The use of biological descriptors of chemical compounds to enrich traditional cheminformatics applications

Alexander Tropsha

University of North Carolina, Chapel-Hill, USA

Cheminformatics approaches such as Quantitative Structure Activity Relationship (QSAR) modeling have been used traditionally for predicting chemical toxicity. In recent years, high throughput biological assays have been increasingly employed to elucidate mechanisms of chemical bioactivity, especially, toxicity and predict effects of chemicals *in vivo*. The data generated in such assays can be considered as biological descriptors of chemicals that can be combined with molecular descriptors and employed in QSAR modeling to improve the accuracy of toxicity prediction. I will discuss several approaches for integrating chemical and biological data for predicting biological effects of chemicals *in vivo* and compare their performance across several data sets. I will show that while no method consistently shows superior performance, the integrative approaches rank consistently among the best yet offer enriched interpretation of models over those built with either chemical or biological data alone. I shall discuss the outlook for such interdisciplinary methods and offer recommendations to further improve the accuracy and interpretability of computational models that predict chemical effects *in vivo*.