BCPT Basic & Clinical Pharmacology & Toxicology

*Basic & Clinical Pharmacology & Toxicology*, 2014, **115**, 50–58 Doi: 10.1111/bcpt.12257

***MiniReview***

# Cancer Biology, Toxicology and Alternative Methods Development Go Hand-in-Hand

**Pekka Kohonen[1], Rebecca Ceder[1], Ines Smit[1], Vesa Hongisto[2], Glenn Myatt[3], Barry Hardy[3], Ola Spjuth[4,5] and Roland Grafström[1,2]**

[1]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden, [2]VTT Technical Research Centre of Finland, Turku, Finland, [3]Leadscope, Columbus, OH, USA, [4]Douglas connect, Zeiningen, Switzerland and [5]Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden

(Received 9 December 2013; Accepted 21 April 2014)

*Abstract:* Toxicological research faces the challenge of integrating knowledge from diverse fields and novel technological developments generally in the biological and medical sciences. We discuss herein the fact that the multiple facets of cancer research, including discovery related to mechanisms, treatment and diagnosis, overlap many up and coming interest areas in toxicology, including the need for improved methods and analysis tools. Common to both disciplines, *in vitro* and *in silico* methods serve as alternative investigation routes to animal studies. Knowledge on cancer development helps in understanding the relevance of chemical toxicity studies in cell models, and many bioinformatics-based cancer biomarker discovery tools are also applicable to computational toxicology. Robotics-aided, cell-based, high-throughput screening, microscale immunostaining techniques and gene expression profiling analyses are common tools in cancer research, and when sequentially combined, form a tiered approach to structured safety evaluation of thousands of environmental agents, novel chemicals or engineered nanomaterials. Comprehensive tumour data collections in databases have been translated into clinically useful data, and this concept serves as template for computer-driven evaluation of toxicity data into meaningful results. Future 'cancer research-inspired knowledge management' of toxicological data will aid the translation of basic discovery results and chemicals- and materials-testing data to information relevant to human health and environmental safety.

Assessing the intrinsic toxicological properties of environmental agents is central to defining hazard within the paradigm of human risk assessment used now for decades [1]. A constantly increasing number of chemicals and engineered nanomaterials (ENMs) emphasize the need of applying novel tools and screening technologies outside of the standard, both lengthy and expensive, rodent toxicological tests traditionally used in such work [1–13]. Especially considered for the future innovation of novel ENMs, safety evaluations should be integrated proactively and efficiently already in the material and product development phase [9,12]. Summarized under a concept termed 'Toxicity Testing in the 21st Century' or 'Tox21', biochemical and cell-based *in vitro* assays coupled with bioinformatics and modelling-driven *in silico* assays are now considered key to transforming toxicology from a previously animal-based testing practice into a computational science built on systems biology [2,4,6,9–11,13]. The resulting novel research field is variably termed 'systems toxicology', 'toxicogenomics' or 'computational toxicology' [3,4,8,14,15]. Such effort typically relies on informatics-driven and modelling-based analyses of results from several experimental systems and data-rich technologies for measuring pools of biological molecules, such as mRNAs, the overall aim being to interdependently analyse toxicity data and profiling results for understanding and predicting toxicity. Optimally, such efforts can define the 'bio-identity' or 'hazard identity' of chemicals or ENMs, serving then as an 'activity' counterpart to the 'physiochemical identity' deduced from structural and physiochemical descriptors [5,6,8,12]. Problematic to advancement of the field, the much needed fusion of high-throughput *in vitro* and *in silico* technologies in toxicology has advanced slowly due to limited comprehensive data collections and a lack of tools and solution-oriented approaches to assess existing data. We argue that the central aims of systems toxicology research can be promoted by making use of data analysis solutions and databases developed for data-rich disciplines such as cancer research, including both high-throughput screening (HTS) and high-content screening (HCS) techniques. We support this claim by outlining herein the paths leading from cancer biology via toxicology to alternative methods development. We address databases that span the cancer biology and toxicology fields and utilize a well-studied chemotherapeutic agent to exemplify the concept. Built primarily on developments in pre-clinical cancer drug discovery in our own laboratories, we also depict a three-step, systems toxicology and 'Tox21'-inspired approach which spans from the high-throughput toxicity assessment of many agents through to genomic profiling analyses of the selected few.

Author for correspondence: Roland Grafström, Institute of Environmental Medicine, Karolinska Institutet, Box 210, SE-171 77 Stockholm, Sweden (fax +468343849, e-mail roland.grafstrom@ki.se).

### Parallels between Cancer Biology, Toxicology and Alternative Methods Research

An obvious parallel is that cancer development *per se* and compound-driven toxicological effects both lead to the deregulation of many normal physiological functions [16]. Thus, the toxicity effects that arise following chemical or ENM exposures both cause and reflect a diseased state at subcellular, cellular and organ levels [2,4,12]. Such studies therefore serve to probe the multiple and complex means whereby detrimental effects and disease processes arise [7,9]. From technology viewpoint, studies of normal and transformed cell line models *in vitro* are central to both cancer biology and toxicology work, paving the way to increasing application of cell models as alternatives to animal-based experiments (fig. 1) [4,9,11,17,18]. Tumour cells are generally more easily grown *in vitro* than normal tissue cell lines, but both are central to the cancer biology and toxicology disciplines, as both require information of the normal state as reference to the transformed state [16–20]. Beneficial to both fields, standard immortalized and tumour-derived monolayer culture models replicate indefinitely, and therefore support a high throughput of data generation [18]. Such models inform on early steps of cell transformation, but they also provide reproducible models for toxicological investigation [16,18,19]. Relevant to cancer research, organotypic culture states commonly enable better modelling of the therapy response to tumour treatments [20]. Similarly, toxicological investigation in organotypic states benefits from the analysis of orientation and multicellular organization by histological and pathological methods, as well as permitting the study of transport of toxicants through complex, multilayered tissue [17]. The mechanisms that drive cancer development lead to loss of differentiation, defects in growth control and resistance to cell killing [16]. Clearly, such knowledge is also imperative for accurate understanding of the usefulness and limitations of commonly applied infinite cell line models to toxicity experiments [19]. Clinically oriented studies for understanding responsiveness of tumour cells to cell killing from exposure to cancer therapeutics provide useful information on the biological fates that cells also encounter from exposure to environmental agents [6,21,22]. Over-expression or knockout of oncogenes or tumour suppressor genes provides much needed information on driving forces underlying transformation, but useful to toxicology, also informs on genes critically regulating survival or killing mechanisms [23]. Considering databases, the large repositories of clinical cancer data, including genomics analyses of tumours and bioinformatics tools, provide templates for what is also increasingly explored in the toxicological sciences [24]. Biomarker discovery in cancer uses data patterns for prediction of patient cancer prognosis, responsiveness to therapy or cancer proneness; it has its counterpart in toxicology where the aim is to understand the mass/number of genes which make up 'the toxome' (genes and pathways causatively involved in toxicity effects) or which have application for modelling how specific toxicity effects can be predicted with subsets of the toxome [10,11,13,25–27]. Thus, for transforming and modernizing toxicology, the usefulness of considering following in the footsteps of the extensive thrust of output generated in other sciences, and especially cancer research, is likely extensive. In reciprocal manners, the cancer biologist would also naturally be able to benefit from the data generation in the toxicology field (fig. 1). For example, such work could gather the *in vitro* toxicological data of known compounds to find new anticancer targets. Alternative methods development represents a broad term that relies broadly on the two pillars constituted by *in vitro* and *in silico* technologies [28]. Typically, protocols that allow for the expression of normal, specific functions in human cell cultures are complementary to the *in silico* tools required for interpretation of toxicogenomics data, and when combined, serve together as a complete non-animal-based novel tool for toxicological research under the 3R principle [5,10,28].

