Carolina Center for Computational Toxicology

Funded by U.S. EPA Cooperative Agreement STAR RD 833825
Progress Report for Project Period 04/01/11 - 03/31/12
Report Date: 02/19/12
Ivan Rusyn, Principal Investigator

The Carolina Center for Computational Toxicology (CCCT) is comprised of three research projects and an administrative core. The major aims and objectives of the CCCT have not changed from the original application. The content of this progress report is organized according to U.S. EPA guidelines and it summarizes significant activities and accomplishments of all four components of the CCCT.

Preliminary Data and Work Progress

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

We made substantial gains in the aims of this Project, with progress in both the mechanistic- and network-based models. Our mechanistic model describing fundamental metabolic function in the liver as well as muscle and adipose compartments was published in *PLoS Computational Biology*. We have continued this work into the modeling of the glutathione pathway and have been comparing it to experimental results from human cells. The first stage of this model is complete and will be submitted with the experimental results in the first half of 2012. Based on discrepancies between model predictions and experimental results, refinements to the model are currently being investigated. The methodology development for the identification of "toxicity modules" has been submitted and extensions of the associated statistical methodology for defining module statistical significance are nearing completion. Our segNET method has been validated and is submission, with its next application to liver eQTL data derived from a mouse population (Project 2). Finally, we have begun the establishment of a novel live-cell computational image analysis system for studying the effects of chemical exposure on the dynamic behaviors of individual cells and cell populations. While at an early stage, this system has significant potential in helping to quantitatively understand the relationship between chemical exposure and the resulting downstream changes in network dynamics and cellular behavior.

Results to Date

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

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.

Mechanistic Modeling: Our mechanistic modeling efforts most recently resulted in the publication of our core 4-compartment whole-body model of liver metabolic function in *PLoS Computational Biology*. Our efforts to add pathways of relevance to chemical toxicity have focused on the incorporation of glutathione transport, synthesis and breakdown into the model. In collaboration with Michael Gamcsik and Jeff MacDonald in the Department of Biomedical

Engineering at UNC/NCSU, we have acquired multiple sets of high-resolution metabolomics data of glutathione dynamics in response to bromobimane exposure and performed model fitting and analysis. Using a NMR-compatible bioreactor housing human hepatocytes, we have acquired high-resolution metabolomics data (NMR time points as frequent as 1/min, with total duration out to 24 hours). Initial results had indicated very good similarity between predictions and results. However, there are observable differences between model predictions and experimental results that suggest the need to incorporate a feedback pathway to achieve complete reconciliation of the results. Interestingly, we have incorporated this feedback into the model and it does not lead to the expected improvements. This suggests that either 1) additional but known aspects of the pathway need to be incorporated into the model to get the proper results, 2) our biological knowledge of this pathway is incomplete and as yet unrecognized regulatory interactions need to be identified and then incorporated into the model, or 3) that both 1) and 2) are in play. We are planning the results of the first model as part of a publication focused on the novel experimental system developed by our collaborators. Follow-on work will focus on improvements to the model that better explain the observed behavior in glutathione dynamics.

Network Modeling and Data Mining: We have been exploring the development of methods for understanding how perturbations, chemical, genetic or otherwise, alter network properties and function. Our methodology for identifying relevant but "fuzzy" patterns from toxicity data such as that provided by ToxCast has matured and been applied in other data-mining contexts. A major component of this work was the incorporation of a measure of statistical significance (the phi coefficient) that aids in the identification of subsets of data (subsets of assays, chemicals and endpoints) that "stand out" from the background of very noisy data. While our submitted work focuses on the use of the phi coefficient for determining significance, we have since extended this work with a bootstrapping approach that allows any parameter to be used as a means of testing significance. This bootstrapping approach is computationally intensive and, when used with frequent item set mining, cannot necessarily be used for data sets of the size of ToxCast or Tox21. However, we have developed modifications that allow us to mine very large data sets while providing a measure of significance. This improvement in the size of the data set that can be searched is done by constraining the search space with user-defined conditions. This work is being written up and will be submitted in the first half of 2012. We also note that this bootstrap approach may potentially be able to help improve the currently used ToxPI metric being supported by the EPA.

The previous period similarly established a diffusion-based approach, segNET, for use in gene prioritization by combining functional gene networks with eQTL data. This approach has now been validated with "gold-standard" yeast genetic association data. We compared it to the work of Sun et al. 2007, and were able to significantly improve the interpretation of the relationships between the genes in the transband(s) with their respective QTL. In effect, this approach allows the dissection of QTL in to 1 or more functional modules that are more easily interpretable and aid in the interpretation of the potential functional linkages between genes. This work has been submitted for publication. We anticipate that it will prove to be a valuable tool in interpreting complex associations and relationships commonly found in toxicology studies and are now revisiting our initial analysis of liver eQTL in BXD recombinant inbred mice begun with Project 2.

