Retracted: Predicting new molecular targets for rhein using network pharmacology

Aihua Zhang1,2, Hui Sun1,2, Bo Yang1,2 and Xijun Wang1,2*

A retraction article was published for this article. It is available from the following link; http://www.biomedcentral.com/1752-0509/8/105.

Abstract

Background: Drugs can influence the whole biological system by targeting interaction reactions. The existence of interactions between drugs and network reactions suggests a potential way to discover targets. The in silico prediction of potential interactions between drugs and target proteins is of core importance for the identification of new drugs or novel targets for existing drugs. However, only a tiny portion of drug-targets in current datasets are validated interactions. This motivates the need for developing computational methods that predict true interaction pairs with high accuracy. Currently, network pharmacology has used in identifying potential drug targets to predicting the spread of drug activity and greatly contributed toward the analysis of biological systems on a much larger scale than ever before.

Methods: In this article, we present a computational method to predict targets for rhein by exploring drug-reaction interactions. We have implemented a computational platform that integrates pathway, protein-protein interaction, differentially expressed genome and literature mining data to result in comprehensive networks for drug-target interaction. We used Cytoscape software for prediction rhein-target interactions, to facilitate the drug discovery pipeline.

Results: Results showed that 3 differentially expressed genes confirmed by Cytoscape as the central nodes of the complicated interaction network (99 nodes, 153 edges). Of note, we further observed that the identified targets were found to encompass a variety of biological processes related to immunity, cellular apoptosis, transport, signal transduction, cell growth and proliferation and metabolism.

Conclusions: Our findings demonstrate that network pharmacology can not only speed the wide identification of drug targets but also find new applications for the existing drugs. It also implies the significant contribution of network pharmacology to predict drug targets.

Background

Developing a new drug is an expensive and time-consuming process that is subject to a variety of regulations such as drug toxicity monitoring and therapeutic efficacy. Lengthy development procedures and the high risk of unexpected side-effects in advanced-stage clinical trials reduce the ability of the drug development process to be innovative. However, the organization of our rapidly growing knowledge on diseases, disease-related genes, drug targets and their structures, and drugs and their chemical structures gives us another exciting way to discover novel areas of drug development. Several networks have recently been constructed to help drug discovery [1]. Meanwhile, finding the potential application in other therapeutic categories of drugs by predicting their targets is an efficient and time-saving method in drug discovery [2]. Additionally, predicting interactions between drugs and target proteins can help decipher the underlying biological mechanisms. Therefore, there is a strong incentive to develop powerful statistical methods that are capable of detecting these potential drug-protein interactions. Various methods have been proposed to address the drug-target prediction problems. One common method is to predict the drugs interacting with a single given protein based on the chemical structure similarity in a classic classification framework. Nevertheless, all the methods did not utilize a wealth of information to assist prediction.
Despite the dramatic increase of global spending on drug discovery and development, the approval rate for new drugs is declining, due chiefly to toxicity and undesirable side effects. Simultaneously, the growth of available biomedical data in the postgenomic era has provided fresh insight into the nature of drug-target pathways. This stagnation in drug approval can be overcome by the novel concept of network pharmacology, which is built on the fundamental concept that drugs modulate the multiple targets. Network pharmacology can be studied with molecular networks that integrate multidisciplinary concepts including cheminformatics, bioinformatics, and systems biology. Network analysis has become a cornerstone of fields as diverse as systems biology. Many studies have successfully reported interesting biological findings from these networks, including the relationships between various statistical properties of a gene and its function at the molecular level based on networks [3]. Network pharmacology can make an impact at several points in the drug-development process: target identification, lead discovery and optimization, mode of action, preclinical efficacy and safety assessment. Therefore, it could facilitate the systematic characterisation of drug targets, thereby helping to reduce the typically high attrition rates in discovery projects. Various approaches have been proposed for this task such as Bayesian Networks, models based on information theory, regression based models, and differential equation models [4,5].

