

# Toward the Replacement of Animal Experiments through the Bioinformatics-driven Analysis of 'Omics' Data from Human Cell Cultures

Roland C. Grafström,<sup>1a</sup> Penny Nymark,<sup>1</sup> Vesa Hongisto,<sup>2</sup> Ola Spjuth,<sup>3,4</sup> Rebecca Ceder,<sup>1</sup> Egon Willighagen,<sup>5</sup> Barry Hardy,<sup>6</sup> Samuel Kaski<sup>7</sup> and Pekka Kohonen<sup>1a</sup>

<sup>1</sup>Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden; <sup>2</sup>Toxicology Department, Misvik Biology Corporation, Turku, Finland; <sup>3</sup>Department of Medical Epidemiology and Biostatistics, Swedish e-Science Research Centre, Karolinska Institutet, Stockholm, Sweden; <sup>4</sup>Department of Pharmaceutical Biosciences and Science for Life Laboratory, Uppsala University, Uppsala, Sweden; <sup>5</sup>Department of Bioinformatics-BiGCat, Maastricht University, Maastricht, The Netherlands; <sup>6</sup>Douglas Connect GmbH, Zeiningen, Switzerland; <sup>7</sup>Helsinki Institute for Information Technology, Aalto University, Department of Computer Science, Helsinki, Finland

**Summary** — This paper outlines the work for which Roland Grafström and Pekka Kohonen were awarded the 2014 Lush Science Prize. The research activities of the Grafström laboratory have, for many years, covered cancer biology studies, as well as the development and application of toxicity-predictive *in vitro* models to determine chemical safety. Through the integration of *in silico* analyses of diverse types of genomics data (transcriptomic and proteomic), their efforts have proved to fit well into the recently-developed Adverse Outcome Pathway paradigm. Genomics analysis within state-of-the-art cancer biology research and Toxicology in the 21st Century concepts share many technological tools. A key category within the Three Rs paradigm is the *Replacement* of animals in toxicity testing with alternative methods, such as bioinformatics-driven analyses of data obtained from human cell cultures exposed to diverse toxicants. This work was recently expanded within the pan-European SEURAT-1 project (Safety Evaluation Ultimately Replacing Animal Testing), to replace repeat-dose toxicity testing with data-rich analyses of sophisticated cell culture models. The aims and objectives of the SEURAT project have been to guide the application, analysis, interpretation and storage of 'omics' technology-derived data within the service-oriented sub-project, ToxBank. Particularly addressing the Lush Science Prize focus on the relevance of toxicity pathways, a 'data warehouse' that is under continuous expansion, coupled with the development of novel data storage and management methods for toxicology, serve to address data integration across multiple 'omics' technologies. The prize winners' guiding principles and concepts for modern knowledge management of toxicological data are summarised. The translation of basic discovery results ranged from chemical-testing and material-testing data, to information relevant to human health and environmental safety.

**Key words:** *alternative methods, adverse outcome pathway, benchmark dose, bioinformatics, databases, in silico, systems toxicology, toxicogenomics.*

**Address for correspondence:** Roland C. Grafström, Institute of Environmental Medicine, Karolinska Institutet, Box 210, S-171 77 Stockholm, Sweden.  
E-mail: roland.grafstrom@ki.se

## Introduction

Toxicological research has been facing an ethical and cost-driven shift away from the use of laboratory animals toward human-based cell culture methodologies. Such *in vitro* assays have the potential to be optimised for high-throughput screening (HTS) assays, allowing entire libraries of potential drugs, or possible toxicants, to be screened for cellular responses (1–3). In parallel, the relatively recently evolved high-content analysis (HCA) in the field of toxicogenomics, has per-

mitted broad genomic analyses in the context of toxicological and pathological effects (4, 5). Rapidly evolving novel bioinformatics tools, in related areas such as in drug discovery and cancer biology, show great promise in helping toxicological HTS and toxicogenomics to become more-widely applicable to safety evaluations, be they for chemicals, drug molecules, nanomaterials, or the complex mixtures within consumer products (4, 6–9).

Detailed molecular cellular interactions and disturbances exerted by environmental agents can potentially be related to biological changes and

<sup>a</sup>These authors were awarded the 2014 Lush Science Prize.

adverse health effects. The Adverse Outcome Pathway (AOP) concept considers toxic agent-induced molecular alterations from within cells, via the cell, organ and organism (10). The concept aims even further than this — to be able to understand toxicity effects within a population (11).

Built on the paradigm that toxicity is a disease of excess and deficiency, for example, when performing toxicogenomics studies, chemical exposures serve to probe the multiple and complex means whereby detrimental effects occur in a dose-dependent manner (11, 12). State-of-the-art protocols that allow for the expression of normal and specific functions in human cell cultures are complementary to the *in silico* methods required for the interpretation of toxicogenomics data. When combined, they serve as a novel non-animal based tool for toxicological research according to the Three Rs principles (8, 13, 14).

