Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Toxicology

Application of transcriptomic data, visualization tools and bioinformatics resources for informing mode of action

Melvin E. Andersen, Salil N. Pendse, Michael B. Black and Patrick D. McMullen

Abstract

Gene expression analyses have proved useful for assessing modes of action (MOAs) with various compounds both following in life exposures to intact animals and in vitro exposures of cells in culture. Various tools have been used/developed over the past decade or so to assist in the analysis of these patterns of differential gene expression – heatmaps, pathway enrichment analysis, benchmark dose estimations and network representations of affected pathways using different ontologies. MOAs assessed from using these gene expression results sometimes have confirmed expectations from conventional toxicity testing and sometimes proved divergent, giving a more comprehensive look at the affected biology with compounds such as styrene or dichloromethane. This chapter discusses modes of action inferred from gene expression studies in relation to dose response modeling, dose-dependent transitions, pathway perturbations and comparing results from one treatment condition with another. Our expectation from even the limited experience to date with differential gene expression for evaluating MOAs is that the continuing application of these tools to both in vivo and in vitro patterns of gene expression will uncover subtle differences in response among even similar compounds and also provide new tools for assessing the biological consequences of exposures to bioactive compounds.

Addresses

ScitoVation LLC, Six Davis Drive PO Box 12878, Research Triangle Park, NC, 27709-2878, United States

Corresponding author: Andersen, Melvin E. (mandersen@scitovation.com)

Current Opinion in Toxicology 2018, **9**:21–27

This review comes from a themed issue on **Risk assessment in toxicology**

Available online 17 May 2018

For a complete overview see the Issue and the Editorial

https://doi.org/10.1016/j.cotox.2018.05.003

Keywords

Mode of action, Transcriptomics, Gene ontology, Pathway enrichment, Pathway visualization, Bioinformatics, Differential gene expression, Benchmark dose modeling, Heatmaps, Information-dependent enrichment analysis (IDEA), High-throughput toxicogenomics, Modified Jaccard index (MJI).

1. Background

Mode of Action (MOA) studies for compounds that increase rodent tumors date back at least to the 1970s with assessments of dose response relationships for biomarkers — e.g., cell proliferation, DNA-reaction products, histopathology, etc. — with compounds such as saccharin and formaldehyde. About 20 years ago, MOA frameworks were formalized to guide data integration to inform human relevance [1]. The majority of MOA applications were developed from data streams collected for exposures at levels known to cause toxic responses in lifetime or at least sub-chronic exposures. Transcriptomics, arriving on the scene in the early part of this century, have provided a more comprehensive tool to assess patterns of cellular alterations. Although gene expression by itself is not a direct read-out of tissue response, this deficiency of gene expression studies has often been addressed by examining differential gene expression (DGE) at doses known to produce specific toxicological responses, a design referred to as phenotypic anchoring. Some of the first MOA applications with DGE [2] were in refining our understanding of the genes, networks and pathways affected by various compounds with known biological targets [2]. Representative studies showing consistency between proposed MOAs and gene expression include dioxin (as an AhR receptor agonist [3]), phthalates (as testicular toxicants [4]), perfluoro acids (as PPARα agonists [5]) and phenobarbital (as a CAR receptor agonist [6]). Subsequent studies were primarily with compounds of known MOAs to assess dose–response and assure the correctness of the MOAs. There has been more limited work reported with compounds such as that with acrylamide to gain clues about MIEs (molecular initiating events) or MOAs [7,8] or with compounds such as dichloromethane (DCM) where gene expression results were different from those expected from earlier MOA research [9]. This chapter highlights a suite of tools that are becoming available for assessing dose–response and target pathways from transcriptomic data, reviews a few examples of the application of these tools with specific compounds in vivo and looks to the future where MOA and MIE studies may be pursued using high-throughput transcriptomic (HTT) studies across dose, time of treatment and various cell types in vitro. The contribution represents more a review of a personal 10-year

excursion into transcriptomics by a practicing toxicologist (MEA) rather than a comprehensive review of approaches for gene expression studies in MOA research. We direct the reader to other more general reviews on gene expression analysis for informing mode of action and risk assessment [10–12].

