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 3:
Development of validated and predictive Quantitative Structure-Toxicity Relationship models that employ both chemical and biological descriptors of molecular structures and take into account genetic diversity between individuals

Preliminary Data and Work Progress

In the current year of the project we have continued to rely on the core strengths of our group that lie in developing robust QSAR modeling methodologies and rigorous models of important toxicity endpoints using data from the literature as well as available from ToxRefDB/ ToxCast™ projects. Specifically, novel models of Ames genotoxicity have been developed and validated. At the same time, we have made significant progress focusing on complex problems in computational chemical toxicology that deal with the incorporation of dose-response data available from short term assays as well as emerging genomic response data into our overall toxico-cheminformatics modeling platform. In the previous years we have established that combining chemical descriptors of toxicants with short-term assay results treated as special biological descriptors of the same chemicals improves the external predictive accuracy of QSAR models. In the current year, we went beyond the simple use of single biological descriptors per compound per assay (i.e., using binary biological descriptors such as “active” or “inactive” definitions) by embracing (when available) more comprehensive multiple biological descriptors of molecules generated from complete dose-response data generated in biological assays. We have refined and finalized special procedures for converting dose-response data into numerical descriptors suitable for QSAR modeling and developed enhanced models of rodent acute toxicity using both chemical and the new dose-response descriptors. In a separate study, we have explored the use of microarray gene expression data as yet another type of biological descriptors that can be combined with conventional chemical descriptors to predict hepatotoxicty of a large dataset (more than 125 entries) of drug molecules. The data was obtained from the Japanese Toxicogenomics Project (TGP: http://toxico.nibio.go.jp/datalist.html). Additional studies are in progress to extend the application of our models to emerging toxicogenomics datasets as well as ToxCast data focusing on rational selection of short-term assays in the context of underlying biological pathways.

Results to Date

In the context of conventional QSAR modeling of chemical toxicity, advanced best practices for the development of conventional QSAR models of chemical toxicity. We have complied and curated the largest, in the public domain, database of more than 7.000 molecules tested in Ames genotoxicity assay. To enable the models of the highest statistical rigor incorporating best practices in the field, we have formed international virtual collaboratory incorporating 14 QSAR modeling groups from eight countries. Each group relied on its own QSAR modeling approaches to develop mutagenicity models using the same modeling set, and we agreed to evaluate the realistic model performance using the same external validation set(s). We have established that consensus QSAR model integrating individual predictions from individual laboratories achieved the highest statistical significance and external predictive power

with the prediction accuracy exceeding 80%. This consensus model was applied for virtual screening of several chemical libraries of environmental agents and a number of compounds predicted to be active in the Ames test have been identified and prioritized for future experimental investigations. Thus, this study presents an example of a fruitful international collaboration between researchers that use different techniques and approaches but share general principles of QSAR model development and validation

We have also continued to advance our analysis of ToxCast Phase I data in collaboration with projects 1 and 2. Through previous studies both in our group and elsewhere we have recognized that global models of ToxCast data using all assays to predict all in vivo end points are unlikely to succeed and have begun to explore modeling approaches that increasingly rely on biological inferences. In a joint study with Profs. Gomez and Rusyn we have analyzed ToxCast in vitro assay data in the context of relevant toxicity pathway information and were able to achieve the improved QSAR models of reproductive toxicity. The in vitro data was obtained from ToxCast Phase I database, which currently lists the results of testing 320 substances in 615 bioassays. To simplify the modeling procedure, we have classified compounds into active or inactive and only used assays that had at least 10 active responses. Thus, 284 (out of 615) assays were selected and used in our study. Animal toxicity endpoints were obtained from ToxRefDB. We selected the following 3 (out of 19) *in vivo* reproductive toxicity endpoints (Multi-generation Rat, MGR) because these endpoints had the most balanced ratio of active/inactive agents: MGR_Rat_Kidney (Kidney microscopic and gross pathologies and weight changes); MGR_Rat_Liver (Liver microscopic and gross pathologies and weight changes); MGR_Rat_Viability (Viability Index). In the modeling workflow, all chemicals were first partitioned into two classes based on whether the results of one in vitro assay and one in vivo toxicity results agree (i.e., the compound is found either active or inactive in both types of assays), or disagree (the compound's classification in vitro versus in vivo disagree). Second, classification QSAR models for the two classes were developed using Random Forest (RF) and two types of Support Vector Machine (SVM) methods. The QSAR models then were used to assign compounds in an external dataset to one of the in vitro/in vivo correlation classes and then predict the associated in vivo toxicity based on a known in vitro response. The results indicate that neither conventional QSAR (chemical descriptors) models, nor ToxCast bioassay data are predictive of rodent reproductive toxicity when used individually. However, when ToxCast data was used as biological descriptors in the hybrid QSAR modeling workflow to predict reproductive toxicity in vivo, the overall prediction accuracy (from consensus prediction by all individual models) was higher than that based on chemical descriptors alone. Furthermore, the use of groups of ToxCast assays organized by their relevance to toxicity pathways showed that equally accurate reproductive toxicity models can be developed based on far fewer assays. This analysis suggests that as more mechanistically relevant bioassay data are generated, and a greater number of compounds are screened, computational toxicology tools could be used to (i) select most relevant HTS assays (cell lines and/or endpoints) and (ii) prioritize chemical agents for further *in vivo* testing.