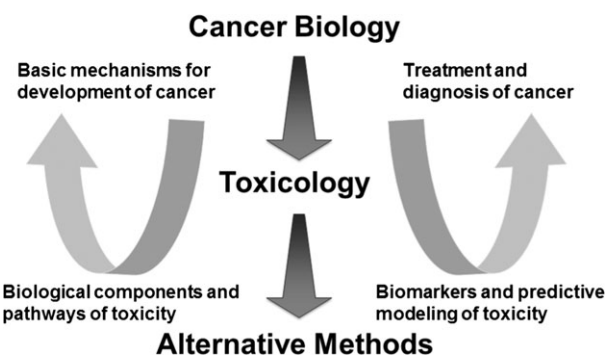


Fig. 1. The fields of cancer biology, toxicology and alternative methods development go hand-in-hand. The figure aims to illustrate schematically an advantageous example from consideration of a conscious information flow among the indicated disciplines. Both basic and clinical research data are able to cross-fertilize the fields. Alternative methods implicate *in vitro* (primarily cell culture experiment-driven) and *in silico* (database and bioinformatics tools-driven) methodologies. The data flow from (and not to) cancer biology is especially emphasized as toxicology and toxicogenomics research using primarily alternative methods constitutes the prime focus of this MiniReview.

### From HTS of Many Agents to Genomic Profiling Analysis of the Selected Few

Based on development work in our laboratories, we present here a cancer research-inspired pipeline that applies high-throughput and high-content profiling technologies integrated with omics profiling for assessing toxicity of chemicals and ENMs. Inspirational to increasing the efficiency of toxicological evaluations, pre-clinical cancer drug development includes the application of a variety of methods and technologies to pinpoint the target specificity and the absence of unwanted side effects [23]. The HTS experiments are mainly carried out to discover new and more selective cancer-associated pathways or processes [16,21]. To study pathway activation or repression caused by the drug candidates, reverse-phase protein lysate arrays (RPPA) or gene expression profiling are commonly employed thereafter [29,30]. Fig. 2 exemplifies cancer research-inspired workflows, equipments, tools and

