activity relationship (QSAR) analyses using the CASE Ultra structural alert mining system yielded predictive models for chemicals likely to induce developmental malformations in zebrafish. The externally validated balanced accuracies of the QSAR models ranged from 0.63-0.74. Several structural alerts related to zebrafish developmental toxicity were identified such as the ortho-dimethoxyaryl group (e.g., as in colchicine), thioamide group (e.g., in ziram), and benzothiazole scaffold (e.g., in 2-mercaptobenzothiazole). In summary, we identified 220 compounds in the Phase II library that elicit adverse developmental responses (including general toxicity) under specific process based mechanisms with POD values interpolated from dose-response curves. We further developed QSAR models to evaluate large chemical libraries, such as Tox21, and prioritize chemicals for further investigation.

37 Rational Design for Chemicals with Reduced Potential to Incur Oxidative Stress

Longzhu Q Shen[1], Karolina E Mellor[1], Evan Beach[1], Fjodor Melnikov[1], Jone Corrales[2], Lauren Kristofco[2], Melissa Mullins[2], Gavin Saari[2], E. Spencer Williams[2], Jakub Kostal[3], Christopher Schaupp[4], Margaret Mills[4], Terrance Kavanagh[4], Nancy Simcox*[4], Adelina Voutchkova-Kostal*[3], Evan Gallagher*[4], Bryan Brooks*[2], Julie B Zimmerman*[1], Paul Anastas*[1]*

[1]Yale University, Center for Green Chemistry & Green Engineering, New Haven, Connecticut, 06511, United States

[2]Baylor University, Department of Environmental Science, Waco, Texas, 76798, United States

[3]The George Washington University, Department of Chemistry, Washington DC, District of Columbia, 20052, United States

[4]University of Washington, Department of Environmental and Occupational Health Sciences, Seattle, Washington, 98105, United States

Beginning as a toxin but later as a more efficient fuel for energy generation, molecular oxygen played a deterministic role in the evolutionary path selection onto the aerobic direction. Apart from the irrefutable benefit, oxygen and its derivatives, collectively called reactive oxygen species (ROS), are highly reactive chemical agents that can cause numerous deleterious damages if their production and transformation in a biological system were not under control. To counteract ROS, living organisms have developed a set of anti-oxidant defense mechanisms to maintain the system at a delicate balance. If the balance were overthrown, oxidative stress may occur and eventually lead to a plethora of clinical syndromes such as arthritis, immunodepression, neurodegeneration and etc. Therefore, unintentional perturbation to ROS balance should be minimized and avoided.

However, the conventional molecular design strategies normally do not take into account of the potency of a molecule to induce ROS generation and certain commercial chemicals have been reported to possess this unwanted competency in literature. This motivates the design of chemicals with reduced potencies to facilitate unintentional ROS production. A consortium of four universities was formed to initiate molecular design network with the aim of deriving design guidelines to minimize adverse effects of commercial chemicals via oxidative stress.

The current presentation is an initial model of this collaborative project. It was built upon the mechanistic understanding of ROS generation and their linkages to oxidative stress damages. This knowledge was used to guide the selection of molecular chemo-physical properties computed at the quantum mechanics level and statistically analyze the hazard-property relationships and ultimately inform benign molecular design guidelines. The model has demonstrated discriminative power to classify molecules with differentiated potencies to incur oxidative stress.

38 Computational Platform for Green Chemistry

Patrick Quigley Leadscope, Glenn Myatt Leadscope, Kevin Cross Leadscope, David Bower Leadscope

The field of green chemistry was established in order to reduce or even eliminate any negative impact from the introduction of new chemical products such as new pesticides, cosmetics, or drugs to human health or the environment. Unfortunately, current approaches for testing chemicals are expensive and time-consuming which makes their routine use throughout product development impractical. New approaches to testing chemicals, including *in vitro* tests and computer models, are being developed that are faster and cheaper. A testing strategy employing a tiered battery of these new tests will eventually enable safety assessment at each stage of the chemical R&D process. These test results will be used to make design decisions that mitigate any identified risks. The objective of this project is to understand how a combination of these tests is able to predict how chemicals may interfere with the endocrine system, the body's hormone signaling system. This poster will outline the development of a reference database containing information on over 10,000 chemicals including specific mechanisms cited in the literature, *in vitro* results from ToxCast and Tox21 and other sources, *in vivo* data from regulatory and commercial databases as well as PBPK data. The results of a number of initial studies will be presented including associations between combination of *in vitro* assays and known mechanisms and *in vivo* endpoints. Using this reference database, a battery of computer models to identify potential Endocrine-Disrupting Chemicals (EDCs) that interfere

with the Estrogen Receptor (ER) signaling pathway was developed. To demonstrate the utility of such a database, assay data from the US EPA ToxCast and cross-agency Tox21 programs for 20 endpoints tested against a library of over 1700 chemicals examining potential perturbations of the ER pathway. A series of predictive models were constructed using the Leadscope Predictive Data Miner software application for ER binding, ER dimerization, DNA binding, RNA transcription, and ER-induced proliferation. Based on the validation results of these models, their predictivity is discussed in detail in relation to any known mechanisms by chemical class of the training compounds.

39 Use of SAR to Assess Impact of Metabolism in ToxCast Modeling of Rat Carcinogenicity

Ann M. Richard[1], Patra Volarath[2], Chris Grulke[3], Richard Judson[1]
[1]National Center for Computational Toxicology, ORD, U.S. EPA, RTP, NC 27711
[2]Center for Food Safety and Applied Nutrition, U.S. FDA, College Park, MD 20740
[3]Lockheed Martin, RTP, NC 27711

ToxCast program objectives rely heavily on the use of *in vitro* high-throughput screening (HTS) data to inform pathway-based models for predicting *in vivo* adverse outcomes. In the absence of metabolic competency for the majority of current ToxCast assays, this paradigm assumes that the chemical undergoing HTS testing is direct-acting, i.e. does not require further metabolic activation for target, assay or pathway activation. When metabolic activation is required, the relationship between the parent (unmetabolized) chemical, and its largely "unactivated" HTS assay profile, with the corresponding *in vivo* outcome is likely to be fundamentally different than for direct-acting chemicals. Thus, mixing direct-acting chemicals with those requiring metabolic activation is expected to obscure the HTS "signal" specific to the *in vivo* response for both direct-acting and metabolically activated chemical subsets, alike. A workflow was developed using combined outputs of the Meteor and Derek structure-based expert systems to predict which chemicals trigger Derek toxicity alerts only after metabolic transformation (i.e., after first processing the parent chemical through Meteor). The workflow was applied to predicting rat carcinogenicity outcomes within ToxCast Phase I chemicals, with results indicating a major influence of metabolic activation in correctly predicting ToxRefDB Phase I rat carcinogens. The workflow was then applied to delineating the subsets of direct-acting and metabolically activated subsets of Phase I rat carcinogens. Univariate correlations of ToxCast HTS results to ToxRefDB *in vivo* rat carcinogenicity outcomes, derived separately within these two subsets, produced dramatically different HTS assay correlations. These results demonstrate that SAR methods can be used to focus modeling efforts into chemical domains where metabolism is or is not likely to occur, thereby increasing the accuracy and interpretability of the resulting HTS to *in vivo* outcome predictive models. This abstract does not reflect EPA or FDA policy.

40 In-silico Study of ToxCast GPCR Assays by Quantitative Structure-Activity Relationships (QSARs) Modeling

Mansouri K.[1], Sipes N.S.[1], Judson R.S.[2]
[1]ORISE Fellow, NCCT, ORD/U.S. EPA/ RTP/ NC/ USA
[2]National Center for Computational Toxicology, ORD/U.S. EPA/ RTP/ NC/ USA

The EPA tested several thousand chemicals in 700 toxicity-related in-vitro HTS bioassays through the ToxCast and Tox21 projects. However, the chemical space of interest for environmental exposure is much wider than this set of chemicals. Thus, there is a need to fill data gaps with in-silico methods, and quantitative structure-activity relationships (QSARs) are a cost effective approach to predict biological activity. The overall goal of this project was to use QSAR predictions to fill the data gaps in a larger environmental database of ~30K structures. The specific aim of the current work was to build QSAR models for multiple ToxCast assays using a subset of 1800 chemicals tested in 18 G-Protein Coupled Receptor (GPCR) assays. These assays are part of the aminergic category which was among the most active within the biochemical assays. Two QSAR modeling strategies were adopted; a classification into active/non-active chemicals, then regression models were built to predict the AC50 potency values of the bioassays for the active chemicals. Multiple software programs were used to calculate constitutional, topological and substructural molecular descriptors from two-dimensional structures. Model-fitting methods included

PLSDA, SVM, kNN and PLS. Genetic algorithms (GAs) were applied as a variable selection technique to select the most predictive molecular descriptors for each assay. N-fold cross-validation (CV) coupled with multi-criteria decision making (MCDM) fitting criterions was used to evaluation the models. Finally, the applicability domain of the models was defined in order to make predictions within established chemical space limits. Using partial least squares discriminant analysis for the bovine non-selective dopamine receptor (bDR_NS) GPCR assay, the classification balanced accuracy reached 96% in fitting and 95% in 5-fold CV, with only 2 latent variables. These results demonstrate the ability of QSAR models to predict bioactivity. This abstract does not necessarily reflect U.S. EPA policy.

41 CERAPP-Collaborative Estrogen Receptor Activity Prediction Project

Mansouri K.[1], Kancherla J.[1], Richard A.[2], Judson R.S.[2]
[1]ORISE Fellow, NCCT, ORD/U.S. EPA/ RTP/ NC/ USA
[2]National Center for Computational Toxicology, ORD/U.S. EPA/ RTP/ NC/ USA

Humans are potentially exposed to tens of thousands of man-made chemicals in the environment. It is well known that some environmental chemicals mimic natural hormones and thus have the potential to be endocrine disruptors. Most of these environmental chemicals have never been tested for their ability to disrupt the endocrine system, in particular, their ability to interact with the estrogen receptor. Regulatory agencies are in need of tools to prioritize thousands of chemicals for a variety of modes of toxicity. This project was intended to be a demonstration of the use of predictive computational models and HTS data to prioritize a large chemical universe for one specific molecular target – the estrogen receptor. CERAPP combined multiple structure-based computational models for prediction of estrogen receptor activity, and used the predicted results of these models to build a unique consensus model. The models were developed in collaboration with 16 groups in the U.S. and Europe using common set of chemical structures divided into training set and a prediction set of ~32k compounds. A total of 42 classification models and 8 regression models were built for binding, agonist and antagonist activity. Docking techniques and several QSAR modeling approaches were employed. All predictions were evaluated using an external validation set from the literature. In order to overcome the limitations of single models, the consensus was built based on rankings of the predictions accuracy scores within the applicability domains of the models. This consensus will be used in prioritization of chemicals for further testing.

