

**DATABASE**

**Open Access**

# CVDHD: a cardiovascular disease herbal database for drug discovery and network pharmacology

Jiangyong Gu, Yuanshen Gui, Lirong Chen\*, Gu Yuan and Xiaojie Xu\*

## Abstract

**Background:** Cardiovascular disease (CVD) is the leading cause of death and associates with multiple risk factors. Herb medicines have been used to treat CVD long ago in china and several natural products or derivatives (e.g., aspirin and reserpine) are most common drugs all over the world. The objective of this work was to construct a systematic database for drug discovery based on natural products separated from CVD-related medicinal herbs and to research on action mechanism of herb medicines.

**Description:** The cardiovascular disease herbal database (CVDHD) was designed to be a comprehensive resource for virtual screening and drug discovery from natural products isolated from medicinal herbs for cardiovascular-related diseases. CVDHD comprises 35230 distinct molecules and their identification information (chemical name, CAS registry number, molecular formula, molecular weight, international chemical identifier (InChI) and SMILES), calculated molecular properties (AlogP, number of hydrogen bond acceptor and donors, etc.), docking results between all molecules and 2395 target proteins, cardiovascular-related diseases, pathways and clinical biomarkers. All 3D structures were optimized in the MMFF94 force field and can be freely accessed.

**Conclusions:** CVDHD integrated medicinal herbs, natural products, CVD-related target proteins, docking results, diseases and clinical biomarkers. By using the methods of virtual screening and network pharmacology, CVDHD will provide a platform to streamline drug/lead discovery from natural products and explore the action mechanism of medicinal herbs. CVDHD is freely available at <http://pkuxxj.pku.edu.cn/CVDHD>.

**Keywords:** Cardiovascular disease, Drug discovery, Network pharmacology, Molecular docking, Virtual screening, Herbal formula, Natural products, Medicinal herbs, Traditional Chinese medicine

## Background

Natural products have been an important source for drug/lead discovery [1,2]. More than half of FDA-approved drugs are natural products or derivatives [3]. The most important approach in current drug discovery is through screening. However, traditional combinatorial chemistry libraries generally show limited structural diversity [4]. Through the natural selection process, natural products have vast diversities both in chemical space and pharmacological space [5-7]. High-throughput screening (HTS) *in silico* is used to reduce the number of chemicals to be tested *in vitro* and *in vivo*. HTS based on molecular docking is an efficient

approach to identify chemicals that could fit into the active site of target proteins. Therefore the virtual screening based on natural product database is a promising approach for drug discovery, especially for complex diseases such as cardiovascular disease.

Cardiovascular disease is regarded to be the main cause of death worldwide [8]. As a complex disease, CVD is the consequence of multiple pathogenic factors and reflects the altered interactions of many interconnected genes and gene products [9]. How to effectively and efficiently reverse these inappropriate interactions in a sick state is a critical problem. However, most drugs for CVD were designed to target a specific target and cannot be very effective [10-12]. In contrast, drugs which target multiple targets will have maximal efficacy and minimal adverse effects [13]. Polypharmacology and network pharmacology are useful approaches for understanding the mechanism and evaluating the

\* Correspondence: [lirongc@pku.edu.cn](mailto:lirongc@pku.edu.cn); [xiaojxu@pku.edu.cn](mailto:xiaojxu@pku.edu.cn)  
Beijing National Laboratory for Molecular Sciences, State Key Lab of Rare Earth Material Chemistry and Applications, College of Chemistry and Molecular Engineering, Peking University, Room A817, No.202, Chengfu Road, Beijing, Haidian District 100871, P. R. China

efficacy of drugs at systems level and could aid in design and development of drugs with higher success rate [14-18]. Meanwhile, biological pathways of diseases or biological process are important networks and usually associate with clinical biomarkers [19]. By investigating the effects of compounds on the biological pathway network, researchers can easily evaluate the potency of compounds and thus move basic biological discoveries into the clinic applications.

The herbal medicines which contain biologically active natural products for CVD have been used for thousands of years in China. However, the ingredients are too complex (usually hundreds of compounds) and the biological mechanisms of herbs are not yet understood clearly [14,16,18,20]. Along with the progress of network pharmacology and the explosive growth of biomedical data, the analysis of action mechanisms of medicinal herbs at the systems level becomes possible [5,15,21-23]. Therefore, we constructed the cardiovascular disease herbal database (CVDHD) that was implemented based on natural products-target proteins interactions and integration of multi-level data to promote the drug discovery from natural products and explore the molecular basis of polypharmacology of medicinal herbs for CVD. CVDHD will also provide an efficient platform for research on traditional Chinese medicine (TCM) and translational research in complementary and alternative medicine.

