Food for Thought ... on Mapping the Human Toxome

Thomas Hartung [1] *and Mary McBride* [2]

[1]CAAT, Johns Hopkins University, Bloomberg School of Public Health, Baltimore, MD, USA, and CAAT-Europe, University of Konstanz, Germany; [2]Agilent Technologies, Government Relations, Life Sciences and Chemical Analysis, Washington, DC, USA

Summary

The report by the National Research Council of the US National Academy of Sciences, Toxicity Testing in the 21st Century: A Vision and a Strategy, has prompted a discussion about renewing regulatory toxicology – especially for chemicals – by harnessing in vitro tests, in silico approaches, and testing in lower organisms. The key change is basing the assessments on mechanisms and toxicant modes of action. Identifying "pathways of toxicity" (PoT), especially on a larger scale, evidently requires omics technologies. When the PoT is known, a test battery allowing higher throughput than the current approach can be constructed. Here, we propose an extension of this concept to mapping the entirety of PoT in humans: the human toxome. Mapping the human toxome will allow us, for the first time, to conclusively identify substances as nontoxic or to identify nontoxic concentrations of substances (i.e., concentrations at which no relevant PoT are triggered). The concept is explained, and opportunities and obstacles are discussed, aiming to promote an initiative which will form the core of a Human Toxicology Project to implement Toxicology for the 21st Century.

Keywords: human toxome, tox-21c, omics, pathway of toxicity, Human Toxicology Project

Introduction

The Toxicology in the 21st Century (Tox-21c) movement, initiated by the 2007 NRC report (NRC 2007; Krewski et al. 2010), has stirred the toxicological community (Hartung and Leist, 2008; Hartung 2008a, 2009a, 2011). Within three years the discussion has moved on from whether or not to change to how and when to do so – from ongoing programs by US federal agencies (Judson et al., 2010b; Knudsen et al., 2011) and the redefinition of the EPA toxicity testing paradigm (Firestone et al., 2010) to the call for a Human Toxicology Project (Seidle and Stephens, 2009; http://htpconsortium.wordpress.com/). This coincides with political requirements to reassess the safety of tens of thousands of existing chemicals (Hartung, 2010e) and the possibly enormous testing needs resulting from these if traditional tests are broadly applied (Rovida and Hartung, 2009; Hartung and Rovida, 2009a,b). Similar pressures have arisen in Europe from cosmetic legislation (Hartung, 2008c) and the recently revised laboratory animal welfare legislation (Hartung, 2010c), as well as worldwide from new testing for nanomaterials (Hartung,

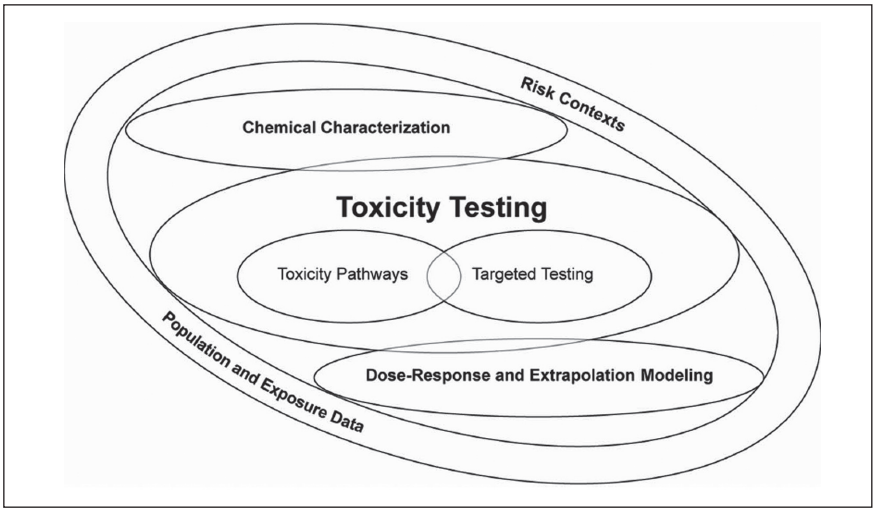

Fig. 1: Toxicity Testing in the 21st Century

Summarizing sketch of the vision set out by the NRC panel (2007) (Fig. 3 from Krewski et al., 2010, reproduced with permission).

The committee's vision for toxicity testing is a process that includes chemical characterization, toxicity testing, and dose response and extrapolation modeling. At each step, population-based and human exposure data are considered, as is the question of what data are needed for decision making.

2010d). The key proposal of Tox-21c is simple: we have to base regulatory toxicology (for environmental chemicals, because this was the mandate of the National Academy of Sciences panel) on mechanisms and modes of action. The term "pathway of toxicity" (PoT) was coined in the NRC report to describe this concept. More recently, the OECD has used the term "adverse outcome pathway" in the context of its QSAR Toolbox and ecotoxicology (Ankley et al., 2010). This is in line with the science of toxicology moving toward a mechanistic understanding. As a logical consequence, we propose the compilation of a comprehensive list of all PoT– that is, the human toxome. This goal is based on the assumption that the number of PoT is finite. We will explain this approach in detail, identify challenges, and lay out steps that are necessary to create a list of all PoT. Mapping the human toxome represents a possible cornerstone of Tox-21c, other components including chemical characterization, targeted testing, *in vitro* to *in vivo* extrapolation, and risk context considerations (Figure 1 represents a summary diagram of the Tox-21c report). Notably, neither Tox-21c, nor the Tox-21c alliance of agencies (Collins et al., 2008) have suggested mapping the entire human toxome yet.

Consideration 1: A mapped human toxome as the basis for a new testing approach allowing identification of non-toxicity

When testing a substance with any system, the negative results pose the principal problem. Absence of evidence is not evidence of absence, i.e., choosing a different test system or a different dosing scheme might still reveal a toxic property. This is the case for animal testing just as it is for alternative approaches. A negative animal test means nothing: a different species, or some other experimental variation, could still yield a positive result. Animals might have a defense mechanism not present in humans or in sensitive human populations, like newborns, who for example lack a functional blood-brain barrier for chemicals. Conventionally, however, we assume that with some additional measures (high dose, species selection, more than one species, structural alerts, etc.) we test enough potential PoT when we use whole organisms. When using less complex systems, such as cell assays, the question is when to stop testing. Do we need one, three, ten, or fifty assays to be certain enough

Tab. 1: Some estimates of the prevalence of toxic effects of substances in animal tests and humans