This abstract does not necessarily reflect U.S. EPA policy

42 Using ToxCast™ in Considering Chemical Activities and Hazard Traits: A Case Study with the Chemical Class Phthalates

Shoba Iyer[1], Nathalie Pham[1], Edward Hackett[1], Bennett Lock[1], Melanie Marty[1], Martha Sandy[1], George Alexeeff[1], Gina Solomon[2], and Lauren Zeise[1]
[1]California Environmental Protection Agency's (CalEPA) Office of Environmental Health Hazard Assessment (OEHHA)
[2]CalEPA

US EPA's Toxicity Forecaster (ToxCastTM) may be a useful tool in State-based efforts for safer consumer products, alternatives analyses, chemical prioritization for biomonitoring or toxicity assessment, and screening level risk assessments. As a preliminary step in exploring the use of ToxCast in California programs, we chose the candidate class of phthalates as a case study due to the well-established toxicological endpoints of some members of this class, and our interest in developing a good understanding of their potential modes of action. In addition, several specific phthalates are on the State of California's Proposition 65 list of known carcinogens and reproductive toxicants. The ToxCast phthalate data can be utilized for comparing bioactivities and investigating structure-activity relationships between chemical anchors and related chemicals, including those that have sparse datasets, and for exploring potentially novel modes of action. In this study, we investigate: 1) correlations between the ToxCast phthalate data and phthalate toxicity information in the scientific literature; 2) ToxCast assay activities between parent phthalates and their monoester metabolites to begin to explore the role of xenobiotic metabolism in toxicity endpoints; and 3) ToxCast assay activity clusters based on phthalate chemical structure and/or size to

investigate the link between chemical structure and potential biological activity. Examining the ToxCast assay data within this candidate class has thus far revealed several links between key early molecular events assayed *in vitro* and chemical-specific hazard traits. For example, several phthalates with endocrine disruption activity as described in the scientific literature have also shown activity in ToxCast assays interrogating the estrogen receptor pathway. We will utilize our results, along with other toxicity data of well-studied phthalates, for conducting several proof-of-concept examples of read-across on lesser-studied phthalates and phthalate substitutes. Ultimately, high-throughput screening data resources like ToxCast may inform us of sensitive upstream toxicity endpoints and can assist in the rapid identification of environmental chemical hazards for screening assessments and prioritization.

43 The Importance of Chemical Data Curation to Achieve Accurate Prediction of ER Binding Affinity for Environmental Chemicals

Sherif Farag, Eugene Muratov, Denis Fourches, Alexander Tropsha
Laboratory for Molecular Modeling, UNC Eshelman School of Pharmacy, UNC-Chapel Hill, Chapel Hill, NC 27599

Earlier studies in our lab and elsewhere have shown that inaccuracies in chemical structure representation even for a small fraction of compounds in a chemical dataset could lead to significant losses in the predictive power of QSAR models. Driven by these observations we have developed and implemented a rigorous chemical data curation workflow [1]. In collaboration with the EPA team (R. Judson, A. Richard, K. Mansouri) we have employed this workflow for curating both the training and virtual screening sets for the Collaborative Estrogen Receptor Activity Prediction Project (CERAPP). The initial CERAPP training set included 1800 compounds, of which 1667 compounds remained after eliminating one compound with covalently bound counter ion and more than 120 duplicates. The CERAPP prediction set was comprised of more than 50000 compounds; the application of our curation procedure resulted in its reduction by almost 40% (from 54785 to 32644 compounds). Compounds were eliminated for a number of reasons such as bad SMILES, inorganics, organometallics as well as compound mixtures compounds of course a large fraction of duplicates (more than 18000 compounds were excluded for this reason). Following this heavy curation, QSAR models for the CERAPP training set have been developed using Random Forest and Support Vector Machines, Dragon descriptors, and five-fold external validation procedure; the external balanced accuracy of these models was on average 80%. These models were applied to screening the curated external CEARPP library resulting in the selection of several hundred high-confidence hits as putative ER actives.

1. Fourches D, Muratov E, Tropsha A. Trust, But Verify: On the Importance of Chemical Structure Curation in Cheminformatics and QSAR Modeling Research. J Chem Inf Model. 2010, 50, 1189–1204

44 Connecting the Dots: Exploring the Chemical Space and Multi-Species Bioactivity of FDA Pesticides

Andrew Defries, Backman Defries
San Francisco, CA

This paper advocates a crucial yet still under-advocated modus operandi in the field of cheminformatics: to enable quick progression of research via online peer scrutiny and public access. Advances in the chemical, biological, and genomic sciences, along with advances in methodologies for computing the information they generate, have enabled us to leverage big data in new and important ways. By using shared (online) computational resources, researchers can focus on questions more precisely within the domain of their sciences, allowing peripheral yet crucial infrastructure to mature “behind the scenes.” Scientists are increasingly using UNIX and R (1) along with open source Bioconductor (2) packages, to attack “big data” problems, foster collaboration on such problems, and disseminate insight into structured data. These computational methods, among others, can (and should) be further integrated into common practice. To demonstrate how, I pair cheminformatics publications with code repositories containing the sources used to generate figures, along with supplementary files to re-constitute the final analysis. This encourages transparency and enables the next iteration of advances by directly linking users with the core problem

(dataset) and toolset (codebase) (3). As such, I will present annotated bioactivity trends as html and PDF reports, and nearest neighbor tables clustered using Tanimoto scoring and a flexible substructure matching tool called fmcsR (4). The results will be accompanied with a public code repository containing a suite of computation and visualization tools, which can be used locally or on distributed systems. Cheminformatics analysis using ChemmineR (5) along with bioassayR, a gateway to the PubChem bioassay database (6), enabled impartial dissection of the relationship between small molecules and biological targets. Through a combination of open source cheminformatics and text mining software we identify intersections of chemical and biological space between FDA approved pesticides and the DrugBank database (7). Our analysis considers FDA approved and small molecule drugs, and several sources including ToxCast. Finally, using this approach to public data, I discuss preliminary phylogenetic analysis of small molecule bioactivity across multiple species.

References

1) R Development Core Team (2008) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org. ISBN 3-900051-07-0.

2) Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome biology 5: R80.

3) Dabbish L, Stuart C, Tsay J, Herbsleb J (2012) Social coding in github: transparency and collab- oration in an open software repository. In: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. ACM, pp. 1277–1286.

4) Wang Y, Backman TW, Horan K, Girke T (2013) fmcsR: mismatch tolerant maximum common substructure searching in R. Bioinformatics 29: 2792–2794.

5) Cao Y, Charisi A, Cheng LC, Jiang T, Girke T (2008) ChemmineR: a compound mining framework for r. Bioinformatics 24: 1733–1734.

6) Wang Y, Suzek T, Zhang J, Wang J, He S, et al. (2014) Pubchem bioassay: 2014 update. Nucleic acids research 42: D1075–D1082.

7) Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, et al. (2014) Drugbank 4.0: shedding new light on drug metabolism. Nucleic acids research 42: D1091–D1097.

Exposure

45 Exploring Exposure Pathways with Chemical/Product Categorical (CPCat) Data

KL Dionisio (US EPA), A Frame (US EPA), M-R Goldsmith (US EPA), JF Wambaugh (US EPA), A Liddell (US EPA), T Cathey (Lockheed Martin), D Smith (US EPA), J Vail (US EPA), RS Judson (US EPA)

We are exposed to thousands of chemicals over our lifetimes. A major challenge to risk assessors is to understand how and when chemical exposures occur, and which “exposure pathways” contribute the most. An informatics-driven approach to assigning “product-use” categories to product “ingredients” will help prioritize which chemicals will be given more scrutiny for a target population (life-stage, gender, ethnicity), identifying (a) human activities that result in increased chemical exposure while (b) reducing the dimensionality of hazard assessment to a tangible subset for risk characterization. How chemicals are used is directly related to their potential exposure routes, and in the sense of near-field source apportionment, to the identification of chemical-specific use-to-receptor exposure pathways. We have developed the Chemical/Products Categories Database (CPCat) to explore these pathways. CPCat is both a front-end interface and back-end relational database comprised of chemical “use category” information, including both consumer- and industrial-process based uses for chemicals, using mined and curated data from 14 major sources from multiple countries, including data from regulatory bodies, manufacturers and retailers. CPCat contains information on 42K unique chemicals mapped to a set of >700 use categories. We provide an example of how CPCat can be used to link life-stage to use-pattern to products and their ingredients for children’s exposure. CPCat can be used to identify chemicals that children would potentially be exposed to, and the major activity/source of such exposures; a database query of the “child” or “baby” keywords results in the identification of 1,300 chemicals, of which the majority are associated with the “child toy” CPCat

category. This information can then be used to rank plausible exposures of chemicals which may be investigated for potential biological activity. The entire database and a set of tools will be made publicly available. The views expressed in this abstract are those of the authors and do not reflect the views or policies of the U.S. EPA.

46 High Throughput Exposure Modeling of Semi-Volatile Chemicals in Articles of Commerce

*Chantel I. Nicolas[1,4], Michael R. Goldsmith[2], Bhavesh K. Ahir[1,4], Barbara A. Wetmore[3], Kathie L. Dionisio[2], Kamel Mansouri[1,4], R. Woodrow Setzer[1], Richard S. Judson[1], James Rabinowitz[1], and John F. Wambaugh[1]**
[1]National Center for Computational Toxicology, U.S. EPA, Research Triangle Park, NC, United States
[2]National Exposure Research Laboratory, U.S. EPA, Research Triangle Park, NC, United States
[3]The Hamner Institutes for Health Sciences, Research Triangle Park, NC, United States
[4]Oak Ridge Institute for Scientific Enrichment, Oak Ridge, TN, United States*

Risk due to chemical exposure is a function of both chemical hazard and exposure. Near-field exposures to chemicals in consumer products are identified as the main drivers of exposure and yet are not well quantified or understood. The ExpoCast project is developing a model that estimates indoor exposures to chemical additives in textiles, such as flame retardants, which are found in flooring, upholstery, and articles of clothing. Physicochemical properties of chemicals largely dictate how they may accumulate in the indoor environment at higher magnitudes than in the outdoor environment, which along with proximity of the sources, is correlated with high indoor exposure rates. Halogenated flame retardants, such as polybrominated diphenyl ethers (PBDEs), are semi-volatile organic compounds that are potentially harmful to humans. Given that chemical emission calculations principally depend on gas-phase concentration (y¬˚) and source surface area, we used a model (Little et al. 2012)1¬ to assess the utility of physicochemical property information in predicting indoor emissions for these additives. Gas-phase concentrations were predicted using a regression model of experimental measurements of 74 chemicals found in 32 flooring materials by Wilke et al. (2004).2 The flooring materials include a range of natural and synthetic floor coverings, installations, and adhesives. A linear regression yielded R2- and p- values of approximately 0.3 and 2.0E-12, respectively, whereby logP and vapor pressure were significant predictors of y¬˚. These results potentially permit the forecasting of gas-phase concentrations of chemicals for which their analytical data in flooring materials are lacking. Predictions generated from using high throughput exposure methods were then combined with ToxCast high-throughput bioactivity screening data as a demonstration of high-throughput risk prioritization. This abstract does not necessarily reflect EPA policy.