Software to integrate and analyse the interactions and their attributes plays an increasingly important role. The most widely used open source network visualization workbench is Cytoscape, a popular bioinformatics package for biological network visualization and data integration, for screening the central nodes of the network, exploiting functional study of the central node genes [6]. Cytoscape software, a web-based network visualization tool, is an open-source software platform for visualizing molecular interaction networks and integrating these interactions with gene expression profiles and other functional genomics data. Data from various sources can be imported into this tool to build networks, and to highlight specific node or edge features. By providing platforms to integrate data with molecular interaction networks, researchers can more rapidly begin interpretation of large data sets collected for a system of interest. Cytoscape’s main functionality is focused on the construction of networks. Currently, the network model is based on genes, proteins and functional relationships between them such as protein-protein, protein-dna regulatory interactions and gene-protein coding relationship [7].

Rhubarb, one of popular traditional Chinese herbal medicine, has been widely used for the treatment of diseases in traditional Chinese medicine. Rhein (Figure 1) is one of the most bioactive components from rhubarb. Earlier studies have identified rhein as a potent inhibitor of hepatic fibrosis induced by carbon tetra-chloride, which is capable of reducing SMA expression, collagen synthesis and deposition in TGF-stimulated hepatic stellate cells. A recent study showed that rhein contributes to induce apoptosis of human cancer, including lung adenocarcinoma A549, myelogenous leukemia HL-60, lung squamous carcinoma CH27, cervical carcinoma HeLa cells, neuroblastoma IMR-32, bladder cancer T24, and hepatoma HepG2 cells [8]. It can inhibits cellular proliferation, induces apoptosis, and prevents metastasis through activation of tyrosine kinases, phosphoinositol 3-kinase, protein kinase C, NF-kappa B, and mitogen-activated protein kinase signaling cascades. It has antitumor properties through the p53 and its downstream p21 pathway, also reduce tumor size, prolong survival, decrease incidence of tumor invasion and neovascularization. Interestingly, rhein blocks the uptake of glucose in tumor cells, causing changes in membrane associated functions to trigger cell death. Although these results suggest that rhein may be a potent antitumor drug, and its therapeutic mechanism were still completely unclear.

Reconstructing networks of biological system entities such as genes, transcription factors, proteins, compounds and other regulatory molecules are very important for understanding the biological processes. Identifying drug targets is a critical step in network pharmacology. Recent years network pharmacology has influenced all areas of life sciences including that of drug mechanism and development, new target discovery [9]. Efficient identification of drug targets is one of major challenges for drug discovery and drug development. Computational integration of different knowledge...

![Figure 1 Chemical structure of rhein](http://www.biomedcentral.com/1752-0509/6/20)
sources is a more effective approach and wealth of SysBiomics data provides unprecedented opportunities for drug target identification [10]. Although a number of computational approaches have been developed to integrate data from multiple sources for the purpose of predicting or prioritizing candidate disease genes, relatively few of them focus on identifying or ranking drug targets. To address this deficit, we construct a biological network to provide a global view of drug interactions and prioritize drug candidate targets. We demonstrate the applicability of integrative network pharmacology approaches to identify potential drug targets and candidate genes by employing information extracted from public databases. In the present investigation, we give an illustrative example to show that the potential drug target identification problem can be solved effectively by our method, which may become an effective strategy for the discovery of new drugs.

Methods

Interaction information was retrieved from NCBI’s Entrez Gene in December 16, 2011. Genes/proteins molecular function, biological processes and cellular component is imported from the Gene Ontology project; while information on biochemical pathways is taken from KEGG. Additional information includes links to databases, such as Reactome, BioCyc, NCI Nature PID, DIP, BIND, HPRD, BioGRID, MINT, and Intact, which represent the major repositories of interactions from multiple organisms for further bioinformatics analysis. The associations between the diseases and genes were from the OMIM (Online Mendelian Inheritance in Man, http://www.ncbi.nlm.nih.gov/omim). The most stable one is the Entrez GeneID, which is the unique identifier for a gene in NCBI’s Entrez Gene database. Cytoscape has being developed for reconstruction and visualization of networks. Nodes (represented as circles) in the interactome correspond to genes, and edges (connecting lines) represent documented interactions. Cytoscape is a desktop Java application and source code for Cytoscape 2.8 are available for download from http://cytoscape.org. Its default annotations are parsed from the GO information available from NCBI http://www.ncbi.nlm.nih.gov/Ftp/.