The vision for transforming toxicology research, termed *Toxicity Testing in the 21st Century*, agrees with the applicability of computational methods. There is a focus on *in vitro* models for establishing information on critical toxicity pathways significant to, and potentially predictive of, the results from animal toxicity studies (11, 12). A further approach is to group data/agents into classes by biomarker signature similarity, by applying advanced clustering algorithms (8, 15). On this basis, part of our work has been to propose protocols that expand and mine toxicity data sets for generating much needed toxicity-predicting tools, biomarker signatures and gene sets that couple strongly and directly to toxic outcomes.

### The Adverse Outcome Pathway (AOP) concept

The AOP concept can be seen to capitalise on the advances in molecular toxicology and bioinformatics for the mechanistic analysis of toxicity pathways, to build overarching toxicity predictive workflows. Both of these strands are brought together in the *Toxicity Testing in the 21st Century* (Tox21) framework advocated in the original National Research Council report (12). Bioinformatics generally excels at mechanistic analysis, but the bridging of mechanistic understanding with prediction of adverse outcomes, and how to go from *in vitro* data to *in vivo* predictions, can be problematic. The development of better cellular and tissue models, and the comprehensive characterisation of toxicity mechanisms by using ‘Big Data’ approaches (e.g. ToxCast and Tox21), aided by large-scale toxicogenomics databases, are needed here. Also, larger scale validation studies for the AOP approach need to be developed, and analysis methods automated, so that complex determinations can be done in a timely and repro-

ducible manner (8, 9, 14). A division of labour across continents is emerging between US and European approaches. European projects have tended to be smaller in scale and focus more on creating model systems for academic research (16, 17). It needs to be recognised that these projects are unlikely to be sufficiently broad in their coverage of the chemical response space to be able to build comprehensive predictive models, even with the AOP approach. Data from large US-based projects are needed for this, including the Connectivity Map (CMap), Library of Integrated Network-based Cellular Signatures (LINCS), and the future ToxCast Phase III transcriptomics projects (1–3, 6, 18). The European projects do, however, generate invaluable validation data sets that need to be properly and carefully managed and documented for data recycling (14).

### ‘Omics’-based methods

A central problem with the current *in vitro* evaluation of potentially toxic agents is that most assays monitor cellular damage in cultured cells, instead of trying to measure indicators (i.e. biomarkers) that better inform toxicity in humans (19). HCA technology, as an ‘omics’-based method, has the potential to be a better indicator of toxicity at a broader level, i.e. at a systems biology level (4, 5). ‘Omics’ data on several levels, and in combination with high-throughput phenotypic screening results, can be integrated to reveal complex physiological effects. Responses that are connected to toxicity can then be isolated from this response-space and used to predict dose-dependent toxicity. Subsequently, models can be validated by using case studies that endorse a ‘learning by doing’ approach to demonstrate the application of toxicity pathways (and AOPs) directly for human health risk assessment (20).

### Combining the methods for a better understanding

An early application of these toxicity pathway models in risk assessment could be realised by gradually incorporating novel technologies in the form of biological read-across within the AOP framework for grouping and classification (8, 20, 21). Connectivity mapping — grouping chemicals together by similarities in gene expression profiles — can be seen as a form of biological read-across, but model-based methods can be combined to enhance performance by putting greater weight on similarities that are toxicity-related (6, 15, 22). These endeavours can be achieved by extending existing paradigms to reveal key issues that need to be addressed, rather than worrying over large-

scale paradigm changes in the regulatory domain. *In vitro* to *in vivo* extrapolation (IVIVE) through toxicokinetics presents a challenge, regardless of the *in vitro* endpoints (20, 23, 24). For example, it has been suggested that a practical starting point would be to measure the free concentration of a substance in solution, i.e. to compare the estimated free concentration *in vivo* (e.g. in plasma or tissue) following consumer exposure with the free concentration of the same substance in the medium *in vitro* (24).

### Systems Toxicology Based on *In Vitro* and *In Silico* Analyses of *In Vitro* High-throughput and High-content Data Underlies the Lush Science Prize Award

Involvement in many EU projects related to alternative methods for safety evaluations of chemicals or nanomaterials, such as SEURAT-1, NANOSOLUTIONS, eNanoMapper, and NANOREG 1 and 2 projects, has served to inspire our team to combine *in vitro*-based and *in silico*-based methods for predictive toxicology purposes. Our aspirations for a systems toxicology approach required the use of *in silico* analyses of *in vitro* high-throughput (HT) and high-content (HC) data (Figure 1). The molecular bases of cell turnover and death induction were addressed, and thus the work had direct toxicological implications (8, 14). Importantly, in agreement with the Lush Prize aims, we applied an animal-free testing philosophy to the generation of novel, broadly applicable means of safety prediction based on toxicity pathway analyses. The multi-faceted nature of our studies meant that:

- Stem cell-like transformed models (or differentiating and non-differentiating progeny of normal and tumour-derived tissues) were variably employed.
- Analyses of transcript and protein profiles were complemented by functional assessments (25).
- The development and dissemination of methods for culturing the models under progressively defined, often serum-free, conditions further contributed toward the generation of increasingly reliable safety evaluations (13).
- Bioinformatics processing tools enabled the sorting of HT screening and HC screening data into visually understandable and interpretable results, and permitted comparison and evaluation of pertinent findings against publicly-stored reference data from various repositories (7, 8, 14) (Figure 2).
- A 'data warehouse' that is under continuous expansion was developed. This warehouse, cou-

pled with the development of novel data storage and management methods for toxicology, serves to address data integration across multiple 'omics' technologies.

