Application of Toxicogenomics in Next-Generation Risk Assessment

Andy Nong, Les Recio, Michael Black, Jeremy Fitzpatrick, Alina Efremenko, Kevin Bronson, Mel Andersen. *ScitoVation.*

Transcriptomic data has proven highly informative in toxicology for determining cellular modes of action (MOA) and determining points of departure (POD) for hazard assessment and comparative potency of congeners. Community standards for best practices in analyzing transcriptomic data have seen significant convergence on principles for statistical analyses, use of benchmark dose (BMD) model fitting, and POD derivation. Analysis workflows for transcriptomic data have matured significantly in the past few years, and these activities can be standardized for most applications in toxicology. By automating much of the analytical processing, efficiency and throughput of transcriptomic data analysis increases significantly, and cost decreases as well. Automation in a computational pipeline also assures that community best practices are applied uniformly to data, ensuring that comparison across compounds and experiments are unbiased in terms of the quality control assessments, statistical analyses, and data normalization.

With increasing interest in using transcriptomic data for comparing the MOA and potency of chemical congeners or as a basis for read-across in cellular response to exposure, use of standard transcriptomic analyses reporting formats and relational databases for comparative analyses becomes essential. While much raw transcriptomic data is publicly available (in databases such as the NIH NCBI Gene Expression Omnibus), the analyzed results and their interpretation are not. This makes comparison with public or legacy data cumbersome as often analyses need to be repeated from raw data for comparing with any new transcriptomic data. Here, we use open-source MySQL tools to capture both raw and final processed data used for interpretation of MOA as well as POD values derived from BMD analysis and ontology pathway enrichment. Such a database is also capable of retaining all the meta-data associated with each data set such that the exact manipulations applied to the raw data can be recapitulated in detail.

ScitoVation has developed visualization tools (MOAviz) and a comparative statistical metric (Modified Jaccard Index) useful for deriving cellular MOAs and comparing similarity of MOA between compounds. MOAviz uses statistically significant differentially expressed genes (DEGs) for ontology enrichment analysis and displays the results in an interactive graphical format that retains the ontologies hierarchical parent to child relationships of cellular pathways (ontologies progress from highly generalized categories, such as metabolism, to ever more refined categories such as lipid metabolism, to final discrete biochemical pathways, such as cholesterol biosynthesis). MOAviz can thus be linked computationally to data in a transcriptomic database to allow for direct interactive sessions examining MOA across dose or time for a compound or for comparison of multiple chemicals.

Our efforts in 2023 will advance on the development of a stakeholder-accessible analyses pipeline from raw data input to standardized output report format, including interactive MOAviz functionality, implementation of the OECD Toxicogenomics Reporting Framework, and a novel platform for evaluating biologically based read-across.

ScitoVation is developing a database schema based on the open-source MySQL framework defined for microarray and RNA-Seq data. This database instance is located in an Amazon Web Services virtual server instance to leverage cloud computing resources for data integrity (i.e., recoverability) and the ability to scale computing resources for computationally intensive processes such as BMD model fitting. In 2022, a standardized pipeline was developed that uses a starting input for microarray data and runs through the computing process to generate BMD values of the significant genes. A basic version of MOAviz will be released in late 2022 and a video describing its application will be presented in 2023.

In 2023, ScitoVation will finalize the development of the transcriptomic pipeline with the integration of MOAviz and standardized reporting tool that is based on the OECD transcript reporting framework (TRF). Minor

refinements to the pipeline scripts and quality of life for user experience will progress in collaboration with external panel of expert users. This interaction will help set up a launch menu system to simplify the execution of a pipeline instance with new data. The pipeline will include a final report of the rankings of gene ontology and POD pathways, in addition to the TRF content. MOAviz interactive bubble maps provides the users an interactive experience that can be modified on demand through different menu commands.

- ScitoVation will implement a system to automatically update Reactome ontology files for enrichment analyses with versioning captured in the database to track the release used in each analysis (necessary to account for any long-term changes observed in MOAviz due to refinements in the ontology and its category elements). The database will be populated with legacy microarray datasets from the Hamner Institutes in vivo experiments conducted for the ACC LRI from 2009 to 2015. This data will be the beginning of building out a comparative MOA/POD system. New databases from various repository such as the Open Toxicogenomics Project-Genomics Assisted Toxicity Evaluation System (TG-GATEs) or the Drug Gene Interaction Database (DGID) will also be added to increase read-across capability.

- ScitoVation will develop a series of case studies to demonstrate the application of the transcriptomic pipeline with the latest research of chemical assessments. These cases will apply and study publicly available RNA-Seq data to explore for the transcriptomic dose-response relevancy of chemical classes.
    - For example, the TGx-DDI transcriptomic biomarker can detect DNA damage-inducing (DDI) chemicals in human cells after exposure. [1] The TGx-DDI biomarker was identified using a dataset comprised of transcriptomic profiles of lymphoblast cells exposed to several chemical agents with well-characterized modes of action for a wide range of DDI mechanisms. The biomarker is specific to the identification of DNA damage. As a case study, ScitoVation will analyze the transcriptomic profiles of TGx-DDI using MOAviz and identify any pathways matching with MJI. We will then validate and compare the mode of action of DDI against the transcriptomic assessment obtained using MOAviz.
    - Different chain length alkylated perfluorinated chemicals were studied in cell cultured liver spheroids. [2] A dataset of transcriptomic profiles was generated to prioritize potency of the chemicals based on lowest transcriptomic point of departure. ScitoVation will investigate the potency of the different alkyl chain lengths using a transcriptomic pathway observed with MOAviz. We will then compare the transcriptomic assessment between this approach against other published methods that have evaluated liver toxicogenomic pathway such as the mode of action for nonalcoholic fatty liver disease. [3]

References

[1] Li HH., et al. (2020). TGx-DDI, a Transcriptomic Biomarker for Genotoxicity Hazard Assessment of Pharmaceuticals and Environmental Chemicals. Frontiers in Big Data.
[2] Reardon A., et al. (2021). Potency Ranking of Per- and Polyfluoroalkyl Substances Using High-Throughput Transcriptomic Analysis of Human Liver Spheroids. Toxicological Sciences. 184(1):154-169.
[3] Subdhi A., et al. (2021). Distinct Hepatic Gene-Expression Patterns of NAFLD in Patients with Obesity. Hepatology Communications. 6(1): 77-89

Implications: This work is designed to establish the utility of gene expression profiling using in vitro models for predicting adverse outcomes and hazard. By identifying a suite of genes with altered expression after exposure to a compound, transcriptomic studies can yield essential information on potency and mode of action. These technologies are being developed as part of the broader multi-tier safety assessment paradigm being refined at ScitoVation.

Current project start and end dates: January 2022 – December 2023

Abstract revision date: February 2023