Results

Module annotation and visualization
Functional annotation for predicted binding targets in our models by looking for enriched terms from multiple functional databases. Since modules often show strong functional coherence, the diverse set of annotations provide a complementary overview of module function. For pathway visualization and analysis of networks, we used open-source Cytoscape version 2.8 software. In Cytoscape, networks are represented as graphs where the nodes are the entities (e.g. genes, proteins) and the edges their interactions (e.g. reactions). For the visualization in the context of biological networks, we developed the different node attribute files and visual style files that can easily be imported into Cytoscape. This benchmark was run on a standard desktop computer (4 GHz Pentium intel with 2 GB of memory running Windows XP).

Constructing regulatory network

We used Cytoscape version 2.8 to model the signaling network. Taking the closely-connected and co-expressed differentially expressed genes, for the network shown in Figure 2 and 3, with 99 nodes and 153 edges. Four of them are significant frequencies and are likely to be responsible for driving the initiation, progression, or maintenance of disease. We find that 3 genes considered linked because they had close relationship and suggested as novel susceptibility genes using the Cytoscape visualization software. The 3 perturbed genes in the network that has strong connections to the genes. This is indicated by the thick edges between these nodes.

Predicting drug targets

Another issue clearly meriting further study is the performance of our proposed method in identifying drug targets. We believe these results are indicative of the multi-level nature of key perturbations, where direct interactions often take place at the protein and/or metabolite level, and therefore do necessarily affect expression level of the encoding genes. Main pathway information, 3 targets are observed in the Cytoscape-based network with significant P-values. Figure 4 denotes the identification of novel molecular targets using network analysis. These closely connected and coexpressed differentially expressed genes and proteins in the networks are regarded as the signatures of the rhein underlying targets. Regulatory sub-network of rhein targets through network pharmacology was constructed, and shown in Figure 5. The detail information of the novel molecular targets were shown in additional Table 1.

Discussion

Network pharmacology approach seeks to comprehend the complexity of organisms by combining many different kinds of data (protein-protein and protein-DNA interactions, protein modifications, biochemistry, etc.) to create predictive models. In the era of SysBiomics, the focus on understanding complex organisms is shifting from studying individual genes and proteins towards the relationships between them [11,12]. These relationships are usually expressed in terms of various kinds of
Figure 2 Solid map on huge interactome of rhein-targets networks, built and visualized with Cytoscape. Edges: interactions. Nodes: specific proteins or genes. Central nodes (Yellow) of the interaction network were used to illustrate the gene expression obtained and represents significant change in expression. The connections between molecules show molecular interactions identified in the interactome. Gene expression illustrated in colored nodes was selected with a \( p < 0.05 \) value.

Figure 3 Illustration and visualization of the interactome network (Circle Layout). Drugs and proteins are linked as per the known drug-target network. Discrete network map with a customized visual Cytoscape Web style. Yellow nodes refer to the significant ontologies of the target. This subnetwork represents a coregulated unit containing 99 nodes and 153 edges.
biological networks that are the focus of many functional genomics studies. Systems biology is characterized by a focus on interaction networks—the biomolecules involved in a particular biological system or process, as well as the relationships between these components. Network pharmacology is used for visualizing and understanding these interactions, interpreting high-throughput experimental data, generating hypotheses and sharing results [13]. These diagrams can be difficult for a user to explore with currently available network display tools—the networks are often too large, on the order of thousands of nodes, and many tools do not provide biological context to the diagram. The increasing complexity of functional genomics data drives the development of methods and tools for data integration and visualization [14].