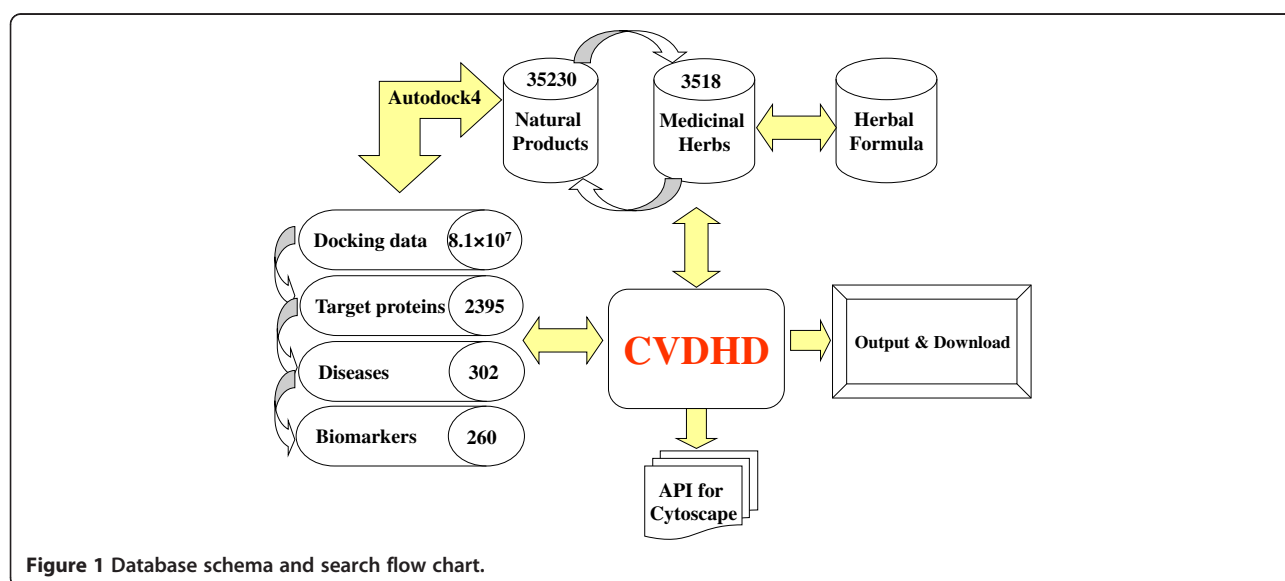
### Construction and content

CVDHD comprises six data entities covering medicinal herbs, natural products, target proteins, docking results between all molecules and target proteins, diseases and clinical biomarkers (Figure 1). The medicinal

herbs (Chinese name, Latin name and pharmacological effects) for cardiovascular-related diseases were collected from Chinese Herbalism [24] and Chinese Pharmacopoeia [25]. The structures and identification information (chemical name, CAS registry number, molecular formula, molecular weight and information of references) of natural products contained in the medicinal herbs were retrieved from CHDD [26] and UNPD [5] which were both developed by our lab in recent years. The absolute configuration of each molecule was generated by Open Babel [27] and the duplicates were deleted according to InChIKey. The molecular properties (AlogP, number of hydrogen bond acceptor and donors, etc.) were calculated by Discovery Studio.

There were two main groups of target proteins in CVDHD. One was the drug targets retrieved from Drug-Bank [28], another was the other proteins of human. Each protein had X-ray or NMR ligand-protein complex structures in RCSB protein data bank (<http://www.rcsb.org/pdb/home/home.do>). These structures were downloaded and treated to be suitable for molecular docking by Autodock4 [29] according to the procedure described previously [5]. For each protein, the binding site was defined as a  $40 \times 40 \times 40$  Å cube centered on the occupied space of the original ligand with a spacing of 0.375 Å between the grid points. The parameters of autodock4 were listed in Additional file 1. The docking score of autodock4 was chosen to evaluate the binding affinity according to binding free energy.

The information of diseases and clinical biomarkers related to CVD was collected from KEGG [30], TTD [31] and literatures manually. Finally, the data was stored in a MySQL (5.0.45) database. CVDHD was



**Figure 1** Database schema and search flow chart.

implemented as a PHP-based web application which was deployed to an Apache Tomcat server (PHP 5.1.6 and HTTPD 2.2.3). CVDHD can be accessed via the internet. Moreover, the API for Cytoscape was reserved to meet the future demands. Therefore, CVDHD can be applied seamlessly to the network pharmacology analysis by using the network analysis software such as Cytoscape and CentiBin.