1. Little et al. Environ. Sci. Technol. 2012, 46, 11171-11178

2. Wilke et al. Indoor Air 2004, 14, 98-107

47 Exposure Modeling of Parabens in Personal Care Products Compared to NHANES Biomonitoring Data and ToxCast *In vitro* Bioactivities

Susan A. Csiszar, University of Michigan, Alexi Ernstoff, Technical University of Denmark, Barbara A. Wetmore, The Hamner Institutes for Health Sciences, Rudy J. Richardson, University of Michigan, Olivier Jolliet, University of Michigan

To understand the potential risks of using chemicals in personal care products (PCPs), we need to incorporate the ToxCast high throughput bioactivity data with estimates of consumer intake. We developed a model to predict population level exposures (in mg/kg/d) due to the use of common personal care products and applied it to a case study of four parabens in the U.S female adult population. We evaluated the model by comparing to NHANES (National Health and Nutrition Examination Survey, CDC) biomonitoring data. Estimated individual chemical intakes were then compared to dosimetry-adjusted Oral Equivalency Doses (OED, mg/kg/d) calculated from the ToxCast assays that yielded the lowest (i.e., most potent) bioactivities (with OED values from Wetmore et al., In Prep.).

To estimate PCP intakes we determined the product intake fraction (PiF), defined as the ratio of the mass of chemical taken in the population per mass of chemical used. The PiF can be calculated and combined for both the use and subsequent environmentally mediated exposures after cosmetics are rinsed-off. The PiF

was found to vary between 0.1-100% depending on product usage, specifically based on whether the product is leave-on or rinse-off. To calculate population level exposure, a Monte Carlo analysis was used to take into account variability in factors such as PCP usage patterns, paraben fractions in products, and chemical uptake into the skin. Our modeled intakes for the 50th, 75th, 90th, and 95th percentile users were generally within a factor of 2 (all within a factor of 4) of NHANES biomonitoring data.

Calculated intakes (e.g., 7.3 mg/kg/day for methyl paraben) were compared to the ToxCast-derived OEDS by calculating the margin of exposure, MOE, defined as the ratio of the OED to the modeled intake. For a high-end individual PCP user, the MOEs were 0.8, 17, 5, and 53 for methyl, ethyl, propyl, and butyl paraben, respectively. This indicates that there may be consumers who are exposed to parabens at biologically active levels. The techniques demonstrated here can be further applied to predict intakes and MOEs for a large number of other PCP ingredients in the ToxCast database.

48 Utilizing ToxCast Data, and Lifestage Physiologically-Based Pharmacokinetic (PBPK) Models to Drive Adverse Outcome Pathways (AOPs)-Based Margin of Exposures (ABMEs)

Hisham El-Masri[1], Linda Adams[1], Nicole C. Kleinstreuer[2], Tamara Tal[1], Stephanie Padilla[1], John Wambaugh[3], Kristin Isaacs[4], Halûk Özkaynak[3]
[1]NHEERL, USEPA
[2]ILS contractor, NTP
[3]NCCT,USEPA
[4]NERL,USEPA

Lifestage Physiologically-based pharmacokinetic models, SHEDS-LITE models, and *In vitro* ToxCast data were used to estimate Adverse Outcome Pathway (AOP)-Based Margin of Exposures (ABME) for a set of chemicals related to vascular disruption. Initially a template Lifestage Physiologically-Based Pharmacokinetic (PBPK) model was developed to include detailed physiological processes (such as tissue volumes, blood flow rates, breast milk composition and flow rate, Etc.) whenever relevant during female growth from fetus to childhood to conception, to pregnancy and for one year after delivery. This model was applied to conduct *In vitro* to *In vivo* (IVIVE) extrapolation to estimate maternal exposure that will yield fetal blood levels equivalent to the estimated EC50s from *in vitro* ToxCast Assay related to an AOP of interest. The resulting *in vivo* dose estimates from the Lifestage model were then compared to real life exposures using models developed in NERL (SHEDS-LITE). An AOP-based margin of exposure (ABME) is then calculated as the ratio of the PBPK-derived *in vivo* estimate to the SHEDs-LITE exposure one. This computational framework was utilized for a list of 4 chemicals related to a putative Vascular Disruption AOP. The estimated ABMEs ranged from (10 to 10,000). The lowest values for Triclosan, Pyridaben, and PFOS indicate exposure risks that are within range to environmental levels while that higher indicated that environmental exposure are well below and risk ones (Fluazinam). The idea of using ABME is novel in linking biological information related to toxicity (using AOPs), HPT *in vitro* data (ToxCast), and age-varying physiological and biochemical information (using experimental clearance data or transport kinetics, partitioning in breast milk) to estimate Margins of Exposures that can be used to help regulators in prioritizing and assessing real-life risk of chemicals in view of their toxicity, dosimetry, and real-life exposure. This abstract does not necessarily reflect EPA policy

49 Incorporating Population Variability and Sensitive Populations into Dosimetry for High-Throughput Toxicity Testing

Barbara A. Wetmore[1], John. F. Wambaugh[2], Lisa M. Almond[3], Harvey J. Clewell III[1], Russell S. Thomas[2]
[1]The Hamner Institutes for Health Sciences, Research Triangle Park, NC, USA
[2]Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA
[3]Simcyp Ltd. (a Certara company), Sheffield, UK

Several efforts have been underway in the U.S. and abroad to assess the ability of HTS assays to evaluate human health effects of chemicals while conserving limited testing resources. Recently, a strategy was introduced that incorporated human dosimetry and exposure information with HTS data. While the results

of this work provided an important *in vivo* context to *in vitro* HTS data, the strategy was limited to dosimetry for a healthy, adult population. Sole reliance on these values could significantly underestimate the risk to children or other sensitive populations due to differences in pharmacokinetics and physiology.

Differences in metabolic enzyme abundances and activities exist across different life stages and ethnic populations and can lead to differences in chemical clearance rates. In this work, the contribution of individual phase I and II metabolic enzymes to the clearance of a subset of ToxCast chemicals was experimentally measured using recombinantly expressed enzymes. These isozyme-specific clearance rates were then incorporated into an *in vitro* to *in vivo* extrapolation (IVIVE) modeling approach which reassembles the data for the 18 isozymes assessed while also accounting for known differences in isozyme abundances and physiology for different life stage, ethnic, or health-based populations. Steady state concentrations (Css) estimated for each chemical-population combination were then used to calculate oral equivalent values: the amount (at mg/kg/day) an individual in a particular population would need to consume to achieve target tissue concentrations that elicit bioactivity in the ToxCast *in vitro* assays. These values can then be directly compared against life stage- or population-specific exposure estimates or predictions. In addition, comparison of the median Css for the general population against the upper 95th percentile for sensitive populations may be used to derive chemical-specific human toxicokinetic adjustment factors. Incorporation of population-specific dosimetry with *in vitro* bioactivity data in this manner provides a viable approach that could potentially be employed within a high-throughput risk assessment framework. This abstract does not necessarily reflect EPA policy.

50 *In vitro* Bioactivity in ToxCast Assays for Fruit and Vegetable Juices

B.A. Wetmore[1], M.B. Black[1], E.L. Berg[2], M.A. Polokoff[2], K.A. Houck[3], R.S. Judson[3], R.S. Thomas[3], and M. E. Andersen[1]

[1]The Hamner Institutes for Health Sciences, Research Triangle Park, NC

[2]BioSeek, a division of DiscoveRx, Inc., S. San Francisco, CA

[3]National Center for Computational Toxicology, Office of Research and Development, USEPA, Research Triangle Park, NC

The ToxCast and Tox21 programs have generated *in vitro* screening data for over 1000 chemicals to aid in hazard identification and setting chemical testing priorities. These data, together with *in vitro* pharmacokinetic data, are used to infer possible toxic responses and external concentrations required to elicit these effects. There is only limited experience in evaluating dose-response for natural products in these assays. In this study, juices were extracted from 30 organically grown fruits and vegetables. These juices were screened in concentration-response format across primary human cell and co-culture assays (BioMAP systems) to compare juice bioactivities against those of ToxCast chemicals and reference compounds. Extensive bioactivities were observed across the BioMAP Diversity 8 Panel, with juices eliciting 4-fold more LEC (i.e., lowest effective concentration) hits than ToxCast chemicals. Unsupervised similarity searches against the BioMAP reference compound database identified similarities between several juices and therapeutic and/or industrial chemicals. Broccoli displayed significant similarity to mitomycin C and the fungicide mancozeb; kale and sweet potatoes displayed similarities to JAK3 and JAK1/2 kinase inhibitors, respectively. To relate *in vitro* concentrations to administered dose, the filtered juice yield per g item was used with the plasma volume of a 75 kg adult to approximate % juice present in systemic circulation after eating, assuming complete uptake and no first-pass metabolism. This value was used as a surrogate for target tissue concentration. To achieve tissue concentrations equivalent to the concentration at which apples elicited LECs in 10% of the assays, one-half cup (or 0.06 kg) apples would need to be consumed. Alternately, assuming a daily exposure to fludioxonil at maximal residue levels across over 20 commonly consumed fruits and vegetables (~3 mg/kg produce), one would need to consume 4.5 kg apples to achieve the fludioxonil oral equivalent (0.18 mg/kg body weight/day) associated with the *in vitro* 10% LEC hit rate concentration. It is important to note that the bioactivities observed do not necessarily lead to adverse effects. These data provide context for assessing the *in vivo* relevance of *in vitro* concentration-response and bioactivity data generated in ToxCast. This abstract does not necessarily reflect EPA policy.

