Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Toxicology

Integrating gene expression biomarker predictions into networks of adverse outcome pathways

J. Christopher Corton

Abstract

Microarray profiling in the context of toxicity testing in animals has been used for years to identify mechanisms of toxicity, derive points of departure using dose–response modeling, and determine human relevance. High-throughput transcriptomic technologies are increasingly being used to screen environmental chemicals in vitro to identify molecular targets and provide mechanistic context for regulatory testing. This review will discuss the use of gene expression biomarkers to make predictions of activity of molecular targets of chemicals that can be linked to adverse outcomes in a number of cellular and tissue contexts. Gene expression biomarkers are built using global gene expression comparisons from cells or tissues exposed to chemicals with known effects on the factor of interest. Incorporating profiles in which the expression of the factor is altered (e.g. in gene-null mice) facilitates the identification of predictive genes. As an example of their use, biomarkers that predict molecular initiating events and key events in liver cancer adverse outcome pathways have been shown to accurately identify chemical–dose combinations in short-term studies that lead to liver cancer in 2-year bioassays. In the near future, batteries of biomarkers that predict modulation of important targets of environmental chemicals could be used to interpret high-throughput transcriptomic screening data.

Addresses

Integrated Systems Toxicology Division, National Health and Environmental Effects Research Laboratory, US Environmental Protection Agency, Research Triangle Park, NC, 27711, USA

Corresponding author: Chris Corton (corton.chris@epa.gov)

Current Opinion in Toxicology 2019, **18**:54–61

This review comes from a themed issue on **Genomic Toxicology**

Edited by **Susan Hester** and **Brian N Chorley**

Available online 25 June 2019

For a complete overview see the Issue and the Editorial

https://doi.org/10.1016/j.cotox.2019.05.006

2468-2020/Published by Elsevier B.V.

Keywords

Toxicogenomics, Biomarkers, Adverse outcome pathways, Nuclear receptors, Liver cancer.

1. Introduction

The 1990s ushered in the era of 'omics,' defined as the measurement and analysis of all components in a system. Gene expression microarrays that measure expression of most, if not all, genes in a genome in particular have been extensively used in conjunction with toxicity testing in animals [1]. Microarray profiling has been most often used in retrospective studies to inform mode of action of a known toxicity and determine human relevance and is increasingly used as a prospective tool to predict effects of long-term exposure such as cancer from short-term exposures [2]. Microarrays are also routinely used to perform dose–response modeling to identify transcriptional points of departure protective of adverse effects [3,4].

While past toxicology studies have heavily used commercially available microarrays (e.g. from Agilent and Affymetrix), investigators have more recently turned to sequencing-based approaches including whole-transcriptome RNA-Seq which allows a deeper interrogation of complex changes in gene expression patterns including differential splicing. However, most of the standard protocols for RNA-Seq are labor-intensive and cost prohibitive for high-throughput use [5]. Alternative versions of RNA-Seq have been developed to profile gene expression changes in treated cells in culture and include PLATE-Seq, which allows samples in 96-well plates to be profiled at a lower cost per sample [6], and Digital RNA with pertUrbation of Genes (DRUG-Seq), a massively parallelized, automated, low-cost next-generation sequencing-based method [7]. Targeted RNA-Seq is an alternative strategy to assess gene expression. The TempO-Seq platform (BioSpyder, Inc.) [8] in particular has been adapted to high-throughput screening (HTS) to allow measurement of expression of genes from lysates of treated cells and is being increasingly used by the Environmental Protection Agency (EPA) [9]. While these high-throughput transcriptomic (HTTr) profiling techniques have been principally used for in vitro profiling, it is possible that they could be adapted to profile large numbers of RNA samples isolated from the archived frozen or formalin-fixed paraffin-embedded tissues from animals or people. These techniques promise to generate ever-increasing amounts of data that will continue to transform toxicity testing.

One of the major challenges of microarray profiling is to accurately identify modulation of specific molecular targets and to put these predictions into the context of chemical toxicities. In this short review, I will describe

an approach to predict molecular effects associated with adverse outcomes using gene expression biomarkers. An example of their use will be discussed, namely how comprehensive biomarker predictions can be used to predict rat liver cancer from gene expression changes after short-term exposures.