In another ongoing project, in collaboration with Project 2, we have expanded our previous analysis and modeling of the qHTS screening data of toxic compounds screened at the NIH Chemical Genomics Center (NCGC). As discussed in Project 2 report (see above) population-based toxicity screening has been conducted against 81 cell lines to test a dataset comprised of 240 compounds (in the previous studies, only 13 cell lines were explored). These compounds were selected from the 1,408 substances of the NTP screening library. Each compound was tested in each cell line using cell viability and caspase assays. Duplicate chemicals were included for quality control (QC) purpose. Data analysis of QC data for these duplicate compounds revealed that the noise-processed qHTS data using procedures developed earlier and reported last year demonstrated higher reproducibility for the viability

assay (Pearson correlation coefficient r=0.86 vs 0.73), but not for the caspase assay (r=0.96 in both cases). This analysis confirms the reliability of newly generated qHTS data.

Using this data, we have applied our standard modeling workflow to develop preliminary models of rat acute toxicity in line with the previous studies using both conventional chemical descriptors as well as noise-treated qHTS data employed as biological descriptors. Interestingly, biological descriptors alone afforded acceptable QSAR models, which were as good as conventional QSAR models only (accuracy ~63% for both). We are in the process of analyzing our models in terms of outliers, both in chemistry space and in the biological descriptor space to understand the possible reasons for relatively low prediction accuracy. We are also investigating the relationships between chemical structure and cell-specificity of biological response with the goal of developing improved models using rationally selected cell lines. Finally, we plan to develop hybrid models utilizing both chemical and biological descriptors concurrently.

Finally, we would like to report on an ongoing exciting investigation where we have completed the first phase of our analysis. QSAR modeling and toxicogenomics are used independently as predictive tools in toxicology. In this study, we evaluated the predictive power of several statistical models for predicting drug hepatotoxicity in rats employing different descriptors of drug molecules, namely their chemical descriptors and toxicogenomic profiles. The records were taken from the Toxicogenomics Informatics Project rat liver microarray database containing information on of 127 drugs (http://toxico.nibio.go.jp/datalist.html). The target property was hepatotoxicity level following 28 days of treatment, established by histopathology and serum chemistry. First, we developed multiple conventional QSAR classification models using a comprehensive set of chemical descriptors (Dragon; Molecular Operating Environment, MOE; and simplex representation of molecular structure, SiRMS) and several classification methods (*k* nearest neighbor, support vector machines, random forests, and distance weighted discrimination). With chemical descriptors alone, external predictivity (Correct Classification Rate, CCR) from 5-fold external cross-validation was 61%. Next, the same classification methods were employed to build models using only toxicogenomic data (24h after dosing) treated as biological descriptors. The optimized models employed only 85 selected genomic descriptors and had CCR as high as 76%. Finally, hybrid models combining both chemical descriptors and transcripts were developed; their CCRs were between 60 and 77%. Although the accuracy of hybrid models did not exceed that of models based on toxicogenomic data alone, the use of both chemical and biological descriptors enriched the interpretation of the models. In addition to finding 85 transcripts that were predictive and highly relevant to the mechanisms of drug-induced liver injury, chemical structural alerts for hepatotoxicity were also identified. We conclude that concurrent exploration of chemical features and acute treatment-induced changes in transcript levels will both enrich the mechanistic understanding of sub-chronic liver injury and afford models capable of accurate prediction of hepatotoxicity from chemical structure and short-term assay results.