Interactions network models are crucially important for disease processes [15]. Many of the important properties of biological systems emerge as a result of the interactions among genes and among their protein products. Genes and the proteins they encode participate in gene-gene, gene-protein, and protein-protein interactions to mediate a wide variety of biological processes. Molecular interaction networks can be efficiently studied using network visualization software. Cytoscape that can generate a putative protein-protein interaction network for target genomes, make the creation of protein-protein interaction network predicting tools possible. Its central organizing principle is a network graph, with biological entities (e.g. genes, proteins) represented as nodes and biological interactions represented as edges between nodes. Data are integrated with the network using attributes, which map nodes or edges to specific data values such as gene expression levels or protein functions. Taken together, these features provide a mechanism for expressing relationships between sets of data while simultaneously visualizing the integrated results.

In this study, we applied Cytoscape to explore targets expression data in the context of biological network information. Of note, Cytoscape successfully provided us with valuable clues for identification of drug-target interactions on a large scale. Rhein, a classic natural product, has been efficiently used for cancer relief in Asia, although its mechanism remains unclear. A
promising approach in drug target discovery involves the integration of available metabolites data through mathematical modeling and data mining. Significant work has been done on drug discovery, however, few papers were discussed with the interaction network. This study was designed to further elucidate the underlying mechanism of rhein from the network pharmacology. Of note, 3 differentially expressed genes were observed. The characteristic functions of the differentially expressed proteins were based on biological processes such as immunity, cellular apoptosis, transport, signal transduction, cell growth and proliferation and metabolism. The detection of these proteins with distinctive regulatory patterns provides evidence that novel biomarkers are actively involved in multifunctional pathways. Proteins of the matrix metalloproteinase (MMP2 and 9) family are involved in the breakdown of extracellular matrix in normal physiological processes, such as immunity, cellular apoptosis, transport, signal transduction, cell growth and proliferation and metabolism. The detection of these proteins with distinctive regulatory patterns provides evidence that novel biomarkers are actively involved in multifunctional pathways. Proteins of the matrix metalloproteinase (MMP2 and 9) family are involved in the breakdown of extracellular matrix in normal physiological processes, such as immunity, cellular apoptosis, transport, signal transduction, cell growth and proliferation and metabolism. The detection of these proteins with distinctive regulatory patterns provides evidence that novel biomarkers are actively involved in multifunctional pathways.

Table 1 Information of the novel molecular targets

| GeneID | Gene name | Disease | Pathway |
|--------|-----------|---------|---------|
| 4313   | MMP2      | Cancer  | GnRH signaling pathway |
|        |           | Choriocarcinoma | Leukocyte transendothelial migration |
|        |           | Torg-Winchester syndrome | Pathways in cancer |
| 4318   | MMP9      | Penile cancer | Pathways in cancer |
|        |           | Metaphysal dysplasias | Leukocyte transendothelial migration |
| 7124   | TNF       | Asthma  | MAPK signaling pathway |
|        |           | Systemic lupus erythematosus | Cytokine-cytokine receptor interaction |
|        |           | Malacia | Natural killer cell mediated cytotoxicity |
|        |           | Pertussis | NOD-like receptor signaling pathway |
|        |           | Type I, II diabetes mellitus | Toll-like receptor signaling pathway |
|        |           | Alzheimer’s disease | RIG-I-like receptor signaling pathway |
|        |           | Amoebiasis | Natural killer cell mediated cytotoxicity |
|        |           | Tuberculosis | T cell receptor signaling pathway |
|        |           | Hepatitis C | Fc epsilon RI signaling pathway |
|        |           | Influenza A | Apoptosis |
|        |           | HTLV-I infection | TGF-beta signaling pathway |
|        |           | Herpes simplex infection | Natural killer cell mediated cytotoxicity |
|        |           | Cardiovascular Diseases | Hematopoietic cell lineage |

This cytokine is involved in the regulation of a wide spectrum of biological processes including cell proliferation, differentiation, apoptosis, lipid metabolism, and coagulation. This cytokine has been implicated in a variety of diseases, including autoimmune diseases, insulin resistance, and cancer.