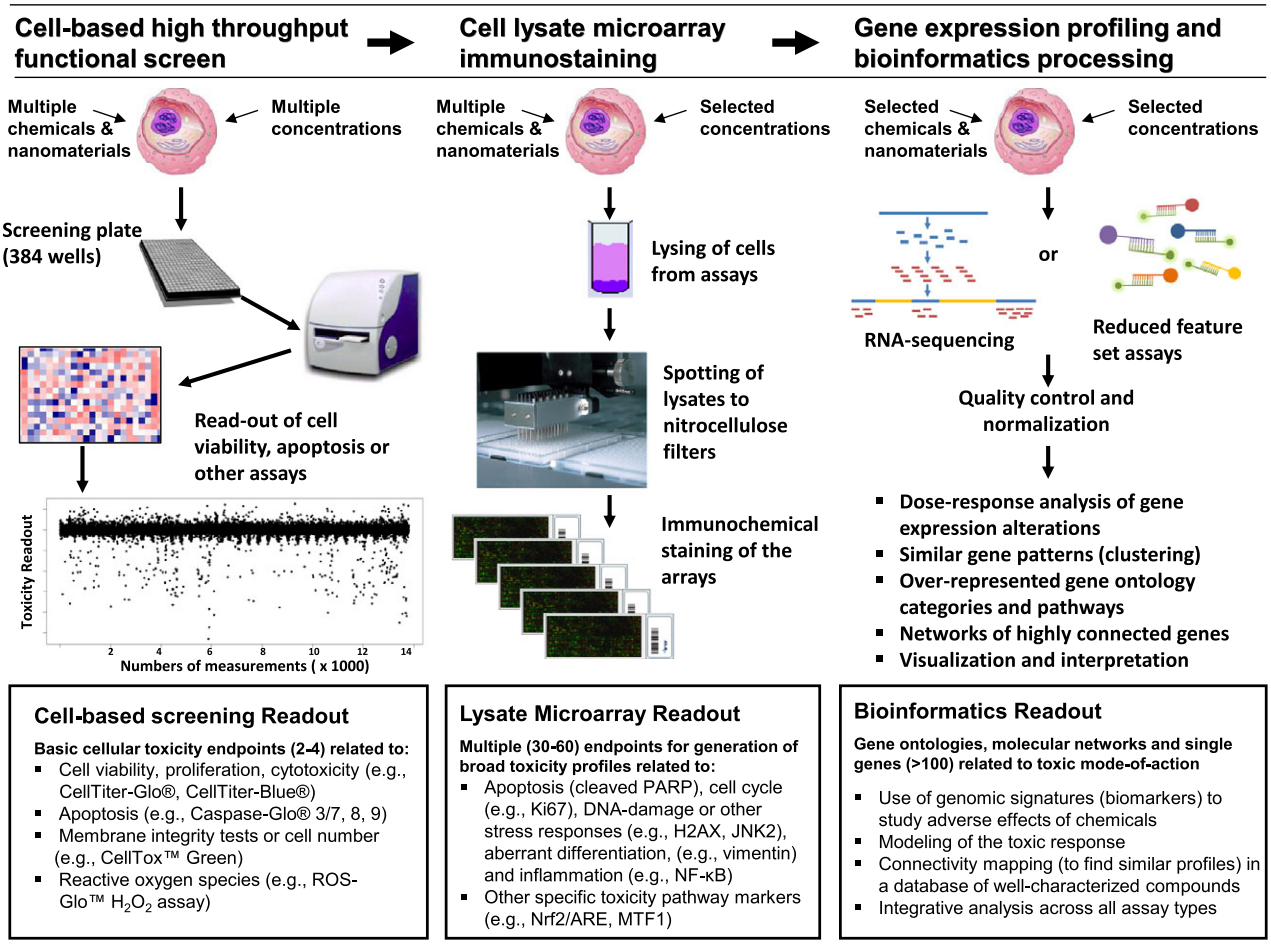


Fig. 2. Systems toxicology: from high-throughput screening of many agents to genomic profiling analysis of the selected few. Systematic application of multitechnological research equipment originally developed for cancer biology and drug screening studies permits environmental health research related to safety classification of chemicals and engineered nanomaterials. The outlined tiered approach provides an example where measurement initially of a limited number of toxicity endpoints establishes dose–response relationships for thousands of chemical and nanomaterial entities. The subsequent steps lead to gradually broader characterization of intoxicating concentrations of selected, potentially class-representative agents to the level of defining their toxic modes-of-action.

assays that are applicable also for modernized toxicological testing under the 'Tox21' concept. The tiered approach to chemical or ENM safety evaluation proceeds from HTS (stage 1 and 2) to 'omics' analysis of selected compounds that require further testing (stage 3).

In stage 1, toxicity tests with high sensitivity and accuracy are used to find chemical/ENM class-representative agents, and typically applied for relative toxic potency ranking of multiple agents. Time lapse-based morphological assessments provide a HCS dimension to the analyses, and moreover, generate a read-out that controls for potential assay interference by the tested agents. Stage 2 broadens the number of toxicity endpoints of still multiple agents in the narrowed concentration intervals considered relevant based on the stage 1 results. Both of these stage analyses depend heavily on robotics-aided experimentations, and therefore represent the HTS tools (fig. 2). In stage 3, class-representative agents with high intrinsic potency for exerting toxic effects are typically subjected to even broader characterization taking into account the concentrations and time points defined in the stage 1 and 2 results. Additionally, if following stage 1 testing, only a low number of compounds has indicated toxicity, or if the toxicity mechanism remains unclear, one could conceptually see stages 2 and 3 being assayed simultaneously, or stage 2 could follow stage 3 to verify mechanisms inferred from genomics profiling [21,30]. Typically, the gene expression profiling analysis of stage 3 would serve to determine broad toxicology-relevant bioidentities of 'the selected few' group. Such work includes the exploration or verification of novel or known pathways of toxicity, MoA mechanisms and in a broader sense, adverse outcome pathways (AOP) [2,4,5,11–13,31,32]. Expanded testing regimes might naturally apply a wider range of cell types and assays, including the consideration of low-dose effects and application of differentiated, metabolically competent cell models [6,7,19,33,34].

The second phase of toxicity testing shown in fig. 2 (middle part) is typically carried out with more endpoints than in stage 1 and involves the use of antibody-based assays that can generate 1000 times more data points using 10,000 times less

sample volume than an ordinary Western blot [29]. It therefore enables monitoring of quantitative proteomic responses for various time-scale and input-dose gradients simultaneously. It is more sensitive than high-throughput tandem mass spectrometry proteomics, enabling for instance quantitative analysis of changes in transcription factor levels [30]. Such sensitivity is important as many toxicity pathways reflect a transcriptional sensor that reacts to stimuli in the intracellular or extracellular environment of the cell [35]. Antibodies for genes and pathways tested in large screening programmes, for example, the US Environmental Protection Agency ToxCast project, are available and could be considered for inclusion in the stage 2 phase.

The third phase of the testing involves the use of sequencing or chip-based technologies for high-throughput genomewide profiling of gene activities, for example., measuring mRNA levels following a toxic insult. Reduced sets of toxicity associated genes can be assayed at higher throughput or lower cost, for example with Luminex® technology [36,37]. Microarrays, such as Affymetrix GeneChips®, form currently the basis for most of the existing gene profiling data, including the 'Japanese Toxicogenomics Project-Genomics Assisted Toxicity Evaluation system' (TG-GATEs) reference database [15,25,38]. Pathway activation can be studied using open source tools or commercial tools, such as the Ingenuity Pathway Analysis (Ingenuity® Systems, www.ingenuity.com) or the freely available InCroMap tool [39,40]. For mechanistic grouping of compounds, various bioinformatics techniques can sort the compounds into clusters by gene or pathway activation level [14]. Network analysis of gene and protein activities may then identify upstream regulators, regulatory nodes or key regulator genes from the data, potentially constituting genomic signatures or biomarkers of toxicity [14,25,26,35,41,42]. Data integration across high-throughput, high-content, pathway-based cellular assays, and omics profiling, gives the final picture on the compound activity (fig. 2).

## Surveying the Landscape of Databases for Cancer Biology and Toxicology

In order to evaluate toxicity in the context of existing knowledge, databases with corresponding results need to be addressed in conjunction with the above data (table 1) [26,38,43–61]. As annotations for genes and probes are updated regularly, it is advantageous to start such analysis with unprocessed or raw data. Gene expression and other omics profiling databases include general databases, such as the European Bioinformatics Institute (EBI) ArrayExpress data repository or domain-specific databases, for example, for toxicology [55] (table 1). The Japanese toxicogenomics project released the Open TG-GATEs with transcriptomics profiles from 170 different compounds, many with liver toxicities, containing both rat and human hepatocyte data and *in vivo* data from short- or long-term (up to 28 days) exposures [38]. The US National Toxicology Program DrugMatrix has transcriptomics, clinical chemistry and histopathology data from 657 chemical compounds in rats, both *in vivo* and *in vitro* [43]. The chemical effects in the biological systems database stores data of interest to toxicologists and environmental health scientists and serves as a search portal [45]. Derived or value-added databases have processed data and usually tools for their analysis and are very useful for targeted analyses. Databases such as Gene Expression Atlas and Human Protein Atlas provide information on almost all the transcripts or proteins, also in relation to their tissue and subcellular localization [46,55].