51 Merging ToxCast™ with ExpoCast™: Incorporating High-Throughput Dosimetry-Adjusted ToxCast Bioactivity Data with High-Throughput Exposure Predictions to Inform Rapid Risk-Based Prioritization

B.A. Wetmore[1], J.W. Wambaugh[2], R. Woodrow Setzer[2], M.A. Sochaski[1], R.S. Judson[2], S. Ferguson[3], K.A. Houck[2], H.J. Clewell III[1], R.S. Thomas[2], and M. E. Andersen[1]
[1]The Hamner Institutes for Health Sciences, Research Triangle Park, NC
[2]National Center for Computational Toxicology, USEPA, Research Triangle Park, NC
[3]NIEHS, Research Triangle Park, NC

Our previous work incorporating dosimetry and exposure with high-throughput screening (HTS) enhanced the utility of HTS data by translating *in vitro* bioactivity concentrations to external exposures required to achieve these levels internally. These oral equivalents were compared against exposure estimates to provide a margin of exposure or activity to exposure ratio (AER). As ToxCast has transitioned to a wider range of chemicals that lack regulatory exposure information, exposure prediction tools become increasingly important. In this study, *in vitro* hepatic clearance and plasma protein binding were measured to estimate dosimetry-adjusted oral equivalents (mg/kg/day) for Phase II chemicals. Chemical exposures were predicted using two ExpoCast™ tools. The first iteration utilized HT far-field mass balance models coupled with an indicator of near-field human exposure, while the second utilized Bayesian methodology to infer ranges of exposures consistent with NHANES urinary biomarker data coupled with HT heuristics to identify chemical descriptors to explain variation across demographic groups. These values were compared against ToxCast oral equivalent values to assess the potential that predicted exposures might elicit bioactivity. These two tools have provided exposure predictions across all Phase 2 chemicals when previously regulatory estimates were only identified for ~5%. In addition to refining the estimates, the latest tool identified five chemical descriptors that explained approximately 50% of the variability across all demographics. Comparison of the minimum oral equivalent to the upper bound exposure predictions from the latest model indicated that 2 and 24 of the 179 chemicals had AERs <1 and < 100, respectively. The availability of HT exposure estimation and bioactivity screening tools provides an opportunity to incorporate a risk-based strategy for use in prioritization of testing. This abstract does not necessarily reflect EPA policy.

52 The ExpoDat High Throughput Exposure Assessment of 229 ToxCast I and II Chemicals: Comparing Near- and Far-Field Exposures to *In vitro* Bioactivities

Susan A. Csiszar (1), Alexi Ernstoff (1,2), Hyeong-Moo Shin (3), Deborah Bennett (3), Jon Arnot (4), Peter Fantke (2), Barbara A. Wetmore (5), Xianming Zhang (6), Olivier Jolliet (1)
(1) Environmental Health Science, School of Public Health, University of Michigan, USA
(2) Quantitative Sustainability Assessment, Technical University of Denmark, Kgs. Lyngby, Denmark
(3) Department of Public Health Sciences, University of California, Davis, CA USA
(4) Arnot Research & Consulting, Inc, Toronto, Canada
(5) The Hamner Institute, Research Triangl Park, NC, USA, (6) Harvard School of Public Health and School of Engineering and Applied Sciences, Harvard University, MA, USA.

This presentation provides high-throughput Tier I (screening-level) risk–based prioritization of 229 ToxCast chemicals. Daily chemical intake rate (iR; mg/kg/day) exposure estimates were compared to effect estimates, i.e. oral equivalent doses (OED; mg/kg/day) calculated from the ToxCast assay yielding the most sensitive *in vitro* bioactivity. We first used the CPCat and the Household Product Database to categorize chemicals in terms of usage and associate the chemical-usage combinations to the relevant exposure pathways. For each chemical and use category, physical-chemical properties and mass balance models were used to calculate the intake fraction (iF, mass taken in by an individual or the population per mass emitted to the environment), or the product intake fraction (PiF, mass taken in per mass of chemical in product), in the case of personal care product applications. The iF and PiF were then multiplied by the estimated quantity emitted, applied, or ingested to determine exposure in mg/kg/day and to compare to OEDs.

We found, using conservative approaches, that chemical intake from consumer product use generally exceeds intakes from environmental emissions, with the greatest intake estimates for direct intakes and for

personal care products. Population-scale average intakes from far-field environmental emissions were usually several orders of magnitude below near-field product related exposures.

When comparing intake estimates with the bioactivity data, we found that about a third of the 229 chemicals have the maximum iR that exceeded the minimum OED for observed biological activity. For these chemicals the following chemical/exposure pathways combinations exceeded minimum OEDs: direct inhalation or ingestion (for 24% of the 229 chemicals), use of personal care product (for 24%), use of product emitting indoors (for 12%), pesticide food residues (for 8%) and exposure to far-field environmental emissions (for 5%). These chemical exposures need refined (higher tier) chemical use and exposure information to understand possibility of public health effects.

Overall, this Tier I approach leads to results that are useful as preliminary high-throughput exposure estimates and may inform testing prioritization strategies. Sales volume data and concentrations of chemicals in products for food or cosmetic items, for example, would help enable more accurate modeling of exposure.

This research is supported by the Long-Range Research Initiative of the American Chemistry Council

Integrating and Applying ToxCast Data

53 An *In vitro* Test System Model to Facilitate the Use of Toxicity and Bioassay Data for Screening-Level Hazard and Risk Assessment

James M. Armitage(1,2), Frank Wania(2), Jon A. Arnot(1,2)
(1) ARC Arnot Research & Consulting Inc., Toronto, Canada
(2) University of Toronto Scarborough, Toronto, Canada

The use of *in vitro* toxicity data for conducting chemical safety evaluations and risk assessment is gaining prominence due to the sheer number of compounds under consideration and the practical, financial, and ethical considerations related to conducting extensive *in vivo* testing. Although it is well understood that the behaviour of chemicals in *in vitro* test systems (e.g., partitioning between different phases) can have an important influence on observed responses, detailed consideration of these factors prior to conducting toxicity tests is not routine. To address these aspects, we recently developed a generic mass balance model to simulate the distribution of neutral organic chemicals in *in vitro* systems given the input parameters entered by the user. For example, in addition to physical-chemical properties, the user is able to define the volume fraction of serum constituents (i.e., serum albumin and lipids) and the properties of the cells (e.g., lipid content). Given the (initial) nominal concentration, the model calculates the freely-dissolved concentration in the bulk medium as well as the concentrations of the chemical in the cells and other phases present in the system. The model also calculates the chemical activity in the system. Chemical activity is rooted in thermodynamic considerations and can be used to distinguish between chemicals exerting baseline toxicity (e.g., membrane disruption) and more specific modes of action (e.g., acetylcholinesterase inhibition). Chemical activity can also be calculated using measured or modeled concentrations in the environment and then compared to chemical activities associated with (adverse) effects determined in toxicity testing. We recently compiled physical-chemical properties for nearly 1200 ToxCast chemicals and used these inputs to relate nominal concentrations (1 nM–50 µM) to freely-dissolved concentrations, cell concentrations and chemical activities under different model parametrizations. Cell concentrations and chemical activities for the same nominal concentration spanned orders of magnitude, highlighting the value of the model for interpreting *in vitro* test data. The model can also be used to rank chemicals with respect to relative potency by relating Effect Concentrations (EC50s) to chemical activities in the system. Both applications of the model are illustrated here and discussed in the context of risk assessment.

54 Using ToxCast Data and the Comparative Toxicogenomics Database to Identify Potential Mechanisms of Human Toxicity

*Bingjie Hu[1], Huijun Wang[1], Timothy Johnson[2], Wendy Bailey[2], Carolyn Mattingly[3], Eric Gifford[1]**

[1]Structural Chemistry, Merck Research Laboratories, Merck & Co., 770 Sumneytown Pike, West Point, PA 19486, USA

[2]Safety Assessment and Laboratory Animal Resources, Merck Research Laboratories, Merck & Co., 770 Sumneytown Pike, West Point, PA 19486, USA

[3]Department of Biological Sciences, North Carolina State University, Raleigh, NC 27695, USA

Many studies have attempted to predict *in vivo* hazards based on the ToxCast *in vitro* assay results with the goal of using these predictions to prioritize compounds for conventional toxicity testing. Most of these conventional studies rely on *in vivo* endpoints observed using preclinical species (e.g., mice and rats). Although the preclinical animal studies provide valuable insights, there can often be significant disconnects between these studies and safety concerns in humans. One way to address these concerns, for an admittedly more limited set of compounds, is to explore relationships between the *in vitro* data from human cell lines and observations from human related studies. The Comparative Toxicogenomics Database (CTD; http://ctdbase.org) is a rich source of data linking chemicals to human diseases/adverse events and pathways. In this study we will explore the relationships between over 500 ToxCast chemicals, their ToxCast *in vitro* test results and their annotations of human diseases/adverse events endpoints as captured in the CTD database. We propose mining these associations to identify potentially interesting, statistically significant, hypotheses of mechanism(s) of toxicity. The correspondence between these associations and associations determined using preclinical *in vivo* studies will also be examined. To the best of our knowledge, this is one of the first studies analyzing the relationships between the ToxCast *in vitro* assays results and the CTD disease/adverse event endpoint annotations. The *in vitro* profiles identified in this analysis may prove useful for prioritizing compounds for toxicity testing, suggesting mechanisms of toxicity, and forecasting potential *in vivo* human drug induced injury.

55 Exploring the Potential Utility of High-Throughput Bioassays Associated with US EPA Toxcast Program for Effects-Based Monitoring and Surveillance

A. Schroeder[1], G. Ankley[2], J. Berninger[3], K. Lee[4], M. Lee[2], K. Houck[5], K. Jensen[2], Z. Jorgenson[6], R. Judson[5], D. Villeneuve[2]

[1]University of Minnesota – Water Resources Center, St. Paul, MN

[2]U.S. EPA Mid-Continent Ecology Division, Duluth, MN

[3]National Research Council

[4]US Geological Survey – Minnesota Water Science Center, Mounds View, MN

[5]U.S. EPA National Center for Computational Toxicology, Research Triangle Park, NC

[6]US Fish and Wildlife Service – Environmental Contaminants , Bloomington, MN

Environmental monitoring and surveillance strategies are essential for identifying potential hazards of contaminant exposure to aquatic organisms. Chemical monitoring is effective for chemicals with well characterized hazards and for which sensitive analytical methods are available. Effects-based monitoring and surveillance utilizes various tools to identify the potential biological consequences of environmental contaminant exposures and provides a useful complement to chemical monitoring. Effects-based monitoring approaches have typically used targeted assays for which a hazard of concern has been identified. However, this tends to focus effects evaluation on a small number of well characterized pathways, rather than facilitating a more open-ended survey of biological activities. High-throughput assays can provide an effective approach for screening a wider range of biological targets. To test this concept, we employed the Attagene subset of assays utilized for the US EPA Toxcast Program to screen selected environmental samples for biological activity. The Attagene assays screen for chemical interactions with over 80 different transcription factors. In the fall and spring of 2012, composite, ambient, water samples were collected at four locations, each with varying proximity to a wastewater discharge, within the St. Louis River Area of Concern, MN. Of the 137 analytes measured by instrumental analysis, 72 of the analytes had been run in Toxcast. At any particular location, the number of analytes detected ranged from 13 to 44 and

greater than half of the chemicals detected at any location had response profiles in Toxcast. The number of gene targets significantly impacted in the Attagene assays varied substantially across locations. However, all of the gene targets identified at each location could be explained by chemicals that have Toxcast response profiles. Response profiles from the same locations sampled in the fall and spring were compared to identify gene targets that may be indicative of seasonality. The gene targets identified using this unsupervised approach can now be used to direct future targeted analyses to examine trends over time at these locations. The results highlight the use of high throughput assays for surveillance and identifying gene targets that can be useful for directing subsequent site-specific monitoring. The contents of this abstract neither constitute, nor necessarily reflect, official US EPA policy.