## Utility and discussion

### Database interface

CVDHD had a powerful and intuitive web interface. The web pages were divided into five sections (Home, Introduction, Browse, Search and Download). A concise retrieval system for natural products was available on the homepage. User can search exactly by its identification information or get a group of compounds by the Latin name of medicinal herb. The introduction page described the database design and main functions. The browse pages contained five lists of medicinal herbs, molecules, proteins, diseases and biomarkers. The various data retrievals of drug discovery and network pharmacology can be achieved on search pages. The download page contained several links for download.

### Lead compound discovery

The lead generation is a crucial step in drug discovery [32]. Nowadays, HTS has become the major paradigm for lead discovery from chemical libraries [32,33]. Lipinski's "rule of five" (molecular weight less than 500 Da, hydrogen bond acceptors less than 10, hydrogen bond donors less than 5, octanol-water partition coefficient less than five) [34] was a simple but powerful rule to estimate the possibility of converting a compound into a drug. CVDHD can easily screen out the molecules which obey "rule of five" or other customized thresholds. For example, user can retrieve the potentially bioactive compounds from *Panax ginseng* (one of the most important medicinal herbs used as a tonic for restoration [35]) after entering the Latin name of this herb and thresholds of molecular properties on the advanced search webpage. The output of this query will be a table containing the molecular information and download link of each compound.

Moreover, if the target proteins of a disease were clear, the receptor-based drug design will accelerate the progress of drug discovery [36]. The HTS results between all molecules and all target proteins were stored in CVDHD and user can access data by simply typing four parameters: the Uniprot accession number of the target protein, the thresholds of docking score ( $pK_i$ ), the top percentage of rank of docking score and the logic value of whether the docking score of compounds was higher than that of original ligand in the X-ray or NMR

structure of the target protein (0 or 1). It may save a lot of time for users in the process of virtual screening.

### Network pharmacology research

Complex diseases such as CVD and diabetes are caused by a variety of genetic and environmental factors [37]. The modern drugs (magic bullets) which act on single target cannot treat complex diseases effectively [38-41]. Therefore, Hopkins AL [15,22] proposed network pharmacology to explore the action mechanism of drugs in the context of biological networks such as biological pathway, gene regulatory network, protein-protein interaction network, drug-target network, etc. The drug-target network (DTN) and its derivatives (drug-drug network (DDN) and target-target network (TTN)) could have important implications in understanding the mechanism and harnessing the vast amount of data from HTS [42-45].

The drug-target network was constructed according to the various binding data between molecules and target proteins. However, there were only a small portion (less than 2%) of natural products which binding data has been reported [5]. Thus, CVDHD used molecular docking to calculate the binding affinity between all molecules and 2395 target proteins. Based on CVDHD, comprehensive research on the mechanism of the medicinal herbs ranging from the level of herbal formula to protein-compound interaction can be achieved. Moreover, user can retrieve all target proteins relating to a pathway of a disease or biological process and finally find out the potential interactions between molecules and proteins to construct DTN, DDN and TTN to unveil the mechanism of TCM. CVDHD may bridge the gap between TCM and modern drugs based on system-level analysis.

### Case study: platelet aggregation pathway based lead discovery

Platelet aggregation plays an important role in arterial thrombosis in coronary heart disease [46]. How to regulate the platelet aggregation of CVD patients is one of the keys of successful treatment. The pathway of platelet aggregation comprised nineteen target proteins [19]. If a compound can block one of these proteins, the platelet aggregation will be inhibited to some extent. That is, the more inhibition of targets, the more effective the compound would be [19,47].

