BIOINFORMATICS APPLICATIONS NOTE Vol. 24 no. 18 2008, pages 2117–2118
doi:10.1093/bioinformatics/btn389

Systems biology

DrugViz: a Cytoscape plugin for visualizing and analyzing small molecule drugs in biological networks

Bing Xiong[1], Ke Liu[1], Jie Wu[1], David L. Burk[2], Hualiang Jiang[1] and Jingkang Shen[1,*]

[1]State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zuchongzhi Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai, China 201203 and [2]Department of Biochemistry, McGill University, 740 Dr Penfield Avenue, Montreal, Quebec, Canada H3A 1A4

Received and revised on May 22, 2008; accepted on July 23, 2008

Advance Access publication July 25, 2008

Associate Editor: Alfonso Valencia

ABSTRACT

Summary: DrugViz is a Cytoscape plugin that is designed to visualize and analyze small molecules within the framework of the interactome. DrugViz can import drug–target network information in an extended SIF file format to Cytoscape and display the two-dimensional (2D) structures of small molecule nodes in a unified visualization environment. It also can identify small molecule nodes by means of three different 2D structure searching methods, namely isomorphism, substructure and fingerprint-based similarity searches. After selections, users can furthermore conduct a two-side clustering analysis on drugs and targets, which allows for a detailed analysis of the active compounds in the network, and elucidate relationships between these drugs and targets. DrugViz represents a new tool for the analysis of data from chemogenomics, metabolomics and systems biology.

Availability: DrugViz and data set used in Application are freely available for download at http://202.127.30.184:8080/software.html

Contact: jkshen@mail.shcnc.ac.cn

1 INTRODUCTION

Over the last decade, genome-wide technologies have been increasingly used to explore biological processes in the whole cells (Vastrik *et al*., 2007). These systems biology approaches are essential to elucidate the interactions between genes, proteins and metabolites. High throughput/content screening and chemical genetics methods are now commonly used to investigate the chemicals involved in cellular functions with the aim of identifying novel drugs or to discover the macromolecular targets of chemical compounds (Yildirim *et al*., 2007). These studies generate large quantities of 'systems' biological/chemical data. Tools are urgently required to facilitate the visualization and analysis of these data to extend our knowledge of biological functions and to examine the relationships between drugs and targets.

Cytoscape, a Java-based extensible software package, was developed to visualize and analyze biological networks, with the aim of helping researchers address biological questions from a network perspective (Shannon *et al*., 2003). This program has been widely adopted for identifying important phenomena hidden in systems biology data. Although numerous plugins have been developed to enhance the functionality of Cytoscape, to the best of our knowledge, there is no plugin that can handle the small molecule structure information that is common in the metabolomics, chemogenomics and drug design fields. DrugViz was developed to incorporate this functionality into Cytoscape, where it could assist researchers in visualizing and examining small molecule interactions within biological networks. In addition to helping identify important relationship between drugs and targets, this plugin will also enable researchers to dissect the essential fragments contributing to biological functions. Ultimately, this may lead researchers to select better targets and design drugs with fewer unwanted side effects on the basis of biological network information.

*To whom correspondence should be addressed.

2 METHODS AND IMPLEMENTATION

To exploit the ability of Cytoscape to handle the 2D structure information of small molecules, an extended SIF file format was adapted by adding a SMILES string of small molecules at the end of each line (Weininger, 1988). SMILES strings are commonly employed in chemoinformatics and give a simple yet sufficient representation of the molecules, simplifying the generation of input files for DrugViz.

Once DrugViz reads in the input file, the SMILES string representing the small molecule is parsed and stored in memory for later searching. Several selection methods were implemented to enable users to select drugs and targets. The user can select interacting drugs based on target nodes and can select interacting targets by drug nodes. Combined with the original selection functions of Cytoscape, one can easily conduct various searches to find targets or drug nodes and then investigate the 2D structure patterns in these drugs by displaying the 2D structures of them. It is also convenient to find drug nodes by utilizing the three structure search methods: isomorphic search, substructure search and the fingerprint-based similarity search. The Chemistry Development Kit (CDK) library was used as a basis for implementing these visualization and 2D structure-based search functions (Steinbeck *et al*., 2003). The structure search results are recorded and enables the user to access these results later in the analysis. In addition to these search methods, a shortest path search algorithm was implemented to facilitate the investigation of closely related drugs or drug targets. To better reveal the relationship in drugs and target, we also implemented a clustering functionality in DrugViz. After the users supply the distance matrices of drugs and targets, DrugViz can perform a two-side clustering and display the hierarchical clustering tree in addition to the drug–target interactions. Together, these capabilities will allow various types of searches in the drug–target network along with appropriate visualization.