56 A ToxBank Integrated Data Analysis of SEURAT-1 Reference Compounds

Glenn Myatt (Leadscope), Pekka Kohonen (Karolinska Institute), Roland Grafstrom (Karolinska Institute), Nina Jeliazkova (Ideaconsult), Markus Hegi (Douglas Connect), Barry Hardy (Douglas Connect)

The SEURAT-1 (Safety Evaluation Ultimately Replacing Animal Testing-1) research cluster is comprised of seven EU FP7 Health projects and is co-financed by Cosmetics Europe. The SEURAT-1 strategy is to adopt a mode-of-action framework to describe repeated dose toxicity to derive predictions of *in vivo* toxicity responses. ToxBank is the cross-cluster infrastructure project which provides a web-accessible shared repository of research data and protocols. Experiments generate dose response data over multiple timepoints using different omics platforms including transcriptomics, proteomics, metabolomics, and epigenetics over different cell lines and a common set of reference compounds (details available at wiki.toxbank.net). Data is also generated from functional assays and bioreactors and supplemented with in silico approaches. This complex and heterogeneous data is consolidated and harmonized through the ToxBank data warehouse in order to perform an integrated data analysis. We describe for 14 reference compounds the meta-analysis of multiple types of time-dependent dose response omics and functional data combined with *in vitro* and *in vivo* background knowledge including consideration of modeling variations in biokinetic responses. Open TG-GATEs human *in vitro* liver data of the reference compounds includes reactive compounds (e.g., acetaminophen, CCl4), mitochondrial disruptors (e.g., Rotenone), promiscuous binders (e.g., valproic acid, amiadarone), nuclear hormone receptor ligands (e.g., tamoxifen, WY14643), selective binders (e.g. fluoxetine) and cardiotoxins (e.g., Doxorubicin, Nifedipine). Adverse events of interest that are represented include cytotoxicity, fibrosis, steatosis, cholestasis and phospholipidosis. Overall we obtained 31,717 differential expression results with 14 compounds from the 45 comparisons, with Doxorubicin for example providing over 5000 results. We evaluate the use of ToxCast and PubChem data in the enrichment analysis, read across and interpretation of the evidence on reference compounds as mapped to biological pathways. The approach includes the use of the ISA-Tab standard to describe experimental metadata and OpenTox services supporting interoperable data integration and analysis. In addition, a proposed data standard for omics and assay data is incorporated in supporting the meta analysis.

57 High Throughput Pharmacokinetic Modeling Using Computationally Predicted Parameter Values: Dissociation Constants

Cory Strope (EPA-ORD-NCCT), Kamel Mansouri (EPA-ORD-NCCT), Jayaram Kancherla (EPA-ORD-NCCT), Caroline Stevens (EPA-ORD-NERL), John Wambaugh (EPA-ORD-NCCT)

Estimates of the ionization association and dissociation constant (pKa) are vital to modeling the pharmacokinetic behavior of chemicals *in vivo*. Methodologies for the prediction of compound sequestration in specific tissues using partition coefficients require a parameter that characterized the charge of the atom type. Current methods for reporting the pKa report only the pH at which the pKa-associated atom will be ionized in 50% of the molecules. Important considerations for *in vivo* usage of the pKa such as (i) the chemical class (i.e., acid/base), or (ii) the interplay between ionization states at other atoms to determine the fraction of a chemical to exist in a particular ionization state are reduced to "missing" information status. We propose a new method that more fully describes the process associated with of reporting pKa values. Further, this new format is designed to support high-throughput applications. We are comparing the ionizable atom types between 815 pharmaceutical and 2200 environmental

compounds, and investigating the performance of several publically and commercially available pKa predictive models on these 3015 chemicals from published sources and the ToxCast library. Finally, the analysis methodology developed herein for efficient estimation of the parameters critical for predicting chemical pharmacokinetics will be publicly accessible as an R package. This abstract does not necessarily represent U.S. EPA policy

58 Elucidating Mechanisms of Toxicity Using Phenotypic Data from Primary Human Cell Systems: A Chemical Biology Approach for Thrombosis-Related Side Effects

Ellen L. Berg, Mark A. Polokoff, Alison O'Mahony, Xitong Li
BioSeek, a division of DiscoveRx
eberg@bioseekinc.com, mpolokoff@bioseekinc.com, aomahony@bioseekinc.com, xli@bioseekinc.com

Here we describe a chemical biology approach for elucidating potential toxicity mechanisms for thrombosis-related side effects in humans. In previous work with the EPA for the ToxCast™program, it was observed that several classes of chemicals, including aryl hydrocarbon receptor (AhR) agonists and estrogen receptor (ER) antagonists, shared an usual activity, that of increasing the level of TF, the primary cellular initiator of coagulation, in a primary human endothelial cell-based model of vascular inflammation, the 3C BioMAP system (Kleinsteuer et al., NBT, 2014, 32:583). Since human exposure to these chemical classes is associated with increased incidence of thrombosis-related side effects, it was of interest to search our large reference database for other compounds that also increase TF levels, to provide a better understand the regulation of tissue factor under these conditions, and potentially a more mechanistic understanding of clinical observations. By searching our reference database of more than 3000 test agents, we identified agents that either increased or decreased the level of tissue factor in the BioMAP 3C system (>20%) at two or more concentrations, without causing overt cytotoxicity. From this effort, mechanisms regulating tissue factor expression under these conditions were identified. Mechanisms for positive regulation of TF that were found include: AhR, H1R, HDAC, hsp90, HIF-1α, IKK2, MEK, NOD2, OSMR, PKC, TR, Jak kinase, and p38 MAPK function, whereas mechanisms involved in negative regulation include Estrogen R, mTOR and vacuolar ATPase (V-ATPase). Given the role of several of these targets in the process of autophagy, and that mTOR and HIF-1 α have been suggested to comprise a functional switch regulating hypoxia induced autophagy in endothelial cells, these data suggest that this process may also involved in thrombosis-related side effects *in vivo*. This effort may help in the construction of an adverse outcome pathway framework for thrombosis related-side effects.

59 The Human Toxome Collaboratorium: Meeting the Challenges of Multi-omic, Multi-institutional Collaboration in the Cloud

Rick A. Fasani^, Carolina Livi, Michael Rosenberg
Agilent Technologies, Inc., 5301 Stevens Creek Blvd., Santa Clara, CA 95051, USA
^rick.fasani@agilent.com

The Human Toxome Project is part of a long-term vision to modernize toxicity testing for the 21st century. In the initial phase of the project, a consortium of six academic and commercial organizations has partnered to comprehensively map pathways of toxicity, using endocrine disruption as a model system along with a range of experimental and computational tools. Of course, effectively gathering, managing, and analyzing the data from high-throughput experiments is a challenge in its own right—doing so for a growing number of -omics technologies, with ever-larger data sets, across multiple institutions complicates the process at every step. To meet the needs of the project, we have created and managed The Human Toxome Collaboratorium, a shared computational environment hosted on third-party cloud services. The Collaboratorium provides a familiar virtual desktop, with a powerful mix of commercial, open-source, and custom-built applications, shared across the trust boundaries of the member organizations. As such, it shares some of the challenges of traditional information technology, but with unique and unexpected constraints that emerge from the cloud. Since its inception three years ago, the Collaboratorium has evolved, and continues to evolve, based on the needs of the group the more efficient use of cloud resources, both for high performance computing and more traditional IT services. The result is a relatively low-cost, secure, modular, and flexible solution.

Here we describe the current architecture of the Collaboratorium, the major lessons we have learned, and the potential application of the Collaboratorium concept to other projects.

60 ToxCast Assay Network (TCAN) Viewer: A Big-Picture Visualization Tool of High-throughput Assays for Environmental Chemicals.

Parth Kothiya, Jimmy Phoung, Matt Martin
National Center for Computational Toxicology, Office of Research and Development, US EPA, RTP, NC

USEPA's ToxCast program provides high-quality screening data on thousands of chemicals. The huge library of compounds and wide diversity of high-throughput assays in ToxCast present daunting practical challenges with respect to big data visualization. Recent developments in R applications to visualize trends in the data have overcome some of the hurdles in High-Throughput Screening data interpretation. However, what is available doesn't enable data comparisons in a way that is friendly or easy. To better incorporate a dynamic viewer to prioritize the landscape of high-throughput screening data, we introduce the GGVIS-based ToxCast Assay Network (TCAN) Viewer, a visualization tool to look across assays and see patterns in the annotation features with respect to a continuous variable (e.g. concentration, time). Each node in the annotation network is a multilayer circle, where each layer indicates color density with respect to a particular dose for each chemical. It can also provide spider plot visualization at each nodes. The goal of the TCAN is to provide visualizations in three different form: chemicals (spider plot), dose responses (multilayer circle) and an assay/gene representation annotation network. The hope is that this viewer would fraction up the active chemicals by the continuous variable and allow for a friendly way to form hypotheses and develop prioritization schemes. This work does not necessarily reflect Agency policy.

61 The New ToxCast Analysis Workflow

Dayne Filer, Parth Kothiya, Richard Judson, Woody Setzer, Keith Houck, Matt Martin
US EPA/ORD/NCCT

US EPA's ToxCast program is generating data in high-throughput screening (HTS) and high-content screening (HCS) assays for thousands of environmental chemicals, for use in developing predictive toxicity models. Currently the ToxCast screening program includes over 1800 unique chemicals and over 700 bioassays. Data collection from multiple assay vendors and collaborators in diverse formats, data standardization, and normalization, coupled with the needs for data transparency and consistency have posed major technical and scientific challenges to the program. We developed a new ToxCast data workflow to allow for consistent cross-technology analysis and data reporting. The new workflow improves upon previous efforts by increasing efficiency, a new curve-fitting algorithm utilizing three robust models that do not rely on outlier detection, and using a more sophisticated hit-calling algorithm. We also provide many new summary statistics for the modeling and hit-calling results, including a new set of flags to identify potential false positives and false negatives. All processing occurs through a novel R package that interacts with a relational database. The new ToxCast workflow standardizes the analysis of highly-heterogeneous chemical assay data sets by accepting heterogeneous data format. Public release the R package and a relational database dump file containing all data and results will enable transparent and reproducible analyses. This work does not necessarily reflect Agency policy.

