

A Structure-Similarity-Based Software for the Cardiovascular Toxicity Prediction of Traditional Chinese Medicine

Feng Lu, Qiong Gu, Ruibo Wu and Jun Xu*

Research Center for Drug Discovery & Institute of Human Virology, School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, China; Jun Xu - Email: junxu@biochemomes.com; *Corresponding author

Received December 22, 2011; Accepted January 07, 2012; Published January 20, 2012

Abstract:

Although natural medicines are generally considered to be safer than synthetic medicines, some reports on toxicity of many herb drugs, for example Traditional Chinese Medicine (TCM), attract more and more common public attention recently. The reasons for these cardiovascular toxic effects are not always satisfactory. Not all these clinical symptoms can be observed regularly. Cardiovascular toxicity concern is one of the hindrances for TCM to enter international markets. Therefore, we create web-based software for people to evaluate TCM cardiovascular toxicity. The software, TCM Cardiovascular Toxicity Prediction (TCMCardioTox), is based on structure similarity search algorithm (<http://rcdd.sysu.edu.cn:8080/home.aspx>). When a user enter a TCM name, the system interprets the name as a number of active component structures, and computes the structure similarities of the components against a number of known toxiphores. If one of the similarity is greater than a given similarity threshold, the TCM will be reported as an alert TCM for medical uses.

Background:

TCM are usually used in the form of an aqueous decoction. An herb soup has many chemical components which are not necessarily well understood. A single purified TCM compound isolated from a TCM can be either too toxic or without therapeutic efficacy [1]. The toxicities of a number of TCM have been reported in many reports [2-9]. However, it is impossible to experimentally measure toxicities for all TCM due to the costs. Developing *in silico* approach to estimate the cardiovascular toxicity of a TCM is an only practical solution for pushing a TCM to international markets.

Principle for the cardiovascular toxicity prediction of TCM:

TCMCardioTox is a web based program implemented in ASP.net and C# on a Microsoft server. The client program of TCMCardioTox is an Internet browser, such as, Internet Explorer (IE), with the installation of Java Run Environment. TCMCardioTox is actually a web-based knowledge-based

system, which collects all the chemical structures that have been identified as human cardiovascular toxins. The main principle for the cardiovascular toxicity prediction is that if one of the chemical components in a TCM has the structural similarity against one of the known cardiovascular toxins in the database of TCMCardioTox system, the TCM will be alerted for cardiovascular toxicity.

TCMCardioTox consists of three components: (1) TCM cardiovascular adverse reactions reported in literatures; (2) TCM and corresponding known chemical structures; (3) chemical structure similarity algorithm. The chemical structure data and cardiovascular toxicity data are stored in a Microsoft Access Database. The chemical structure data are saved in the form of SD format.

When TCMCardioTox receives a TCM query, it will translate the TCM name into a set of chemical structures known as its

compositions. These structures will be sent to a software module which checks them against a set of known cardiovascular toxins by substructure matching. At least one of the structures is matched; the TCM will be reported as potential a cardiovascular toxin. Otherwise, these structures will be sent to structural similarity algorithm to calculate structure similarities between the TCM chemical structures and the known cardiovascular toxic structures in the TCM CardioTox knowledge base. If one of the structure similarities is greater than a given threshold (such as, 80%), the TCM will be reported for cardiovascular toxicity. If a TCM is reported for cardiovascular toxicity, TCMCardioTox will reported the reasoning details in a spreadsheet, such as, the referenced structures, literatures, and experimental descriptions (if it is available) etc. However, if the spreadsheet is not displayed, it means that the TCM is not found being cardiovascular toxic.

Features Functions:

The principle of predicting cardiovascular toxicity in TCMCardioTox is to figure out if a TCM contains compounds that are structurally similar to known cardiovascular toxic molecules or contains key toxic substructures. The more cardio-toxic molecules existing in a TCM, the TCM will be more potentially cardiovascular toxic.

When users retrieve a TCM in TCMCardioTox, similarity search will be done in a few seconds. The results depend on the similarity threshold. Usually, the threshold can be set to 85%. Higher similarity threshold will result in fewer hits (the number of cardiovascular toxicity hits means more risky). If a user wants to see more possible hits, he/she can lower the similarity threshold. The detailed principle is articulated in (Figure 1).

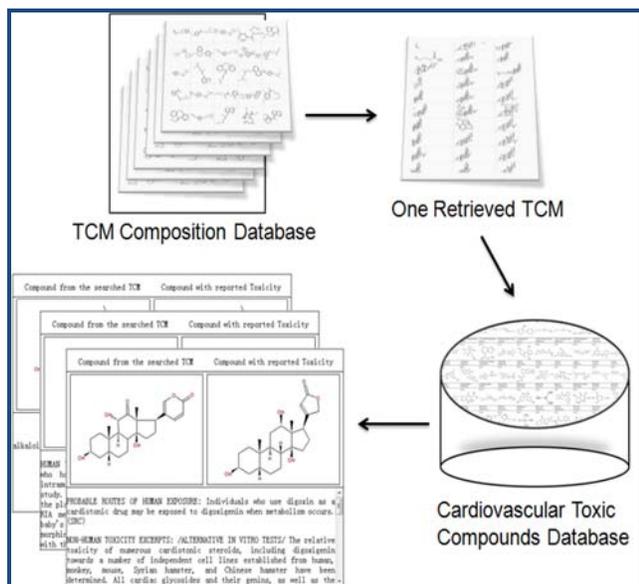


Figure1: Flowchart for the principle of TCM cardiovascular toxicity. First, a TCM is translated in to a set of composition structures; then the similarity algorithm computes the structure similarities between TCM structures and the structures of cardiovascular toxic compounds; meanwhile, a substructure match algorithm checks if the TCM structures contain pre-defined cardiovascular toxic fragments (toxin

substructures); finally, the system reports the results based upon the similarity threshold and the substructure matching.

