The 17th European Symposium on Quantitative Structure-Activity Relationships & Omics Technologies and Systems Biology, Uppsala, Sweden, 21-26 September 2008.

Advances in Toxico-Cheminformatics: Supporting a New Paradigm for Predictive Toxicology

Ann Richard[1], Maritja Wolf[2], Thomas Transue[2], Clarlynda Williams-Devane[3]

[1]National Center for Computational Toxicology, US EPA, Mail Drop D343-03, Research Triangle Park, NC 27711, richard.ann@epa.gov; [2]Lockheed Martin -Contractor to the US EPA, Research Triangle Park, NC 27711; [3]NC State University Bioinformatics Program – US EPA Student COOP, Research Triangle Park, NC 27711

EPA's National Center for Computational Toxicology is building capabilities to support a new paradigm for toxicity screening and prediction through the harnessing of legacy toxicity data, creation of data linkages, and generation of new high-throughput screening (HTS) data. The DSSTox project is working to improve public access to quality structure-annotated chemical toxicity information in less summarized forms than traditionally employed in SAR modeling, and in ways that facilitate both data-mining and read-across. Both DSSTox Structure-Files and the dedicated on-line DSSTox Structure-Browser are enabling seamless structure-based searching and linkages to and from previously isolated, chemically indexed public toxicity data resources (e.g., NTP, EPA IRIS, CPDB). Most recently, structure-enabled search capabilities have been extended to chemical exposure-related microarray experiments in the public EBI Array Express database, additionally linking this resource to the NIEHS CEBS toxicogenomics database. The public DSSTox chemical and bioassay inventory has been recently integrated into PubChem, allowing a user to take full advantage of PubChem structure-activity and bioassay clustering features. The DSSTox project is providing cheminformatics support for EPA's ToxCast™ project, as well as supporting collaborations with the National Toxicology Program (NTP) HTS and the NIH Chemical Genomics Center (NCGC). Phase I of the ToxCast™ project is generating HTS data from several hundred biochemical and cell-based assays for a set of 320 chemicals, mostly pesticide actives with rich toxicology profiles. Reference *in vivo* toxicity data for these compounds have been incorporated into a toxicity relational database, ToxRef DB, and the newly developed ACToR system will house all the data being collected and generated, as well as provide informatics and analysis support for ToxCast™. Public data release of HTS, genomics, and reference toxicity data, through ACToR, DSSTox, and PubChem, is a central precept of the ToxCast™ effort. Incorporating structure-based data-mining capabilities and traditional SAR concepts into this new data-rich world poses conceptual and practical challenges, but holds great promise for improving predictive toxicology capabilities.

This work was reviewed by EPA and approved for publication, but does not necessarily reflect EPA policy.

References:

Dix, D.J., Houck, K.A., Martin, M.T., Richard, A.M., Setzer, W., Kavlock, R.J. (2007) The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Tox. Sci.* 95:5-12.

Richard A.M. (2006) Future of Predictive Toxicology: An Expanded View of "Chemical Toxicity" - Future of Toxicology Perspective. *Chem. Res. Toxicol.* 19:1257-1262.

Richard, A., Yang, C., Judson, R. (2008) Toxicity Data Informatics: Supporting a New Paradigm for Toxicity Prediction. *Tox. Mech. Meth*. 18:103-118.

Yang, C., Richard, A.M., and Cross, K.P. (2006) The art of data mining the minefields of toxicity databases to link chemistry to biology. *Curr. Comput.-Aided Drug Design* 2(2):135-150.

EPA DSSTox. (2008) U.S. Environmental Protection Agency's Distributed Structure-Searchable Toxicity (DSSTox) Database Network. Available on the Internet at: http://www.epa.gov/ncct/dsstox/