Prediction of *in vivo* hepatotoxicity effects using *in vitro* transcriptomics data

JA Franzosa[1,2], J Liu[1,2], RS Thomas[1], JF Wambaugh[1], KA Houck[1], I Shah[1]
[1]NCCT/ORD/US EPA, [2]ORISE

High-throughput *in vitro* transcriptomics data support molecular understanding of chemical-induced toxicity. Here, we evaluated the utility of such data to predict liver toxicity. First, in vitro gene expression data for 93 genes was generated following exposure of metabolically competent HepaRG cells to 1060 environmental chemicals from the US EPA ToxCast library. The empirical relationship between these data and rat chronic liver endpoints from animal studies in the Toxicity Reference Database (ToxRefDB) was then evaluated using machine learning techniques. Chemicals were classified as positive (242) or negative (135) based on observed hepatic histopathologic effects, and divided into three categories: hypertrophy (183), injury (112) and proliferative lesions (101). Hepatotoxicants were classified on the basis of the bioactivity of 93 genes (descriptors) using six machine learning algorithms: linear discriminant analysis, naïve Bayes, support vector classification, classification and regression trees, k-nearest neighbors, and an ensemble of classifiers. Classification performance was evaluated using 10-fold cross-validation testing, and in-loop, filter-based, feature subset selection. The best balanced accuracy for prediction of hypertrophy, injury and proliferative lesions were 0.81 ± 0.07, 0.79 ± 0.08 and 0.77 ± 0.09, respectively. Gene specific perturbation of xenobiotic metabolism enzymes (CYP7A1/2E1/4A11/1A1/4A22) and transporters (ABCG2, ABCB11, SLC22A1) were most predictive of hypertrophy. The bioactivity descriptor for increased expression of NRF2, a major regulator of cell defense pathways, was highest across chemical inducers of liver injury. The results also showed that specific genes, eg. increased MYC, IL-6, and TGF-beta and decreased SULT2A1 expression, were most predictive of proliferative lesions. In all, our findings demonstrate the utility of *in vitro* transcriptomics data to predict adverse hepatic outcomes *in vivo,* and support integration of 'omics technology into hazard prioritization paradigms. *This abstract does not represent EPA policy.*