When analysing the toxicological potency of a chemical, knowing the phenotype of the cellular model system is important as different cellular models vary in their susceptibility to toxic treatment. Genomics of Drug Sensitivity in Cancer and Cancer Therapeutics Response Portal (see table 1) enable connections to be made between, for example, genomic alterations and drug resistance, yielding MoA information about the drugs themselves [48,51]. Connectivity mapping has therefore been suggested as a very useful tool for toxicity testing and for facilitating biological read-across [26]. The connectivity map (CMap) has thousands of genomewide expression profiles of chemical perturbations, mainly using US Food and Drug Administration approved drugs, on three cancer cell lines [44]. Investigators searching the CMap enter over- and under-expressed genes into the search engine, ranking chemotherapeutic agents on whether they regulate the same genes, either in an opposite or in similar fashion. The Mode of Action by NeTwoRk Analysis Mantra 2.0 database has clustered the CMap database and annotated the MoA of each compound, enabling determination of an unknown compound's MoA by referring to the neighbouring compounds in the network [58]. Connectivity mapping has been implemented for the TG-GATEs data set in the liver toxicity map service; the Toxygates interface to the TG-GATEs data also enables ranking of compounds based on the genes that they regulate [54,57].

In order to take full advantage of the data provided by the pipeline cellular and pathway-based (stage 1 and 2, fig. 2) assays need to be integrated with transcriptomics data as well as with traditional toxicity data. The EPA Interactive Chemical Safety for Sustainability (iCSS) dashboard which covers over 800 assays enables easy access to the ToxCast program data that include many pathway activity assays, and more than 1800 compounds [6,50,59]. The Library of Integrated Network-Based Cellular Signatures (http://www.lincsproject.org/) measures pathway activities with more than 100 antibody-based assays, as well as gene expression data from more than 10,000 compounds and over 15 cellular models sharing data in standardized format [61]. The US Environmental Protection Agency aggregated computational toxicology resource (AcToR) database, PubChem and chEMBL databases as well as the Comparative Toxicogenomics Database connects chemicals with gene expression changes and disease or toxicity associations [50,52,56,59]. But the toxicology field would benefit from even further developed integrated tools such as the Mantra 2.0 or the cBio Cancer genomics portal that integrates gene expression information, genomic alterations, cancer survival analysis and antibody-based pathway assays from more than 20 different cancer types and 1000 cellular models [24,58].

*Table 1.*

Selected examples of databases and tools for systems toxicology.

| Name | Description | Organisms | URL |
|---|---|---|---|
| Data repositories (raw profiling data) | | | |
| Gene Expression Omnibus | US public repository of gene expression data | Any | http://www.ncbi.nlm.nih.gov/geo/ |
| ArrayExpress | European public repository of gene expression data | Any | http://www.ebi.ac.uk/arrayexpress/ |
| ToxBank Data Warehouse | Warehouse for SEURAT-1 (www.seurat-1.eu) cluster data | Human, rat | http://www.toxbank.net/data-warehouse |
| diXa data warehouse | Warehouse for public toxicogenomics data by the diXa project | Human, rat | http://wwwdev.ebi.ac.uk/fg/dixa |
| Open TG-GATEs repository | Japanese toxicogenomics database (*in vivo, in vitro*, 170 compounds) | Human, rat | http://toxico.nibio.go.jp/ |
| DrugMatrix | US toxicogenomics database (*in vivo, in vitro*, 638 compounds) | Rat | https://ntp.niehs.nih.gov/drugmatrix |
| Chemical Effects in Biological Systems | Repository for public data from US National Toxicology Program | Any | http://cebs.niehs.nih.gov |
| Added value databases and tools (general and cancer research) | | | |
| Gene Expression Atlas | Gene expression patterns under different biological conditions | Any | http://www.ebi.ac.uk/gxa/ |
| Human Protein Atlas | Map protein expression in normal human tissues, cancer and cells | Human | http://www.proteinatlas.org/ |
| cBio Cancer genomics portal | Visualization, analysis and download of cancer genomics data sets | Human | http://www.cbioportal.org/public-portal/ |
| Molecular Signatures Database | Collection of annotated gene sets (>5000) for enrichment analysis | Human, other | http://www.broadinstitute.org/gsea/msigdb |
| KEGG | Kyoto encyclopedia of Genes and Genomes, pathway database | Any | http://www.genome.jp/kegg/ |
| InCroMap | Software tools to perform pathway and gene ontology analysis of processed data | Any | http://www.ra.cs.uni-tuebingen.de/software/InCroMAP/ |
| Added value databases and tools (chemical genomics and toxicogenomics) | | | |
| Genomics of Drug Sensitivity in Cancer | Identify molecular features of cancers that predict response to anti-cancer drugs (over 1000 cell lines, 142 compounds) | Human | http://www.cancerrxgene.org/ |
| Cancer Therapeutics Response Portal | Identify relationships between genomics and small-molecule sensitivities (242 cell lines, 354 compounds) | Human | http://www.broadinstitute.org/ctrp/ |
| Connectivity Map | Find similar compounds to your profile (3 cell lines, 1309 compounds) | Human | http://www.broadinstitute.org/cmap/ |
| Mantra 2.0 | Mode of Action by NeTwoRk Analysis for MoA annotated CMap analysis | Human | http://mantra.tigem.it/ |
| Comparative Toxicogenomics Database | Curated chemical, gene and disease connections and tools to analyze chemicals, genes and gene signatures (over 10000 compounds) | Human, other | http://ctdbase.org/ |
| Liver Toxicity Map | TG-GATEs similarity search (158 compounds, *in vivo/in vitro* Liver) | Human, rat | http://tcm.zju.edu.cn/ltmap/ |
| Toxygates | TG-GATEs data portal: Explore, analyze and rank compounds based on gene activity (170 compounds, *in vivo/in vitro* liver/kidney) | Human, rat | http://toxygates.nibio.go.jp/toxygates/ |
| Chemical Safety information, high-throughput screening data and tools (chemical biology and toxicology) | | | |
| AcTor | US online warehouse of publicly available chemical toxicity data | Any | http://actor.epa.gov/ |
| chEMBL | Bioactive small molecules with structures, properties and activities | Any | https://www.ebi.ac.uk/chembl/ |
| PubChem | Information on the biological activities of small molecules | Any | https://pubchem.ncbi.nlm.nih.gov |
| Toxicity Prioritization Index | A visual analytics tool for ranking compounds using multiple toxicity data | Human, other | http://www.epa.gov/ncct/ToxPi/ |
| iCSS Dashboard | Interactive Chemical Safety for Sustainability (iCSS) Dashboard is a tool to explore ToxCast and Tox21 high throughput screening data | Human, other | http://actor.epa.gov/dashboard/ |
| NCI-60 DTP | Tumor Cell Line Assay cytotoxicity data on over 100000 compounds | Human | http://dtp.nci.nih.gov/ |
| DrugBank | Drug data with comprehensive drug target information | Human, other | http://www.drugbank.ca/ |
| Opentox | Interoperable predictive toxicology framework | Any | www.opentox.org |
| Liver Toxicity Knowledge Base | US FDA drug induced liver injury centralized resource (137 most-DILI-concern drugs, 85 less-DILI-concern drugs, 65 no-DILI-concern drugs) | Human | http://www.fda.gov/. . ./LiverToxicityKnowledgeBase/ucm226811.htm |