Health effect	Prevalence in animal tests	Estimated prevalence for humans
Cancer	50-60% 54%, 35% in both rat and mouse (ToxRefDB)	5-20% (various expert estimates) (about 100 listed by IARC, 300 by NTP)
Mutagenicity	29% (marketed drugs)	Unknown
Reproductive toxicity	64% (theoretical from prevalence and species difference) 61% (new chemicals EU plus EPA HPV list) 87% any pathology (ToxRefDB, i.e. mainly pesticides) 19% in two species (ToxRefDB)	2-3% (expert estimate) (about 270 listed in California proposition 65)
Acute toxicity	13% (new chemicals EU have LD_{50} <2g/kg) 36% (chemicals in RTECS have LD_{50} <100mg/kg)	Unknown 20-40% of candidate drugs any toxicity (43% with correlate in rats)
Chronic toxicity	88% any pathology (ToxRefDB) 53% <300mg/kg (EU new chemicals)	Unknown
Skin sensitization	35-40%	Unknown (about 7,000 chemicals that sensitize humans identified)
Skin corrosion	3%	Unknown, likely similar
Skin irritation	7%	40-56% of classified substances positive in human patch tests
Eye irritation	21% (EU new chemicals)	Unknown

Sources for above estimates: Ames and Swirsky Gold, 2000; Basketter et al., 2004; Bremer et al., 2007; Bulgheroni et al., 2009; Hartung, 2009a, 2010b; Hoffmann et al., 2005; Jírová et al., 2010; Knudsen et al., 2009; Kola and Landis, 2004; Martin et al., 2009a,b; Olson et al., 2000; Scott et al., 2010; Snyder and Green, 2001.

that the chemical is safe? A definitive answer could be given if we had a list of all human PoT and a respective test battery reflecting these. Then we could, for the first time, be confident that a substance does not trigger any relevant PoT. Similarly, we could establish concentrations of substances (*in vitro* no-effect levels – IVNOEL) at which no PoT is triggered. Notably, the triggering of a PoT does not necessarily indicate harm, but a *potential* for harm. It is a tremendous risk to continue to automatically consider each and every change in a biological system to be a hazard when we now move to larger test batteries (which is inevitable for Tox-21c). We will have to learn which combinations of PoT create a harmful effect and whether there are also "pathways of defense" (PoD) that must be taken into consideration. With an increasing knowledge of PoT and PoD we can refine the alerts that indicate the requirement for further testing. Identifying nontoxic substances, however, is the real challenge in toxicology. Paracelsus was right about the dose making the poison, which is also the basis of IVNOEL, but he was wrong about every substance being a poison. The majority of chemicals are nontoxic. Even in overly sensitive animal tests, the majority of substances have no effects (Tab. 1).

As preliminary as Table 1 is, it shows several things. First, a large proportion of substances is not toxic in animals, with the notable exceptions of cancer, reproductive, and, in part, chronic toxicity studies. For the former two expert estimates suggest a much lower prevalence in humans. For chronic toxicity, cross-species concordance was shown to be only 68% (Martin et al., 2009b), i.e. interspecies differences are similar to those found in cancer and reproductive toxicity studies. However, these are numbers were all obtained with high-dose treatments. We have discussed the limitations of these tests elsewhere (Hartung, 2008b; Hartung and Daston, 2009). The point here is that even with tests designed to be over-predictive (few false negatives by accepting false positives), a large number of substances do not show a given hazard. Where data are available, it appears that in humans the proportion of toxic substances is far lower, which is only to be expected considering the precautionary approach of testing high doses, multiple endpoints, and multiple species. Thus, under normal use scenarios an even larger proportion of substances are nontoxic in humans. Furthermore, selection of substances for consumer products will favor the nontoxic substances. Last, for the toxic substances that make it into application, the general goal is to identify doses/concentrations that are nontoxic.

Consideration 2: How many PoT are there and is the number finite?

Mel Anderson, one of the proponents of Tox-21c, often answers this question with "132," adding, after a pause, "As a toxicologist I am used to working with false accuracy." At this moment any number is pure speculation; however, as the number of cellular targets and metabolic pathways is finite, the number of PoT should be, too. Evolution cannot have left too many Achilles heels given the number of chemicals surrounding us and the astonishingly large number of healthy years

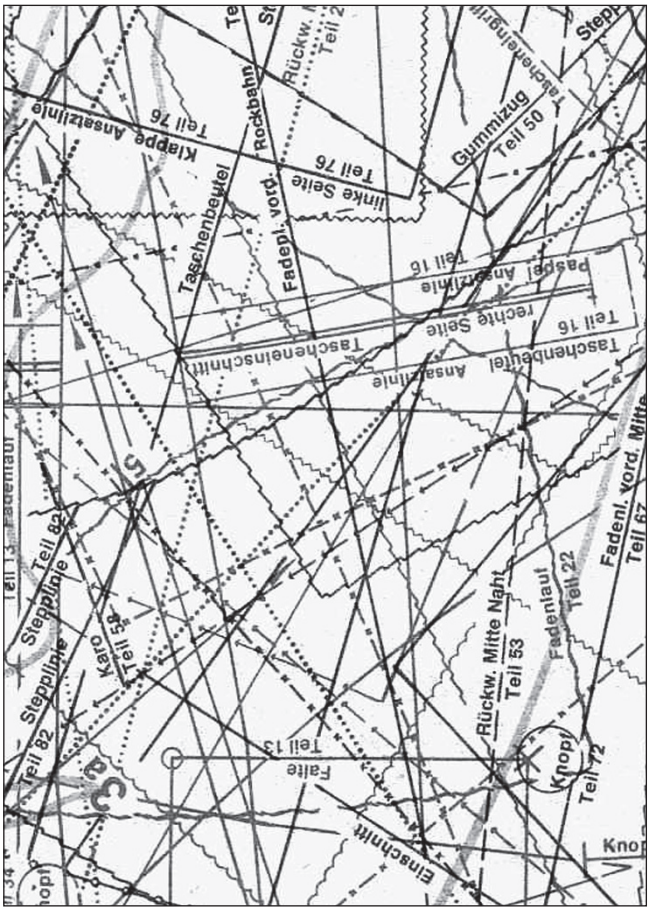

Fig. 2: Illustration of complex pathways, here a sewing pattern (Flux USA, 2009)

we enjoy on average. How many PoT there are depends very much on the definition of PoT – what is a PoT on its own, what are variants, what are groups, etc.? Most likely we still focus too much on linear pathways. As we increasingly learn that processes in living organisms are networked, so we will likely learn that PoT are mostly perturbations of the network, not a one way chain of events. Figure 2 might serve as an illustration, though it represents not a biological system but a sewing pattern, since nobody knows yet how to represent such systems properly. The well-known book *An Introduction to Systems Biology – Design Principles of Biological Circuits* by Uri Alon (Alon, 2007) is an excellent illustration of the complexity of biological pathways. There is no reason to assume that many PoT are simple. We may have to define PoT as critical constellations in the network brought about by the chemical effector. The enormous redundancy and buffering possibilities in an organism is demonstrated by the astonishing number of viable knockout mice that have only subtle phenotypes though lacking an entire gene.