2. Building and characterization of gene expression biomarkers

Since the first use of microarrays in 1995 [10], a large number of computational methods have been used to identify groups of differentially expressed genes (often called biomarkers, signatures, and classifiers) that allow the prediction of a predefined set of changes in molecular, cellular, and tissue endpoints that in many cases can be linked to health or disease [2]. One definition of a biomarker is 'a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention' [11]. To be useful, a gene expression biomarker should ideally have three properties. First, the biomarker must yield highly accurate predictions in the context of its use (e.g., a cell line or tissue). Second, the biomarker must be able to yield these predictions using a minimal number of samples from the experimental system tested with a parsimonious set of genes (i.e., tens rather than thousands) which ensures feasibility and reduces noise. Third, the biomarker must be insensitive to differences in sample processing or assay technologies (e.g., different methods and platforms for assessing gene expression) including platforms that differ from those used to generate the biomarker, to ensure its generalizability.

To define predictive gene sets using machine learning methods, a microarray data set can be divided into a training set used to identify the predictive genes and a test set used to estimate the classification accuracy. The classifier (biomarker) is developed using the training set and applied to each sample in the test set. Although this general approach is widely used, most of these classification methods do not incorporate *a priori* biological information to identify biomarker genes. The use of biological information adds a layer of validation and prioritization that can be exploited for biomarker discovery [12].

Our laboratory has recently used a relatively straightforward weight-of-evidence approach to build gene expression biomarkers that relies on the known biology of the factor to be predicted. The biomarkers built so far and used to predict effects in animal tissues are shown in Table 1. We have also characterized a number of biomarkers that are useful for predicting transcription factor (TF) modulation in human cells in vitro [13–15], and these are discussed in detail in a recent review [16].

While we have focused on major TFs modulated by xenobiotics in a number of tissue contexts, our approach can in principle be used to identify predictive gene sets for pathways and phenotypic effects as well if the system can be appropriately perturbed. All of the biomarkers were built starting with lists of statistically filtered genes from microarray comparisons of tissues after exposure to chemicals known to activate the TF of interest (Figure 1). As examples, the biomarkers for aryl hydrocarbon receptor, constitutive androstane receptor (CAR), and peroxisome proliferator–activated receptor α (PPARα) were made using similar approaches. Genes that exhibited consistent changes in gene expression (either increased or decreased expression) in most if not all of the exposures (e.g. agonists) in the livers of wild-type mice were first identified. Second, we used a genetic filter to exclude genes if they exhibited similar changes after chemical exposure in livers from TF-null mice [17–22]. Third, the genes were filtered for those with fairly robust expression (usually an absolute average fold-change of $\geq$1.5-fold). The resulting biomarker genes exhibited patterns of expression similar to the activators in wild-type mice that were used to create the biomarker. In each case, there is a consistent hierarchy of gene expression changes from highly induced or suppressed to those with more modest changes. The resultant gene lists were generally in the range of 50–150 genes, consistent with one of the biomarker requirements.

The preliminary biomarker gene lists were evaluated by a number of methods to confirm regulation by the TF including an assessment of the biological pathways the genes fall into, an analysis of the upstream activators predicted to regulate the genes, and importantly, performing a reexamination of archived chromatin immunoprecipitation (ChIP)-Seq data sets in which genes were identified that were physically associated with the TF [14,15,22]. These approaches have consistently shown that the TF regulates most of the biomarker genes.

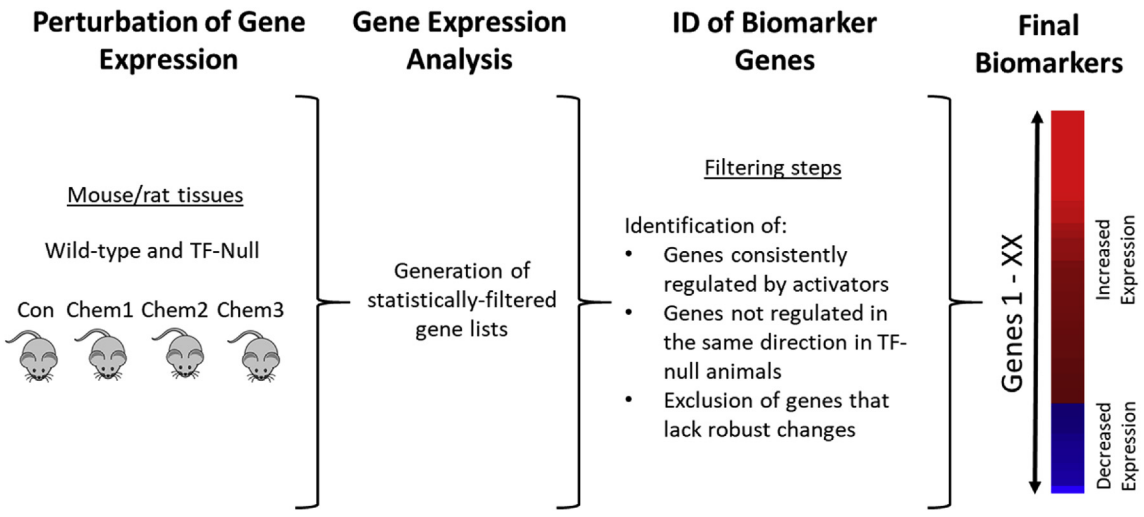
There are numerous methods for comparing the biomarker to other gene lists allowing predictions to be made of TF activity. We have used the rank-based nonparametric Running Fisher algorithm [23]. The approach finds, in an unsupervised manner, microarray comparisons with expression patterns of biomarker genes with statistically significant positive or negative correlation corresponding to activation or suppression of the TF. The Running Fisher test is analogous to the Gene Set Enrichment Analysis method [24,25]. The Running Fisher algorithm differs from Gene Set Enrichment Analysis in the assessment of the statistical significance because p-values are computed by a Fisher exact test rather than by permutations (see the study by Kuperschmidt et al. [23] for details). After the genes are ranked by absolute fold-change, the Running Fisher test