### Cancer Research-Inspired Analysis of Toxicity Data Using Doxorubicin as Example

The European SEURAT-1 (Safety Evaluation Ultimately Replacing Animal Testing 1) including the ToxBank consortium (www.toxbank.net) has created a cross-cluster data warehouse to facilitate storage and analysis of all data generated by the more than 70 research groups that form the SEURAT-1 cluster [28,53]. The data warehouse uses the Investigation/Study/Assay (ISA) tab-delimited (TAB) format as a general purpose framework with which to collect and communicate complex metadata (i.e. sample characteristics, technologies used, type of measurements made) from experiments employing a combination of technologies [62]. Publicly available data from the ToxCast and the TG-GATEs projects are used as a benchmark [6,10,38,43,45,53]. The SEURAT-1 cluster has selected reference compounds, including doxorubicin and compiled extensive information on them in a semantic mediawiki (wiki.toxbank.net) [53]. Focusing on *in vitro* data, fig. 3 illustrates the use of multiple databases for storage of toxicity results in standardized formats and for generating hypotheses about a compound's toxicity profile.

Adriamycin hydrochloride (Doxorubicin hydrochloride, CAS Nr. 23214-92-8) is being screened as part of the Tox21 project, and results are made available via the EPA iCSS dashboard and the PubChem database as well as the EBI chEMBL database of bioactive drug-like small molecules also contain information related to doxorubicin toxicity [6,49,56,59]. Cytotoxicity studies in 59 cell lines from the US National Cancer Institute human tumour cell line assay developmental therapeutics programme (NCI-60 DTP) indicate that doxorubicin (NSC 123127) has an average LC50 of 30.55 μM, TGI of 2.48 μM and GI50 of 0.15 μM [22].

