Abstract

In-silico screening is crucial in drug discovery, including herbal-medicines discovery. Given one or more plants, we want to know their target diseases. Likewise, for one or more kinds of diseases, we desire to know which plants can potentially (help to) cure them. Ijah is a web-server that provides searching and predicting services for the interaction between plants and diseases. Ijah searches over its huge aggregating database that contains information about plants-to-compounds, compounds-to-proteins and proteins-to-diseases, extracted from several separate databases. In addition, Ijah estimates unknown interactions using several distinct algorithms, then reports prediction confidence levels. Finally, Ijah visualizes the pharmacological network from plants to compounds to proteins to diseases in the form of multi-partite graphs. Ijah is publicly available at http://ijah.apps.cs.ipb.ac.id at no cost.

Website address: http://ijah.apps.cs.ipb.ac.id
Website name: Ijah Webserver
Authors: Vektor Dewanto (vektor.dewanto@gmail.com), Wisnu Ananta Kusuma (w.ananta.kusuma@gmail.com), Ajmal Kurnia (ajmalkuria@apps.ipb.ac.id), Haekal Zidni Barkan (hzbarkan@gmail.com)
Author Affiliation: Computer Science Department, Bogor Agricultural University, Indonesia

Input Data:
The Ijah webserver takes as input, plants (or compounds) in the drug-side and disease (or proteins) in the target-side. It accomodates 3 use cases as follows.
Use-case 1: having plants (or compounds) and diseases (or proteins), users want to know 3 things: a) which compounds are contained in those plants (or which plants contain those compounds), b) which proteins are associated with those disease (or which diseases are associated with such proteins), and c) the interaction among those compounds and proteins, for which Ijah gives both factual and predicted results.
Use-case 2: having only plants (or compounds), users want to know: a) which compounds are contained in those plants (or which plants contains those compounds) and b) top-n proteins based on the degree of interactions with those compounds, along with diseases associated with those proteins.
Use-case 3: having only diseases (or proteins), users want to know: a) which proteins are associated with those diseases (or which diseases are associated with such proteins), and b) top-n compounds based on the degree of interactions with those proteins, along with plants containing those compounds. The Ijah webserver also accomodates command-line and programming interface in order to allow programmatic access for rapid and efficient data retrieval.

Output:
The Ijah webserver outputs interactions from plants to compounds to proteins to diseases. Those interactions are visualized using a multipartite graph and formatted texts. The thickness of an edge indicates the confidence level that the interaction between two vertices exists in the graph. The Ijah webserver also provides downloads for most parts of its database.

Processing method:
The database of Ijah webserver is essentially a partial mirror of several major and related databases, such as KnapSack [5], DrugBank [3], and Uniprot [1]. We built database crawlers to autonomously extract essential information from relevant databases. These crawlers are parts of our programming interface, which are publicly available.

The Ijah web-server estimates unknown interaction between a (natural) compound and a protein. To this end, it has two strategies, namely: a) running its own predictors, i.e. an extension of BLM-NII [4] and b) querying drug-target prediction servers, e.g. DINIES [7], Superpred [6] and SwissTargetPrediction [2]. Prediction results from several distinct predictors are then combined via majority voting that determines the confidence level that interaction exists. The prediction results are then stored in the database so that

the same query later will be as simple as table look-up.

Keyword:
drug-target prediction, herbal medicine, webserver, In-silico screening for drug discovery

References

[1] A. Bateman and et.al. UniProt: a hub for protein information. *Nucleic Acids Res.*, 43(Database issue):D204–212, Jan 2015.

[2] David Gfeller, Aurélien Grosdidier, Matthias Wirth, Antoine Daina, Olivier Michielin, and Vincent Zoete. Swisstargetprediction: a web server for target prediction of bioactive small molecules. *Nucleic Acids Research*, 42(Webserver-Issue):32–38, 2014.

[3] C. Knox, V. Law, T. Jewison, P. Liu, S. Ly, A. Frolkis, A. Pon, K. Banco, C. Mak, V. Neveu, Y. Djoumbou, R. Eisner, A. C. Guo, and D. S. Wishart. DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res.*, 39(Database issue):D1035–1041, Jan 2011.

[4] Jian-Ping Mei, Chee-Keong Kwoh, Peng Yang, Xiao-Li Li, and Jie Zheng. Drug-Target Interaction Prediction by Learning From Local Information and Neighbors. *Bioinformatics*, 29(2):238–245, November 2012.

[5] K. Nakamura, N. Shimura, Y. Otabe, A. Hirai-Morita, Y. Nakamura, N. Ono, M. A. Ul-Amin, and S. Kanaya. KNApSAcK-3D: a three-dimensional structure database of plant metabolites. *Plant Cell Physiol.*, 54(2):e4, Feb 2013.

[6] Janette Nickel, Björn-Oliver Gohlke, Jevgeni Erehman, Priyanka Banerjee, Wen Wei Rong, Andrean Goede, Mathias Dunkel, and Robert Preissner. Superpred: update on drug classification and target prediction. *Nucleic Acids Research*, 42(Webserver-Issue):26–31, 2014.

[7] Yoshihiro Yamanishi, Masaaki Kotera, Yuki Moriya, Ryusuke Sawada, Minoru Kanehisa, and Susumu Goto. DINIES: drug-target interaction network inference engine based on supervised analysis. *Nucleic Acids Research*, 42(Webserver-Issue):39–45, 2014.