- The toxicogenomic dimensions of our research developments were clearly aligned with the ambitions of Tox21, and furthermore, our methods have resulted in user-friendly translational outcomes consisting of open-source data collections for informatics-driven toxicity and pathway-based assessment of chemicals and nanomaterials.
- Novel bio-identity descriptions of the agents were generated by infusing 'omics' results into Quantitative Structure–Activity Relationships (QSARs).

### Some long-term benefits of the prize-winning work

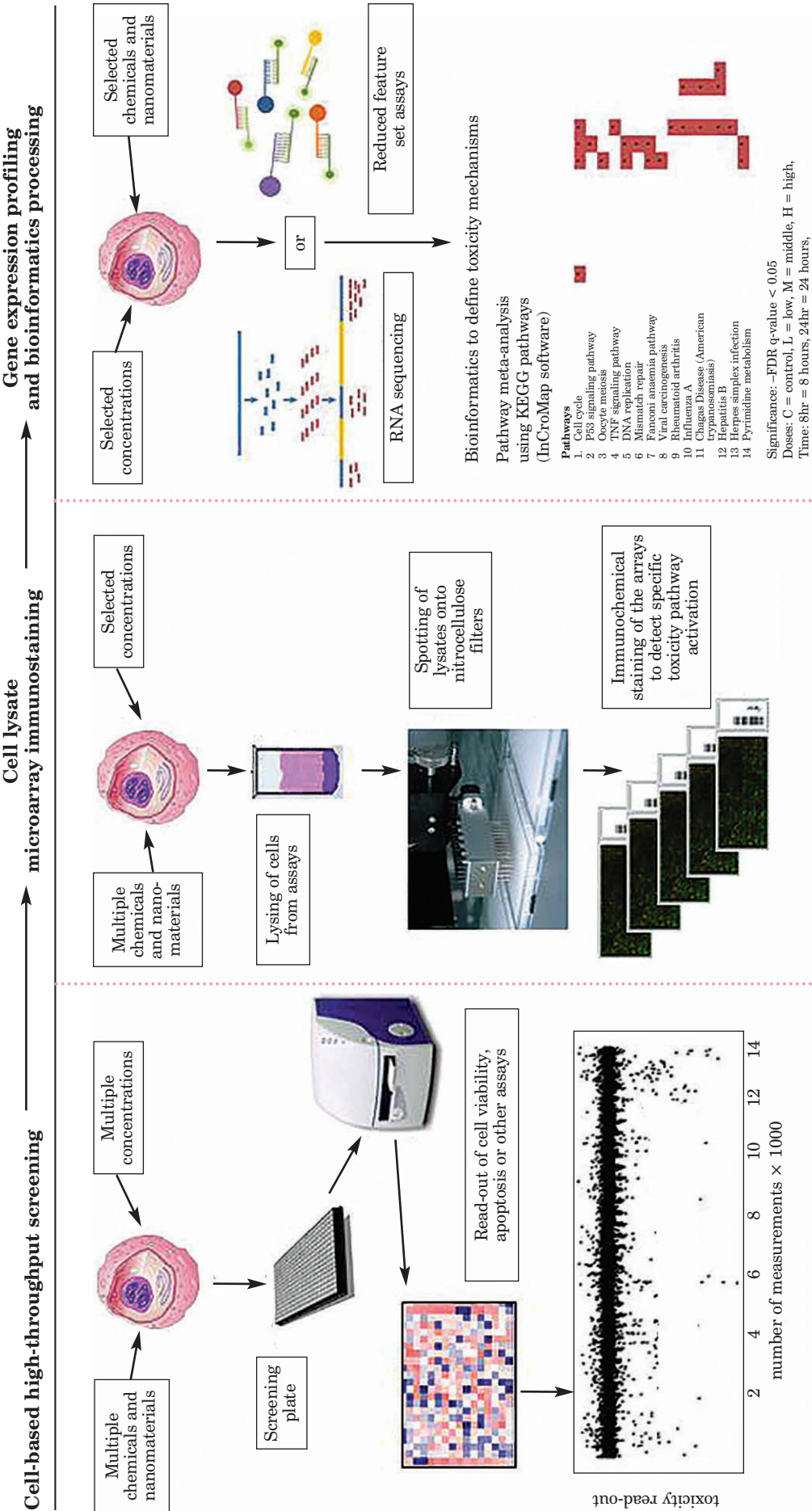
The ToxBank data warehouse will enable researchers to avoid toxicity testing in animals through the adoption of bioinformatics-driven analyses of database results (e.g. data retrieved from the Comparative Toxicogenomics Database). Moreover, the association of toxicity effects of specific compounds to individual biomarker genes/gene ontologies that reflect/couple relevant toxicity pathways to cellular functions were also derived. Both basic and clinical research data are required to cross-fertilise the *in vitro* (e.g. primary cell culture and experiment-driven) and *in silico* (e.g. database and bioinformatics-driven) methodologies. The tiered approach has gradually led to a broader characterisation of the toxic concentrations of selected, and potentially class-representative, agents to the level of defining their relevant toxic modes-of-action. The utility of bioinformatics technology applied to toxicity pathway discovery work has been enhanced and supported by comprehensive use of existing biological pathways, toxicogenomics and other 'omics'-related databases and tools for systems toxicology. Furthermore, we have successfully demonstrated to the ToxBank and SEURAT cluster of 70 research groups the novel application of bioinformatics technology to *in vitro*-based relevant toxicity pathway discovery work — the case in point being the analysis of doxorubicin toxicity (see Figure 2). We aimed for a broad scientific audience, to create an easily understood and comprehensive methodology of applying functional assay-derived and gene expression profiling data for relevant toxicity pathway discovery.

