Virtual Screening of Chemicals for Estrogen and Androgen Activities

K. Mansouri*, N. Kleinstreuer**, C. Grulke***, A.M. Richard***, Imran Shah***, Antony J. Williams***, Richard S. Judson***

* ScitoVation LLC, Research Triangle Park, NC, USA (currently at ILS, Research Triangle Park, NC, USA)

** NIH/NIEHS/DNTP/NICEATM, Research Triangle Park, NC, USA

***National Center for Computational Toxicology, U.S Environmental Protection Agency, Research Triangle Park, NC, USA

Endocrine disrupting chemicals (EDCs) can alter hormone synthesis, transport, and metabolism pathways. EPA evaluates EDC bioactivity in part through the use of high-throughput screening (HTS) *in vitro* approaches and computational modeling. EPA led two worldwide consortiums to screen chemicals *in silico* for potential estrogenic and androgenic activity. The Collaborative Estrogen Receptor (ER) Activity Prediction Project (CERAPP) predicted activities for 32,464 chemicals and the Collaborative Modeling Project for Androgen Receptor (AR) Activity (CoMPARA) generated predictions on 55,450 unique structures. Scientists from 35 international groups contributed structure-based models and results for activity prediction to the projects, with various methods used to predict binding, agonism, and antagonism activities. Models were trained on a set of 1746 chemicals having ToxCast/Tox21 HTS *in vitro* assay results integrated into computational networks. The models were validated using curated literature data (~7,000 results for ER and ~11,000 results for AR). CERAPP and CoMPARA models were then each combined into consensus models that achieved 92% predictive accuracy. These consensus models were implemented in the free and open-source application OPERA to avoid running every single model on new chemicals. This was used to screen the entire EPA DSSTox database of ~750,000 chemicals, and these ER and AR activity predictions are freely available on the EPA CompTox Chemistry dashboard. *This work does not reflect EPA or NIEHS policy.*