2. Dose response

The MOA for formaldehyde tumor formation in the front of the rat nose has been an active area of investigation for well over 30 years [13]. With short-term exposures, tissue responses in the nasal epithelium included toxicity and enhanced cell proliferation. The dose response for cell proliferation and tumor incidence rise steeply with inhaled concentration. Combining results from two bioassays [14,15], the incidence of tumors on lifetime inhalation of formaldehyde was 0/122, 0/27, 0/126, 3/113, 22/34 and 157/182 at 0, 0.7, 2, 6, 10 and 15 ppm, respectively. Benchmark dose calculation at the 10% response level (BMD_{10}) for cell proliferation and tumor incidence were, respectively, 4.9 and 6.4 ppm [16].

An early contribution quantifying dose response in gene expression studies was calculation of benchmark doses (BMDs) for individual genes and for enriched categories of genes [17]. Gene expression changes in the rat nasal epithelium following acute formaldehyde exposure were analyzed using a one-way ANOVA to identify genes that showed significant dose–response behavior. The dose–response data for these genes were fit to a series of four statistical models and the least complex model that best described the data was selected. Genes were matched to associated Gene Ontology (GO) categories. GO enrichment analysis identifies relevant groups of genes that function together, reducing a catalog of gene changes to a much smaller number of biological functions, so that it is possible to understand the meaning of the changes in gene expression in terms of cellular processes (http://www.geneontology.org/).

Average BMD and benchmark dose lower confidence limit (BMDL) were then calculated for each GO category that contained some minimum number of elements from the category. This analysis identified doses at which individual cellular processes, not simply genes, were significantly altered. For the formaldehyde exposures, the BMD estimates for the GO categories related to cell proliferation and DNA damage for the short-term exposures were similar to those measured for cell proliferation and tumors. The gene expression BMD software and algorithm have been updated and are available for download (www.sciome.com/bmdexpress/). The tool can organize the gene expression dose response by GO categories, KEGG or Reactome pathways or most any other user defined category [18].

3. Time and dose dependent changes

In addition to the single day formaldehyde exposures [19], a time series of exposures – 1, 4 and 13 weeks across 5 concentrations – assessed the manner in which tissue responses were altered with prolonged injury [20]. While the major pathways from Go Ontology were similar across times, there were groups of genes differentially affected depending on the length of exposure. There were smaller groups of genes differentially affected at 1 and 4 weeks that could be analyzed separately. At 1 week, there was upregulation of genes related to inflammatory pathways; at 4 weeks there was downregulation of pathways for TGF-β, Wnt and cytoskeleton remodeling. These time-dependent changes indicated adaptation to injury with repeated exposures that cause cytotoxicity and the persistence of changes in biological pathway components likely to be related to cancer. The patterns of intermediate responses differed for the marginally carcinogenic concentration (6 ppm) and the two overtly tumorigenic exposures (10 and 15 ppm).

4. Dose-dependent transitions

The consequences of high exposures were clear and relatively stable over exposures causing cancer. In contrast to the groups of genes affected at overtly toxic exposures, there were also consistent changes in a smaller set of genes at exposures below those causing overt toxicity, i.e., at 0.7 and 2 ppm rather than 6 ppm and above. There were too few genes affected at these exposures to do enrichment analysis. Among the genes affected were *Hmox1* (heme oxygenase, a protein that produces CO), *Areg* (Amphiregulin -a member of the epidermal growth factor family), *Slc7a11* (a cysteine-glutamate transporter), *Tnfrsf12a* (a receptor that affects angiogenesis), *Maff* (a protein that partners with Nrf2 in stress pathways), *Fos1* (a regulatory protein for proliferation), *Srxn1* (sulfiredoxin) and *Trxn1* (thioredoxin). The latter two code for proteins that reduce sulfur and can participate in synthesis of another vasodilator - H_2S. These genes were grouped into a clade of "sensitive response genes" and the group had BMDs in the range of 1 ppm, below the BMD for cell proliferation or apoptosis. The dose and time responses of the various pathways provided a picture of tissue responses occurring at doses lower than those affecting overt cytotoxicity and regenerative cell proliferation. The functions of these genes indicated a mild response to cellular stress at the lower concentrations, with evidence for vasodilation and preparation for proliferation, and then a transition in cellular responses from stress sensing at the lower exposures to overt toxicity and proliferation at 6 ppm and higher. These qualitative dose-dependent changes in patterns of response indicated that back extrapolation of high dose responses for low dose risk assessment are likely to be biologically inaccurate and