### Discussion

Approaches for a Toxicity Testing in the 21st Century-inspired, pathway-based and *in vitro*-



Figure 1: A systems toxicology approach to toxicity pathway-based safety assessment



The systematic application of multi-technological research equipment originally developed for cancer biology and drug screening studies permits environmental health research related to the safety classification of chemicals and engineered nanomaterials. The tiered approach outlined 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 characterisation of the toxic concentrations of selected, potentially class-representative agents to the level of defining their toxic modes-of-action, useful to AOP-based safety assessment. For example, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses depict molecular pathways influenced by the treatments. Adapted and developed further from Kohonen et al. (14).

**Figure 2: The concept for database and bioinformatics tool-driven analyses of chemicals and engineered nanomaterials**

#### Data curation and retrieval



#### Analysis in databases



#### Hazard prediction (e.g. doxorubicin)

1. Cardiovascular diseases
2. Digestive system diseases
3. Neoplasms
4. Nervous system diseases

*Transcriptomics profiles of human hepatocytes treated with a selected agent (the chemical doxorubicin is used as an example) from the Open TG-GATES repository are found in databases, for example the ToxBank data warehouse. Differentially-expressed genes are extracted. The top 100 up-regulated and down-regulated genes (8-hour time point, 10 $\mu$ M concentrations) allow for connectivity mapping to the genomic profiles of other agents with similar modes-of-action in the Connectivity Map database (6). Analysis of the top 20 connectivity-retrieved agents relative to the Comparative Toxicogenomics Database generates hypotheses about the disease association; cardiovascular disease, which is the top-ranked adverse effect by p-value, is a known side-effect of doxorubicin treatment. Adapted and developed further from Kohonen et al. (8, 14).*

based risk assessment have been widely reviewed (2, 8, 11, 17, 20, 26), and the framework based on AOPs has been described in detail (10, 27). Case studies focusing on P53-activating DNA damaging agents, as an example, point the way toward practical applications of the AOP concepts (20, 28). Efforts are under way to accelerate progress

toward the implementation of these concepts to embrace the Three Rs and, eventually, to realise the One R: *Replacement* (1, 3, 8, 16, 26, 30). This research is being cross-fertilised by, and performed in collaboration with, drug development and cancer research programmes (1, 3, 8, 13, 25). The development of novel cellular and organotypic models, as well as of toxicity pathway-oriented reporter systems, has been reviewed elsewhere (8, 13, 16, 29, 30). These are, in many cases, able to predict drug-induced injury with moderate success (19). Toxicogenomics has performed very well to describe mechanisms and pathways associated with toxicity (4, 7), and, with the advent of large-scale collections spanning thousands to tens of thousands of whole-genome profiles, it has become an increasingly predictive science (4, 5, 31–33). Statistical modelling techniques, thus far applied to predict drug responses, will be able to encapsulate mechanisms associated with toxicity that are shared by many compounds (34–36).

Considerations of statistical and biological plausibility, as well as the recent modelling of existing large-scale toxicogenomics data sets, indicate that even the current largest data sets (e.g. Open TG-GATES [Toxicogenomics Project-Genomics Assisted Toxicity Evaluation System] and DrugMatrix) may not be sufficient enough to encapsulate all toxicity mechanisms. This can be demonstrated in practice by using increasing numbers of compounds in predictive modelling, whereby the accuracy of the model steadily increases but this increase does not exhibit a levelling-off effect — this implies that even better models would have been built by using more compounds (5). The minimal effective number of compounds for comprehensive mechanistic modelling would therefore be above 200; optimally, it would reach thousands. It had been inconceivable to profile such large numbers of compounds with whole-genome transcriptomics technologies until the introduction of the Connectivity Map Project (6), and its successor LINCS, which has generated, to date, over 1.5 million profiles covering over 100,000 conditions that are publicly available and searchable (18).

