The power of an ontology-driven developmental toxicity database for data mining and computational modeling

A. Richard [1], E. Julien[2], E. Busta[3], M. Wolf[4], C. Yang[3]

1. NCCT (D343-03), US EPA, RTP, NC 27711, USA
2. LifeSci Research Service, Rockville, MD 20847, USA
3. CFSAN (HFS-275), US FDA, College Park, MD 20740, USA
4. Lockheed-Martin, Contractor to the US EPA, RTP, NC 27711, USA

Modeling of developmental toxicology presents a significant challenge to computational toxicology due to endpoint complexity and lack of data coverage. These challenges largely account for the relatively few modeling successes using the structure–activity relationship (SAR) paradigm. The development of new *in vitro* profiling approaches, employing screening assays or lower organisms to evaluate developmental toxicity requires anchoring to the results of *in vivo* studies. The International Life Science Institute (ILSI) recently released a public ontology-driven DevToxDB with data extracted from published studies. This project heavily influenced other efforts, and the data content in the ILSI DevToxDB was combined into an ontology-based database along with other public data from EPA ToxRefDB, FDA Center for Food Safety and Applied Nutrition and Center for Drug Evaluation Research and the National Toxicology Program. All chemical structures are indexed in DSSTox, providing the ability to assess chemical coverage and diversity of these largely non-overlapping data inventories, which include drugs, pesticides, industrial chemicals, food ingredients and contact substances. The use of a common toxicological ontology provides a logical means to group and aggregate biological effects in subsequent computational analyses. Chemoinformatics methods are used to compare the chemical space of each data source, and to explore associations with biological endpoints through the toxicology ontology perspective. Chemicals in these distinct datasets cause common phenotypes as well as distinct, non-overlapping phenotypes in the same target organs. Incorporation of the ILSI DevTox database into other public database efforts should provide a rich foundation for spurring innovation in SAR modeling of developmental endpoints. *This abstract does not necessarily represent policies of FDA or EPA.*