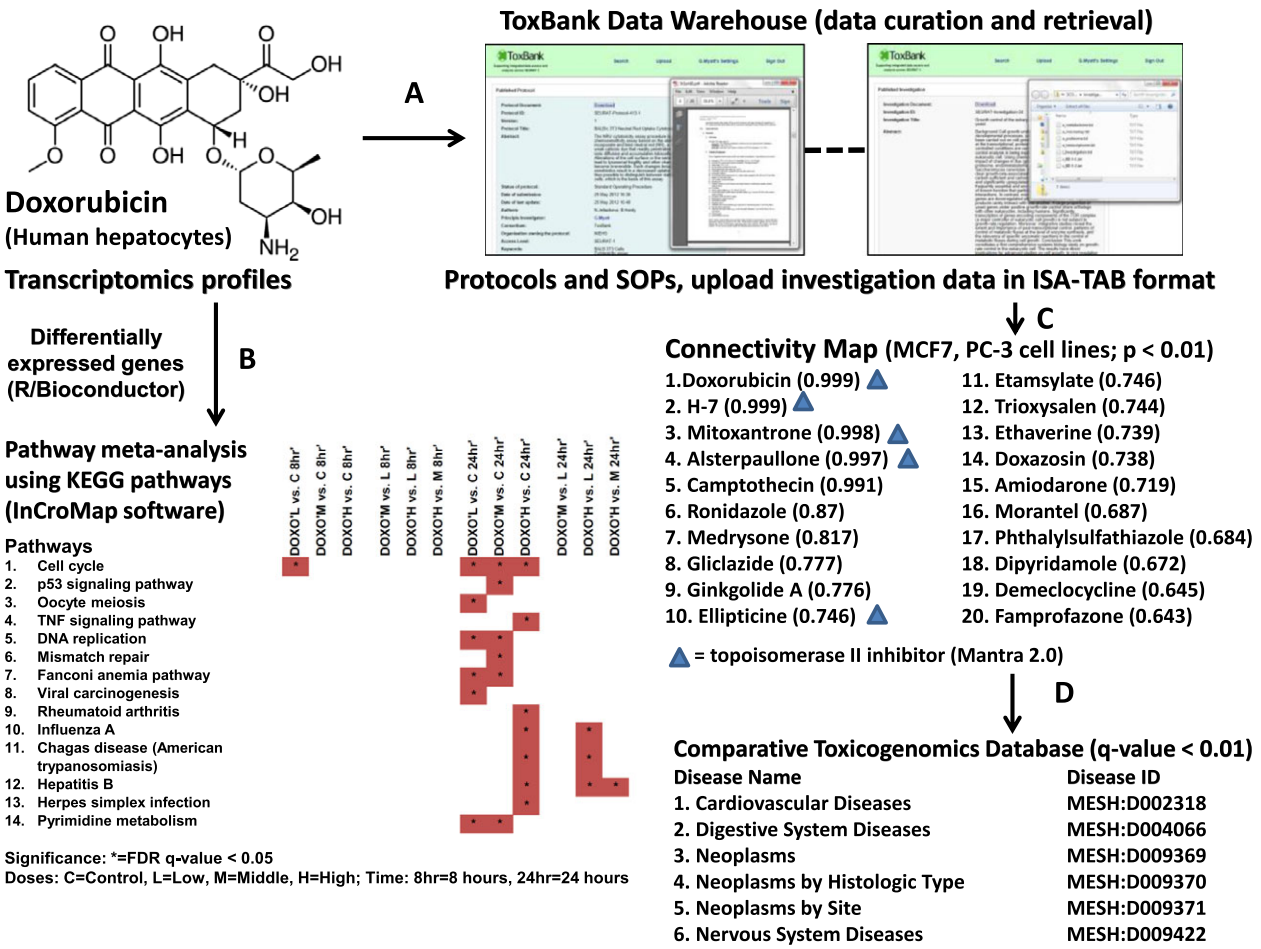


Fig. 3. Database and bioinformatics tool-driven analyses of doxorubicin toxicity. (A) Transcriptomics profiles of human hepatocytes treated with doxorubicin from the Open TG-GATEs repository are found in databases, for example the ToxBank data warehouse. (B) Differentially expressed genes can be extracted for different treatment concentrations relative the control (see legend). Kyoto Encyclopedia of Genes and Genomes pathway analyses depict molecular pathways influenced by the treatments. (C) Top 100 up- and down-regulated genes (8 hr time point, 10 μM concentrations) allows for connectivity mapping to genomic profiles of other agents with similar modes-of-action. Analysis with Mantra 2.0 implicates broad association to topoisomerase inhibitors. (D) Analysis of the top 20 connectivity-retrieved agents relative the Comparative Toxicogenomics Database generates hypotheses about the disease association; cardiovascular disease is the primer indication, being a known side effect of doxorubicin treatment. Table 1 has further details on databases and the software used in the analyses.

Gene expression data profiles of doxorubicin effects in human hepatocytes exist in the TG-GATEs data, and such results are also retrievable from the ToxBank, Toxygates or the DiXa data warehouses (see table 1) (fig. 3A) [53,54]. Bioinformatics tools variably identify a number of key pathways at different doses (fig. 3B). Interestingly, an analysis with the most significantly altered genes in the CMap service identifies doxorubicin itself and other topoisomerase inhibitors such as mitoxantrone and camptothecin (fig. 3C). Anthracyclines and related substances (ATC code L01DB01) are also enriched as a class, showing that commonly used cancer cell models and primary liver cells can have very similar profiles. Analysis of the CMap-enriched compounds using the Mantra 2.0 tool likewise indicates many of the identified connections are topoisomerase II inhibitor compounds (see fig. 3C) and CMap results submitted into the CTD service, point to 'Cardiovascular Diseases' as the most strongly enriched disease. Therefore, results are in line with doxorubicin causing cardiomyopathy and being a topoisomerase inhibitor that intercalates with DNA and induces oxidative DNA damage (fig. 3D) [63]. Thus, publicly available tools and databases help generate a correct hypothesis of systemic toxicity and define the mode of action of a toxicant.

## Concluding Statements

Toxicity assessment can be seen both as a data-driven activity and concept-driven activity. Connectivity mapping with gene expression or HTS data is an example of data-driven activity [10,13,44,54,57]. Differently, the identification of molecular initiating events and key events that lead to an adverse outcome is a concept-driven activity that facilitates evaluation of evidence for toxicity [12,13,31,32,64]. The AOP Wiki currently developed under the Organization for Economic Co-operation and Development guidelines will allow users to cooperate in documenting and evaluating information underlying AOPs [13,31,64]. Integrating mechanistic understanding with data from HTS and toxicogenomics efforts will facilitate AOP development and use so that compounds and ENMs can be assigned to various classes based on the cellular toxicity pathway activities that they trigger.

Ongoing reductions in costs for sequencing and multiplexing will make the use of high-content information technologies, especially transcriptomics, increasingly attractive. Differently, the cost of data interpretation is instead bound to increase [36,65]. Considering a tiered approach for toxicity testing (fig. 2), the numbers of compounds entering stages 3 and 2 in parallel is likely to increase, which might aid data interpretation as the results can be directly integrated. Development of well-standardized and documented bioinformatics workflows is key for integration of various omics and HTS data. Solutions to standardization of data and metadata descriptions sufficiently well for biomarker development and modelling come from implementing standardized file formats such as ISA-TAB, ontologies, for experimental factors, chemical structural and ENM descriptors, and also standard operating procedures for accurate models and classifiers of toxicity [47,53,62,66,67]. Constituting a useful example of such efforts, OpenTox has provided an extensive specification for an open interoperable standards-based predictive toxicology framework involving components for data, algorithms, compounds, biological features, models, validation and reporting which may be used to develop such workflows [47,68].

In conclusion, recent cancer genomics and systems biology studies have generated results and tools that can inspire to systems toxicology/toxicogenomics-based investigation of the toxicity effects of drugs, chemicals and ENMs; tutorials on this topic are available at the ToxBank web site (www.toxbank.net). Key tasks will be to define the 'bioidentity' or 'hazard identity' of such agents from applying HTS, HCS and bioinformatics tools and databases. The fusion of *in vitro* and *in silico* methods in toxicology serves to analyse toxicity effects from the subcellular levels up to the intact organism or population level, and overall, focuses to define molecular initiating events, pathways of toxicity, MoAs and AOPs. Bioinformatics technologies and modelling approaches are central to the data interpretation and knowledge management via databases and give so far unprecedented opportunity for rapid translational interpretation and application of research findings from alternative methods experiments. Overall, existing data, expertise and tools from cancer biology may fill central knowledge gaps that exist at the moment in toxicology, arguing for benefits from an information flow between the respective research fields. We argue for consideration of future *in vitro* and *in silico* tiered strategies that can aid environmental health research as well as constituting a safe-by-design testing protocol for synthesized agents of diverse origin and type.

### Acknowledgements

The authors have received funding from FP7-Theme HEALTH-2010-Alternative-Testing, through SEURAT/ToxBank and Cosmetics Europe under Grant Agreement nr: 267042, under grant agreement no: NMP-LA-2013-310584 from EU FP7 NMP-2012-NANoREG, under no: 309329 for EU FP7 NMP-2012-NANOSOLUTIONS and under NMP4-SL-2013-604134 for EU FP7 eNanoMapper. Further sources are grants from the Swedish Cancer and Allergy Fund, the Swedish Research Council, the Swedish Fund for Research without Animal Experiments, Finnish Foundation's Post-doc Research grant award to PK, and from Karolinska Institutet. Valuable comments to the manuscript was obtained by Professor Bengt Fadeel.

### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest are apparent to the authors.

## References

1 Classen CD (ed.) Casarett & Doull's Toxicology: The Basic Science of Poisons, 8th edn. McGraw-Hill, New York, 2013.
2 Collins FS, Gray GM, Bucher JR. Toxicology. Transforming environmental health protection. Science 2008;**319**:906–7.
3 Casciano DA, Sahu SC. (eds). Handbook of Systems Toxicology. Wiley-Blackwell, John Wiley and Sons, Chichester, UK, 2011.