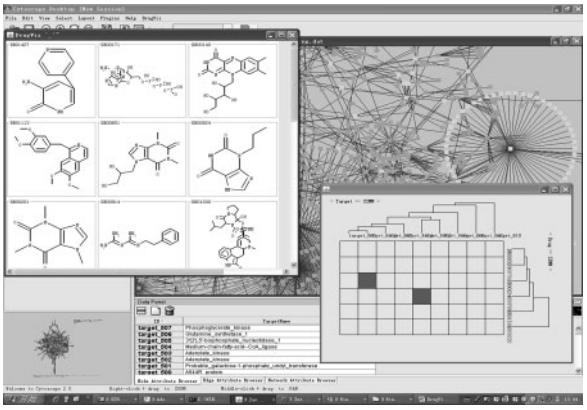

Fig. 1. The Screenshot of the sample DrugBank application conducted by DrugViz plugin of Cytoscape.

3 APPLICATION

To demonstrate the utility of the DrugViz plugin, we analyzed a sample set of small molecule drugs and their targets. The drug–target interaction data were derived from the DrugBank (http://www.drugbank.ca, (Wishart *et al*., 2006). Only the small molecular drugs approved by FDA were considered in the analysis. Target protein sequences were formatted for a blast sequence similarity search. We built a target–target similarity network by making a bond between two targets if they had an *E*-value lower than 1E-10 and had a alignment sequence longer than 150 aa. After filtering the target–target similarity network by removing targets not interacting with FDA-approved small molecule drugs, the target–target network was found to contain a total of 809 targets and 3137 bonds. The drug–target network was found to have 2697 nodes and 3822 bonds. By analyzing this network, the following questions can be addressed (Fig. 1):

(1) How many drugs interact with a target, and what are their 2D structures?
Answer: Select the target node and select the menu item 'Select BindingDrugs'; then, click the "Display Structs".
(2) Do similar targets have similar patterns in their interacting drugs?
Answer: First select the target node and use the Cytoscape node selection method 'first neighbors of selected nodes'; then, select the 'Select BindingDrugs' in the DrugViz menu.
(3) Which drugs are similar to this one and what are the targets for these similar drugs?
Answer: Click the 'Searching…' to conduct a similarity or substructure search; then, click the 'select BindingTargets'.
(4) Are two drugs related in a biological pathway, and if so, what is the pathway?
Answer: One can import a pathway dataset into Cytoscape; then, load a drug–target network file. Select two drug molecule nodes of interest and calculate the path by selecting 'Shortest Path (Target)' in DrugViz.
(5) What are the relationships between selected drugs, targets and drug–targets?
Answer: One can supply the distance matrices of drugs and targets and click 'clustering' menu item. The results are displayed as a hierarchical tree for drugs and targets. DrugViz will also label the drug–target interaction as a green cell on the cross of the drug–target grid. This clearly identifies the relationships between drugs, targets and drug–targets.

4 CONCLUSION

The Cytoscape plugin DrugViz was designed and implemented to enhance chemogenomics and systems biology research related to small molecules. The plugin utilizes algorithms and methods from the chemoinformatics field to facilitate the investigation of the 2D structure information of small molecules in drug–target networks, which takes advantage of both chemoinformatics and bioinformatics tools to enable researchers to conduct various queries for selecting and searching. It provides a visualization framework for examining the 2D structures of small molecules together with a two-side clustering on drugs and targets. By analyzing the network of targets and drugs, users can gain nonobvious information about the relationship between drugs and their targets, and identify hidden patterns between them. Furthermore, by studying the relationships between the small molecules in a biological network or pathway, researchers may be able to identify proteins that can serve as drug targets for treating certain diseases.

ACKNOWLEDGEMENTS

We thank Dr A.M. Berghuis (McGill University, Canada) for providing suggestions and comments.

Funding: National Natural Science Foundation of China (Grant 30600784 to B.X.); Science and Technology Commission Of Shanghai Municipality (Grant 06zr14101).

Conflict of Interest: none declared.

REFERENCES

Shannon,P. *et al*. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*., **13**, 2498–2504.
Steinbeck,C. *et al*. (2003) The Chemistry Development Kit (CDK): an open-source Java library for Chemo- and Bioinformatics. *J. Chem. Inf. Comput. Sci.*, **43**, 493–500.
Vastrik,I. *et al*. (2007) Reactome: a knowledge base of biologic pathways and processes. *Genome Biol*. **8**, R39.
Weininger,D. (1988) SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci*., **28**, 31–36.
Wishart,D.S. *et al*. (2006) DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res*., **34**, D668–D672.
Yildirim,M.A. *et al*. (2007) Drug-target network. *Nat Biotechnol*., **25**, 1119–1126.