

Protein Target Prediction of Toxic Molecules Identifies Toxicological Relationships Between Proteins

F. Nigsch¹

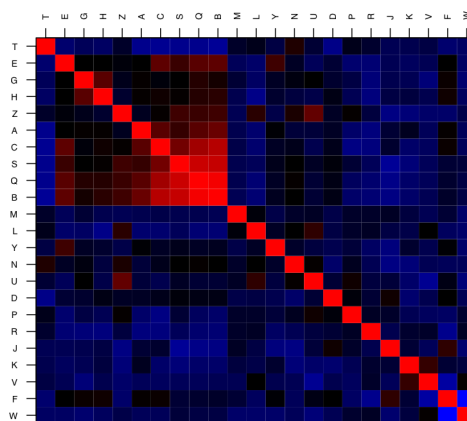
J.B.O. Mitchell¹

¹ Unilever Centre for Molecular Science Informatics, Department of Chemistry, Lensfield Road, Cambridge CB2 1EW, UK

Computational methods for protein target prediction can be used for the in silico identification of potential off-target activities. These off-target activities may often be the origin of clinically observed toxic effects or adverse drug reactions.

Based on a method that we recently applied successfully to a similar but easier classification problem, we built a model encompassing a larger dataset for protein target prediction of toxic molecules. [1] Our model uses the Winnow algorithm as underlying classification framework and circular fingerprints as molecular descriptors. A protein target dataset with 90,000 molecules spanning 233 activity classes was obtained by selecting all relevant classes from the MDL Drug Data Report (MDDR). Prior to the application of the protein target prediction model, we validated it using a 15-fold Monte Carlo cross-validation, each of which using a 50:50 split. We retained the 3 top-ranking predictions and found that in 82 percent of all cases the correct target was predicted within these three predictions. The first prediction was the correct one in almost 70 percent of cases.

This model was then applied to predict the protein targets of 150,000 molecules with experimentally determined toxicities contained in the MDL Toxicity Database. The resulting associations allowed us to determine proteins that are related with respect to their toxicities, as well as to cluster toxicities which are related with respect to the proteins likely to cause these toxicities. For both highly correlated protein clusters and also the top-ranking proteins for each toxicity class, we were able to confirm the significance of our results by independent evidence from published literature.



1. Nigsch, F.; Mitchell, J.B.O. How to Winnow Actives from Inactives: Introducing Molecular Orthogonal Sparse Bigrams (MOSBs) and Multiclass Winnow. *J. Chem. Inf. Model.* **2008**, ASAP article