*In vitro*-only methods are also crucial to increasing throughput, as for ethical, practical and cost reasons, such large studies can only be carried out with HT model systems. A further dimension is the tissue coverage of different model systems, and the complexity of the model systems that can be profiled in such a high-throughput manner. The LINCS project uses up to 18 relatively simple model systems that, nevertheless, span most human tissues. Consideration should also be given to the number of doses involved: LINCS covers only a narrow range of concentrations per compound, mostly a single concentration with many biological replicates across the various cell types. Regardless, even using single con-

centrations per compound can enable the creation of relevant cross-compound toxicity pathway descriptions. Thus, a two-stage plan of attack for comprehensive toxicogenomic space mapping emerges. This can comprise, for instance, cross-compound modelling by using many compounds, which is verified with the help of focused *Toxicity Testing in the 21st Century* case studies (28), where a small number of compounds are profiled with a large number of biological replicates (up to six) and five to eight concentrations that allow traditional benchmark modelling by using scores from machine learning models or gene set analysis that have been formulated by using large-scale data sets (i.e. Big Data). Existing and future HT transcriptomic platforms employed by the LINCS and the Tox21 Phase III projects will permit profiling with sufficient numbers of doses and biological replicates at relatively low cost, but operating with a reduced feature set of 800–1,500 genes (2, 8, 18). The increasingly complex, and thus necessarily lower throughput, biological *in vitro* models can also be brought in at this stage, to generate a ‘best of both worlds’ predictive infrastructure.

## Concluding Remarks

Currently, the most widely employed animal-free risk assessment methods involve the use of chemical structures for read-across from well-characterised to information-poor compounds, while *in vitro*-derived data have been mainly used to support these assessments (21). ‘Omics’ profiling yields a very information-rich and unbiased view of cellular processes, and has been used to discover mechanistically similar compounds by using the so-called connectivity mapping approach (6, 18). Connectivity mapping, being very similar to read-across in concept, can be equally applied to toxicogenomics data and will likely find more widespread applicability in the risk assessment context (8, 15). Case studies outlining *in vitro*-only risk assessment usually apply benchmark dose (BMD) modelling to define, in a statistically robust manner, safe concentrations for chemicals (20, 37). Subsequently, toxicokinetics studies relate concentrations used in cell cultures to blood or organ concentrations *in vivo* in humans, in order to derive safe levels of exposure from various routes (11). For *in vitro*-only risk assessment, it is also important to perform these investigations without resorting to animal use (23, 24). BMD modelling can also be applied to genes and gene sets, and, perhaps, also to toxicogenomics-derived predictive models (2, 20, 28, 37).

Full dose–response modelling requires large numbers of doses that can be expensive to measure for whole-genome transcriptomics data, but tiered testing strategies can help by identifying doses of interest with high-throughput cellular or reporter assays (8, 38). A more comprehensive framework

also considers expected levels of exposure and matches those with information requirements increasingly filled with *in vitro* data, as the precision and predictivity of these assays increases (11, 39). To further the development of *in vitro*-only approaches, large data masses need to be integrated in an open knowledge management environment, by using the latest bioinformatics and modelling technologies (14, 40, 41). We consider the future of *in vitro*-only toxicity prediction to be bright, although challenging, as it can only be further developed if there is a greater realisation of the importance of combining the many interdisciplinary efforts and methods discussed above.

Ensuring the safety of chemical compounds requires an understanding of the dose–response relationships that underlie the induction of toxic effects. Thus, the focus of our team is to:

- a) generate analysis schemes that can reduce, transform and decompose large data masses into functional, predictive models;
- b) consider potentially small sets of gene components, rather than single distinguishable biomarker genes or gene sets;
- c) predict dose-dependency over a range of cell types and toxicity endpoints;
- d) generate ‘omics’-based scoring measures that can also serve to rank the strength of the compound safety assessment, and finally;
- e) introduce a solid genomics-driven basis for defining compound safety, both in drug development and for environmental chemicals, i.e. to describe the ‘omics’ changes that couple most strongly to toxic effects.

Constituting a key challenge for the future, searching for the ‘magic bullet’, i.e. single biomarker genes or one or few endpoints that reflect fully the AOP, may not necessarily be the correct approach. Instead, highly information-rich ‘omics’ assays provide more complex endpoints able to measure, in parallel, the masses of genes that are variably involved with the numerous mechanisms that are altered under toxicological challenge.