We retrieved the potentially active compounds from CVDHD with the following parameters: the thresholds of docking score, the top percentage of rank of docking score were set to 7.0 and 0.10, respectively. Meanwhile, the docking score of compounds should be higher than that of original ligand. The drug-target network (Additional file 2: Table S1) was constructed by linking the compound and target protein if the docking score exceeded the

**Table 1 Top potential compounds for inhibition of platelet aggregation according to degree centrality of drug-target network**

Rank	Compound	Degree	CAS NO.	Chemical name
1	CVDHD016293	6	62218-13-7	(+)- $\alpha$ -viniferin
2	CVDHD027110	6	N/A	hypericinate
3	CVDHD007020	6	54352-30-6	tricotonyltetramine
4	CVDHD012997	6	N/A	blumeanine
5	CVDHD001236	5	N/A	kadlongilactone B
6	CVDHD012285	5	225662-66-8	hypericin radical cation
7	CVDHD002703	5	50838-55-6	trisjuglone
8	CVDHD032784	5	55954-61-5	pseudohypericine
9	CVDHD029497	5	7034-04-0	tricotonyltetramin
10	CVDHD002390	4	N/A	20(R)-21,24-Cyclo-3beta,25-dihydroxydammar-23(24)-en-21-one
11	CVDHD014854	4	N/A	longipedlactone F
12	CVDHD034764	4	N/A	celastrolone B alpha
13	CVDHD010515	4	123522-98-5	isowithametelin
14	CVDHD014739	4	N/A	eremodimer B
15	CVDHD025068	4	N/A	bismorphine B
16	CVDHD017384	4	1253379-22-4	dievodiamine
17	CVDHD021374	4	N/A	lycochinine B
18	CVDHD025464	4	511-98-8	solasodanol
19	CVDHD019403	4	129748-10-3	incarvillateine
20	CVDHD019431	4	548-04-9	hypericin
21	CVDHD029988	4	129225-31-6	aminopropylcanavalmine
22	CVDHD009071	4	77646-14-1	3 $\beta$ ,17 $\alpha$ -Cinchophylline
23	CVDHD015834	4	N/A	bismorphine A
24	CVDHD011497	4	N/A	dihydro-18,19 3 $\beta$ ,17 $\beta$ cinchophylline

threshold values. Cytoscape 2.8.3 [48] was adopted to draw the DTN and calculate the degree and betweenness centrality of each node (compound or target). The larger degree and the higher betweenness, the more effect on the inhibition of platelet aggregation of that compound will be according to the network theory. Table 1 listed the top rank of potential compounds which may be the lead compounds for inhibition of platelet aggregation for further study. However, it is a simple model to use degree centrality to evaluate the efficacy of natural products. Users can use other models to predict the efficacy in the context of drug-target network.

### Conclusions

CVDHD is aimed to integrate medicinal herbs, natural products, CVD-related target proteins, docking results, diseases and clinical biomarkers to be a comprehensive database for drug discovery from natural products isolated from medicinal herbs. First, CVDHD is a chemical library of natural products and ready for virtual screening. All 3D structures of natural products and

information of binding site of target proteins can be accessed on the website. Second, drug/lead discovery for single target or a group of proteins related to a disease or biological process can be achieved. Finally, it is also a research platform for network pharmacology of medicinal herbs and TCM. The molecule and herb search can be applied to get compounds contained in herbs. Moreover, CVDHD is useful for study of network pharmacology of CVD-related natural products. The herbs contain a variety of ingredients and the compounds would interact with multiple cellular targets. By identifying the associations between bioactive compounds and cellular target proteins, CVDHD may bridge the gap between the systems level (herbal formula or medicinal herb) and molecular biology (compounds and proteins).

### Availability and requirements

CVDHD is freely available at <http://pkuxxj.pku.edu.cn/CVDHD> and there are no restrictions for academic use. The database will be updated annually.

## Additional file

**Additional file 1: Autodock4 parameters.**

**Additional file 2: Table S1.** Drug-target network of potential lead compounds and target proteins of platelet aggregation pathway.

### Abbreviations

CVD: Cardiovascular disease; NP: Natural product; DTN: Drug-target network; DDN: Drug-drug network; TTN: Target-target network; HTS: High-throughput screening; TCM: Traditional Chinese medicine.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

XJX and LRC conceived the study. JYG and YSG constructed the database, performed computational analyses and interpreted the results with the help of GY. JYG designed the website and wrote the manuscript. All authors read and approved the final manuscript.

### Acknowledgements

This work was financially supported by National Key Special Project of Science and Technology for Innovation Drugs (Grant No. 2012ZX09501001-004 and 2013ZX09402202) and the Enterprise Academician Workstation of Jiangsu Province, PR China (Grant No. BM2011027). The calculations were performed on TianHe-1(A) at National Supercomputer Center in Tianjin, PR China.