We have similarly been exploring the use of the segNET approach as well as other methods to better understand and predict how perturbations affect network behavior. While in the context of cancer drug treatment, we have been investigating the identification of "network signatures" that can be used to help identify the state of a cell as well as perhaps predicting how it will respond to chemical (in this case, chemotherapeutic) exposure. Initial work on defining

such a signature and how the kinase network responds to drugs has been accepted in the journal *Cell*.

A web version of our tool for spectra alignment (web-PCANS) and downstream STOCSY analysis has been undergoing testing by outside collaborators and their feedback is being incorporated. We will be submitting a manuscript describing this tool as an "Application Note" to the journal *Bioinformatics* in the first quarter of 2012. We will be making the tool publicly available before submission of the paper.

Activities for Subsequent Reporting Period

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

We will continue our work on extending the mechanistic modeling to pathways of relevance to toxicity. First efforts are focused on finishing the additions to the glutathione model that will allow us to reconcile the observed discrepancies between model and experiment. Numerous other pathways and targets of modeling interest are possible and will be pursued based on their overall importance as well as on the availability of necessary detailed data.

As described earlier, our latest work on defining toxicity modules has led to the development of a bootstrap approach for determining significance. We think there is significant potential to apply to, or extend on the current ToxPi methodology for chemical ranking. This is an area of significant interest and we have already started to explore this application.

Our work on prediction of where chemicals interact with network biomolecules has been on hold while other research projects were completed. With these projects completed, we will be finishing this aspect of the project. This work has initially utilized chemical-protein interactions from the STITCH database as a source of training and test data and was 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. The next major piece of work is to compare these predictions with known chemical-protein interactions taken from outside our training data. This assessment should provide the information we need to evaluate under which conditions this approach provides the greatest predictive benefit.

We have recently been active in the development of computational methods for live-cell image analysis. We believe this work has significant promise for helping to quantify aspects of network dynamics and cellular behavior in a living system. We particularly see it as a potential opportunity to link how a cell and its underlying regulatory and signaling network actively change to chemical perturbation. The methods we have developed have been published in both PLoS ONE as well as in Cell (in press), primarily in the context of quantifying aspects of cell motility. Their application to understanding chemical toxicity should similarly provide improved insight as to modes of chemical action and their effect on network dynamics.

Publications Arising From this Project in Year 4

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

Papers

1. Choi, K., Huang, J., and Gomez, S.M. Network-based dissection of functional modules from genetic association data. (*submitted*)
2. Staab, J.S., Choi, K., and Gomez, S.M. Identification of "toxicity modules" in noisy high-throughput chemical testing data sets. (*submitted*)

3. Duncan, J.S., Whittle, M.C., Nkamura, K., Abell, A.N., Midland, A.A., Zawistowksi, J.S., Johnson, N.L., Granger, D.A., Jordan, N.V., Darr, D., Usary, J., Major, B., He, X., Hoadley, K., Sharpless, N.E., Perou, C.M., Gomez, S.M., Jin, J., Frye, S.V., Earp, H.S., Graves, L.M., and Johnson, G.L. (2012) Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple negative breast cancer. (*accepted in Cell*).
4. Xu, K., Morgan, K.T., 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. *PLoS Computational Biology*. 7:e1002272. (*from Revision Under Review to Published*)
5. Abell, A.N., Jordan, N.V., Huang, W., Prat, A., Midland, A.A., Johnson, N.L., Granger, D.A., Mieczkowski, P.A., Perou, C.M., Gomez, S.M., Li, L., and Johnson, G.L. (2011) MAP3K4/CBP-Regulated H2B Acetylation Controls Epithelial-Mesenchymal Transition in Trophoblast Stem Cells. *Cell Stem Cell*. 8:525-37.
6. Houser, J., Ford, E., Errede, B., and Elston T.C. Mathematical modeling reveals positive roles for negative regulators of transcription in the yeast mating response. In revision for M*olecular Systems Biology*.

Posters/Abstracts/Presentations

1. Gomez, S.M. Predictive modeling of chemical-perturbed regulatory networks in systems toxicology. Third Toxicogenomics Integrated with Environmental Sciences (TIES) international conference. UNC-Chapel Hill, September 15-16, 2011.