## Acknowledgements

The research outlined in this article was funded, in part, by the following: the Swedish Cancer and Allergy Fund, the Swedish Research Council, the Swedish Fund for Research without Animal Experiments and the Karolinska Institutet.

## References

1. Kleinstreuer, N.C., Yang, J., Berg, E.L., Knudsen, T.B., Richard, A.M., Martin, M.T., Reif, D.M.,



- Judson, R.S., Polokoff, M., Dix, D.J., Kavlock, R.J. & Houck, K.A. (2014). Phenotypic screening of the ToxCast chemical library to classify toxic and therapeutic mechanisms. *Nature Biotechnology* **32**, 583–591.
2. Andersen, M.E., McMullen, P.D. & Krewski, D. (2015). Developing tools for defining and establishing pathways of toxicity. *Archives of Toxicology* **89**, 809–812.
3. Attene-Ramos, M.S., Huang, R., Michael, S., Witt, K.L., Richard, A., Tice, R.R., Simeonov, A., Austin, C.P. & Xia, M. (2015). Profiling of the Tox21 chemical collection for mitochondrial function to identify compounds that acutely decrease mitochondrial membrane potential. *Environmental Health Perspectives* **123**, 49–56.
4. Chen, M., Zhang, M., Borlak, J. & Tong, W. (2012). A decade of toxicogenomic research and its contribution to toxicological science. *Toxicological Sciences* **130**, 217–228.
5. Gusenleitner, D., Auerbach, S.S., Melia, T., Gómez, H.F., Sherr, D.H. & Monti, S. (2014). Genomic models of short-term exposure accurately predict long-term chemical carcinogenicity and identify putative mechanisms of action. *PLoS One* **9**, e102579.
6. Lamb, J. (2007). The Connectivity Map: A new tool for biomedical research. *Nature Reviews Cancer* **7**, 54–60.
7. Afshari, C.A., Hamadeh, H.K. & Bushel, P.R. (2011). The evolution of bioinformatics in toxicology: Advancing toxicogenomics. *Toxicological Sciences* **120**, Suppl. 1, S225–S237.
8. Kohonen, P., Ceder, R., Smit, I., Hongisto, V., Myatt, G., Hardy, B., Spjuth, O. & Grafström, R. (2014). Cancer biology, toxicology and alternative methods development go hand-in-hand. *Basic & Clinical Pharmacology & Toxicology* **115**, 50–58.
9. Spjuth, O., Bongcam-Rudloff, E., Hernández, G.C., Forer, L., Giovacchini, M., Guimera, R.V., Kallio, A., Korpelainen, E., Kańduła, M.M., Krachunov, M., Kreil, D.P., Kulev, O., Łabaj, P.P., Lampa, S., Pireddu, L., Schönherr, S., Siretskiy, A. & Vassilev, D. (2015). Experiences with workflows for automating data-intensive bioinformatics. *Biology Direct* **10**, 43.
10. Villeneuve, D.L., Crump, D., Garcia-Reyero, N., Hecker, M., Hutchinson, T.H., LaLone, C.A., Landesmann, B., Lettieri, T., Munn, S., Nepelska, M., Ottinger, M.A., Vergauwen, L. & Whelan, M. (2014). Adverse outcome pathway (AOP) development I: Strategies and principles. *Toxicological Sciences* **142**, 312–320.
11. Krewski, D., Westphal, M., Andersen, M.E., Paoli, G.M., Chiu, W.A., Al-Zoughool, M., Croteau, M.C., Burgoon, L.D. & Cote, I. (2014). A framework for the next generation of risk science. *Environmental Health Perspectives* **122**, 796–805.
12. National Research Council (2007). *Toxicity Testing in the 21st Century: A Vision and a Strategy*, 216pp. Washington, DC, USA: The National Academies Press.
13. Ceder, R., Merne, M., Staab, C.A., Nilsson, J.A., Höög, J.O., Dressler, D., Engelhart, K. & Grafström, R.C. (2007). 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. *ATLA* **35**, 621–639.