4 Krewski D, Westphal M, Al-Zoughool M, Croteau MC, Andersen ME. New directions in toxicity testing. Annu Rev Public Health 2011;**32**:161–78.

5 Tralau T, Riebeling C, Pirow R, Oelgeschläger M, Seiler A, Liebsch M *et al.* Wind of change challenges toxicological regulators. Environ Health Perspect 2012;**120**:1489–94.

6 Tice RR, Austin CP, Kavlock RJ, Bucher JR. Improving the human hazard characterization of chemicals: a Tox21 update. Environ Health Perspect 2013;**121**:756–65.

7 Dourson M, Becker RA, Haber LT, Pottenger LH, Bredfeldt T, Fenner-Crisp PA. Advancing human health risk assessment: integrating recent advisory committee recommendations. Crit Rev Toxicol 2013;**43**:467–92.

8 Low Y, Sedykh A, Fourches D, Golbraikh A, Whelan M, Rusyn I *et al.* Integrative chemical-biological read-across approach for chemical hazard classification. Chem Res Toxicol 2013;**26**:1199–208.

9 Nel AE. Implementation of alternative test strategies for the safety assessment of engineered nanomaterials. J Intern Med 2013;**274**: 561–77.

10 Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I *et al.* Incorporating new technologies into toxicity testing and risk assessment: moving from 21st century vision to a data-driven framework. Toxicol Sci 2013;**136**:4–18.

11 Bouhifd M, Hogberg HT, Kleensang A, Maertens A, Zhao L, Hartung T. Mapping the Human Toxome by Systems Toxicology. Basic Clin Pharmacol Toxicol 2014;**115**:24–31.

12 Stone V, Pozzi-Mucelli S, Tran L, Aschberger K, Sabella S, Vogel U *et al.* ITS-NANO—prioritising nanosafety research to develop a stakeholder driven intelligent testing strategy. Part Fibre Toxicol 2014;**11**:9.

13 Sturla SJ, Boobis AR, Fitzgerald RE, Hoeng J, Kavlock RJ, Schirmer K *et al.* Systems toxicology: from basic research to risk assessment. Chem Res Toxicol 2014;**27**:314–29.

14 Afshari CA, Hamadeh HK, Bushel PR. The evolution of bioinformatics in toxicology: advancing toxicogenomics. Toxicol Sci 2011;**120**(Suppl. 1):S225–37.

15 Chen M, Zhang M, Borlak J, Tong W. A decade of toxicogenomic research and its contribution to toxicological science. Toxicol Sci 2012;**130**:217–28.

16 Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;**144**:646–74.

17 Freshney RI. Culture of Animal Cells: A Manual of Basic Technique, 6th edn. Wiley-Blackwell, Hoboken, NJ, 2010.

18 Pfragner R, Freshney RI. Culture of Human Tumor Cells (Culture of Specialized Cells). Wiley-Blackwell, Hoboken, NJ, 2003.

19 Ceder R, Merne M, Staab CA, Nilsson JA, Höög JO, Dressler D *et al.* The application of normal, SV40 T-antigen-immortalised and tumour-derived oral keratinocytes, under serum-free conditions, to the study of the probability of cancer progression as a result of environmental exposure to chemicals. Altern Lab Anim 2007;**35**: 621–39.

20 Breslin S, O'Driscoll L. Three-dimensional cell culture: the missing link in drug discovery. Drug Discov Today 2013;**18**:240–9.

21 Iljin K, Ketola K, Vainio P, Halonen P, Kohonen P, Fey V *et al.* High-throughput cell-based screening of 4910 known drugs and drug-like small molecules identifies disulfiram as an inhibitor of prostate cancer cell growth. Clin Cancer Res 2009;**15**:6070–8.

22 Shoemaker RH. The NCI60 human tumour cell line anticancer drug screen. Nat Rev Cancer 2006;**6**:813–23.

23 Boehm JS, Hahn WC. Towards systematic functional characterization of cancer genomes. Nat Rev Genet 2011;**12**:487–98.

24 Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;**2**:401–4.

25 MAQC Consortium. The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. Nat Biotechnol 2010;**28**:827–38.

26 Smalley JL, Gant TW, Zhang SD. Application of connectivity mapping in predictive toxicology based on gene-expression similarity. Toxicology 2010;**268**:143–6.

27 Sung J, Wang Y, Chandrasekaran S, Witten DM, Price ND. Molecular signatures from omics data: from chaos to consensus. Biotechnol J 2012;**7**:946–57.

28 Scholz S, Sela E, Blaha L, Braunbeck T, Galay-Burgos M, García-Franco M *et al.* A European perspective on alternatives to animal testing for environmental hazard identification and risk assessment. Regul Toxicol Pharmacol 2013;**67**:506–30.

29 Spurrier B, Ramalingam S, Nishizuka S. Reverse-phase protein lysate microarrays for cell signaling analysis. Nat Protoc 2008;**3**: 1796–808.

30 Östling P, Leivonen SK, Aakula A, Kohonen P, Mäkelä R, Hagman Z *et al.* Systematic analysis of microRNAs targeting the androgen receptor in prostate cancer cells. Cancer Res 2011;**71**: 1956–67.

31 Vinken M. The adverse outcome pathway concept: a pragmatic tool in toxicology. Toxicology 2013;**312**:158–65.

32 Wilson VS, Keshava N, Hester S, Segal D, Chiu W, Thompson CM *et al.* Utilizing toxicogenomic data to understand chemical mechanism of action in risk assessment. Toxicol Appl Pharmacol 2013;**271**:299–308.

33 Feliu N, Fadeel B. Nanotoxicology: no small matter. Nanoscale 2010;**2**:2514–20.

34 Vandenberg LN, Colborn T, Hayes TB, Heindel JJ, Jacobs DR Jr, Lee DH *et al.* Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. Endocr Rev 2012;**33**:378–455.

35 Andersen ME, McMullen PD, Bhattacharya S. Toxicogenomics for transcription factor-governed molecular pathways: moving on to roles beyond classification and prediction. Arch Toxicol 2013;**87**: 7–11.

36 Peck D, Crawford ED, Ross KN, Stegmaier K, Golub TR, Lamb J. A method for high- throughput gene expression signature analysis. Genome Biol 2006;**7**:R61.

37 Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009;**10**:57–63.