Most importantly, toxicology is not alone in aiming to identify pathways – all the life sciences are on the same quest under the label of systems biology. It is the logical next step after the introduction of the high-content technologies (mainly omics) to bring order into the observed changes by defining the underly-

ing pathways. We will see whether such definitions of pathways across sectors will work or whether we need specific additions, e.g., the addition of reference compounds and where they interfere with the pathways for PoT.

As a working hypothesis, it is fair to assume that at least the number of important PoT is limited. But what is "important" in this context? We might define important PoT as pathways that are involved in the action of known toxicants in relevant test systems. This definition is open to additions from future data, but follows the primary goal of regulatory testing that "something like this must never happen again."

Consideration 3: The lack of a PoT concept

At this moment even the proponents of Tox-21c have no clear idea of what a PoT is, how to annotate it, validate it, or translate it into testing (Hartung, 2009b). Most think of cellular events, but is carbon monoxide poisoning a cellular event? Most think of a chain of reactions starting with a molecular target, but does this describe narcosis and excess lethality by volatile organic compounds? Is there a pathway of corrosion? And where we have pathways, are the key points the involved proteins, the metabolites or the induced genes? The NRC report defined toxicity pathways as biologic pathways that, when sufficiently perturbed, can lead to adverse health outcomes. But does destruction of functional integrity, e.g., of cell membranes by reactive substances, fall under this definition?

Bumgartner and Yeung (2009) stated "Presently, the words 'pathway' and 'network' are used almost interchangeably. However, in a given use, the constructs these words represent can be vastly different (e.g., literature relationships, physical interactions, or coupled chemical reactions)." They suggest the following terminology and definitions:

- *Molecular or biochemical pathway:* A set of coupled chemical reactions or signaling events. Nodes are molecules (often substrates) and edges represent chemical reactions. We also include conformational changes as the result in downstream signaling via other chemical reactions in this definition.
- *Physical interaction network:* A graphical representation of molecular binding interactions such as a protein-protein interaction network. Nodes are molecules; edges represent physical interactions between molecules.
- *Correlation or co-expression network:* A graphical representation that averages over-observed expression data. Nodes are molecules (typically mRNAs); edges represent correlations between expression levels of connected nodes.
- *Bayesian expression network (Bayes nets):* A directed, graphical representation of the probabilities of one observation given another. In our use, nodes represent mRNA molecules; edges represent the probability of a particular expression value, given the expression values of the parent nodes.
- *Knowledge-based network:* A graphical representation of relationships between genes or molecules as inferred from external knowledge. An example would be a literature-based network in which the nodes represent the presence of co-citation in a Pubmed abstract.

It seems that these distinctions are helpful and at the same time illustrate the layers of complexity of interconnections to be considered when trying to understand the human toxome.

The increasing identification of signatures of toxicity in omics approaches is somewhat suggestive of the existence of distinct PoT. Signatures of toxicity are changes – in either genes, proteins, metabolites or whatever the omics technology measures – which are associated with a certain toxic effect. The hope is that these measured changes can be associated with PoT and thus might be understood, and real contributors to the signature might be distinguished from noise and unrelated epiphenomena. However, a lot of the signature might actually be provided by the stress response and by PoD. Only if we identify the PoT can we use them to explain, by their presence or absence, the different reactions of different cells or organisms.

The definition of PoT will directly correspond with the annotation of PoT. There is some similarity with the effort of mapping the human genome (HUGO), but here annotation was easy, i.e., a sequence of four bases. CAAT (Daneshian et al., 2010) will host a series of workshops starting later this year to develop consensus on PoT identification, definition, validation, annotation, and sharing. Uri Alon (Alon, 2007, fig. 5.5 on page 82) shows 199 different types of pathway interactions between only four nodes of a network. It is frightening to imagine the connections between thousands of metabolites, genes and proteins… However, we will not have to start from scratch, as pathway mapping and visualization tools are increasingly being optimized in other areas of the life sciences. A very special challenge will be that we not only have to represent the PoT or their network, but also the kinetics and locations of these events, as a PoT represents a spatio-temporal event.

It is important to realize that PoT mapping is not a fancy new name for alternative methods. *In vitro* tests have limitations, just as do animal models (Hartung, 2007). Each in vitro test in use today is, like each animal test, a black box of many unknown PoT, some relevant to humans while others are not. Each cellular test is more or less complete in reflecting the PoT of relevance. The concept of PoT promises to annotate them to cellular tests or design even PoT-specific tests (such as reporter gene assays, biomarkers, etc.). This might be expanded to lower organisms as well as subcellular systems. This makes use of a given species or a given cell independent of the overall reaction. We can then reduce the information to the perturbation of a relevant PoT, which, in certain settings, is linked to hazard.

Consideration 4: How to identify PoT?

At this time, the technologies that most lend themselves to PoT identification (van Vliet, 2011) are mass spectrum-based metabolomics and transcriptomics. Transcriptomics is arguably the most developed omics technology; prices for gene chips have come down considerably, the chips are highly standard-

ized – annotated to the human and other genomes and even across technological platforms. Therefore, some of the former concerns regarding validation of such technologies (Corvi et al., 2006) are vanishing. As validation means assuring the reliability of the technology, difficulties in this regard would question the usefulness of this technology as a basis for PoT identification. Next-generation sequencing is enabling improvements in novel transcript discovery. However, changes in mRNA expression only indicate possible phenotypical and metabolic changes. We do not know from expression alone whether observed changes are translated into protein structures/functions. What is important is that, independent of the technology, gene expression is increasingly being linked to gene function and to the interactions of different genes (i.e., we are making sense of gene expression changes). Here, proteomics are obviously already a step further, but functionality and impact – even of expressed proteins – is complicated. Although there has been dramatic progress within the field of proteomics, standardization and running costs still lag behind those of other omics. Our expectations are especially high for MS-based metabolomics: both the levels of standardization and the running costs are good, we are measuring actual metabolic changes, and the restricted number of metabolites and known biochemical pathways aid interpretation. Metabolic phenotyping has been successfully used *in vivo* (van Ravenzwaay et al., 2007, 2010), especially for cancer signatures, and *in vitro* (Cuperlovic-Culf et al., 2010), including our own work in (developmental) neurotoxicity within toxicology (van Vliet, 2008). The first attempts to move from signatures to PoT identification using human embryonic stem cells are in progress (Cezar et al., 2007; West et al., 2010). NMR-based metabolomics has many valuable applications in toxicology, as demonstrated by Jeremy Nicholson and his group in London (Nicholson et al., 2002; Coen et al., 2003, 2004), but is less suited for PoT identification as metabolite identification is more difficult than with MS.