62 Communication and Data Integration Features within the Annotations for the ToxCast Assay Library

Jimmy Phuong, Keith Houck, Lisa Truong, Nisha Sipes, Parth Kothiya, Richard Judson, Matthew Martin

High-throughput screening (HTS) assays have increased the pace of chemical-biological data collection for faster throughput and increasing our understanding of chemical-biological interactions. The EPA's ToxCast Program has amassed millions of data points and is continually seeking new chemicals and assays to screen. This comes with a growing information demand, so we initially provided 37 annotations to capture important metadata that can help data users understand the technologies and biology behind the data. We also developed methods incorporating the annotations for differentiating and aggregating the data. One

important feature is the annotation structure, which connects the assay endpoints (the analyses of the assay readouts and the way the data is released) to where the data came from and the annotations carried with them. This stretches to the assay component level (the raw assay readouts), the assay level (the experimental protocols), and their assay source level (the entity performing the assays). Furthermore, most of the annotation terms used are controlled vocabularies from the BioAssay Ontology and various other biomedical ontologies, providing definitions and hierarchical relationships such as for gene families, cellular pathway events, and technology groups. The goal for the annotations is to act as a sustainable and central source for metadata, continued support for analytical and dashboard developments in areas like the data processing pipeline, creation of data viewers and decision making tools, and enable linkages to other HTS data repositories. This abstract does not necessarily reflect US EPA policy.

63 Environmental Probabilistic Hazard Assessments of ToxCast Phase I and II Data Related to Oxidative Stress Assays

Lauren A. Kristofco[1], Jone Corrales[1], E. Spencer Williams[1], Margaret Mills[2], Evan P. Gallagher[2], Terrance J. Kavanagh[2], Nancy Simcox[2], Longzhu Q Shen[3], Evan Beach[3], Fjodor Melnikov[3], Jakub Kostal[4], Adelina Voutchkova-Kostal[4], Julie B. Zimmerman[3], Paul Anastas[3], Bryan W. Brooks[1]

[1]Baylor University, Department of Environmental Science, Waco, Texas, 76798, United States

[2]University of Washington, Department of Environmental and Occupational Health Sciences, Seattle, Washington, 98105, United States

[3]Yale University, Center for Green Chemistry & Green Engineering, New Haven, Connecticut, 06511, United States

[4]The George Washington University, Department of Chemistry, Washington DC, District of Columbia, 20052, United States

The potential hazards that commercial chemicals can pose to human health and the environment through unintended biological activity has prompted a drive for more detailed study in recent years. Of the 700+ commercial chemicals introduced to the US market each year, over 85% are approved for manufacturing despite the lack of toxicity data. Our research group has examined approaches to utilize available toxicological data to predict thresholds of untested chemicals, and to select assays for assessing toxicity when multiple model systems exist. Here we employed chemical toxicity distributions (CTDs) to perform probabilistic hazard assessments (PHAs) using data from the U.S. EPA's ToxCast program. PHAs may predict toxicological potencies of similar chemicals, prioritize chemicals for additional toxicity testing, support read-across for regulatory purposes, and identify characteristics for sustainable molecular design. This PHA approach was used to specifically look at ToxCast Phase II data associated with oxidative stress. CTDs were compared at 5th centiles first for all pesticides and pharmaceuticals and then for classes of pesticides and pharmaceuticals. A PHA of an oxidative stress (H2AX) assay demonstrated decreasing assay response with increasing assay duration, based on 0.77 and 0.263 ug/L for 24hrs and 72hrs, respectively. Conversely, a stress kinase (c-Jun) assay did not display such a relationship with 5th centiles of 0.082, 0.550 and 1.099 ug/L for 1hr, 24hrs and 72hrs, respectively. Moreover, oxidative stress assays specific to AP-1 and Nrf2 were both more sensitive to pesticides than pharmaceuticals. Strengths and weaknesses of this approach continue to be explored.

64 Bioactivity of Food-Relevant Chemicals in ToxCast

Agnes Karmaus[1,2,3] and Keith Houck[3]

[1]ILSLI NA Technical Committee on Food and Chemical Safety Summer Fellow

[2]Oak Ridge Institute for Science Education Fellow

[3]NCCT/ORD/US EPA, RTP, NC

ToxCast screening program has evaluated 1,892 chemicals across hundreds of *in vitro* assays to prioritize chemicals needing more extensive toxicological evaluation and to identify targets for potential toxicity pathways. To identify bioactive food-relevant chemicals, the ToxCast chemical inventory was mined for food-relevant chemicals including direct food additives, Generally Recognized As Safe (GRAS) ingredients, Food Contact Substances (FCS), and indirect food additives. Publicly available FDA sources including the

Everything Added to Food in the US (EAFUS), GRAS and FCS databases identified ~8,000 food-relevant chemicals. The 1,892 ToxCast phase II chemicals included 806 food-relevant chemicals, which had median AC50 values generally ranging between 71-100 µM across assays. The food-relevant chemicals can be separated into three categories: 291 direct food ingredients including GRAS chemicals, 247 indirect food additives or food contact substances, and 268 pesticides. Pesticides elicited effects in an average of 33 assays per chemical with 5 chemicals having AC50 values below 5 µM. Meanwhile, direct food additives elicited effects on an average of only 10 assays with 18 chemicals having AC50 values below 5 µM, revealing greater biochemical selectivity. Although biological activities in the *in vitro* assays do not necessarily lead to adverse effects *in vivo*, the results suggest that high-throughput screening can identify food-relevant chemicals that elicit bioactivity and the cellular pathways that are most frequently targeted. This work was sponsored by the ILSI NA Food and Chemical Safety Committee 2012 Summer Fellowship Program. This abstract does not necessarily reflect US EPA policy.

65 Profiling 774 Chemicals for Disruption of Male Reproductive Development: Hierarchical Relationship across Apical Endpoints (ToxRefDB) and Chemical-Effects Clusters for *In vitro* Bioactivity (ToxCast)

Maxwell C.K. Leung†, Jimmy Phuong†, Nancy C. Baker‡, Nisha S. Sipes*†, Gary R. Klinefelter§, Matt T. Martin†, Keith W. McLaurin†, Ann M. Richard†, R. Woodrow Setzer†, Sally D. Perreault§, Richard S. Judson†, and Thomas B. Knudsen†*

**Oak Ridge Institute for Science Education, Oak Ridge, Tennessee, 37831*

†National Center for Computational Toxicology and §Toxicity Assessment Division, National Health and Environmental Effects Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711

‡Lockheed Martin, Research Triangle Park, North Carolina 27713

Compilations of information and data for large numbers of chemicals through analysis of the open literature and public databases, together with newer data-mining and visualization techniques, enable chemical profiling and classification based on complex phenotypic and chemical-effects relationships. The objective of this study was to evaluate such relationships for male reproductive developmental toxicity, utilizing U.S. EPA Toxicity Reference Database (ToxRefDB) and Toxicity Forecaster Database (ToxCastDB). Stepwise filtering of 3,479 guideline studies in rodent/rabbit species for 774 chemicals evaluated for adverse effects specific to the male reproductive system returned 281 chemicals across 1,653 studies with 23 different male reproductive endpoints. The hierarchical aspects of Testicular Dysgenesis Syndrome (TDS; i.e. testicular atrophy > sperm effects > tumors > malformations) were revealed in 63 chemicals following early lifestage exposure in prenatal developmental, one-generational, and/or multigenerational reproductive studies. Data for 48 chemicals in ToxCastDB were used to construct a bipartite network to visualize the complex chemical-effects relationships on 126 molecular targets. Primary clusters were arranged by chemical classes (e.g. phthalates, conazoles, carbamates, and phenol compounds) and *in vitro* bioactivity profiles (e.g. nuclear receptors, G-protein coupled receptors, and cytochrome-P450 reductases). Although a TDS-like hierarchy was evident in ToxRefDB, the chemical-effects network in ToxCastDB was pleiotropic and did not support a simple, unified mode of action for TDS. Rather, the available HTS data suggest that pleiotropic effects may contribute to TDS-like symptoms during male reproductive development. [Disclaimer: this work does not reflect EPA policy]

66 Scientific Text Extraction Using FIDDLE: A Foundation for Accurate Literature Mining

Jason Phillips, Brian E. Howard, Ruchir Shah

(jason.phillips, brian.howard, ruchir.shah)@sciome.com

Sciome, LLC

A principal aim of the ToxCast high-throughput screening program is to increase the efficiency with which potentially dangerous chemicals can be evaluated with regard to their impact on human health. With similar intent, recent research has demonstrated the usefulness of text-mining and natural language processing to extract valuable information about the likely biological impact of chemicals from latent information and

relationships hidden in the scientific literature. In practice, however, an important technical obstacle to the wider applicability of these methods has been the difficulty of accurately extracting chemical names and other scientifically relevant text from a large corpus of scientific documents. This is difficult not only due to the fact that chemical nomenclature is so variable and imprecise, but also because the portable document format (PDF), in which most scientific texts are published, sacrifices word order and natural text flow in favor of a more graphically-oriented representation. Here we introduce the 'Flow-Intelligent Document Decoder for Literature Extraction' (FIDDLE), a novel algorithm and a tool under development, which can accurately extract text from scientific PDF documents. In addition to preserving the correct word order across columns, and even around figures and tables, FIDDLE is also able to accurately sectionalize scientific documents in an automated fashion. When key scientific terms (e.g. chemical names) are mentioned in different document sections (e.g. title, abstract, methods, references, etc.) they may carry greatly different implications. Therefore, the ability to correctly divide documents into meaningful sections is a critical capability in the context of literature mining. When combined with a custom built, dictionary-based chemical name recognizer, FIDDLE is able to accurately extract ToxCast/Tox21 chemical names from full text PubMed manuscripts at a rate several times greater than would otherwise be possible using only the associated MeSH terminology or the text of the titles and abstracts. Our approach offers the ability to perform full text extraction of scientifically relevant keywords (e.g. chemical name or phenotypic endpoint) and in turn greatly enriches text and literature mining capabilities with wide ranging applicability in the environmental and health sciences.

Regulatory Implications and Endocrine Disrupting Chemicals

67 Integrating ToxCast Assays into an Androgen Receptor (AR) Pathway Model

N Kleinstreuer, ILS/NICEATM, D Filer, EPA/ORD/NCCT, M Martin, EPA/ORD/NCCT, K Houck, EPA/ORD/NCCT, D Zang, ILS/NICEATM, D Allen, ILS/NICEATM, W Casey, IH/NIEHS/DNTP/NICEATM, R Judson, EPA/ORD/NCCT

The Tox21 and ToxCast programs include multiple *in vitro* assays conducted in a high-throughput screening (HTS) format that are relevant to the AR pathway and can be used to identify substances with potential androgenic/anti-androgenic activity *in vivo*. Here we used a number of assays that map to the androgen receptor (AR) pathway to build a mathematical model that attempts to delineate true AR pathway activity from technology-specific assay interference. This battery of 9 assays (5 from ToxCast and 4 from Tox21) probes perturbations of the AR pathway at multiple points (receptor binding, cofactor recruitment, gene transcription and protein production) in multiple cell types. We compiled a list of putative AR reference chemicals, covering agonists, antagonists, selective androgen receptor modulators (SARMs), and inactive chemicals, from the ICCVAM (2003) and OECD (2010) reference chemical lists. The model showed 96% (23/24) concordance across the reference set, including successfully identifying multiple SARMs with both agonist and antagonist activity. However, Fluoranthene, a SARM, was active in the cofactor recruitment assays but none of the other AR pathway assays, and was therefore mispredicted by the model as acting via an assay-specific interference pathway. All chemicals in the ToxCast library known to target AR were identified by the model as agonists or anatagonist. We will discuss a variety of patterns of assay activity and pathway predictions across 1846 ToxCast chemicals, and identify those prioritized to be active against the AR pathway. Where available, we will compare predictions to toxicity data from the literature and look for potential trends relating to use case and exposure scenarios. This project was funded in whole or in part with Federal funds from the NIEHS, NIH under Contract No.HHSN27320140003C and does not represent EPA or NIEHS policy.