38 Uehara T, Ono A, Maruyama T, Kato I, Yamada H, Ohno Y *et al.* The Japanese toxicogenomics project: application of toxicogenomics. Mol Nutr Food Res 2010;**54**:218–27.

39 Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Comput Biol 2012;**8**:e1002375.

40 Wrzodek C, Eichner J, Büchel F, Zell A. InCroMAP: integrated analysis of cross-platform microarray and pathway data. Bioinformatics 2013;**29**:506–8.

41 Staab CA, Ceder R, Jägerbrink T, Nilsson JA, Roberg K, Jörnvall H *et al.* Bioinformatics processing of protein and transcript profiles of normal and transformed cell lines indicates functional impairment of transcriptional regulators in buccal carcinoma. J Proteome Res 2007;**6**:3705–17.

42 Ceder R, Haig Y, Merne M, Hansson A, Zheng X, Roberg K *et al.* Differentiation-promoting culture of competent and noncompetent keratinocytes identifies biomarkers for head and neck cancer. Am J Pathol 2012;**180**:457–72.

43 Ganter B, Tugendreich S, Pearson CI, Ayanoglu E, Baumhueter S, Bostian KA *et al.* Development of a large-scale chemogenomics database to improve drug candidate selection and to understand mechanisms of chemical toxicity and action. J Biotechnol 2005;**119**:219–44.

44 Lamb J. The Connectivity Map: a new tool for biomedical research. Nat Rev Cancer 2007;**7**:54–60.

45 Waters M, Stasiewicz S, Merrick BA, Tomer K, Bushel P, Paules R *et al.* CEBS–Chemical Effects in Biological Systems: a public data repository integrating study design and toxicity data with microarray and proteomics data. Nucleic Acids Res 2008;**36**(Database issue):D892–900.
46 Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M *et al.* Towards a knowledge-based Human Protein Atlas. Nat Biotechnol 2010;**28**:1248–50.
47 Willighagen EL, Jeliazkova N, Hardy B, Grafström RC, Spjuth O. Computational toxicology using the OpenTox application programming interface and Bioclipse. BMC Res Notes 2011;**4**:487.
48 Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW *et al.* Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature 2012;**483**:570–5.
49 Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res 2012;**40**(Database issue):D1100–7.
50 Judson RS, Martin MT, Egeghy P, Gangwal S, Reif DM, Kothiya P *et al.* Aggregating Data for Computational Toxicology Applications: the U.S. Environmental Protection Agency (EPA) Aggregated Computational Toxicology Resource (ACToR) System. Int J Mol Sci 2012;**13**:1805–31.
51 Basu A, Bodycombe NE, Cheah JH, Price EV, Liu K, Schaefer GI *et al.* An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. Cell 2013;**154**:1151–61.
52 Davis AP, Murphy CG, Johnson R, Lay JM, Lennon-Hopkins K, Saraceni-Richards C *et al.* The Comparative Toxicogenomics Database: update 2013. Nucleic Acids Res 2013;**41**(Database issue): D1104–14.
53 Kohonen P, Benfenati E, Bower D, Ceder R, Crump M, Cross K *et al.* The ToxBank data warehouse: supporting the replacement of in vivo repeated dose systemic Toxicity testing. Mol Inform 2013;**32**:47–63.
54 Nyström-Persson J, Igarashi Y, Ito M, Morita M, Nakatsu N, Yamada H *et al.* Toxygates: interactive toxicity analysis on a hybrid microarray and linked data platform. Bioinformatics 2013;**29**: 3080–6.
55 Rung J, Brazma A. Reuse of public genome-wide gene expression data. Nat Rev Genet 2013;**14**:89–99.
56 Willighagen EL, Waagmeester A, Spjuth O, Ansell P, Williams AJ, Tkachenko V *et al.* The ChEMBL database as linked open data. J Cheminform 2013;**5**:23.
57 Xing L, Wu L, Liu Y, Ai N, Lu X, Fan X. LTMap: a web server for assessing the potential liver toxicity by genome-wide transcriptional expression data. J Appl Toxicol 2013; (in press).
58 Carrella D, Napolitano F, Rispoli R, Miglietta M, Carissimo A, Cutillo L *et al.* Mantra 2.0: an online collaborative resource for drug mode of action and repurposing by network analysis. Bioinformatics 2014; (in press).
59 Judson R, Houck K, Martin M, Knudsen T, Thomas RS, Sipes N *et al.* In Vitro and Modelling Approaches to Risk Assessment from the U.S. Environmental Protection Agency ToxCast Programme. Basic Clin Pharmacol Toxicol 2014;**115**:69–76.
60 Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res 2014;**42**(Database issue): D199–205.
61 Vempati UD, Chung C, Mader C, Koleti A, Datar N, Vidovic D *et al.* Metadata Standard and Data Exchange Specifications to Describe, Model, and Integrate Complex and Diverse High-Throughput Screening Data from the Library of Integrated Network-based Cellular Signatures (LINCS). J Biomol Screen 2014; (in press).
62 Sansone SA, Rocca-Serra P, Field D, Maguire E, Taylor C, Hofmann O *et al.* Toward interoperable bioscience data. Nat Genet 2012;**44**:121–6.
63 Ewer MS, Ewer SM. Cardiotoxicity of anticancer treatments: what the cardiologist needs to know. Nat Rev Cardiol 2010;**7**:564–75.
64 Organisation for Economic Co-operation and Development, Environment Directorate Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology. Guidance document on developing and assessing adverse outcome pathways, Series on Testing and Assessment No. 184, ENV/JM/MONO(2013)6. Inter-organization Programme for the Sound Management of Chemicals, Paris, p. 45.
65 Sboner A, Mu XJ, Greenbaum D, Auerbach RK, Gerstein MB. The real cost of sequencing: higher than you think!. Genome Biol 2011;**12**:125.
66 Malone J, Holloway E, Adamusiak T, Kapushesky M, Zheng J, Kolesnikov N *et al.* Modeling sample variables with an Experimental Factor Ontology. Bioinformatics 2010;**26**:1112–8.
67 Thomas DG, Gaheen S, Harper SL, Fritts M, Klaessig F, Hahn-Dantona E *et al.* ISA-TAB-Nano: a specification for sharing nanomaterial research data in spreadsheet-based format. BMC Biotechnol 2013;**13**:2.
68 Hardy B, Douglas N, Helma C, Rautenberg M, Jeliazkova N, Jeliazkov V *et al.* Collaborative development of predictive toxicology applications. J Cheminform 2010;**2**:7.