Certainly, other existing and emerging technologies will feed into PoT identification. For example, more than 400 kinases represent key targets of drug development; excess pharmacology, i.e., overstimulation of the pharmacological target as a common mode of action, as well as non-specific effects on other kinases, will likely represent important PoT (which could require phosphoproteomics, etc.).

The main challenge lies in the bioinformatics of PoT identification, as *in vitro* assays, reference substances, and the measurement technologies are available. It appears that no single measurement technology is sufficient for pathway identification. The combination of data from different sources and their integration into one result represents the next generation of pathway identification tools, which is well on its way in the respective industries. Early attempts were made to identify plant metabolic pathways (Oksman-Caldentey and Saito, 2005).

Combining data from different platforms (primarily transcriptomics and metabolomics) and assays into a coherent approach that appropriately weighs and evaluates the different data sources will be a challenging task. An important part of this integration will be the development of visualization tools that display the combined data in an easily understood format. The recently announced collaboration between Agilent Technologies and Strand Life Sciences portends an awareness of the critical importance of developing such integrative approaches for handling the enormous amounts of disparate data that will be combined in the new toxicity testing approach. To do this, Agilent is already building enhancements, both to the commercial bioinformatics software (GeneSpring), and to the public-domain Cytoscape network analysis and visualization platform. The first tool scheduled to emerge from this partnership will be a version of GeneSpring designed to help users perform statistical analyses of and visualize data from genomics, metabolomics, and proteomics together for the first time using a familiar interface.

The enhancements need to be built on a novel, flexible architecture, engineered specifically to provide a broad foundation for joint analysis and visualization of orthogonal data. Several key processes critical to pathway-based orthogonal analysis, including shuttling of different kinds of data between software applications, facilitating new custom visualizations, enabling statistical analyses involving pathway databases, and providing workflow and help facilities in order to ensure that the software is accessible to users with different levels of experience, must be considered. As such, this provides an ideal environment in which to develop new software tools for any application relying on joint analysis in the context of pathways. The tools need to be developed to use transcriptomic and metabolomic data to aid data management of primary data, visualization, analysis, and annotation of pathways of toxicity closely tied to the needs of toxome mapping.

Another interesting option to identify PoT is finding inter-individual differences (see below) in reactions to toxicants. Where we can link these to differences in gene expression and genetic variability, we gather evidence for components of critical PoT. One example is the difference in response to radiation, which has been linked to cellular responses (Smirnov et al., 2009). Similarly, metabolic phenotyping was used to identify inter-individual differences linking diet and blood pressure (Holmes et al., 2008). Population variability studies in toxicology are rare; an example was given most recently by O'Shea et al. (2011), but the results described here have not yet been traced back to genetic differences and underlying PoT.

Consideration 5: Identification of pathways of interaction of substances with cells by genetic variation of cellular test systems

Our genetic make-up determines our reaction to substances, including, but not restricted to, chemicals. If we create a panel of similar cells, which differ in individual, groups of, or many genes, and carry out the same test on substances of interest, differences in reaction might be traced back to the genetic peculiarities. Endpoints to be assessed could simply be cytotoxicity tests or specific cell responses. By identifying abnormal cell

responses and tracing them back to the respective genetic makeup, pathways of interaction of the substance with the cell system can be identified or supposed pathways verified.

A panel of genetically different cells can be obtained among others by:

– the combination or comparison of cells from different donor humans or animals
– the combination or comparison of cells from donors with or without a certain disease
– the induction of mutations in cells from one or more donors
– the random or targeted insertion of genetic material and disruptors of genetic materials in the genome of cells from one or more donors
– the recombination of genetic material of different donors
– the construction of artificial cells

This panel of cells can be brought into contact with test substances and cellular responses can be assessed. Abnormal responses, such as increased or decreased responses compared to the majority of cells or historic controls, are used to identify those with a genetic makeup relevant for the identification of pathways of toxicity or defence. This includes the survival of an otherwise lethal concentration of the substance. In case of dividing cells, this might include favored growth in the presence of the substance. Methods allowing identification or isolation of those cells with a genetic makeup causing a different response to the test substance can include, but are not limited to, cell image analysis and cell sorting.

It is crucial to identify the genetic variation linked to the variation in the response. This can be done by sequencing single nucleotide polymorphisms (SNP), or otherwise obtained information on the genetic makeup of the respective cell. If the cells differ in multiple aspects of their genetic makeup, consensus patterns of various cell variants can be used.

This approach allows the identification of genes impacting on the response of cells to substances based on knowledge of pathways connecting these genes, their proteins, or their metabolites and binding partners. This is especially relevant for the identification of PoT or PoD and to manipulate or alter cell responses, such as drug pathways. The latter allows deducing ways to identify new substances by designing test systems representative of the pathway identified. The former allows the identification of PoT and thus the deduction of tests for them, as well as the verification of the presence of these critical pathways in a given test system.

Consideration 6: How to validate PoT?

First we should be clear that validating PoT is different from validating a test based on PoT. The critical point here is determining the scientific validity of the PoT, not ring trials demonstrating reproducibility, and this is not (yet) about the results obtained for test substances. Validating a PoT means

Tab. 2: Evidence required for validating a pathway of toxicity (PoT)

Mode of PoT validation	Value of evidence	Limitation of evidence	Overall value
Orthogonal technology identifies component of the same PoT	Shows that pathway is triggered	Does not show that it is a critical PoT	Pos: ++ Neg: -
Inhibition of PoT	Shows that PoT is essential for toxic effect (depending on specificity of intervention)	Negative findings do not exclude a role of PoT as alternative PoT might be involved	Pos: +++ Neg: -
Substances with similar mode of action or toxic effect trigger PoT	Supports that PoT is relevant for the hazard	Neither pos. nor neg. findings prove or exclude	Pos: + to +++ (depending on number of examples) Neg: -
(Similar) substances with no toxic effect do not trigger PoT	Supports the role of PoT for toxic effect	Pos. findings (triggering) might indicate that other essential PoT are not triggered	Pos: - Neg: + to +++ (depending on number of examples and structural similarity)
PoT is activated at concentrations which represent thresholds of toxic effects	Supportive, but activation at lower concentrations might indicate other PoT is necessary	No activation at concentrations at which toxic effects occur make involvement unlikely	Pos: + Neg: +++
Strength of PoT activation and toxic effect correlate	Supportive	The limiting factor might be another PoT	Pos: + Neg: -
PoT is activated before toxic effect	Supportive	Largely excludes a role of the PoT	Pos: + Neg: +++

