Carolina Center for Computational Toxicology

Funded by U.S. EPA Cooperative Agreement STAR RD 833825
Progress Report for Project Period 04/01/10 - 03/31/11
Report Date: 04/21/11
Ivan Rusyn, Principal Investigator

Project 1:
Predictive Modeling of Chemical-Perturbed Regulatory Networks in Systems Toxicology

Preliminary Data and Work Progress

A substantial amount of progress has been made in the aims of this project, with a significant amount of effort in this past period being focused on the development and extension of our mechanistic model for metabolic function that includes the blood, liver, muscle, and adipose tissues as primary compartments. The revision of the manuscript describing the core model is in review at *PLoS Computational Biology*. We have also initiated the extension of this model into more toxicity-relevant areas by incorporating a model for glutathione transport, synthesis and breakdown. In collaboration with Michael Gamcsik and Jeff MacDonald in the Department of Biomedical Engineering at UNC/NCSU, we are now acquiring high-resolution metabolomics data of glutathione dynamics in response to bromobimane exposure. These measurements are being acquired within a NMR-compatible bioreactor housing human hepatocytes and will provide additional model validation as well as provide insight into a potential feedback mechanism thought to exist within the glutathione pathway. The work with Project 2 on trying to identify genes of high-relevance within eQTL data is ongoing and has now been split into two projects/papers: one focused on methodology validation and the second on application to liver eQTL data derived from a mouse population. The methodology development for the identification of “toxicity modules” is also nearly completed, with a statistical measure having been developed to help assess the “significance” of observing a particular module within a given data set. Its use as part of a predictive methodology is currently under investigation.

Results to Date

The goals of our project are focused on the development of tools and methodologies for understanding chemical-induced perturbation in the context of biological networks.

Our mechanistic modeling efforts have progressed significantly, with the manuscript describing our core 4-compartment whole-body model for glycogen regulation under review at *PLoS Computational Biology*. The underlying reactions are now fully described using Michaelis-Menton kinetics. Simulation results have been validated against existing rodent data at the cellular and whole-liver levels. In addition, an extensive sensitivity analysis has been performed to help establish parameter sensitivity. Finally, the simulation code for the model is being made publicly available as an SBML (Systems Biology Markup Language) representation that can also be utilized by other investigators interested in using the simulation.

We have begun adding additional pathways of relevance to chemical toxicity. We have initiated this effort with the incorporation of a model of glutathione transport, synthesis and breakdown. We are validating and investigating novel aspects into the glutathione model through collaboration with Michael Gamcsik and Jeff MacDonald in the Department of Biomedical Engineering at UNC/NCSU. Using a NMR-compatible bioreactor housing human hepatocytes, we are acquiring high-resolution metabolomics data (NMR time points as frequent as 1/min, with total duration out to 24 hours) of glutathione dynamics in response to bromobimane exposure. Aspects relevant to the model include the identification and modeling of a potential feedback mechanism thought to exist internal to the glutathione pathway. This

pathway has been integrated into the liver metabolism model previously developed and is thus a strong step towards the further development of this system as a model of broad utility to the toxicology community. Significant additional work based on this model is planned as is described in the “Activities for Subsequent Reporting Period” section below.

In collaboration with Ingenuity Systems Inc., we have investigated the effect of fasting on hepatic gene expression and the potentially confounding effects that may result during the interpretation of toxicogenomics data. The goal of this work was to identify a set of fasting-response transcripts that could serve as a flag for potential fasting effects. We were able to identify 11 candidate flag genes including *Ech1*, *Dci*, *Acot2, Acaa1, Scd1, Pklr, Pygl, Csad, Fasn, Acly*, and *Dak* that can act as a signal for possible fasting side effects. This gene set is being incorporated into the Ingenuity IPA system as an aid to toxicogenomic data interpretation. This work was presented at the 2011 Society of Toxicology Annual Meeting.

The data mining methodology based on frequent itemset approaches was developed during the previous reporting period and has been applied to the analysis of ToxCast data. As a brief review, this approach is related to biclustering approaches, but leverages conditional constraints during the search process, provides the capability for the complete enumeration of relevant assay combinations, and now also incorporates uncertainty into the assays discovered. The goal of this approach is to enable the determination of sets of assays having the greatest functional relevance (and potential predictive power) for one or more particular disease endpoints.

This approach has now been significantly refined, with the ability to incorporate uncertainty or “fuzziness” into the results also more fully developed. An important addition to this approach is the development of a measure of “significance” for these modules. Currently based on the phi coefficient, this statistic gives an idea of the amount of consistency that exists within the data supporting a specific combination of chemicals, assays and disease endpoints. We are further investigating other statistical approaches, including resampling-based methods, for establishing the significance of modules and that can then be used in prioritization. Overall, this approach appears to have significant utility as it provides a useful means of dealing with the inherent noise and uncertainty associated with high-throughput screening of biological systems.

The previous period established a diffusion-based approach for use in gene prioritization by combining functional gene networks with eQTL data. This approach has been refined so as to incorporate “negative” information regarding relationships between genes, as well as a novel voting scheme that helps to identify the most probable associations between a set of regulated genes and a set of potential regulator genes. This methodology is currently undergoing validation using established “gold-standard” data taken from yeast genetics studies. With Project 2, we have begun to apply this approach to the analysis of liver eQTL data in BXD recombinant inbred mice.