Table 1 Characteristics of gene expression biomarkers used for analysis of rodent liver microarray profiles.

Key event predicted	Domain of applicability	Number of genes	Chemical or other filter	Genetic filter	Predictive accuracy, activation	Predictive accuracy, suppression	Reference
AhR	Mouse liver	63	2,3,7,8-tetrachlorodibenzo-p-dioxin, benzo[a]pyrene	Wild-type vs. AhR-null	95%	Not performed	[18]
CAR	Mouse liver	83	CITCO, phenobarbital, TCPOBOP	Wild-type vs. CAR-null	97%	Not performed	[17]
PPARα	Mouse liver	131	WY-14,643, fenofibrate, perfluorohexane sulfonate	Wild-type vs. PPARα-null	98%	Not performed	[19]
Nrf2	Mouse liver	81	CDDO-Im	Wild-type vs. Nrf2-null and Wild-type vs. Keap1-null	96%	Not performed	[22]
STAT5b	Mouse liver	144	Male vs. female	Wild-type vs. STAT5b-null	99%	97%	[20]
AhR	Rat liver	63	3-Methylcholanthrene, beta-naphthoflavone, hexachlorobenzene, leflutamide, omeprazole, 2,3,7,8-tetrachlorodibenzo-p-dioxin	None	91%	Not performed	[36]
CAR	Rat liver	113	clotrimazole, nifedipine, phenobarbital, phenytoin	None	92%	Not performed	[40]
PPARα	Rat liver	58	clofibrate, fenofibrate, WY-14,643	None	98%	Not performed	[40]
ER	Rat liver	35	ethinyl estradiol, estradiol-3-benzoate, beta-estradiol, tamoxifen	Wild-type vs. ERα-null	96%	Not performed	[40]
Genotoxicity	Rat liver	7	2-nitrofluorene, aflatoxin B1, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, dimethyl nitrosamine	None	92%	Not performed	[40]
SREBP	Rat and mouse liver	99	None	Wild-type vs. constitutively active SREBP1a, SREBP1c or SREBP2	95% (mice)	94% (mice)	[38, 41]
				Wild-type vs. Scap-null	93% (rats)	98% (rats)	

AhR, aryl hydrocarbon receptor; CAR, constitutive androstane receptor; ER, estrogen receptor; PPARα, peroxisome proliferator–activated receptor α; SREBP, sterol regulatory element binding protein.

Figure 1



Construction of biomarkers from microarray data generated in animal tissues. Biomarker genes are identified from the differentially expressed genes in the tissues of wild-type and TF-null mice exposed to chemicals that modulate the TF. After generation of statistically filtered gene lists, the biomarker genes are derived using a number of filtering steps. Final biomarkers include a list of ~50–150 genes and associated fold-change values which are the average expression across the activator comparisons. Post hoc analysis of the genes includes pathway analysis and queries of ChIP coupled with DNA sequencing (ChIP-Seq) data sets to identify primary targets of the TF. TF, transcription factor.

compares one gene list with another to determine the level of correlation of overlapping genes. This normalized ranking approach enables comparability across data from different studies, platforms, and analysis methods by removing dependence on absolute values of fold-change.