that we show a PoT is, in fact, relevant to the toxic effect of a given known toxicant, in the sense of predicting a hazard to humans. Two basic approaches come to mind, i.e., orthogonal mapping technologies and inhibition strategies. The former would require that additional technologies (for example, proteomics for a PoT identified by the integration of metabolomics and genomic data) would identify changes in elements of the PoT. The latter requires inhibiting a postulated component of the PoT to block toxic action or the expression of the signature of toxicity. Inhibitors might include silencing RNA technologies, genetic knockout strategies, pharmacological inhibitors, etc. This will be case-dependent and will usually require expert knowledge. Linking of inter-individual differences in toxic vulnerability mapped on genetic differences, such as single nucleotide polymorphisms (SNP), is an interesting option. A third reassuring – though not proven – aspect would be the identification of the same PoT for compounds with a supposedly similar mode of action/toxic effect or the absence of identification of the given PoT for structurally similar substances that do not show the toxic effect. Due to redundancies and possible similar effects of different PoT, this approach does not provide ultimate proof. Supportive evidence can come also from the correlation of PoT activation and toxic effect with regard to dose, strength, and timing. The value of different approaches is summarized in Table 2. In the end, this is an application of the Koch-Dale and Bradford-Hill criteria for mediation of an effect. However, in this specific case, it is assumed from the beginning that parallel PoT can be at work.

Consideration 7: What would PoT-based testing look like?

The pharmaceutical industry is moving towards pathway-based drug discovery (Fishman and Porter, 2005). Pathway knowledge is typically converted, after target validation, in a first step to high-throughput assays. Automated and robotized testing allows the screening of thousands of substances. The US Tox-21c alliance between EPA, NIEHS-NTP, NHGRI-NCGC, and, most recently, FDA, does exactly this (notably, however, with off-the-shelf assays available as a result of pharmacological screening and not on the basis of PoT identification, validation, and test construction). Already, this delivers impressive results, with several hundred assays run per substance in full concentration response curves and replicates for less than $ 20,000, showing the potential of such data generation.

A recent impressive example was the evaluation of eight possible dispersants to be used in the Gulf oil spill disaster (Judson et al., 2010a). The Deepwater Horizon oil spill has led to the use of >4 million liters of oil spill dispersants (surfactants and solvents). In this emergency situation it was necessary to assess the potential toxicity of the dispersants. A series of *in vitro* high-throughput assays on eight commercial dispersants was carried out. This allowed a regulatory decision on which dispersant to use in less than four weeks at costs that represent a small fraction of what is typically required.

Once PoT have been identified, the construction of a test system is usually not that difficult. A study we initiated at ECVAM on developmental toxicology, in a collaboration with Michael Schwarz, Tuebingen, Germany (Uibel et al., 2010), might serve as a test case. Astonishingly, interference with only a small number of canonical pathways across species appears to be responsible for most developmental disturbances. An NAS report from 2000 (NRC, 2000) lists only five crucial PoT, i.e., the Wnt/-catenin, the TGF-, the Notch, the Hedgehog, and the receptor kinase/ras pathway. A reporter gene assay was developed for the Wnt/-catenin pathway using murine embryonic stem cells. Several known human teratogens could be detected in a concentration dependent manner, including retinoic acid, lithium, and, most intriguingly, the potency of different retinoic acid derivatives was correctly reflected. After adding hepatocytes as a metabolizing system, even cyclophosphamide, which requires metabolic activation, was picked up. The assay – termed ReProGlow – shows the potential of relatively simple reporter assays once a PoT is known.

Our vision is to produce a suite of *in vitro*, subcellular, and *in silico* tools which comprehensively represent the human toxome. The difference to the current high-throughput testing of the Tox-21c alliance is that this approach is not limited to existing assays that may reflect a variety of unknown PoT. On the contrary, assays would be chosen or constructed which reflect known PoT and allow for a clear query and responses that demonstrate whether these PoT are perturbed or triggered. "Perturbed" would be used to describe that the test substance interferes with a physiological pathway, while "triggered" would be used to describe that a pathway leading to damage is activated. It will probably be necessary to extend this concept to PoD. Most importantly, however, interpretation of results would not be correlative, such as in the most interesting ToxPi approach (Reif et al., 2010), but based on PoT annotation to certain hazards.

Consideration 8: (Pre-)validated *in vitro* systems as starting points for mapping PoT

Where should we start to map PoT? There is broad consensus on using known human toxicants to map relevant PoT. However, the choice of cell systems is most important. Experience from validation shows that many of these tests have reproducibility issues and only very few are really predictive of the hazards of a larger group of substances. Of the assays promising enough to enter formal validation, roughly one-third fail pre-validation, and another third fail final validation. It seems ill-advised to choose just any cell system to map pathways. In turn, it appears advisable to make use of those models that have withstood the (pre-)validation process. Tests which come to mind are:

- Human artificial skin models (skin irritation and corrosion, genotoxicity, phototoxicity, and skin penetration)
- Human blood monocytes (inflammation, skin sensitization)
- MCF-7 cells (endocrine disruption)
- HepaRG cells (liver toxicity)

– Human blood lymphocytes (genotoxicity)
– 3T3 mouse fibroblasts (acute toxicity, cancer, phototoxicity)

These have robust, standardized protocols and are of known reproducibility. We also have the respective laboratories that are proficient in performing the tests and the reference compounds for which correct predictions are made. Most importantly, these methods have associated prediction models, i.e., thresholds and algorithms for establishing whether an effect is indicative of hazard. This is very rare for test systems and represents an enormous advantage of the established alternative methods. About $ 300 million went into their development and validation. This represents a capital investment that should be utilized. Some of these tests also have, to some extent, achieved regulatory acceptance; thus it is only logical that underlying PoT identified in these models should be more acceptable for further regulatory uses.

In conclusion, by using well-known human toxicants and reliable (human) cell systems to start with, we have a high likelihood of identifying relevant PoT. This is especially important at a moment where the concepts for PoT are only emerging. The approach is, however, flexible, as it can be extended to more substances as well as to other predictive cell systems. It will be most promising to expand from generally known toxicants to substances that have shown toxicities in human clinical trials. The current inclusion of such drugs and clinical data in the Tox-21c alliance program is a good call and should also be adopted for new PoT identification.