14. Kohonen, P., Benfenati, E., Bower, D., Ceder, R., Crump, M., Cross, K., Grafström, R.C., Healy, L., Helma, C., Jeliaskova, N., Jeliaskov, V., Maggioni, S., Miller, S., Myatt, G., Rautenberg, M., Stacey, G., Willighagen, E., Wiseman, J. & Hardy, B. (2013). The ToxBank data warehouse: Supporting the replacement of *in vivo* repeated dose systemic toxicity testing. *Molecular Informatics* **32**, 47–63.
15. Smalley, J.L., Gant, T.W. & Zhang, S.D. (2010). Application of connectivity mapping in predictive toxicology based on gene-expression similarity. *Toxicology* **268**, 143–146.
16. Daston, G., Knight, D.J., Schwarz, M., Gocht, T., Thomas, R.S., Mahony, C. & Whelan, M. (2015). SEURAT: Safety Evaluation Ultimately Replacing Animal Testing — recommendations for future research in the field of predictive toxicology. *Archives of Toxicology* **89**, 15–23.
17. Knudsen, T.B., Keller, D.A., Sander, M., Carney, E.W., Doerrer, N.G., Eaton, D.L., Fitzpatrick, S.C., Hastings, K.L., Mendrick, D.L., Tice, R.R., Watkins, P.B. & Whelan, M. (2015). FutureTox II: *In vitro* data and *in silico* models for predictive toxicology. *Toxicological Sciences* **143**, 256–267.
18. Duan, Q., Flynn, C., Niepel, M., Hafner, M., Muhlich, J.L., Fernandez, N.F., Rouillard, A.D., Tan, C.M., Chen, E.Y., Golub, T.R., Sorger, P.K., Subramanian, A. & Ma'ayan, A. (2014). LINCS Canvas Browser: Interactive web app to query, browse and interrogate LINCS L1000 gene expression signatures. *Nucleic Acids Research* **42**, Web Server issue, W449–W460.
19. Chen, M., Bisgin, H., Tong, L., Hong, H., Fang, H., Borlak, J. & Tong, W. (2014). Toward predictive models for drug-induced liver injury in humans: Are we there yet? *Biomarkers in Medicine* **8**, 201–213.
20. Adeleye, Y., Andersen, M., Clewell, R., Davies, M., Dent, M., Edwards, S., Fowler, P., Malcomber, S., Nicol, B., Scott, A., Scott, S., Sun, B., Westmoreland, C., White, A., Zhang, Q. & Carmichael, P.L. (2015). Implementing Toxicity Testing in the 21st Century (TT21C): Making safety decisions using toxicity pathways, and progress in a prototype risk assessment. *Toxicology* **332**, 102–111.
21. Berggren, E., Amcoff, P., Benigni, R., Blackburn, K., Carney, E., Cronin, M., Deluyker, H., Gautier, F., Judson, R.S., Kass, G.E., Keller, D., Knight, D., Lilienblum, W., Mahony, C., Rusyn, I., Schultz, T., Schwarz, M., Schüürmann, G., White, A., Burton, J., Lostia, A.M., Munn, S. & Worth, A. (2015). Chemical safety assessment using read-across: Assessing the use of novel testing methods to strengthen the evidence base for decision making. *Environmental Health Perspectives* [E-pub ahead of print.]
22. Parkkinen, J.A. & Kaski, S. (2014). Probabilistic drug connectivity mapping. *BMC Bioinformatics* **15**, 113.
23. Wambaugh, J.F., Wetmore, B.A., Pearce, R., Strobe, C., Goldsmith, R., Sluka, J.P., Sedykh, A., Tropsha, A., Bosgra, S., Shah, I., Judson, R., Thomas, R.S. & Woodrow Setzer, R. (2015). Toxicokinetic triage for environmental chemicals. *Toxicological Sciences* **147**, 55–67.
24. Wetmore, B.A. (2015). Quantitative *in vitro*-to-*in vivo* extrapolation in a high-throughput environment. *Toxicology* **332**, 94–101.
25. Ceder, R., Haig, Y., Merne, M., Hansson, A., Zheng,