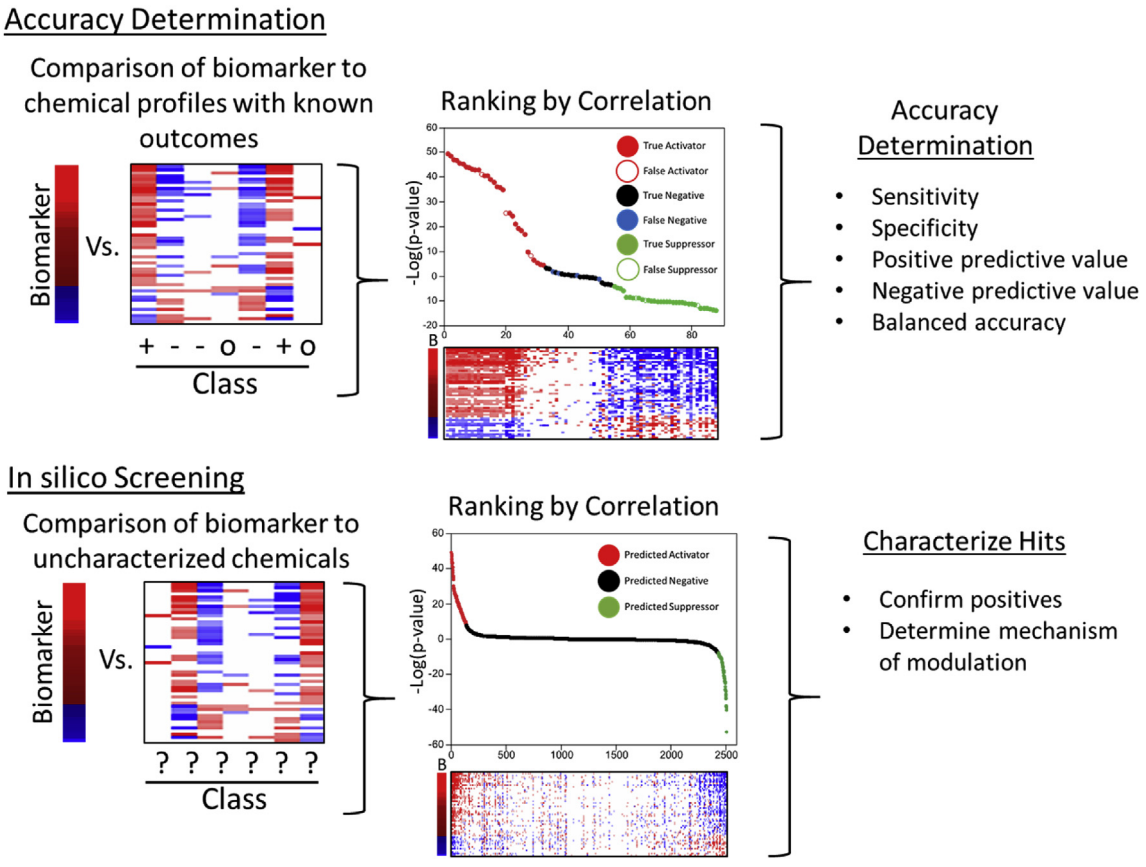
Determination of predictive accuracy of the biomarkers has been facilitated by access to a large commercially available genomic database which includes lists of statistically filtered genes derived from archived raw microarray files found in Gene Expression Omnibus or ArrayExpress [23]. While the database is continually growing, there are ~137,000 gene lists derived from ~22,400 studies (as of November, 2018). These gene lists include those derived after exposure to >4000 chemicals. They also include those derived from genetic perturbations in nullizygous mice and gene knockdown by siRNA or overexpression in human cells. We have annotated the microarray comparisons for characteristics that facilitate their use in deriving biomarker gene sets and in determining predictive accuracy. For example, the biomarker for PPARα was generated from microarray comparisons in the database of the livers from wild-type and PPARα-null mice treated with 3 PPARα agonists and then compared with a large set of profiles from mouse liver after exposure to chemicals known to exhibit/not exhibit PPARα activity [19] (Figure 2, top). The biomarkers used with the Running Fisher test were very accurate at predicting TF modulation, being in the range of 92–98% balanced accuracy (Table 1) which at least for one biomarker was shown to be independent of the platform used to assess gene expression [26]. Remarkably, these accuracies were superior to those derived using 7 machine learning methods [17–19,22].