The dominant paradigm in drug discovery is the concept of designing maximally selective drug targets. However, many effective drugs act via modulation of multitargets rather than single targets. Advances in systems biology are revealing that integrated network biology holds the promise of expanding the current opportunity space for drug targets [16]. Identification of drug targets is one of the major tasks in drug discovery [17]. Under these circumstances, there is an urgent need to integrate phenotypic and chemical indexes together and develop new methods to predict drug-target interactions on a large scale. With the development of systems biology and the emergence of network pharmacology approach, it has been possible to integrate multidimensional information and heterogeneous data in drug studies [18]. Our method benefits from current knowledge such as the known drug-target interactions, more importantly, extends the candidate target proteins to a genome-wide scale, which greatly enlarges the number of known targets. Together with known drug-target interactions, such information makes it possible to relate pharmacological space with genomic space. Thus, we believe that combining the integration of multi-dimensional information in pharmacological space and
The identified targets were found to encompass a variety of biological processes related to immunity, cellular apoptosis, transport, signal transduction, cell growth and proliferation and metabolism. Perturbed proteins tend to be highly coexpressed and functionally coherent and we have used this property for predicting drug targets and associating novel functions to drug. The findings demonstrate that the network target-based methods are of importance for elucidating the inter-relationship between complex diseases and drug interventions through the network target paradigm estimating.

Acknowledgements
This work was supported by grants from the Key Program of Natural Science Foundation of State (Grant No. 90709019), the National Specific Program on the Subject of Public Welfare (Grant No. 200807014), National Key Subject of Drug Innovation (Grant No. 2009ZX09502-005), and National Program on Key Basic Research Project of China (Grant No. 2005CB523406).

Author details
1 National TCM Key Lab of Serum Pharmacoechemistry, Heilongjiang University of Chinese Medicine, Hepping Road 24, Harbin 150040, China. 2 Key Pharmacometabolomics Platform of Chinese Medicines, Hepping Road 24, Harbin 150040, China.

Authors' contributions
XWJ and AHZ conceived of the study. AHZ and HS carried out the data analysis, simulations, and drafted the manuscript. HS analyzes the results. BY carried out extensive revisions to the manuscript. All authors read and approved the final content.

Competing interests
The authors declare that they have no competing interests.

Received: 12 January 2012 Accepted: 21 March 2012 Published: 21 March 2012

References
1. Gilchrist M, Thorisson V, Li R, Hust AG, Kord M, Roach JC, Kennedy K, Hai T, Boloun H, Adren K. Systems biology approaches identify AT3 as a negative regulator of Toll-like receptor. Nature 2006, 441:173-178.
2. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. Nat Chem Biol 2008, 4:682-690.
3. Li S, Zhang B, Zhang N. Network target for screening synergistic drug combinations with application to traditional Chinese medicine. BMC Syst Biol 2011, 5(Suppl 1):S10.
4. Fang K, Zhao H, Sun C, Lam CM, Chang S, Zhang K, Panda G, Godinho M, Martins dos Santos VA, Wang J. Exploring the metabolic network of the epidemic pathogen Burkholderia cenocepacia J2315 via genome-scale reconstruction. BMC Syst Biol 2011, 5:83.
5. Overtom IM, Graham S, Gould KA, Hindi J, Botting CH, Shirran S, Barton GJ, Coote PJ. Global network analysis of drug tolerance, mode of action and virulence in methicillin-resistant S. aureus. BMC Syst Biol 2011, 5:68.
6. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003, 13:2498-2504.
7. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campolo I, Creech M, Gross B, Hanspers K, Islerin R, Kelley R, Killooye S, Lotia S, Maere S, Morris J, Ono K, Pavlovic V, Pico AR, Vallaya A, Wang PL, Adler A, Conklin BR, Hood L, Kuiper M, Sander C, Schmulevich I, Schwikowski B, Warner GJ, Ideker T, Bader GD. Integration of biological networks and gene expression data using Cytoscape. Nat Protoc. 2007, 2:3366-3382.
8. Chang CY, Chan HL, Lin HY, Wu TD, Kao MC, Song MZ, Lin YJ, Lin CW. Rhein induces apoptosis in human breast cancer cells. Evid Based Complement Alternat Med 2012, 2012:952504.
9. Janga SC, Tzakos A. Structure and organization of drug-target networks: insights from genomic approaches for drug discovery. Mol Biol Evol 2009, 26:1536-1548.
10. Huthmacher C, Hoppe A, Bulik S, Holzhueter HG. Antimalarial drug targets in Plasmodium falciparum predicted by stage-specific metabolic network analysis. BMC Syst Biol 2010, 4:120.
11. Collins SR, Miller KM, Maas NL, Roguey A, Fillingham J, Chu CS, Schulzner M, Gebbia M, Recht J, Shales M, Ding H, Xu H, Han J, Ingvandottir K, Cheng B, Andrews B, Boone C, Berger SL, Hieter P, Zhang Z, Brown GW, Ingles CJ, Emili A, Allis CD, Toczyski DP, Weisman JS, Greenblatt JF, Krogan NJ. Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature 2007, 446:806-810.
12. Wang X, Yang B, Zhang A, Sun H, Yan C. Potential drug targets on insomnia and intervention effects of Jujuboside A through metabolic pathway analysis as revealed by UPLC/EI-SYNAPT-HDMS coupled with pattern recognition approach. J Proteomics 2012, 75:1411-1427.
13. Morrow JK, Tian L, Zhang S. Molecular networks in drug discovery. Crit Rev Biomed Eng 2010, 38(2):143-56.
14. Kim HU, Sohn SB, Lee SY: Metabolic network modeling and simulation for drug targeting and discovery. *Biotechnol J* 2012, 7:330-342.