68 Predicting Skin Sensitization Using ToxCast Assays

N Kleinstreuer, ILS/NICEATM, J Strickland, ILS/NICEATM, D Allen, ILS/NICEATM, W Casey, NICEATM/DNTP/NIEHS/NIH/HHS

Allergic contact dermatitis (ACD) is an adverse health effect from repeated exposure to skin-sensitizing chemicals and products. To minimize ACD, regulatory authorities require tests, like the murine local lymph node assay (LLNA), to identify potential skin sensitizers. The Organisation for Economic Co-operation and

Development (OECD) established an Adverse Outcome Pathway (AOP) for skin sensitization. Many organizations, including the OECD and NICEATM, are pursuing integrated testing strategies using novel *in vitro* and in silico approaches to reduce or replace animal use. The U.S. EPA's ToxCast project includes high-throughput screening (HTS) assays in human primary skin cells and other systems that map to key events in the AOP (e.g., oxidative stress, cytokines). We built a cross-validated random forest model using ToxCast data and a balanced training set of 60 chemicals with *in vivo* LLNA data. The model predicted LLNA results with 80% accuracy, representing the performance against all chemicals when they appear in external test sets. The assays with highest variable importance included known AOP targets (e.g., Nrf2, T-cell proliferation) as well as targets outside the current AOP (e.g., Coll III, PPAR, PXR, ER). Compounds mispredicted by the model were found to be structurally similar, and we will discuss potential enrichment of this approach by incorporating molecular descriptors. Well-characterized AOPs like skin sensitization provide opportunities to use ToxCast HTS data to identify critical biological targets and develop efficient testing strategies that minimize animal use in regulatory testing. This project was funded in whole or in part with Federal funds from the NIEHS, NIH under Contract No.HHSN27320140003C.

69 A Real-Time Evaluation of Human-Based Approaches to Safety Testing: What We Can Do Now

Robert A Coleman, Katya Tsaioun & Kathy Archibald, Safer Medicines Trust; Keith Houck, ORD/EPA

Despite ever-increasing efforts in early safety assessment in all industries, there are still many drugs that prove toxic in humans. While greater use of human *in vitro* test methods may serve to reduce this problem, the formal validation process applied to such tests represents a major hurdle to their adoption. We contend that what is really needed to justify the adoption of any new test is not a demonstration of an ability to identify all potential safety issues in one approach, but a clear demonstration that it is adding value or is superior to whatever is currently in use – 'pragmatic validation'. A study based on such pragmatic validation, comparing the value of a range of human-based *in vitro* test methods with established regulatory tests, is currently underway and preliminary data will be presented. Importantly, all the *in vitro* tests have undergone significant evaluation already through two phases of the EPA ToxCast Program and are already used to various degrees by many pharmaceutical, agrichemical and cosmetic companies internally. The pragmatic validation approach we are testing now is designed to reduce to practice the application of these tests, bring them to a common denominator and provide the guidance to industry and regulators on the appropriate context of use. The study employs a range of marketed drugs that passed regulatory safety testing but were subsequently withdrawn, having caused serious toxicity in human subjects. Each of these drugs is paired with a negative control, i.e. a structurally and/or functionally similar marketed drug that does not exhibit such toxicity. This study is now being conducted as a distinct part of US EPA's Phase 3 ToxCast *in vitro* profiling program. Data are to be made publicly available when testing is completed, at such time they will be submitted for detailed analysis to compare the performance of the new *in vitro* tests with the regulatory regime that secured the original marketing approval. On completion, the outcome of this unique study will be presented to the regulatory authorities with the aim of developing appropriate documents for use by the industry and published in peer-reviewed media. This abstract does not necessarily reflect US EPA policy.

70 Rank-Based nScore for Analysis of Compound Activity Signatures: Application to BioMAP Phenotypic Data

Xitong Li, Mark A. Polokoff and Ellen L. Berg
BioSeek, a division of DiscoveRx
xli@bioseekinc.com, mpolokoff@bioseekinc.com, eberg@bioseekinc.com

One challenge in applying *in vitro* methods to profile chemical compounds for mechanism-of-action and safety assessment decisions is that the phenotypic screening data are not directly comparable for different compounds or the same compounds at different assay concentrations. It is attractive to develop a consensus profile to capture the essential features of a compound even though the screening data vary at different concentrations. Here we describe a non-parametric method for building consensus compound

profiles using a normalized metric nScore that is based on the ranking of the screening endpoints. For this we used a previously reported reference data set (Berg, JBS, 2013, 18:1260) consisting of 88 compounds, representing 28 target/pathway mechanisms, assayed at several concentrations that were tested in a panel of 8 BioMAP systems, previously employed for the Environmental Protection Agency's ToxCast™program (Kleinstreuer, NBT, 2014, 32:583). The log ratio data of 84 features for each compound/concentration pair were ranked and nScores were calculated to build a consensus profile for each compound by calculating the median of the nScores across all concentrations. We also built nScore profiles for the 28 mechanisms by combining data from multiple compounds with the same target mechanism. We then evaluated profiles generated from additional compounds by performing similarity searches using Spearman correlation against these nScore profiles. Using the BioMAP nScore consensus profiles, built for 88 compounds and 28 mechanism classes, we performed similarity searches to identify similarities for the Aryl Hydrocarbon Receptor (AHR) antagonist, CH223191, and the AHR agonist, FICZ. Our results demonstrated that the CH223191 profile was significantly negatively correlated to FICZ profile, consistent with the known mechanisms. We also tested the profiles for EP R agonists and PDE IV inhibitors, two classes that were difficult to separate using a SVM-based classifier method (Berg, JBS, 2013, 18:1260), and found that these were more accurately classified by the nScore method. Our results indicate that BioMAP nScore profiling is an effective method for profiling compounds, finding similar agonists or antagonists, and building mechanism-of-action models.

71 Evaluating the Performance of the ToxCast Estrogen Receptor (ER) and Androgen Receptor (AR) Assays Using the ICCVAM List of Validation Chemicals

M. Sue Marty, The Dow Chemical Company, Dan Wilson, The Dow Chemical Company, Barun Bhhatarai, The Dow Chemical Company, Richard Becker, American Chemistry Council

In vitro assays are imperfect predictors of biological activity because all assays are subject to some rate of false positive and false negative responses. The goal of this study was to evaluate the performance of the ToxCast ER- and AR-related assays by comparing results with the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) list of recommended validation chemicals for ER and AR assays. For each of the assay types, mixed agonists/antagonists were not included in the evaluation. For the ER binding assays, 32 ICCVAM chemicals were evaluated in three ToxCast binding assays using human, mouse and bovine ER. Across these binding assays, the balanced accuracy (BA) was 0.82 with the human ER binding assay meeting the greatest proportion of expected responses. For the ER transactivation assays (ERTA), 43 chemicals were evaluated in up to 16 ERTA assays. Across the ERTA, the BA was 0.73. Across the two ER antagonist assays, there were fewer chemicals evaluated (27) and a lower BA (0.60). For the AR-related assays, BA could not be determined for AR binding assays due to a paucity of negative control compounds in the ICCVAM data set. Sensitivity for three AR binding assays (human, rat and chimpanzee; 24 chemicals) was 0.74. Across the six AR transactivation assays (ARTA), an evaluation of 42 chemicals yielded a BA of 0.71. As with ER, AR antagonism had only two assays with 23 chemicals evaluated. The BA for AR antagonist assays was 0.69. It is important to note that some ToxCast assays target specific aspects of receptor interactions (e.g., binding, dimerization, etc.) and therefore, would detect only a subset of ICCVAM chemicals. A more thorough comparison of ToxCast assay performance is hampered by the finite number of chemicals common to both databases. Despite this limitation, it appears that an optimized ToxCast battery may be useful for prioritizing compounds for further assessment based on their interaction with ER and AR signaling assays.

72 Assessing the Potential Endocrine Disruptors within the ToxCast Compound Library by Profiling Their Impacts on Activities of Multiple Nuclear Receptors

Alexander Medvedev[1], Matt Moeser[1], Liubov Medvedeva[1], Elena Martsen[1], Alexander Granick[1], Lydia Raines[1], Ming Zeng[1], Keith A. Houck[2], Sergei S. Makarov[1]
[1]ATTAGENE, PO Box 12054, Research Triangle Park, NC 27709
[2]US Environmental Protection Agency, 109 TW Alexander Dr., D343-03, Research Triangle Park, NC 27711

Rationale. Nuclear receptors (NRs) are ligand-regulated transcription factors that modulate gene expression

by hormonal or metabolic stimuli. The human NRs share significant homology yet regulate distinct sets of genes. The activity of many NRs can be readily manipulated by various exogenous ligands, making NRs of particular interest to drug developers and toxicologists. Also, many exogenous NR ligands are shown to bind and regulate more than a single NR. Thus, comprehensive evaluation of environmental compounds as potential NR ligands is required for assessing their bioactivity.

Experimental design. To assess the effects of environmental compounds on multiple NRs we developed a multiplexed reporter assay (trans-FACTORIAL™) that enables a quantitative assessment of a compound's impact on multiple one-hybrid GAL4-NR reporter constructs in a single well of test cells. Here we used the trans-FACTORIAL assay to profile the activities of over 1,800 of Phase 1 and II environmental toxicants and known endocrine disrupting compounds against 24 human nuclear receptors.

Results. We found that the profiling in trans-FACTORIAL assay produced highly reproducible NR response patterns ("NR signatures") that quantitatively characterized the compounds' impact on NR activities. The most frequently affected NRs were as follows: PXR > ER-alpha > PPAR-gamma > ROR-beta > RXR-beta > RAR-alpha > PPAR-alpha > HNF4. Importantly, the NR signatures of ToxCast compounds formed distinct clusters, thus suggesting predictive capabilities of the NR signatures. For example, the NR profiling of eight structurally-related organotins produced very similar NR signatures with the similarity scores above 0.85. Unsupervised hierarchical clustering analysis revealed that the tight cluster of organotins had an adjacent tight cluster of four compounds including 2,4,6-Tris(tert-butyl)phenol and 2,6-Di-tert-butyl-4-ethylphenol that exhibited organotin-like NR signatures, suggesting that bioactivities of these compounds may be similar to those of the rexinoid-like group of endocrine disruptors.