Received: 30 July 2013 Accepted: 12 December 2013

Published: 18 December 2013

### References

- Newman DJ, Cragg GM: Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod* 2012, **75**(3):311–335.
- Harvey AL: Natural products in drug discovery. *Drug Discov Today* 2008, **13**(19–20):894–901.
- Chin YW, Balunas MJ, Chai HB, Kinghorn AD: Drug discovery from natural sources. *AAPS J* 2006, **8**(2):E239–E253.
- Hong J: Role of natural product diversity in chemical biology. *Curr Opin Chem Biol* 2011, **15**(3):350–354.
- Gu JY, Gui YS, Chen LR, Yuan G, Lu HZ, Xu XJ: Use of natural products as chemical library for drug discovery and network pharmacology. *PLoS One* 2013, **8**(4):e62839.
- Lopez-Vallejo F, Giulianotti MA, Houghten RA, Medina-Franco JL: Expanding the medicinally relevant chemical space with compound libraries. *Drug Discov Today* 2012, **17**(13–14):718–726.
- Yongye AB, Waddell J, Medina-Franco JL: Molecular scaffold analysis of natural products databases in the public domain. *Chem Biol Drug Des* 2012, **80**(5):717–724.
- Cases M, Mestres J: A chemogenomic approach to drug discovery: focus on cardiovascular diseases. *Drug Discov Today* 2009, **14**(9–10):479–485.
- Barabasi AL, Gulbahce N, Loscalzo J: Network medicine: a network-based approach to human disease. *Nat Rev Genet* 2011, **12**(1):56–68.
- Lu JJ, Pan W, Hu YJ, Wang YT: Multi-target drugs: the trend of drug research and development. *PLoS One* 2012, **7**(6):e40262.
- Petrelli A, Giordano S: From single- to multi-target drugs in cancer therapy: when aspecificity becomes an advantage. *Curr Med Chem* 2008, **15**(5):422–432.
- Csermely P, Agoston V, Pongor S: The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol Sci* 2005, **26**(4):178–182.
- Mencher SK, Wang LG: Promiscuous drugs compared to selective drugs (promiscuity can be a virtue). *BMC Clin Pharmacol* 2005, **5**:3.
- Zhao J, Jiang P, Zhang WD: Molecular networks for the study of TCM Pharmacology. *Brief Bioinform* 2010, **11**(4):417–430.
- Hopkins AL: Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol* 2008, **4**(11):682–690.
- Tao WY, Xu X, Wang X, Li BH, Wang YH, Li Y, Yang L: Network pharmacology-based prediction of the active ingredients and potential targets of chinese herbal radix curcumae formula for application to cardiovascular disease. *J Ethnopharmacol* 2013, **145**(1):1–10.
- Zhang AH, Sun H, Yang B, Wang XJ: Predicting new molecular targets for rhein using network pharmacology. *BMC Syst Biol* 2012, **6**:20.
- Li J, Lu C, Jiang M, Niu XY, Guo HT, Li L, Bian ZX, Lin N, Lu AP: Traditional chinese medicine-based network pharmacology could lead to New multicomponent drug discovery. *Evid-Based Compl Alt Med* 2012, **2012**:149762.
- Gu JY, Li Q, Chen LR, Li YY, Hou TJ, Yuan G, Xu XJ: Platelet aggregation pathway network-based approach for evaluating compounds efficacy. *Evid-Based Compl Alt* 2013, **2013**:425707.
- Li S, Zhang B: Traditional chinese medicine network pharmacology: theory, methodology and application. *Chin J Nat Med* 2013, **11**(2):110–120.
- Berger SI, Iyengar R: Network analyses in systems pharmacology. *Bioinformatics* 2009, **25**(19):2466–2472.
- Hopkins AL: Network pharmacology. *Nat Biotechnol* 2007, **25**(10):1110–1111.
- Paolini GV, Shapland RHB, van Hoom WP, Mason JS, Hopkins AL: Global mapping of pharmacological space. *Nat Biotechnol* 2006, **24**(7):805–815.
- Chinese Herbalism Editorial Board, State Administration of Traditional Chinese Medicine of the People's Republic of China: *Chinese Herbalism*. Shanghai: Shanghai scientific & Technical Publishers; 1999.
- Chinese Pharmacopoeia Commission: *Chinese Pharmacopoeia*. Beijing: China Medical Science Press; 2010.
- Qiao XB, Hou TJ, Zhang W, Guo SL, Xu SJ: A 3D structure database of components from chinese traditional medicinal herbs. *J Chem Inf Comp Sci* 2002, **42**(3):481–489.
- O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR: Open label: an open chemical toolbox. *J Cheminformatics* 2011, **3**:33.
