Regulatory Toxicology and Pharmacology 72 (2015) 292–309

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

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

Technical guide for applications of gene expression profiling in human health risk assessment of environmental chemicals

Julie A. Bourdon-Lacombe [a], Ivy D. Moffat [a,*], Michelle Deveau [a], Mainul Husain [b], Scott Auerbach [c], Daniel Krewski [d], Russell S. Thomas [e], Pierre R. Bushel [f], Andrew Williams [b], Carole L. Yauk [b]

[a] Water and Air Quality Bureau, Health Canada, Ottawa, ON, Canada
[b] Environmental Health Science and Research Bureau, Health Canada, Ottawa, ON, Canada
[c] Biomolecular Screening Branch, Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, United States
[d] McLaughlin Centre for Population Health Risk Assessment, University of Ottawa, Ottawa, ON, Canada
[e] National Centre for Computational Toxicology, U.S. Environmental Protection Agency, Research Triangle Park, NC, United States
[f] Biostatistics and Computational Biology Branch, Division of Intramural Research, National Institute of Environmental Health Sciences, Research Triangle Park, NC, United States

ARTICLE INFO

Article history:
Received 2 April 2015
Available online 2 May 2015

Keywords:
Human health risk assessment
Toxicogenomics
Chemical mode of action
Hazard assessment
Risk characterization

ABSTRACT

Toxicogenomics promises to be an important part of future human health risk assessment of environmental chemicals. The application of gene expression profiles (e.g., for hazard identification, chemical prioritization, chemical grouping, mode of action discovery, and quantitative analysis of response) is growing in the literature, but their use in formal risk assessment by regulatory agencies is relatively infrequent. Although additional validations for specific applications are required, gene expression data can be of immediate use for increasing confidence in chemical evaluations. We believe that a primary reason for the current lack of integration is the limited practical guidance available for risk assessment specialists with limited experience in genomics. The present manuscript provides basic information on gene expression profiling, along with guidance on evaluating the quality of genomic experiments and data, and interpretation of results presented in the form of heat maps, pathway analyses and other common approaches. Moreover, potential ways to integrate information from gene expression experiments into current risk assessment are presented using published studies as examples. The primary objective of this work is to facilitate integration of gene expression data into human health risk assessments of environmental chemicals.

1. Introduction

Chemical risk assessment agencies worldwide are facing challenges that require new toxicity testing approaches. Major limitations of current approaches include the high cost and length of time required for tests that rely on the observation of adverse clinical or pathological effects in whole animals. As a result, human health risk assessments have been performed for only a small fraction of chemicals in commerce. To date, chemical substances inventories in Canada, the United States, and Europe contain over 23,000 (Health Canada, 2003), 84,000 (U.S. EPA, 2013), and 107,000 (ECHA, 2011) compounds, respectively. In contrast, just over 1100 compounds are regulated under U.S. legal statutes (Dernbach, 1997), and occupational exposure limits from around the world have only been derived for approximately 6000 chemicals (Brandys and Brandys, 2008). Thus, there is an urgent need for faster and more cost-effective testing strategies capable of consistently predicting chemical toxicity, and the doses at which adverse effects occur in humans.

Gene expression profiling in the context of a toxicology study (also referred to as toxicogenomics) has been identified as a promising method to alleviate some of the current constraints on human health risk assessment of chemicals. Emerging science has demonstrated the utility of gene expression profiling in identifying likely health hazards and in deciphering chemical modes of action (U.S. EPA, 2009). In the long-term, gene expression profiling may be used in chemical screening to guide further testing approaches as well as to derive points of departure (PoDs) for chemicals with limited data (Thomas et al., 2013a). This is part of the larger vision for "toxicity testing in the 21st century", in which recent advances in molecular biology are used to make more informed decisions relating to potential health risks of chemical

* Corresponding author at: Water and Air Quality Bureau, Health Canada, 269 Laurier Avenue, Ottawa, ON, Canada.
E-mail address: Ivy.Moffat@hc-sc.gc.ca (I.D. Moffat).

http://dx.doi.org/10.1016/j.yrtph.2015.04.010
0273-2300/

exposures (Krewski et al., 2010, 2011; NRC, 2007). Toxicogenomics data have also been identified as occupying a prominent place in the next generation of risk science, as envisioned by the US Environmental Protection Agency (Krewski et al., 2014).

Toxicogenomics studies offer rich datasets that can provide valuable information on chemical toxicity relevant to human health risk assessment. Important applications to current risk assessment practices include: (1) improving confidence in selecting critical endpoints through building and supporting mechanistic information; (2) enhancing understanding of whether adverse effects observed in animals are likely to occur in humans via similar modes of action; (3) guidance in selecting appropriate risk assessment approaches (such as threshold or non-threshold approaches); and (4) supporting read-across for chemical groupings. However, applications of toxicogenomics data in risk assessment have been limited, in part because there is very little information available that would allow a risk assessor with limited background in genomics to critically evaluate data quality and suitability for toxicological risk assessment.