Consideration 9: How to implement mapping of the human toxome?

If we consider, for the moment, this exercise to be technically feasible, the question is: who has the incentive to tackle this? Certainly regulators and regulated communities, first of all. We see some efforts to identify PoT, mainly in the US EPA (ToxCast), the NIEHS (within the National Toxicology Program), NCGC (the high-throughput testing program) and FDA (the Critical Path Initiative). The efforts of the US EPA, in particular, have been highlighted several times in this series of articles (Hartung, 2010d,e). Similarly, the FDA has most recently embraced this strategy (Hamburg, 2011):

"We must bring 21st century approaches to 21st century products and problems. Toxicology is a prime example. Most of the toxicology tools used for regulatory assessment rely on high-dose animal studies and default extrapolation procedures and have remained relatively unchanged for decades, despite the scientific revolutions of the past half-century. We need better predictive models to identify concerns earlier in the product development process to reduce time and costs. We also need to modernize the tools used to assess emerging concerns about potential risks from food and other product exposures. … With an advanced field of regulatory science, new tools, including functional genomics, proteomics, metabolomics, high-throughput screening, and systems biology, can replace current toxicology assays with tests that incorporate the mechanistic underpinnings of disease and of underlying toxic side effects. This should allow the development, validation, and qualification of preclinical and clinical models that accelerate the evaluation of toxicities during drug development. … Ultimately, investments in regulatory science can lead to a new era of progress and safety. Because such investments will promote not only public health but also the economy, job creation, and global economic competitiveness, they have major implications for the nation's future."

We fully agree.

No such governmental initiatives exist in Europe. Some research funding and national activities, such as the Dutch Toxicogenomics Centre, represent an exception, rather than the rule. None of the projects currently underway, however, have taken up the challenge to create a public database and start comprehensive, systematic mapping. In the current economy, a large-scale, international program similar to HUGO is unlikely to emerge in the near term. The regulated industries are under very different pressures. In Europe, cosmetic and chemical industries are under exceptional pressure to adopt novel approaches, but only very few of the global companies have expertise in promoting a paradigm shift towards risk assessment. Incentives are low and deadlines short (Rovida, 2010), so that little momentum was gained for a complete revision of toxicity testing methods. Agrochemical companies might be incentivized differently – novel legislation in Europe is eliminating many established substances from the market, creating the need for the development of substitutes. At the same time, pesticides receive the most intense toxicological assessment of all products. The new legislation is moving risk management to a hazard-driven regulation; this means that the presence of a hazard is sufficient to ban products independent of exposure considerations. Therefore, it might be more attractive for this industry to improve the predictivity of tests. Thus, a mechanism-based approach, replacing the largely precautionary approaches with respective over-labeling, should be appealing. The pharmaceutical industry is another industry that could benefit from adopting a new test paradigm, and, most importantly, the industry is used to pathway-based approaches. The contribution of toxicology to early identification and down-selection of drug discovery targets, as well as early identification (i.e., before pre-clinical trials) of drug candidates with low efficacy and/or high human toxicity, are incentives to promote change. At the same time, PoT-based approaches promise to deliver results faster and require less test material, both of which are crucial in the time-to-market-driven, costly development process of new drugs. Many drug companies would like to frontload toxicology. Here, PoT-based tests appear to be a perfect match, as they can likely be automated, offering far better throughput at early stages. Pathway knowledge is of general interest for drug discovery: not only is the excess stimulation of therapeutic targets a common PoT, but what is unwanted (PoT) in the healthy system can turn into a helpful intervention in the diseased system. In short, there is no clear distinction between PoT and drug modes of action. Elucidation of PoT might even result in novel drug targets. Possibilities to generate intellectual property rights (IPR) in

the process of mapping the human toxome should be explored as an incentive.

Another type of industry with obvious incentive is the technology provider industry. Multi-billion dollar companies serving the life science research market have emerged. A project with the visibility and lighthouse function of mapping the human toxome must appeal to those spearheading the development of the respective technologies. This holds true also, to some extent, for the contract research industry.

The critical step is coordinating a larger consortium and linking it to the development of the necessary concepts. Central steering needs to be established, incorporating the ideas of opinion leaders and needs of stakeholders, and especially regulators, who ultimately have to accept changing to the novel approaches developed. Regulators, therefore, need a seat at the table and need to be able to input into the processes. The governance of the consortium effort needs to be established, as does the quality assurance (validation), comparison to the current approaches, and possible transition. We have recently discussed the possible role of retrospective evaluations in the spirit of evidence-based medicine (Hartung, 2010a). These concepts were discussed at the conference "21st Century Validation for 21st Century Tools" in Baltimore in July, 2010 and prompted the creation of an evidence-based toxicology collaboration (EBTC) on March 10, 2011 as an official satellite to the US Society of Toxicology (SOT) meeting. This promises to generate a partnership between agency representatives, individuals from the corporate sector, and those promoting the paradigm shift in toxicology. The generous support of an anonymous private donor allows CAAT to run the secretariat for EBTC for the next five years and to become a crucial partner for Tox-21c and the PoT mapping project.

Conclusions

The identification and use of PoT can revolutionize toxicity testing. Although modern toxicology has identified many modes of action, these have remained largely as isolated mechanisms that cannot be broadly applied to sufficient numbers of toxicants to warrant the establishment of dedicated toxicity tests. Currently, toxicity testing typically involves studying adverse health outcomes in animals subjected to high doses of toxicants with subsequent extrapolation to expected human responses at lower doses. Currently, humans are potentially exposed to more than 80,000 chemicals for which no toxicity data exists. *This is unacceptable*. At the same time, products traded at $ 10 trillion per year are regulated based on animal tests (Bottini and Hartung, 2009, 2010) and business decisions are taken based on animal tests which are less than 60% predictive between different laboratory animal species, let alone between rodents and humans. The challenge we face as scientists is to turn around the testing paradigm of regulatory safety assessments from phenotypical tests to tests based on a mechanistic understanding identified on the basis of known human toxicants.

The vision for toxicity testing in the 21st century welcomes innovation arising from our rapidly evolving understanding of systems biology and a host of molecular, informational, and computational tools, that provide the potential to identify PoT (along with the respective *in vitro* bioassays or *in silico* modeling (Hartung and Hoffmann, 2009)) to evaluate the effects of tens of thousands of chemicals at concentrations relevant to human exposure levels. A comprehensive list of PoT, the mapped human toxome, can become a cornerstone of this new regulatory toxicology. The overall impact of this project will put into practice the vision described in the 2007 NRC report and transform the way in which toxicity testing and risk assessments are conducted. The project will advance regulatory toxicology by piloting a more efficient and relevant toxicological assessment by PoT knowledge sharing (through a public database) leading to PoT-based integrated testing strategies.