- Knox C, Law V, Jewison T, Liu P, Ly S, Frolkis A, Pon A, Banco K, Mak C, Neveu V, et al: DrugBank 3.0: a comprehensive resource for 'Omics' research on drugs. *Nucleic Acids Res* 2011, **39**:D1035–D1041.
- Jiang XH, Kumar K, Hu X, Wallqvist A, Reifman J: DOVIS 2.0: an efficient and easy to use parallel virtual screening tool based on AutoDock 4.0. *Chem Cent J* 2008, **2**:18.
- Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M: KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 2012, **40**(D1):D109–D114.
- Zhu F, Shi Z, Qin C, Tao L, Liu X, Xu F, Zhang L, Song Y, Liu XH, Zhang JX, et al: Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery. *Nucleic Acids Res* 2012, **40**(D1):D1128–D1136.
- Keseru GM, Makara GM: Hit discovery and hit-to-lead approaches. *Drug Discov Today* 2006, **11**(15–16):741–748.
- Bleicher KH, Bohm HJ, Muller K, Alanine AI: Hit and lead generation: beyond high-throughput screening. *Nat Rev Drug Discov* 2003, **2**(5):369–378.
- Lipinski CA, Lombardo F, Dominy BW, Feeney PJ: Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 1997, **23**(1–3):3–25.
- Shin HR, Kim JY, Yun TK, Morgan G, Vainio H: The cancer-preventive potential of Panax ginseng: a review of human and experimental evidence. *Cancer Causes Control* 2000, **11**(6):565–576.
- Marrone TJ, Briggs JM, McCammon JA: Structure-based drug design: computational advances. *Annu Rev Pharmacol* 1997, **37**:71–90.
- Cho DY, Kim YA, Przytycka TM: Chapter 5: network biology approach to complex diseases. *PLoS Comput Biol* 2012, **8**(12):e1002820.
- Gu JY, Yuan G, Zhu YH, Xu XJ: Computational pharmacological studies on cardiovascular disease by Qishen Yiqi Diwan. *Sci China Ser B* 2009, **52**(11):1871–1878.
- Pujol A, Mosca R, Farres J, Aloy P: Unveiling the role of network and systems biology in drug discovery. *Trends Pharmacol Sci* 2010, **31**(3):115–123.
- Drews J: Case histories, magic bullets and the state of drug discovery. *Nat Rev Drug Discov* 2006, **5**(8):635–640.
- Morphy R, Kay C, Rankovic Z: From magic bullets to designed multiple ligands. *Drug Discov Today* 2004, **9**(15):641–651.
- Janga SC, Tzakos A: Structure and organization of drug-target networks: insights from genomic approaches for drug discovery. *Mol Biosyst* 2009, **5**(12):1536–1548.
- Vogt I, Mestres J: Drug-target networks. *Mol Inform* 2010, **29**(1–2):10–14.
- Yildirim MA, Goh KI, Cusick ME, Barabasi AL, Vidal M: Drug-target network. *Nat Biotechnol* 2007, **25**(10):1119–1126.
- Mestres J, Gregori-Puigjane E, Valverde S, Sole RV: The topology of drug-target interaction networks: implicit dependence on drug properties and target families. *Mol Biosyst* 2009, **5**(9):1051–1057.



46. O'Donnell CJ, Larson MG, Feng DL, Sutherland PA, Lindpaintner K, Myers RH, D'Agostino RA, Levy D, Toftler GH: **Genetic and environmental contributions to platelet aggregation - the framingham heart study.** *Circulation* 2001, **103**(25):3051–3056.
47. Li Q, Li XD, Li CH, Chen LR, Song J, Tang YL, Xu XJ: **A network-based multi-target computational estimation scheme for anticoagulant activities of compounds.** *PLoS One* 2011, **6**(3):e14774.
48. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T: **Cytoscape 2.8: new features for data integration and network visualization.** *Bioinformatics* 2011, **27**(3):431–432.

doi:10.1186/1758-2946-5-51

**Cite this article as:** Gu et al.: CVDHD: a cardiovascular disease herbal database for drug discovery and network pharmacology. *Journal of Cheminformatics* 2013 **5**:51.

Publish with **ChemistryCentral** and every scientist can read your work free of charge

*“Open access provides opportunities to our colleagues in other parts of the globe, by allowing anyone to view the content free of charge.”*

W. Jeffery Hurst, The Hershey Company.

- available free of charge to the entire scientific community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
<http://www.chemistrycentral.com/manuscript/>



**ChemistryCentral**