Title: Using ToxCast *in vitro* Assays in the Hierarchical Quantitative Structure-Activity Relationship (QSAR) Modeling for Predicting *in vivo* Toxicity of Chemicals

Authors: Zhu, Hao[1]; Zhang, Liying[1]; Sedykh, Alexander[1]; Tang, Hao[1]; Richard, Ann[2]; Rusyn, Ivan[3]; Tropsha, Alexander [1]
1. UNC at Chapel Hill, NC, USA.
2. NCCT, US EPA, RTP, NC, USA.

Key words: computational toxicology, QSAR, ToxCast

The goal of chemical toxicology research is utilizing short term bioassays and/or robust computational methods to predict *in vivo* toxicity endpoints for chemicals. The ToxCast program established at the US Environmental Protection Agency (EPA) is addressing this goal by using ca. 600 *in vitro* assays to create bioactivity profiles for a set of 320 compounds with known in vivo toxicity measured in ca. 80 assays. The analysis of this data requires new computational approaches to link chemical structures, *in vitro* responses and *in vivo* toxicity effects. We have employed a novel hierarchical QSAR approach to develop predictive models of three ToxCast *in vivo* multi-generation rat toxicity endpoints: i.e., kidney and liver pathologies, and animal viability index. This approach relies on the relationships between *in vitro* and *in vivo* assay results as follows: First, all chemicals are partitioned into two classes based on whether the results of the *in vitro* and *in vivo* assays agree (i.e., the compound is found either active or inactive in both types of assays) or disagree (the compound's annotations *in vitro* versus *in vivo* disagree). Second, classification QSAR models for these two classes are developed using Random Forest and Support Vector Machine methods. The resulting QSAR models are 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 the known *in vitro* response. All the ToxCast bioassays were then ranked based on the external predictivity of the associated models for each *in vivo* toxicity endpoint. The prediction accuracy for all models was in the range of 61-73% for all three *in vivo* endpoints, while that achieved by conventional QSAR models was only 50-65% for the same external set. Our models could be used to guide the future toxicity studies on the EPA-10K compounds by selecting *in vitro* assays and prioritizing compounds for *in vivo* toxicity evaluation. Abstract does not reflect EPA policy.