15. McEachin RC, Chen H, Sartor MA, Saccone SF, Keller BJ, Prossin AR, Cavalcanti JD, McNinis MG: A genetic network model of cellular responses to lithium treatment and cocaine abuse in bipolar disorder. *BMC Syst Biol* 2010, 4:158.

16. Yao X, Hao H, Li Y, Li S: Modularity-based credible prediction of disease genes and detection of disease subtypes on the phenotype-gene heterogeneous network. *BMC Syst Biol* 2011, 5:79.

17. Wang L, Zhou GB, Liu P, Song JH, Luang Y, Yan XJ, Xu F, Wang BS, Mao JH, Shen ZK, Chen SJ, Chen Z: Dissection of mechanisms of Chinese medicinal formula Realgar-Indigo naturalis as an effective treatment for promyelocytic leukemia. *Proc Natl Acad Sci USA* 2008, 105:4826-4831.

18. Morse DL, Gillies RJ: Molecular imaging and targeted therapies. *Biochem Pharmacol* 2010, 80(5):731-8.

19. Xia Z, Wu LY, Zhou X, Wang ST: Semi-supervised drug-protein interaction prediction from heterogeneous biological spaces. *BMC Syst Biol* 2010, 4(Suppl 2):S6.

20. Shen C, Huang Y, Liu Y, Wang G, Zhao Y, Wang Z, Teng M, Wang Y, Flockhart DA, Skaar TC, Yan P, Nephew KP, Huang TH, Li L: A modulated empirical Bayes model for identifying topological and temporal estrogen receptor α regulatory networks in breast cancer. *BMC Syst Biol* 2011, 5:67; doi:10.1186/1752-0509-5-67.

21. Zhao S, Li S: Network-based relating pharmacological and genomic spaces for drug target identification. *PLoS One* 2010, 5:e11764.

22. Gu J, Chen Y, Li S, Li Y: Identification of responsive gene modules by network-based gene clustering and extending: application to inflammation and angiogenesis. *BMC Syst Biol* 2010, 4:47.

23. van Laarhoven T, Nabuurs SB, Marchiori E: Gaussian interaction profile kernels for predicting drug-target interaction. *Bioinformatics* 2011, 27(21):3036-3043.

24. Burga A, Casanueva MO, Lehner B: Predicting mutation outcome from early stochastic variation in genetic interaction partners. *Nature* 2011, 480:250-253.