User Interface:

TCM CardioTox is a web based system, which can be accessed through an internet browser at <http://rcdd.sysu.edu.cn:8080/home.aspx>.

A query for TCMCardioTox can be a string for a TCM name, a compound structure drawing, or a SDF file for a number of chemical structures that are in a TCM herb as active components.

For example, *Gambirplant* stem is a TCM herb. TCMCardioTox predicts this herb that has cardiovascular toxicity for arrhythmia, hypotension and even carcinogenicity. Animal model experiments showed that this herb could cause arrhythmia. Some TCM cardiovascular adverse effects [2-9] testing results are listed in **Table 1**.

Discussion:

Current version of TCMCardioTox predicts cardiovascular toxicity based only structure similarity to toxins; this is a ligand based prediction. The advantage of this approach is simple and fast. The disadvantage is that it cannot explain the mechanism of actions.

The correct predictions rely on the coverage of the knowledge base in the system and the structure similarity algorithm. TCMCardioTox allows an administrative user to expand the knowledge base. But, the similarity algorithm is not changeable at this time. Currently, the structure similarity algorithm is based upon a simple structural fragment comparison using Tanimoto schem. In the future, three dimensional structure similarities can be implemented for alternative solutions.

In order to reach more precise prediction, the target structure based models are necessary. This requires the knowledge of mechanism of actions. The structure-based models demands for high performance computing, therefore, the new version of TCMCardioTox will need parallel computing technology, such as CUDA-based GPU technology.

Acknowledgment:

This work was funded in part of the major scientific and technological special project by the Ministry of Science and Technology of China (2009DFA31530, 2010ZX09102-305), Guangdong Recruitment Program of Creative Research Groups, the Fundamental Research Funds for the Central Universities, the National Natural Science Foundation of China (No.81001372, 81173470). The editors are appreciated for their valuable comments and continuously encouragement.

References:

- [1] Chan K *et al. Chemosphere*. 2003 **52**: 1361 [PMID: 12867165].
- [2] *MMWR Morb Mortal Wkly Rep*. 1993 **42**: 633 [PMID: 8350855].
- [3] Chou WC *et al. Int J Cardiol*. 1995 **49**: 173 [PMID: 7628889].
- [4] Ernst E *et al. Can J Cardiol*. 2003 **19**: 18 [PMID: 12813616].
- [5] Phua DH *et al. Clin Toxicol (Phila)*. 2008 **46**: 1067 [PMID: 18763152].

- [6] Valli G *et al.* *J Am Coll Cardiol.* 2002 **39**: 1083 [PMID: 11923030].
- [7] Villegas JF *et al.* *Heart Dis.* 2001 **3**: 169 [PMID: 11975788].
- [8] Cheng TO *et al.* *Int J Cardiol.* 2007 **121**: 9 [PMID: 17363091].
- [9] Izzo AA *et al.* *Int J Cardiol.* 2005 **98**: 1 [PMID: 15676159].

Edited by P Kanguane

Citation: Lu *et al.* *Bioinformation* 8(2): 110-113 (2012)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

Supplementary material:

Table 1. TCMCardioTox prediction examples and explanations

English Name of TCM	Reported Cardiovascular Adverse Effects	Predicted Cardiovascular Toxicity by TCMCardioTox
Redroot Gromwell	Arrhythmia	Hypotension, Methemoglobinemia, Arrhythmias, Bleeding
Hellebore	Arrhythmia, Hypotension	Hypotension, Cardiogenic shock, Hypovolemia, Bradycardia
Dan-shen	Arrhythmia, Over anticoagulation	Bleeding, Cardiovascular defects, Arrhythmia
Safflower	Arrhythmia	Arrhythmia
Toad Secretion	Arrhythmia	Arrhythmia, Hypertension
Siberian Thorowax	Arrhythmia	Arrhythmia, Hypotension
Licorice	Arrhythmia	Pseudohyperaldosteronism, Thrombocytopenia
Chinese Angelica	Hypertension	Cardiac failure, Cardiovascular damage, Hypertension
Chinese Ephedra	Hypertension	Hypertension
Ginseng	Hypertension	Slight hypotensive effect, Hypocalcaemia, Pseudohyperaldosteronism
Forbes Wildginger	Hypotension	Hypotension
Common Threewingnut Root	Hypotension	Arrhythmias, Hypotension
Fangchi	Hypotension	Hypotension
Weeping Forsythia	Insufficient anticoagulation	Hemorrhage
Aromatic Turmeric	Over anticoagulation	Hypotension
Ginkgo Leaf	Over anticoagulation	Anticoagulation
Gastrodia	Allergic purpura	Methemoglobinemia, Hypotension