Given a chemical, it would often be desirable to predict the points within a cell where it could potentially interact and/or cause alteration of network behavior. Based on earlier protein interaction prediction approaches, we have developed a methodology for the prediction of chemical-protein interactions. This work has initially utilized chemical-protein interactions from the STITCH database as a source of training and test data and has been able to achieve ROC scores greater than 0.8 when applied to interactions known to occur in Humans, as well as mouse, rat and yeast species. We are currently assessing the ability of this prediction method to aid in the interpretation of high-throughput data by initially investigating predictions in the context of ToxCast data. We note that this area of work is highly complementary to both the network and mechanistic modeling efforts of this project in helping to provide more explicit relationships between chemicals and the likely points within a biological system where they are potentially able to initiate some effect. Such approaches will also be of particular value in the assessment and interpretation of high-throughput chemical screens.

Activities for Subsequent Reporting Period

Our modeling efforts have now achieved a core multi-compartment mechanistic description of basic liver-focused metabolism. The incorporation of basic glutathione dynamics into this model represents only our first step in further extending our modeling efforts to areas of direct relevance to chemical toxicity. We will continue to incorporate toxicity relevant pathways/mechanisms as we move forward and as appropriate data becomes available.

We also see an opportunity for further developing the model towards current EPA areas of focus. In particular, recent work by the EPA has introduced the idea of estimating the toxicity of chemicals by the degree to which they alter relevant pathways in a dose-dependent manner (e.g., Judson et al, 2011). For this approach to be successful, the ability to estimate *in vivo* dose from *in vitro* concentration is essential. An important aspect in this regard is the need for adequate models capable of estimating such doses. Two major classes of models have been considered for these efforts: 1) physiologically based pharmacokinetic (PBPK) and 2) reverse dosimetry or reverse toxicokinetic (RTK) models. While PBPK models are a potential basis for these estimates, they require a significant amount of parametric detail. In addition, where these models provide the most detail is generally not a good match for the types of data being collected from high-throughput/content screens. As a result, 1-compartment RTK models have been used initially to test the feasibility of these dose estimation efforts. The extreme simplicity of these models introduces their own complications with regard to accuracy as well as overall applicability.

We thus see the opportunity to use this model as a more appropriate “hybrid” between PBPK models and simple 1-compartment models. In particular, our model incorporates the majority of toxicity-relevant organ systems, and can be further extended. In addition, it can be composed of both high-level generalizations of physiological processes (e.g. generalized “clearance rates”) that are well established in PBPK models as well as detailed sub-models where appropriate. Detailed sub-models can be similar to that of glutathione or glycogen metabolism already completed and can be readily defined when appropriate data is available. This is of particular importance as high-throughput screens used in ToxCast (or Tox21) generally provide measurements at this level of molecular detail (i.e. specific genes/proteins/metabolites) and not measurements typically found in PBPK or RTK models. Thus we have the potential to link these assays more directly with how they are used in toxicity assessment and will pursue this in the subsequent reporting period.

We will continue with expanding the chemical-protein interaction work as there is significant potential to use this as a method for linking high-throughput data with particular biological pathways/networks. This can also be used for identifying points of chemical perturbation within particular processes of our developing mechanistic models.

With project 2 we will continue to explore the network perturbation and modeling studies to explore the relationships between genetic variation and toxicity effects on biological pathways. Furthermore, we will continue in investigating a resampling based approach to establish the “significance” of toxicity modules for use in prioritization of both chemicals as well as identification of assays having potentially strong predictive value.

Publications Arising From this Project in Year 3

Papers

1. Xu, K., Morgan, K.T., Todd, A., Elston, T.C. and Gomez, S.M. (2011) A whole-body model for glycogen regulation reveals a critical role for substrate cycling in maintaining blood glucose homeostasis. (Revision in review at PLoS Computational Biology)
2. Gulyani, A., E. Vitriol, R. Allen, J. Wu, D. Gremyachinskiy, S. Lewis, B. Dewar, L. Graves, B. Kay, B. Kuhlman, T.C. Elston and K. Hahn. (2011) A Src family biosensor based on an engineered scaffold enables sensitive quatitation of Src dynamics at the leading edge. (Under review for Nature Chemical Biology).

3. Garcia, G.J.M., Picher, M., Zuo, P., Okada, S.F., Lazarowski, E.R., Button, B., Boucher, R.C., and Elston, T.C. (2011) Computational model for the regulation of extracellular ATP and adenosine in airway epithelia. In: *Purinergic Regulation of Respiratory Diseases* (M. Picher and R.C. Boucher, eds.), Chapter 3, Springer Publishing.
4. Doolittle, J.M., and Gomez, S.M. (2011) Mapping protein interactions between Dengue virus and its human and insect hosts. PLoS Neglected Tropical Diseases, 5(2):e954.
5. Doolittle, J.M. and Gomez, S.M. (2010) Structural similarity-based predictions of protein interactions between HIV-1 and Homo sapiens. Virology J, 7:82 (*from in press to published*).
6. Strychalski, W., Adalsteinsson, D., and Elston, T.C. (2010) Simulating biochemical networks in complex moving geometries. SIAM J Sci Comput 32(5): 3039–3070 (*from in press to published*).

Posters/Abstracts

1. Zhu H, Liying Z, Staab J, Sedykh A, Tang H, Gomez S, Rusyn I and Tropsha A. Incorporation of ToxCast *in vitro* assay data and relevant toxicity pathway information improves the external prediction accuracy of Quantitative Structure-Activity Relationship (QSAR) Models of Chemical Hepatotoxicity. Society of Toxicology Annual Meeting, Washington DC, March 6-10, 2011.
2. Xu K, Kirchberg C, Gomez SM and Morgan K. Development of a transcript profile for fasting as a confounding variable in toxicogenomic studies. Society of Toxicology Annual Meeting, Washington DC, March 6-10, 2011.