Once the biomarker predictive accuracy has been determined and there is high confidence that the procedures can be used to identify TF modulators, the same approach described previously can be used to identify novel uncharacterized conditions that modulate the TF in the microarray compendium (Figure 2, bottom). These approaches have allowed us and others to identify (1) the chemical mode of action related to steatosis or liver cancer [27,28], (2) known and novel chemicals that activate or suppress the TF [17–19,21,22,29,30], (3) the frequency of the responses across the chemicals (e.g. the TF STAT5b was found to be frequently modulated by diverse stressors [20]), (4) genes that when modulated activate or suppress the TF in the absence of chemical exposure [18,20], and (5) interactions between the TFs (e.g. activation of CAR and Nrf2 [30] or suppression of STAT5b and activation of Nrf2 [30] are tightly coupled). Our work shows that the biomarkers can be useful tools to not only identify chemical modulators but also reveal novel aspects of the biology of the factors being examined.

3. Integration of biomarker predictions into the adverse outcome pathway framework

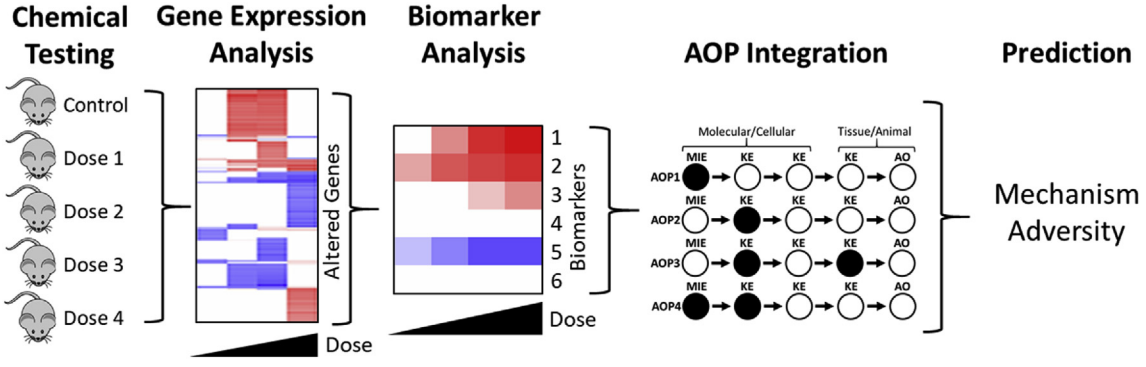
Pathway information organized by mode of action (MOA) [31] or more recently by adverse outcome pathways (AOPs) [32–34] plays a central role in compiling evidence of effects in animals and assessing human relevance. The chemical-agnostic AOP starts

Figure 2



Gene expression biomarkers: testing for accuracy and use in in silico screening. Biomarker genes are imported into the BaseSpace Correlation Engine (Illumina) environment, in which internal protocols rank-order the genes based on their fold-change. Comparisons are made between the biomarker and each microarray in the database using a pair-wise rank-based algorithm (the Running Fisher test). The results of the test, including the direction of correlation and p-value for each comparison in the compendium, are used to determine predictive accuracy of chemicals known to activate (+), suppress (–), or not affect (o) the TF and to perform screening tests to identify novel chemicals that could activate or suppress the TF which can be confirmed using other techniques and further characterized.

Figure 3



Putting biomarker predictions into the adverse outcome pathway (AOP) framework. In ongoing and future studies, statistically filtered gene lists derived from short-term animal studies are compared to a large battery of validated biomarkers, each of which can accurately predict the activation or suppression of a TF, pathway, or phenotype (e.g. cell proliferation). MIE/KE predictions are then used to populate AOP network models to predict specific AOs. Various scoring systems could be used to predict chemical doses that would cause toxicity and their underlying mechanism.

with the molecular initiating event (MIE) followed by a series of key events (KEs) that lead to an adverse outcome (AO). The overlay of chemical-specific information such as absorption, disposition, metabolism, and excretion, prediction of chemical concentrations at the site of the MIE, and quantification of MIE perturbation onto a discrete AOP construct then permits the derivation of the corresponding chemical-specific MOA for risk assessment [33].

