Regulatory Toxicology and Pharmacology 61 (2011) 141–153

Contents lists available at SciVerse ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Commentary

Current and future use of genomics data in toxicology: Opportunities and challenges for regulatory applications

Amber K. Goetz [a,*], Bhanu P. Singh [b], Michael Battalora [c], Joseph M. Breier [d], Jason P. Bailey [e], Amechi C. Chukwudebe [f], Erik R. Janus [g]

[a] Syngenta Crop Protection, LLC, Greensboro, NC 27419, USA
[b] DuPont Haskell Global Centers for Health and Environmental Sciences, Newark, DE 19714, USA
[c] DuPont Crop Protection, Newark, DE 19714, USA
[d] Bayer CropScience LP, Research Triangle Park, NC 27709, USA
[e] Dow AgroSciences LLC, Indianapolis, IN 46268, USA
[f] BASF Corporation, Research Triangle Park, NC 27709, USA
[g] Steptoe & Johnson LLP, Washington, DC 20036, USA

ARTICLE INFO

Article history:
Received 18 April 2011
Available online 5 August 2011

Keywords:
Toxicogenomics
Risk assessment
Cancer
Gene expression
Microarray

ABSTRACT

Toxicogenomics is the application of toxicology, genetics, molecular biology and environmental health to describe the response of organisms to environmental stimuli. The field of toxicogenomics has developed over the past 15 years mainly due to advances in toxicology, molecular genetics and cell biology. Its prospective use to resolve crucial data gaps and data inconsistencies could improve risk assessment by providing additional data to increase the understanding of mechanisms and modes of action (MOA) and enhance the reliability of dose–response extrapolation. Thus, toxicogenomics holds promise for advancing the scientific basis of risk assessments. However, one of the current issues is how genomic/transcriptional data is being used to further describe a MOA for oncogenicity and, in turn, its potential uses in cancer risk assessment. This commentary identifies how toxicogenomics could be used on a case by case basis to add information to a MOA addressing both the opportunities and challenges this technology holds. In addition, some pitfalls to avoid in the generation and interpretation of toxicogenomic data and validation issues that need to be addressed before toxicogenomics can be used in the risk assessment process and regulatory decisions are discussed.

1. Introduction

Well before the completion of the draft sequence of the human genome in 2000, toxicologists were integrating gene expression profiling into animal models of chemical safety testing (Lockhart et al., 1996). These efforts helped establish the field of toxicogenomics[1]. Given that toxicity testing has traditionally focused on apical endpoints, the goal of toxicogenomics is to provide mechanistic insight into toxicological effects through the integration of new technologies into chemical safety testing (Fig. 1). While toxicogenomics provides an abundance of gene expression data, there is no regulatory framework available to provide guidance on the analysis and interpretation of these data. In addition, there still exists the challenge of reproducing gene expression analyses and consistent conclusions, largely due to incomplete or limited reporting of experimental and analytical conditions used in toxicogenomic experiments (Coombes et al., 2007; Ioannidis et al., 2009; Shi et al., 2008, 2010). An additional challenge to interpreting toxicogenomic data is the fact that a change in gene expression does not necessarily reflect a change in protein expression or indicate an adverse event.

Abbreviations: BMD, benchmark dose; CAR, constitutive activated receptor; CFSAN, FDA center for food safety and applied nutrition; CTD, Comparative Toxicogenomics Database; DBP, dibutyl phthalate; eGOn, explore gene ontology; EPA, Environmental Protection Agency; EST, estimated sequence tags; FDA, Food and Drug Administration; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; LO(A)EL, lowest observed (adverse) effect level; MAQC, MicroArray Quality Control; MIAME, Minimum Information About a Microarray Experiment; MOA, mode of action; NCEA, National Center for Environmental Assessment; NO(A)EL, no observed (adverse) effect level; NRC, National Research Council; OFAS SAR, Office of Food Additive Safety Structure–Activity Relationship; OPP, Office of Pesticide Programs; ORD, Office of Research and Development; POD, point of departure; PPARα, Peroxisome Proliferator-Activated Receptor alpha; PPDC, Pesticide Program Dialogue Committee; RfC, reference concentration; RfD, reference dose; RT-PCR, Real Time Polymerase Chain Reaction; TGx, toxicogenomics; ToxML, leadscope XML database; WOE, weight of evidence.

* Corresponding author. Address: Syngenta Crop Protection Inc., P.O. Box 18300, 410 Swing Road, Greensboro, NC 27419-8300, USA. Fax: +1 336 632 7581.
E-mail address: amber.goetz@syngenta.com (A.K. Goetz).

[1] While toxicogenomics also includes investigations of global protein expression or metabolite expression in response to chemicals, this paper concentrates on global gene expression profiling (transcriptomics).

