Finding Key Members in Compound Libraries by Analyzing Networks of Molecules Assembled by Structural-Similarity

Zsolt Lepp, Chunfei Huang, Takashi Okada

Characterization of chemical libraries is an essential task in everyday chemoinfomatics practice. This study describes some potential uses of network visualization and analysis methods to identify distinguished members of compound libraries. Molecules were ordered into networks by their structural similarity defined by molecular fingerprints. Various properties of such networks were examined. It was shown, that the correlation methods used to calculate the similarity between two structures radically determined the topology of networks. From the same set of molecules, the Russel-Rao and the Baroni-Urbani methods created sparser and denser networks, respectively, than using the Tanimoto method. Central nodes, corresponding to central compounds in the libraries, were determined for some example data sets. It was shown by the case of Adenosine A1, A2 and dual antagonists, that the methods used to identify central nodes could be divided into two groups; (1) centrality methods, exemplified by the centroid centrality, which could pick up structures that were the most similar to the largest number of other molecules, and group (2) exemplified by netweenness centrality, that could identify molecules that had intermediate structures between some homogeneous subsets of the library. The latter method gave significantly higher ranks to dual Adenosine antagonists, hinting the suitability of this measure to identify molecules with multiple activities. In the frame of the study, a Jchem plug-in has been developed to the Cytoscape network visualization software, which makes the visual observation of molecular networks more convenient. The plug-in is included in the supplementary information of the article for free usage.

KEYWORDS molecular similarity, molecular fingerprint, similarity network, complex network analysis, centrality, minimum spanning tree, threshold network, adenosine antagonists, CRF inhibitors, MDDR, compound library, correlation measures