The present manuscript provides an overview of criteria that can be used to assess the quality of toxicogenomics data, along with guidance on data interpretation. Recommendations for inclusion of these data in human health risk assessment are also provided, along with examples of their application. As this work is intended to guide non-specialists in using published genomics data to inform risk assessment, only high-level concepts have been presented, with references provided for those seeking more information on technical aspects. The manuscript is also divided into multiple sections to facilitate finding specific information when working through a toxicogenomics paper. The overarching objective of this work is to facilitate and promote the use of toxicogenomic data in human health risk assessment.

2. The basics: why gene expression profiling?

Every biological system (including cells, tissues, and whole organisms) must cope with changes in its environment, including exposure to toxic substances. A first line of defense in response to an environmental challenge can include alterations in gene expression, which generally translate into an increase or decrease in specific proteins required to carry out important tasks related to the maintenance of homeostasis. Gene expression profiles provide a snapshot of the system's overall response to a toxicant, which can be related to the mode of action (MoA) of the toxicant, and can be captured by measuring levels of messenger RNA (mRNA or protein coding RNA) in the system. These changes correspond to the molecular alterations that will give rise to phenotypic changes at higher levels of organization. Although recent research has demonstrated that non-coding RNAs are also important regulatory and structural molecules involved in biological responses (Bhan and Mandal, 2014; Cech and Steitz, 2014), the focus of this guide is primarily on mRNA.

In this article, we refer to gene expression profiling or toxicogenomics as the large scale measurement of changes in gene expression relative to control cells or tissues following a toxicological challenge. Gene expression profiling/toxicogenomics examines all of the genes in the system, or of a substantive portion of them, and takes into consideration that the human, rat and mouse genomes contain over 38,000, 29,000 and 33,000 genes, respectively (NCBI, 2011). Information on the identity of affected genes, the dose levels at which their function is altered, and the relationships among these genes is subsequently used to understand how a chemical is perturbing the system and to predict the adverse effects that may ensue.

It is important to note that toxicogenomics studies can vary considerably with respect to the biological questions under investigation as well as the technologies used to measure gene expression. Technical and biological considerations in toxicogenomics are briefly discussed below.

2.1. Technical considerations

Various technologies are used in gene expression profiling (McBride, 2015). Among those most frequently applied include DNA microarrays, large scale real-time quantitative polymerase chain reaction (RT-qPCR or qPCR) experiments, and, more recently, RNA sequencing (RNA-seq). These platforms each have their own inherent advantages and disadvantages (Table 1). In general, the data generated from these well-established toxicogenomic methodologies have been shown to be reproducible and concordant across platforms (Black et al., 2014; Shi, 2006; Wang et al., 2009; Yauk and Berndt, 2007; SEQC/MAQC-III Consortium, 2014). A brief overview of the concepts behind these technologies is presented in Fig. 1. Although the technical concepts underlying each method are different, the general concept is to identify changes in transcript abundance from exposed samples relative to controls.

Table 1
Criteria determined to be mandatory (*) or important in evaluating the overall quality of toxicogenomics experiments.

In vitro	In vivo animal	In vivo human
• Cytotoxicity was assessed, and at least some of the concentrations are below those inducing overt toxicity*	• Control animals were handled alongside treated animals using identical procedures (e.g., controls in oral gavage experiments received vehicle only) and at similar times*	• Appropriate control and exposed populations were identified*
• Unexposed control cells were cultured at the same time as the treated cells using identical cell culture procedures*	• A minimum of three biological replicates (animals) were used per group (in order to reach desired power)*	• A sufficiently large sample size was used to ensure adequate power*
• A minimum of three experimental replicates (plates) were used (in order to reach desired power)*	• If temporality is considered, time-matched controls were used*	• The exposure scenarios considered do not involve appreciable exposures to additional chemicals*
• The appropriate cell type was used, and there is a rationale for the chemical concentrations, exposure duration and harvest time selected	• The appropriate animal model and tissue was used, and there is a rationale for the doses selected	• The level of exposure was addressed and duration of the exposure is known
• If dose–response data is considered for risk modeling (including estimation of the BMD), a minimum of three doses plus control was used	• If dose–response is considered for risk modeling (including estimation of the BMD), a minimum of three doses plus control was used Ideally, at least one of these doses should be near the NOAEL	• The data take into consideration potential confounding effects*
• Tests to assess various toxicities (e.g., markers of genotoxicity or oxidative stress) were done using the same biological samples	• Tests to assess various toxicities (e.g., histopathology, biomarkers of disease) were done using the same biological samples	• There is a rationale for the tissue sample selected for dosimetry (blood, urine, other)
		• If dose–response data is considered for modeling (including estimation of the BMD), the data were divided into ordered categories of exposure, or continuous exposure data were available for individual study participants