A goal of a number of laboratories is to develop methods to project high-throughput screening predictions onto the MIEs and KEs in networks of AOPs including those found in the AOPWiki (https://aopwiki.org/). In the context of in vitro HTTr profiling, biomarker predictions could be used in testing programs as 'Tier 0' assays defined as those assays carried out before ToxCast Tier 1 screening or before Tier 1 screening in short-term animal tests (Figure 3). The AOs predicted from the integrated responses could then be confirmed by medium-throughput organotypic tissue assays or model organisms such as zebrafish. Efforts are currently underway in the laboratory to build, test, and use for screening high-value human biomarkers that can be developed from existing microarray data [13]. Given the accuracy of biomarkers developed using a combination of expression profiling after chemical exposure and gene modulation (Table 1), it is worthwhile to consider comprehensive strategies for building a large number of biomarkers that can predict MIEs/KEs that may include integrating Crispr-Cas9 knockout of MIE/KE factors in human cell lines (similar to wild-type vs. knockout mouse studies [17–20,22].

A central challenge in regulatory toxicology is comprehensive measurement of the MIEs and KEs in AOPs within prospective short-term studies as a means to predict potential health hazards, inform risk assessment, and eventually replace or eliminate the need for rodent bioassays including 2-year cancer bioassays [35]. A growing body of evidence in the literature demonstrates that environmentally relevant chemicals and human pharmaceuticals may cause liver cancer in rodents through a limited number of known AOPs [35]. We hypothesized that measurement of MIEs/KEs in the most frequently observed liver cancer AOPs in short-term assays using gene expression biomarkers will allow early identification of chemicals and their associated doses that are likely to be tumorigenic in the liver in two-year bioassays [36].

We tested this hypothesis using the rat in vivo TG-GATES study which includes data from rats exposed to ~130 chemicals at three doses and 8 time points up to 28 days [37]. This rich data set includes not only full-genome assessment of gene expression in the liver but also typical measures of animal testing. There were 390 chemical–dose pairs for 77 chemicals scored positive or negative for liver tumorigenicity, enabling the association of short-term measures with liver cancer. Gene expression biomarkers were built using microarray comparisons from the livers of rats treated with prototypical activators of the responses from rat liver microarray data for the most common MIEs (genotoxicity, aryl hydrocarbon receptor, CAR, estrogen receptor [ER], PPARα). Additional gene sets were used to assess cell proliferation or activation of the oxidant-induced TF Nrf2 (to predict oxidative stress). Phenotypic measures were used to assess the cytotoxicity MIE using alanine transaminase and aspartate transaminase levels; liver-to-body weight ratios were used to predict cell proliferation and cell hypertrophy. Assessment of these MIEs/KEs overlapped with most of the liver cancer MOA categories [35]. The ability of the biomarkers to accurately predict MIE activation ranged from 91% to 98% (Table 1). The chemical–dose pairs were ranked using the Toxicological Priority Index (ToxPi) based on their ability to activate MIEs/KEs. Chemicals administered at tumorigenic doses clearly gave the highest ranked scores; the method had a predictive accuracy of ~96%. At least for the 77 chemicals examined, assessment of the 6 AOPs appeared to identify all of the liver tumorigens.

Although the approach of using gene expression biomarkers to predict MIEs/KEs is promising, there are a number of limitations. First, there is uncertainty as to the domain of applicability for their use. While our group has had success in using at least one of the biomarkers made using mouse liver microarray data to predict responses in rat liver [38], it is very likely that the biomarkers will not be useful in predicting responses in other tissues. For example, the ER biomarker constructed in ER-positive MCF-7 cells did not appear useful to predict responses in other ER-positive human cell lines [14]. Second, the ability to build and test biomarkers will be limited by the number of studies examining transcriptional responses of reference compounds with well-known set of targets. Finally, there is some uncertainty as to how well biomarkers constructed using microarray data will be able to predict responses using RNA-Seq or targeted-Seq data. However, we have recently shown that the TGx-DDI biomarker that identifies DNA damaging agents and was built using microarray data can accurately predict responses using gene expression profiles derived from other microarray platforms, reverse transcriptase-polymerase chain reaction, or nCounter data [39].

4. Conclusions

Safety assessment of chemicals will continue to benefit from microarray-based approaches to understand MOA and human relevance. We have developed procedures to identify chemical MOA using gene expression biomarkers. The biomarker genes were identified using a

weight-of-evidence approach that incorporates knowledge of TF-dependent gene regulation. Using the Running Fisher test, the biomarkers were very accurate at predicting TF-dependent effects. Used in conjunction with common measures of animal studies, a set of rat liver biomarkers that predict MIEs/KEs were accurate at predicting liver cancer from 3 to 28d exposures. These results demonstrate that the methods can be specifically used to predict rat liver cancer from short-term exposures. The general approach could be applied to predicting other AOs if the associated MIEs/KEs are known and can be quantitated. Our approaches hold promise in future HTTr studies in which a large battery of characterized biomarkers can be used to predict chemical-induced effects on TFs, pathways, and phenotypes. Our approach to identify MOA and predict cancer has the potential of derisking compounds in early development pipelines by identification of doses that would not cause cancer and identification of MOAs associated with pathologies deemed irrelevant to humans [40].