- X., Roberg, K., Nees, M., Iljin, K., Bloor, B.K., Morgan, P.R., Fadeel, B. & Grafström, R.C. (2012). Differentiation-promoting culture of competent and noncompetent keratinocytes identifies biomarkers for head and neck cancer. *American Journal of Pathology* **180**, 457–472.
26. Tralau, T., Oelgeschläger, M., Gürtler, R., Heinemeyer, G., Herzler, M., Höfer, T., Itter, H., Kuhl, T., Lange, N., Lorenz, N., Müller-Graf, C., Pabel, U., Pirow, R., Ritz, V., Schafft, H., Schneider, H., Schulz, T., Schumacher, D., Zellmer, S., Fleur-Böl, G., Greiner, M., Lahrssen-Wiederholt, M., Lampen, A., Luch, A., Schönfelder, G., Solecki, R., Wittkowski, R. & Hensel, A. (2015). Regulatory toxicology in the twenty-first century: Challenges, perspectives and possible solutions. *Archives of Toxicology* **89**, 823–850.
27. Villeneuve, D.L., Crump, D., Garcia-Reyero, N., Hecker, M., Hutchinson, T.H., LaLone, C.A., Landesmann, B., Lettieri, T., Munn, S., Nepelska, M., Ottinger, M.A., Vergauwen, L. & Whelan, M. (2014). Adverse outcome pathway development II: Best practices. *Toxicological Sciences* **142**, 321–330.
28. Clewell, R.A., Sun, B., Adeleye, Y., Carmichael, P., Efremenko, A., McMullen, P.D., Pendse, S., Trask, O.J., White, A. & Andersen, M.E. (2014). Profiling dose-dependent activation of p53-mediated signaling pathways by chemicals with distinct mechanisms of DNA damage. *Toxicological Sciences* **142**, 56–73.
29. Simmons, S.O., Fan, C.Y. & Ramabhadran, R. (2009). Cellular stress response pathway system as a sentinel ensemble in toxicological screening. *Toxicological Sciences* **111**, 202–225.
30. Leist, M., Hasiwa, N., Rovida, C., Daneshian, M., Basketter, D., Kimber, I., Clewell, H., Gocht, T., Goldberg, A., Busquet, F., Rossi, A.M., Schwarz, M., Stephens, M., Taalman, R., Knudsen, T.B., McKim, J., Harris, G., Pamies, D. & Hartung, T. (2014). Consensus report on the future of animal-free systemic toxicity testing. *ALTEX* **31**, 341–356.
31. Natsoulis, G., Pearson, C.I., Gollub, J.P., Eynon, B., Ferng, J., Nair, R., Idury, R., Lee, M.D., Fielden, M.R., Brennan, R.J., Roter, A.H. & Jarnagin, K. (2008). The liver pharmacological and xenobiotic gene response repertoire. *Molecular Systems Biology* **4**, 175.
32. Davis, A.P., Grondin, C.J., Lennon-Hopkins, K., Saraceni-Richards, C., Sciaky, D., King, B.L., Wiegers, T.C. & Mattingly, C.J. (2015). The Comparative Toxicogenomics Database's 10th year anniversary: Update 2015. *Nucleic Acids Research* **43**, Database Issue, D914–D920.
33. Igarashi, Y., Nakatsu, N., Yamashita, T., Ono, A., Ohno, Y., Urushidani, T. & Yamada, H. (2015). Open TG-GATES: A large-scale toxicogenomics database. *Nucleic Acids Research* **43**, Database Issue, D921–D927.
34. Blei, D.M. (2014). Build, compute, critique, repeat: Data analysis with latent variable models. *Annual Review of Statistics & its Application* **1**, 203–232.
35. El-Hachem, N., Grossmann, P., Blanchet-Cohen, A., Bateman, A.R., Bouchard, N., Archambault, J., Aerts, H.J. & Haibe-Kains, B. (2015). Characterization of conserved toxicogenomic responses in chemically exposed hepatocytes across species and platforms. *Environmental Health Perspectives* [Epub ahead of print.]
36. Costello, J.C., Heiser, L.M., Georgii, E., Gönen, M., Menden, M.P., Wang, N.J., Bansal, M., Ammad-uddin, M., Hintsanen, P., Khan, S.A., Mpindi, J.P., Kallioniemi, O., Honkela, A., Aittokallio, T., Wennerberg, K.; NCI DREAM Community, Collins, J.J., Gallahan, D., Singer, D., Saez-Rodriguez, J., Kaski, S., Gray, J.W. & Stolovitzky, G. (2014). A community effort to assess and improve drug sensitivity prediction algorithms. *Nature Biotechnology* **32**, 1202–1212.
37. Thomas, R.S., Clewell, H.J., 3rd, Allen, B.C., Wesselkamper, S.C., Wang, N.C., Lambert, J.C., Hess-Wilson, J.K., Zhao, Q.J. & Andersen, M.E. (2011). Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. *Toxicological Sciences* **120**, 194–205.
38. Thomas, R.S., Philbert, M.A., Auerbach, S.S., Wetmore, B.A., Devito, M.J., Cote, I., Rowlands, J.C., Whelan, M.P., Hays, S.M., Andersen, M.E., Meek, M.E., Reiter, L.W., Lambert, J.C., Clewell, H.J., 3rd, Stephens, M.L., Zhao, Q.J., Wesselkamper, S.C., Flowers, L., Carney, E.W., Pastoor, T.P., Petersen, D.D., Yauk, C.L. & Nong, A. (2013). Incorporating new technologies into toxicity testing and risk assessment: Moving from 21st century vision to a data-driven framework. *Toxicological Sciences* **136**, 4–18.
39. Embry, M.R., Bachman, A.N., Bell, D.R., Boobis, A.R., Cohen, S.M., Dellarco, M., Dewhurst, I.C., Doerr, N.G., Hines, R.N., Moretto, A., Pastoor, T.P., Phillips, R.D., Rowlands, J.C., Tanir, J.Y., Wolf, D.C. & Doe, J.E. (2014). Risk assessment in the 21st century: Roadmap and matrix. *Critical Reviews in Toxicology* **44**, Suppl. 3, 6–16.
40. Ma'ayan, A., Rouillard, A.D., Clark, N.R., Wang, Z., Duan, Q. & Kou, Y. (2014). Lean Big Data integration in systems biology and systems pharmacology. *Trends in Pharmacological Sciences* **35**, 450–460.
41. Zhu, H., Zhang, J., Kim, M.T., Boison, A., Sedykh, A. & Moran, K. (2014). Big data in chemical toxicity research: The use of high-throughput screening assays to identify potential toxicants. *Chemical Research in Toxicology* **27**, 1643–1651.