Activities for Subsequent Reporting Period

In project 3, we will continue to advance workflows for addressing the entire chemical structure – in vitro – in vivo data continuum. We have advanced and professed two major workflows for data analysis: (1) using relationships between in vitro and in vivo biological screening profiles classify compounds into concordant or discordant classes followed by class-specific QSAR modeling to distinguish toxic from non-toxic molecules, and (2) using hybrid chemical/biological descriptors including, when possible, dose-response data. We expect that Phase 2 ToxCast data becomes available in the next few months so that we will continue to employ our workflows for the analysis of this growing and unique data collection. We will especially focus on two aspects of data analysis and predictive modeling: (1) using dose-response data as rich descriptors of biological response to combine with chemical descriptors and enable model interpretation in terms of dose-dependent combination of chemical features

that determines chemical effects in vivo; and (2) using pathway mapping information on short term biological assays to improve in vivo toxicity prediction using functionally inferred biological descriptor selection. Our underlying hypothesis is that we are unlikely to succeed using global "kitchen sink" models to relate chemical structure and in vitro data to toxicity endpoints. We suggest that rigorous and interpretable models of chemical toxicity will emerge for classes of chemical compounds tested in selected, mechanistically relevant biological assays. We also plan to expand out computational toxicogenomics studies using gene expression profiles as additional biological descriptors of molecules. Finally we plan to enable all our models at our ChemBench web portal (chembench.mml.unc.edu), which will allow unrestricted public access to models and their use by researchers worldwide.

Publications Arising From this Project in Year 3

Papers

1. Sedykh A, Zhu H, Tang H, Zhang L, Richard A, Rusyn I, and Tropsha A. (2011). Use of in vitro HTS-Derived Concentration-Response Data as Biological Descriptors Improves the Accuracy of QSAR Models of in vivo Toxicity. Environ. Health Perspect. 119:364-370.
2. Sushko I, Novotarskyi S, Körner R, Pandey AK, Cherkasov A, Li J, Gramatica P, Hansen K, Schroeter T, Müller KR, Xi L, Liu H, Yao X, Öberg T, Hormozdiari F, Dao P, Sahinalp C, Todeschini R, Polishchuk P, Artemenko A, Kuz'min V, Martin TM, Young DM, Fourches D, Muratov E, Tropsha A, Baskin I, Horvath D, Marcou G, Muller C, Varnek A, Prokopenko VV, Tetko IV. (2010). Applicability domains for classification problems: benchmarking of distance to models for ames mutagenicity set. J Chem Inf Model. 50:2094-2111.
3. Low Y., Uehara T., Minowa Y., Yamada H., Ohno Y., Urushidani T., Sedykh A., Muratov E., Fourches D., Zhu H., Rusyn I., Tropsha A. Predicting Drug-induced Hepatotoxicity Using QSAR and Toxicogenomics Approaches. (2011) Chem. Res. Tox. Submitted.

Posters/Abstracts

1. Zhu H., Zhang L., Staab J., Sedykh A., Tang H., Gomez S., Rusyn I., 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 50th Annual Meeting, Washington DC, March 2011.
2. Pu D., Zhu H., Zhao G., Strickland J., Salicru E., and Tropsha A. Quantitative Structure-Activity Relationship Modeling of Skin Sensitizers Tested in Local Lymph Node. Society of Toxicology 50th Annual Meeting, Washington DC, March 2011.
3. Low Y., Uehara T., Minowa Y., Yamada H., Ohno Y., Urushidani T., Sedykh A., Fourches D., Zhu H., Rusyn I., Tropsha A. Predictive value of chemical and toxicogenomic descriptors for drug-induced hepatotoxicity. Society of Toxicology 50th Annual Meeting, Washington DC, March 2011.
4. Zhang L., Zhu H., Afantitis A., Melagraki G., Sarimveis H., Rusyn I., and Tropsha A. Quantitative Structure-Activity Relationship (QSAR) Modeling of Estrogen Receptor (ER) Binding Affinity and Virtual Screening for Potential Endocrine Disrupting Compounds (EDCs). Society of Toxicology 50th Annual Meeting, Washington DC, March 2011.
5. Afantitis A., Melagraki G., Sarimveis H., Zhang L., Zhu H., and Tropsha A. Combinatorial QSAR Modeling of Toxicity Data Using 2D & 3D Chemical Descriptors. Book of abstract, 18th European Symposium on Quantitative Structure Activity Relationships, Rhodes, Greece, September 2010.