Acknowledgements

This study was carried out as part of the EPA High-throughput Testing project within the chemical safety for sustainability (CSS) program. The information in this document has been funded by the U.S. Environmental Protection Agency. The views expressed in this article are those of the author and do not necessarily reflect the statements, opinions, views, conclusions, or policies of the United States EPA. The author declares he has no actual or potential competing financial interests. The author thanks Drs. Sheau-Fung Thai and Urmila Kodavanti for critical review of the manuscript and Mr. Chuck Gaul for assistance in creating the figures.

Conflict of interest statement

The author has no conflicts of interest.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Yauk C: **The value of genomics data for regulatory decision making – a summary of case studies from Health Canada**. *CurrOpin Toxicol* 2019 [This issue].

2. Waters MD, Jackson M, Lea I: **Characterizing and predicting carcinogenicity and mode of action using conventional and toxicogenomics methods**. *Mutat Res* 2010, **705**: 184–200.

3. Thomas RS, *et al.*: **A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure**. *Toxicol Sci* 2007, **98**:240–248.

4. * Phillips JR, *et al.*: **BMDExpress 2: enhanced transcriptomic dose-response analysis workflow**. *Bioinformatics* 2019, **35**: 1780–1782.
This paper describes an updated, greatly improved version of BMDExpress that has been widely used for dose–response modeling of transcriptomic data.

5. Merrick BA: **Next generation sequencing data for use in risk assessment**. *Curr Opin Toxicol* 2019, **18**:18–26.

6. Bush EC, *et al.*: **PLATE-Seq for genome-wide regulatory network analysis of high-throughput screens**. *Nat Commun* 2017, **8**:105.

7. ** Ye C, *et al.*: **DRUG-seq for miniaturized high-throughput transcriptome profiling in drug discovery**. *Nat Commun* 2018, **9**:4307.
The paper describes a high-throughput method for gene expression analysis that significantly reduces the cost/sample and simplifies the procedures involved.

8. * Yeakley JM, *et al.*: **A trichostatin A expression signature identified by TempO-Seq targeted whole transcriptome profiling**. *PLoS One* 2017, **12**:e0178302.
Describes the targeted RNA-Seq method for analysis of the full genome or the "1500+" genes.

9. Shah I, Harrill J, Setzer RW, Haggard D, Karmaus A, Martin MT, Thomas RS: **Predicting chemical mechanisms of action using high-throughput transcriptomic data**. *Toxicologist* 2018, **162**: 1689.

10. Schena M, *et al.*: **Quantitative monitoring of gene expression patterns with a complementary DNA microarray**. *Science* 1995, **270**:467–470.

11. **Biomarkers and surrogate endpoints: preferred definitions and conceptual framework**. *Clin Pharmacol Ther* 2001, **69**: 89–95.

12. Moreau Y, Tranchevent LC: **Computational tools for prioritizing candidate genes: boosting disease gene discovery**. *Nat Rev Genet* 2012, **13**:523–536.

13. Corton JC: *Identification of potential chemical carcinogens in compendia of gene expression profiles*. The Toxicologist; 2018: 1690.

14. Ryan N, *et al.*: **Moving toward integrating gene expression profiling into high-throughput testing: a gene expression biomarker accurately predicts estrogen receptor alpha modulation in a microarray compendium**. *Toxicol Sci* 2016, **151**:88–103.

15. Rooney JP, *et al.*: **Identification of androgen receptor modulators in a prostate cancer cell line microarray compendium**. *Toxicol Sci* 2018.

16. Corton JC, Kleinstreuer NC, Judson RS: **Identification of potential endocrine disrupting chemicals using gene expression biomarkers**. *Toxicol Appl Pharmacol* 2019, **380**: 114683.

17. * Oshida K, *et al.*: **Identification of chemical modulators of the constitutive activated receptor (CAR) in a gene expression compendium**. *Nucl Recept Signal* 2015, **13**. e002.
Describes the use of microarray profiles from CAR-null mice to identify biomarker genes predictive of CAR activation.

18. Oshida K, *et al.*: **Screening a mouse liver gene expression compendium identifies modulators of the aryl hydrocarbon receptor (AhR)**. *Toxicology* 2015, **336**:99–112.