0273-2300/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.yrtph.2011.07.012

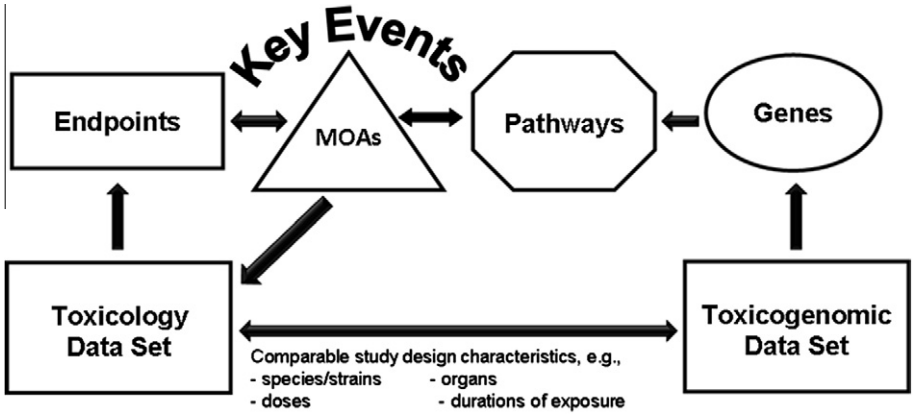

Fig. 1. Scheme for integrating toxicogenomics and traditional toxicity testing to identify key events, dysregulated pathways, and candidate modes and mechanisms of action. Toxicogenomic data can be analyzed for differentially expressed genes and categorized into adaptive or dysregulated pathways. Toxicity data can provide information about affected endpoints. Together, the data can inform the mechanisms of action, including MOAs by relating the endpoints and pathways involved (including upstream and downstream effects). In turn the proposed MOA can inform what additional endpoints may be necessary to fully describe the toxicity of the chemical. This approach requires similar study parameters for the toxicity and toxicogenomic studies (e.g., species, dose levels, tissues, duration of exposure).

Toxicogenomic data must be correlated to other biological events in order to understand the toxicological meaning behind the changes in gene expression.

Early on in toxicogenomics the concept of phenotypic anchoring was used to support the correlation between the expression profile of a chemical with a known apical endpoint or phenotypic effect (e.g., an observable effect on histopathology, clinical chemistry or hematology). The intent was to demonstrate characteristic gene expression patterns within specified dose and time parameters for a particular chemical. However, the analysis, interpretation, and integration of transcriptional profiling into toxicity testing can take several different approaches (Afshari et al., 2010). For example, specific hypothesis-driven studies may measure the expression of a defined set of genes that are considered to be involved in a compound's mechanism of action. The gene expression profile is used to accept or reject the hypothesis. The alternative, more common approach makes no assumptions about a chemical's specific MOA and utilizes the full transcriptome to identify and characterize relevant transcriptional changes related to the treatment. The resultant gene expression data is used to describe a possible mechanism of action (functional change) and generates additional hypotheses which can be explored and validated by further testing. The aim of applying transcriptional data to understand a toxicological MOA is to eventually move towards class prediction (Raghavan et al., 2005). It is postulated that using a comprehensive toxicology database of reference profiles would allow the pathway(s) perturbed by an uncharacterized chemical to be ascertained by defining which expression patterns in the database its profile most strongly resembles, in a manner analogous to fingerprinting. The establishment of genomic chemical signatures is presently a work in progress (Martin et al., 2007), as is the establishment of data repositories, including the DrugMatrix database by Iconix and the Comparative Toxicogenomics Database (CTD[2]). (As of early August 2010, CTD has listed 246,514 chemical-gene relationships and 265,748 chemical-disease relationships as taken from over 22,000 references). Given that often the researcher has limited understanding of the gene expression data generated in an experiment, one must keep in mind that the goal of this endeavour is to inform mechanistic toxicology, not drive it.

In practice, gene expression data often becomes more of a puzzle than a source of insight due to the sheer volume of data generated. Incorporating additional 'omics' technologies such as proteomics and metabolomics in conjunction with transcriptomics could help further understand the responses of an organism to toxicants (Ankley et al., 2006; Villeneuve and Garcia-Reyero, 2011). However, to use these data proficiently it will take time and validation, well-trained experts and advanced capabilities for data analysis. Interlaboratory evaluations of genomic signatures have demonstrated several challenges a regulatory reviewer is likely to face with transcriptomics datasets coming from different labs, models, strains, to different durations of exposure (Fielden et al., 2008). Currently the wider adoption of using transcriptomics data as a predictive tool is hindered by the lack of cross-lab validation of experiments, analysis and interpretation. Indeed, this will be a major challenge for regulatory agencies to overcome in establishing sound, uniform, and transparent methods to interpret and apply available genomic data in risk-based decision making. The field of toxicogenomics has made headway in understanding the meaning of gene expression data. This has been facilitated in part by developments in the field of systems biology (Alon, 2007).

Several initiatives have improved data quality, in particular the establishment of the *Minimum Information About a Microarray Experiment* (Brazma et al., 2001) has advanced data generation, analysis and reproducibility, and improvements to this process are ongoing. However, the statistical approaches used to analyze and interpret the data to identify relevant pathways sometimes fail to take the redundancy of biological pathways into consideration and could lead to a misinterpretation of the data (Khatri and Draghici, 2005). There are several tools available that can characterize association(s) of gene expression patterns with underlying biology (Afshari et al., 2010); yet these tools can bias data interpretation depending on the approach used and the amenability of the data to the applied approach (see below). Thus, there is still a need for major advancements in data interpretation, and this may in fact be predicated on greater cross-disciplinary collaboration, a process that could ultimately lead to a larger suite of tools for specific regulatory needs (Zhou et al., 2009).

Since toxicological data is not generated in a vacuum, a consequential issue from a regulatory perspective is to define when, and, more importantly, how it is appropriate to incorporate toxicogenomic information into the risk assessment process. Regulators need to be assured these new tools maintain or enhance current levels of human health and environmental protection. In this regard, a workshop was held on the validation of toxicogenomic technologies by the U.S. National Research Council in 2005. According to the workshop report (NRC, 2007b), validation should

[2] More information about this community-supported resource, including the data, can be found at http://ctd.mdibl.org.