Abstract

Despite technological advances and improved understanding of biological systems, drug discovery remains an inefficient and arduous task, with the high attrition of candidate molecules. Side effects (adverse reactions) is one of the key factors contributing to the rejection of candidate molecules with therapeutic potential. Hence, accurate prediction of phenotypic side effects is an important problem in drug discovery. The action of drugs needs to be seen from the systems perspective knowing that cellular mechanisms form a web of interactions with intricate cross-talks among biomolecules. Availability of data capturing molecular interaction of drugs, and their phenotypic side effects have facilitated systems-level models aimed at prediction of potential side effects. Towards the goal of predicting side effects, objectives set in this thesis were driven by the idea of creating holistic models using empirical data, and devising mathematical as well as computational strategies.

We integrated data from existing resources such as DrugBank and SIDER for systems-level investigations of side effects, and developed an integrative Generalized Canonical Correlation Analysis model which facilitates consolidation of various drugs features. We concluded that models implementing chemical profiles show more consistent accuracy than those based on target profiles. Further we constructed a graph theoretical model of biological space to account for associations among drug targets, and by comparing the performance of various network metrics inferred that simple network parameters are comparable to intricate parameters. Our studies performed for identification of minimal 'known side effects' set as a predictor for a class of adverse reactions suggest that, partial information of side effects profile could be used as a factor for arriving at the remaining side effects. Finally, towards the goal of obtaining drug features that contribute the most to side effects prediction, we developed a partial canonical correlation analysis model that facilitates enumeration of contribution from individual drug features. Our systems-level investigations offer insights into mechanisms of adverse drug reactions and provide data-driven methods for their prediction.

...