Conclusion. This study demonstrated that the trans-FACTORIAL assay is uniquely suited for high-content assessment of putative environmentally relevant NR ligands and indicated the utility of the multi-endpoint NR signatures for the classification of the compounds according to their impact on the NR superfamily.

The views expressed in this abstract are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.

73 Chemical Security Analysis Center: Assessing the Risk and Hazard of Chemical Terrorism; Predictive Toxicology Program

Janet Moser[1,2]; Ray Jablonski[1]; William Ashman[1,2]; Paul Grasso[1]
[1]Chemical Security Analysis Center, Department of Homeland Security, Aberdeen Proving Ground, MD
[2]Battelle Memorial Institute, Columbus, OH

The U.S. Department of Homeland Security (DHS) has established the Chemical Security Analysis Center (CSAC) to identify and assess vulnerabilities and response to potential chemical threats to the homeland. The CSAC primary goals are to provide a robust knowledge repository of chemical threat information, a science based threat and risk analysis, and a technical based reachback capability. To achieve these goals, CSAC must consider thousands of chemicals when devoting limited resources to assess hazard and risk to human populations. Accurate human toxicity estimates are critical; inaccurate toxicity estimates lead to unreliable outcomes of predictive chemical terrorism assessments of risk, hazard, and countermeasure efficacy and safety. In contrast with the efforts of existing Tox21 programs, DHS is concerned with: 1) an intentional release/attack and a one-time chemical exposure; 2) different chemicals of concern and chemical prioritization (e.g., chemicals designed specifically to have detrimental effects on humans – emerging and novel threats); 3) single chemicals and chemical combinations; 4) primarily acute toxicity effects, both lethal and non-lethal; 5) different exposure scenarios (e.g., novel chemical dissemination methods, larger exposure amounts); and 6) rapid chemical identification and response to attacks. In addition, DHS has high interest in applications to countermeasure development. CSAC recently instigated a predictive toxicology program that leverages existing Tox21 programs to develop acute toxicity prediction models that combine *in vitro* toxicity testing with computational toxicology, applied specifically to DHS needs. Initial efforts are focused on investigating the objectives and current status of existing programs, forming collaborative relationships with participating organizations, and identifying shortfalls in existing programs in light of DHS needs. To this end, CSAC has conducted a market survey, collecting over 3,000 references and setting up a comprehensive structural framework for a searchable computational toxicology

library in which references are placed. *In vivo* and *in vitro* databases containing information on DHS chemicals of interest have been identified. Once gaps in existing programs are characterized, CSAC will develop/revise tools to address DHS-specific needs, test and validate developed models with representative chemicals from one selected toxidrome (acetyl cholinesterase inhibitors), and potentially develop models to address other toxidromes of interest.

74 Prediction of Binding Affinity and Efficacy of Thyroid Hormone Receptor Ligands Using QSAR and Structure Based Modeling Methods.

Regina Politi (1), Ivan Rusyn (2), and Alexander Tropsha (1)
(1) Laboratory for Molecular Modeling, Division of Chemical Biology and Medicinal Chemistry, University of North Carolina, Chapel Hill, NC, 27599, United States
(2) Department of Environmental Sciences and Engineering, University of North Carolina, Chapel Hill, NC, 27599, United States

The thyroid hormone receptor (THR) is an important member of the nuclear receptor family. Activation of THR by environmental chemicals may cause the disruption of the endocrine system leading to adverse effects. We have developed Quantitative Structure-Activity Relationship (QSAR) models to facilitate the prioritization of putative endocrine disrupting chemicals (EDC) acting via THR for the experimental validation. We have assembled the largest database of binding affinities available at the time of this study for ligand binding domain (LBD) of THRβ. Both continuous and classification QSAR models were developed with an external accuracy of R2=0.55 and CCR=0.76, respectively. In addition, for the first time a QSAR model was developed to predict binding affinity of antagonists inhibiting the interaction of coactivator with AF-2 domain of THRβ (R2=0.70). Furthermore, molecular docking studies were performed for THRβ ligands (57 agonists and 15 antagonists of LBD, 210 antagonists of AF-2 domain, supplemented by putative decoys/non-binders) using several THRβ structures retrieved from the Protein Data Bank. We found that two agonist-bound THRβ conformations could effectively discriminate their corresponding ligands from presumed non-binders. Moreover, one of the agonist conformations could discriminate agonists from antagonists. Finally, we conducted virtual screening of a chemical library screened by EPA as part of the Tox21 program to identify potential THRβ-mediated EDCs using both QSAR models and docking. We concluded that the library is unlikely to have any EDC that would bind to THRβ. Models developed in this study can be employed to either identify environmental chemicals interacting with THR or, conversely, eliminate the THR-mediated mechanism of action for chemicals of concern.

75 Potential Applications of ToxCast Data in Safety Evaluation of FDA-Regulated Products: Comparative Analyses between ToxCast Data and FDA Databases

Huixiao Hong, Hui Wen Ng, Mikyung Lee, Minjun Chen, Zhichao Liu, Zhenqiang Su, Roger Perkins, Weida Tong*
Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, FDA
**Correspondence: Huixiao.Hong@fda.hhs.gov; 870-543-7296*

The Food and Drug Administration (FDA) analyzes the benefits and risks of FDA-regulated products when reviewing applications and also continues safety monitoring of products once they reach the market. FDA has long constructed and used databases as an integral part of safety evaluation and monitoring. The U.S. Environmental Protection Agency's (EPA) ToxCast project evaluated a diverse set of chemicals, including both environmental chemicals and drugs, using a broad panel of high-throughput *in vitro* assays. ToxCast data have been demonstrated to be efficient to characterize the toxicological profiles of environmental chemicals. However, applicability of ToxCast data in safety evaluation of the FDA-regulated products such as drugs has not been thoroughly assessed. We explored the potential of utilizing ToxCast data in safety evaluation of the FDA-regulated products through comparative analyses between ToxCast data and some frequently used FDA's databases. Specifically, data in FDA's estrogenic activity database (EADB) of the endocrine disruptor knowledgebase (EDKB), liver toxicity knowledgebase (LTKB) and substance registration system (SRS) were compared with ToxCast data. Three types of comparative analyses were conducted: coverage of applicability domains, concordance of qualitatively categorizing chemicals and consistency of

quantitative measurements. The chemical space of ToxCast is similar to those of FDA's databases, showing the potential of applying ToxCast data in safety evaluation of the FDA-regulated products. The data from the 18 estrogen receptor (ER) pathway-related ToxCast assays had discrepancies between the assays, indicating that both false positives and false negatives exist in all of the 18 assays, thus warranting further investigations on effective ways of utilizing ToxCast data. We found ToxCast data were generally concordant in assignment of actives and non-actives (with ligand binding data especially concordant) with the FDA's EADB data that were curated from literature. The estrogenic activity measures of ToxCast were moderately well correlated with the data in EADB. Our comparative analysis demonstrates the potential of utilizing ToxCast data in FDA regulatory sciences.

76 Assessing Steroidogenesis Disruption by Mixtures of Chemicals on the Basis of ExpoCast and ToxCast

Frederic Y. Bois, E. Lemazurier, INERIS, DRC/VIVA, France

We applied a recently developed mammalian steroidogenesis model (Quignot & Bois, 2013) to the assessment of endocrine disruption (ED) effects by mixtures of chemicals. All chemicals for which we had ToxCast aromatase or CYP19A1 significant dose-responses and ExpoCast exposure data (n = 325) were used. Those chemicals were partitioned into direct acting ED (acting at the enzymatic catalysis level) and indirect acting ED (acting at the DNA transcription level). Using Monte Carlo simulations, we generated 10 millions random mixtures of those chemicals, at random steady-state exposure levels (using the confidence limits of the exposure estimates). The relative perturbation of ovarian estradiol concentration was predicted for all mixtures following exposure, using a combined minimal pharmacokinetic model and the SBML coded steroidogenesis model. Mixtures of chemicals at average exposure levels had only minimal impact on estradiol predictions. However, taking the variability and uncertainty of exposures into account indicated that about 70% of the mixtures led to less than 5% decrease in estradiol levels, the remaining 30% leading to effects up to nearly 100% inhibition of estradiol production. We can also estimate the impact of specific chemicals on those results.

We discuss several current limitation of the approach and needs to improve it: The relatively limited exposure database we used; The mix of uncertainty and variability in the exposure estimates; The artificial generation of exposure mixtures; The lack of *in vitro* and *in vivo* pharmacokinetic data; The limited extrapolation capabilities of the steroidogenesis model, due to a lack of central nervous system feedback. We are working on improving our approach in all those directions. Still, we are able to make predictions of the joint effect of mixtures of several hundred of chemicals at the organ level, a step closer to risk assessment for ED.

77 Endocrine Bioactivity and Exposure-Based Prioritization and Screening*

Patience Browne[1], Richard Judson[2], Alan Dixon[1], Nicole Kleinstreuer[4], Warren Casey[3], Kevin Crofton[2], Scott Lynn[1], John Wambaugh[2], Russell Thomas[2]

[1]US EPA, Office of Chemical Safety and Pollution Prevention, Washington DC

[2]US EPA, Office of Research and Development, Research Triangle Park, NC

[3]National Toxicology Program, Interagency Center for the Evaluation of Alternative Toxicological Methods, Research Triangle Park, NC

[4]ILS/NICEATM, Research Triangle Park, NC

The U.S. EPA is using computational toxicology and exposure tools (i.e., high throughput screening (HTS) assays, and computational models), along with other data streams to develop an integrated bioactivity and exposure-based approach for prioritization and screening the universe of EDSP chemicals. Endocrine bioactivity of chemicals will be quantified using HTS *in vitro* estrogen, androgen, and thyroid receptor assays, along with other potential molecular targets, and compared with Tier 1 screening and other *in vivo* assay results obtained from published and publically available peer-reviewed studies. The ability of the agency to screen the thousands of environmental chemicals in the EDSP universe is substantially limited by the capacity of the current Tier 1 screening methods. Rapid screening of 1000s of chemicals using EPA's ToxCast high throughput assays allows the agency to focus attention on chemicals with the greatest potential endocrine bioactivity. In addition, high throughput models of human and ecological exposure (e.g. ExpoCast) allows the agency to identify chemicals with the greatest potential for exposures that could result in bioactive concentrations. This pivotal shift in the EDSP framework allows focus to be placed on chemicals that pose the greatest likelihood of exposures leading to endocrine bioactivity in humans or wildlife. EPA's proposed integrated bioactivity exposure based approach for ranking chemicals in the EDSP universe for further screening and testing will be the subject of an independent external scientific peer review December 2-5 of 2014 (http://www.epa.gov/scipoly/sap/meetings/2014/).

*The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the US Environmental Protection Agency.