The proposed project represents a nucleus for the Human Toxicology Project. In contrast to the currently used phenomenological "black box" animal testing, pathways of toxicity (PoT) will be identified in human *in vitro* systems to provide more relevant, accurate, and mechanistic information for the assessment of human toxicological risk. The goal is to map the entirety of the human toxome. The concentration at which a substance triggers a PoT can then be extrapolated to a relevant human blood or tissue concentration and, finally, a corresponding dose by (retro-) PBPK (physiology-based pharmacokinetic) modeling, thereby informing human risk assessment. Perhaps more importantly, if a substance does not trigger any PoT, for the first time it may be possible to establish the lack of toxicity, i.e., safety, of a substance at a given concentration.

To conclude with Freeman Dyson (Princeton), and his 1995 book, *The Scientist as a Rebel: "The great advances in science usually result from new tools rather than from new doctrines."* The map of the human toxome promises to be such a new tool.

References

Alon, U. (2007). *An introduction to systems biology – design principles of biological circuits* (301). London: Chapman & Hall/CRC.

Ankley, G. T., Bennett, R. S., Erickson, J. et al. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem. 29*, 730-741.

Ames, B. N. and Swirsky Gold, L. (2000). Paracelsus to parascience: the environmental cancer distraction. *Mutat. Res. 447*, 3-13.

Basketter, D. A., York, M., McFadden, J. P. and Robinson, M. K. (2004). Determination of skin irritation potential in the human 4-h patch test. *Contact Dermat. 51*, 1-4.

Bottini, A. A. and Hartung, T. (2009). Food for thought … on economics of animal testing. *ALTEX 26*, 3-16.

Bottini, A. A. and Hartung, T. (2010). The economics of animal testing. *ALTEX, Spec. Issue 27*, 67-77.

Bremer, S., Pellizzer, C., Hoffmann, S. et al. (2007). The development of new concepts for assessing reproductive toxicity applicable to large scale toxicological programs. *Curr. Pharm. Des. 13*, 3047-3058.

Bulgheroni, A., Kinsner-Ovaskainen, A., Hoffmann, S. et al.

(2009). Estimation of acute oral toxicity using no adverse effect level (NOAEL) from the 28-day repeated dose toxicity studies in rats. *Reg. Tox. Pharmacol. 53*, 16-19.

Bumgartner, R. E. and Yeung, K. Y. (2009). Methods for the interference of biological pathways and networks. In J. McDermott, R. Samudrala, R. E. Bumgartner et al. (eds.), *Computational Systems Biology* (225-245). New York: Humana Press.

Cezar, G. G., Quam, J. A., Smith, A. M. et al. (2007). Identification of small molecules from human embryonic stem cells using metabolomics. *Stem Cells Develop. 16*, 869-882.

Coen, M., Lenz, E. M., Nicholson, J. F. et al. (2003). An integrated metabonomic investigation of acetaminophen toxicity in the mouse using NMR spectroscopy. *Chem. Res. Toxicol. 16*, 295-303.

Coen, M., Ruepp, S. U., Lindon, J. C. et al. (2004). Integrated application of transcriptomics and metabonomics yields new insight into the toxicity due to paracetamol in the mouse. *J. Pharmac. Biomed. Anal. 35*, 93-105.

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

Corvi, R., Ahr, H.-J., Albertini, S. et al. (2006). Validation of toxicogenomics-based test systems: ECVAM-ICCVAM/NICEATM considerations for regulatory use. *Environ. Health Persp. 114*, 420-429.

Cuperlovic-Culf, M., Barnett, D. A., Culf, A. S. et al. (2010). Cell culture metabolomics: applications and future directions. *Drug Discov. Today 15*, 610-621.

Daneshian, M., Leist, M. and Hartung, T. (2010). Center for alternatives to animal testing – Europe (CAAT-EU): a transatlantic bridge for the paradigm shift in toxicology. *ALTEX 27*, 63-69.

Firestone, M., Kavlock, R., Zenick, H. et al. (2010). The U.S. Environmental Protection Agency strategic plan for evaluating the toxicity of chemicals. *J. Toxicol. Environ. Health 13*, 139-162.

Fishman, M. C. and Porter, J. A. (2005). Pharmaceuticals: a new grammar for drug discovery. *Nature 437*, 491-493.

Flux USA (2009). Available at: http://fluxusa.blogspot.com/2009/01/schnittbogen.html

Hamburg, M. A. (2011). Advancing regulatory science. *Science 331*, 987.

Hartung, T. (2007). Food for thought … on cell culture. *ALTEX 24*, 143-147.

Hartung, T. (2008a). Towards a new toxicology – evolution or revolution? *ATLA 36*, 635-639.

Hartung, T. (2008b). Food for thought … on animal tests. *ALTEX 25*, 3-9.

Hartung, T. (2008c). Food for thought … on alternative methods for cosmetics safety testing. *ALTEX 25*, 147-162.

Hartung, T. and Leist, M. (2008). Food for thought … on the evolution of toxicology and phasing out of animal testing. *ALTEX 25*, 91-96.

Hartung, T. (2009a). Toxicology for the twenty-first century. *Nature 460*, 208-212.

Hartung, T. (2009b). A toxicology for the 21st century: Mapping the road ahead. *Toxicol. Sci. 109*, 18-23.

Hartung, T. and Daston, G. (2009). Are in vitro tests suitable for regulatory use? *Toxicol. Sci. 111*, 233-237.

Hartung, T. and Hoffmann, S. (2009). Food for thought on … in silico methods in toxicology. *ALTEX 26*, 155-166.

Hartung, T. and Rovida, C. (2009a). Chemical regulators have overreached. *Nature 460*, 1080-1081.

Hartung, T. and Rovida, C. (2009b). That which must not, cannot be... a reply to the EChA and EDF responses to the REACH analysis of animal use and costs. *ALTEX 26*, 307-311.

Hartung, T. (2010a). Evidence based-toxicology – the toolbox of validation for the 21st century? *ALTEX 27*, 241-251.

Hartung, T. (2010b). Lessons learned from alternative methods and their validation for a new toxicology in the 21st century. *J. Toxicol. Env. Health 13*, 277-290.

Hartung, T. (2010c). Comparative analysis of the revised Directive 2010/63/EU for the protection of laboratory animals with its predecessor 86/609/EEC – a t^4 report. *ALTEX 27*, 285-303.

Hartung, T. (2010d). Food for thought … on alternative methods for nanoparticle safety testing. *ALTEX 27*, 87-95.