19. Oshida K, *et al.*: **Identification of modulators of the nuclear receptor peroxisome proliferator-activated receptor alpha (PPARalpha) in a mouse liver gene expression compendium**. *PLoS One* 2015, **10**. e0112655.

20. Oshida K, *et al.*: **Disruption of STAT5b-regulated sexual dimorphism of the liver transcriptome by diverse factors is a common event**. *PLoS One* 2016, **11**. e0148308.

21. Oshida K, Waxman DJ, Corton JC: **Chemical and hormonal effects on STAT5b-dependent sexual dimorphism of the liver transcriptome**. *PLoS One* 2016, **11**. e0150284.

22. Rooney J, *et al.*: **Activation of Nrf2 in the liver is associated with stress resistance mediated by suppression of the growth hormone-regulated STAT5b transcription factor**. *PLoS One* 2018, **13**. e0200004.

23. Kupershmidt I, *et al.*: **Ontology-based meta-analysis of global collections of high-throughput public data**. *PLoS One* 2010, **5**.

24. Lamb J: **The Connectivity Map: a new tool for biomedical research**. *Nat Rev Canc* 2007, **7**:54–60.

25. Lamb J, *et al.*: **The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease**. *Science* 2006, **313**:1929–1935.

26. * Corton JC, Williams A, Yauk CL: **Using a gene expression biomarker to identify DNA damage-inducing agents in microarray profiles**. *Environ Mol Mutagen* 2018, **59**(9):772–784.
Paper shows that the TGx-DDI biomarker can be used with the Running Fisher test to accurately predict DNA damage.

27. Peffer RC, *et al.*: **Sedaxane-use of nuclear receptor transactivation assays, toxicogenomics, and toxicokinetics as part of a mode of action framework for rodent liver tumors**. *Toxicol Sci* 2018, **162**:582–598.

28. Rooney JP, *et al.*: **From the cover: genomic effects of androstenedione and sex-specific liver cancer susceptibility in mice**. *Toxicol Sci* 2017, **160**:15–29.

29. Rooney JP, Oshida K, Kumar R, Baldwin WS, Corton JC: **Chemical activation of the constitutive androstane receptor leads to activation of oxidant-induced Nrf2**. *Toxicol Sci* 2019, **167**(1):172–189.

30. Boobis AR, *et al.*: **IPCS framework for analyzing the relevance of a cancer mode of action for humans**. *Crit Rev Toxicol* 2006, **36**:781–792.

31. Ankley GT, *et al.*: **Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment**. *Environ Toxicol Chem* 2010, **29**:730–741.

32. Edwards SW, *et al.*: **Adverse outcome pathways-organizing toxicological information to improve decision making**. *J Pharmacol Exp Ther* 2016, **356**:170–181.

33. Vinken M: **Application of adverse outcome pathway framework for regulatory decision making**. *Curr Opin Toxicol* 2019 [This issue].

34. Cohen SM: **Evaluation of possible carcinogenic risk to humans based on liver tumors in rodent assays: the two-year bioassay is no longer necessary**. *Toxicol Pathol* 2010, **38**:487–501.

35. ** Rooney J, *et al.*: **Adverse outcome pathway-driven identification of rat liver tumorigens in short-term assays**. *Toxicol Appl Pharmacol* 2018, **356**:99–113.
First example of using gene expression biomarkers for molecular initiating events to predict cancer.

36. Igarashi Y, *et al.*: **Open TG-GATEs a large-scale toxicogenomics database**. *Nucleic Acids Res* 2015, **43**:D921–D927.

37. Slikker Jr W, *et al.*: **Emerging technologies for food and drug safety**. *Regul Toxicol Pharmacol* 2018, **98**:115–128.

38. Rooney J, Chorley B, Corton JC: **A gene expression biomarker identifies factors that modulate sterol regulatory element binding protein**. *Comput Toxicol* 2019. In press.

39. Corton JC, Williams A, Yauk CL: **Using a gene expression biomarker to identify DNA damage-inducing agents in microarray profiles**. *Environ Mol Mutagen* 2018, **59**(9):772–784.

40. Rooney J, Hill 3rd T, Qin C, Sistare FD, Corton JC: **Adverse outcome pathway-driven identification of rat liver tumorigens in short- term assays**. *Toxicol Appl Pharmacol* 2018, **356**:99–113.

41. Corton JC: **Frequent modulation of the sterol regulatory element binding protein (SREBP) by chemical exposure in the liver of rats**. *Comput Toxicol* 2019, **10**:113–129.