Hartung, T. (2010e). Food for thought … on alternative methods for chemical safety testing. *ALTEX 27*, 3-14.

Hartung, T. (2011). From alternative methods to a new toxicology. *Eur. J. Pharmaceutics Biopharmaceutics 77*, 338-349.

Hoffmann, S., Cole, T. and Hartung, T. (2005). Skin irritation: prevalence, variability, and regulatory classification of existing in vivo data from industrial chemicals. *Regul. Toxicol. Pharmacol. 41*, 159-166.

Holmes, E., Loo, R. L., Stamler, J. et al. (2008). Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature 453*, 396-400.

Jírová, D., Basketter, D., Liebsch, M. et al. (2010). Comparison of human skin irritation patch test data with in vitro skin irritation assays and animal data. *Contact Dermatitis 62*, 109-116.

Judson, R. S., Martin, M. T., Reif, D. M. et al. (2010a). Analysis of eight oil spill dispersants using rapid, in vitro tests for endocrine and other biological activity. *Environ. Sci. Technol. 44*, 5979-5985.

Judson, R. S., Houck, K. A., Kavlock, R. J. et al. (2010b). In vitro screening of environmental chemicals for targeted testing prioritization: the ToxCast project. *Environ. Health Perspect. 118*, 485-492.

Knudsen, T. B., Martin, M. T., Kavlock, R. J. et al. (2009). Profiling the activity of environmental chemicals in prenatal developmental toxicity studies using the U.S. EPA's ToxRefDB. *Reprod. Toxicol. 28*, 209-219.

Knudsen, T. B., Houck, K. A., Sipes, N. S. et al. (2011). Activity profiles of 309 ToxCast chemicals evaluated across 292 biochemical targets. *Toxicol. 282*, 1-15.

Kola, I. and Landis, J. (2004). Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov. 3*, 711-715.

Krewski, D., Acosta, D. A., Andersen, M. et al. (2010). Toxicity testing in the 21st century: A vision and a strategy. *J. Toxicol. Environ. Health 13*, 51-138.

Martin, M. T., Mendez, E., Corum, D. G. et al. (2009a). Profiling the reproductive toxicity of chemicals from multigeneration studies in the toxicity reference database. *Toxicol. Sci. 110*, 181-190.

Martin, M. T., Judson, R. S., Reif, D. M. et al. (2009b). Profiling chemicals based on chronic toxicity results from the U.S. EPA ToxRef database. *Environ. Health Perspect. 117*, 392-399.

Nicholson, J. F., Connelly, J., Lindon, J. C. and Holmes, E. (2002). Metabonomics: a platform for studying drug toxicity and gene function. *Nature Rev. Drug Discov. 1*, 153-161.

NRC (2000). *Scientific frontiers in developmental toxicology and risk assessment* (327). New York: The National Academies Press.

NRC (2007). Committee on Toxicity Testing and Assessment of Environmental Agents, National Research Council: *Toxicity testing in the 21st century: A vision and a strategy* (pp. 196). New York: The National Academies Press.

Oksman-Caldentey, K-M. and Saito, K. (2005). Integrating genomics and metabolomics for engineering plant metabolic pathways. *Curr. Opin. Biotechnol. 16*, 174-179.

Olson, H., Betton, G., Robinson, D. et al. (2000). Concordance of the toxicity of pharmaceuticals in humans and animals. *Regulat. Toxicol. Pharmacol. 32*, 56-67.

O'Shea, S. H., Schwarz, J., Kosyk, O. et al. (2011). In vitro screening for population variability in chemical toxicity. *Toxicol. Sci. 119*, 398-407.

Reif, D. M., Martin, M. T., Tan, S. W. et al. (2010). Endocrine profiling and prioritization of environmental chemicals using ToxCast data. *Environ. Health Perspect. 118*, 1714-1720.

Rovida, C. and Hartung, T. (2009). Re-evaluation of animal numbers and costs for in vivo tests to accomplish REACH legislation requirements. *ALTEX 26*, 187-208.

Rovida, C. (2010). Food for thought … why no new in vitro tests will be done for REACH by registrants. *ALTEX 27*, 175-183.

Scott, L., Eskes, C., Hoffmann, S. et al. (2010). A proposed eye irritation testing strategy to reduce and replace in vivo studies using bottom-up and top-down approaches. *Toxicol. In Vitro 24,* 1-9.

Seidle, T. and Stephens, M. L. (2009). Bringing toxicology into the 21st century: a global call to action. *Toxicol. In Vitro 23*, 1576-1579.

Smirnov, D. A., Morley, M., Shin, E. et al. (2009). Genetic analysis of radiation-induced changes in human gene expression. *Nature 459*, 587-591.

Snyder, R. D. and Green, J. W. (2001). A review of the genotoxicity of marketed pharmaceuticals. *Mutat. Res. 488*, 151-169.

Uibel, F., Mühleisen, A., Köhle, C. et al. (2010). ReProGlo: A new stem cell-based reporter assay aimed to predict embryotoxic potential of drugs and chemicals. Reprod. *Toxicol. 30*, 103-112.

van Ravenzwaay, B., Cunha, G. C., Leibold, E. et al. (2007). The use of metabolomics for the discovery of new biomarkers of effect. *Toxicol. Lett. 172*, 21-28.

van Ravenzwaay, B., Cunha, G. C., Fabian, E. et al. (2010). The use of metabolomics in cancer research. In W. C. S. Cho (ed.), *An omics perspective on cancer research* (141-166). Berlin: Springer.

van Vliet, E., Morath, S., Linge, J. et al. (2008). A novel in vitro metabolomics approach for neurotoxicity testing, proof of principle for methyl mercury chloride and caffeine. *Neurotox. 29*, 1-12.

van Vliet, E. (2011). Current standing and future prospects for the technologies proposed to transform toxicity testing in the 21st century. *ALTEX 28*, 17-44.

West, P. R., Weir, A. M., Smith, A. M. et al. (2010). Predicting human developmental toxicity of pharmaceuticals using human embryonic stem cells and metabolomics. *Toxicol. Appl. Pharmacol. 247*, 18-27.

Acknowledgement

The most valuable comments of colleagues and friends, especially the collaborators of the Tox-21c movement, are gratefully appreciated.

Correspondence to

Prof. Thomas Hartung, MD, PhD
Johns Hopkins University
Bloomberg School of Public Health
Doerenkamp-Zbinden Chair for Evidence-based Toxicology
Center for Alternatives to Animal Testing (CAAT)
615 N. Wolfe St. W7035
Baltimore, MD, 21205, USA
e-mail: Thartung@jhsph.edu