Method article

INPUT: An intelligent network pharmacology platform unique for traditional Chinese medicine

Xianhai Li a, Qiang Tang b, Fanbo Meng a, Pufeng Du c,*, Wei Chen a,b,*

a Innovative Institute of Chinese Medicine and Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu 611137, China
b School of Basic Medical Sciences, Chengdu University of Traditional Chinese Medicine, Chengdu 611137, China
c College of Intelligence and Computing, Tianjin University, Tianjin 300072, China

ABSTRACT

The application of network pharmacology has greatly promoted the scientific interpretation of disease treatment mechanism of traditional Chinese medicine (TCM). However, the data required by network pharmacology analysis were scattered in different resources. In the present work, by integrating and reorganizing the data from multiple resources, we developed the intelligent network pharmacology platform unique for traditional Chinese medicine, called INPUT (http://cbcb.cdutcm.edu.cn/INPUT/), for automatically performing network pharmacology analysis. Besides the curated data collected from multiple resources, a series of bioinformatics tools for network pharmacology analysis were also embedded in INPUT, which makes it become the first automatic platform able to explore the disease treatment mechanisms of TCM. With the built-in tools, researchers can also analyze their own in-house data and obtain the results of pivotal ingredients, GO and KEGG pathway, protein-protein interactions, etc. In addition, as a proof-of-principle, INPUT was applied to decipher the antidepressant mechanism of a commonly used antidepressant ingredient. In summary, INPUT is a powerful platform for network pharmacology analysis and will facilitate the researches on drug discovery.

© 2022 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

With more than 2,000 years of knowledge accumulation and clinical practice, traditional Chinese medicine (TCM) has become an important part of modern medicine [1]. The characteristics of TCM are multiple components, multiple targets, and multiple pathways, which coincide with the multiple factors and complex pathological characteristics of diseases [2,3]. In recent decades, owing to its accurate curative effect, small side effects, and low cost, TCM has been widely used in clinics and attracted worldwide attentions [4–6]. For example, the Huangqin decoction has been shown excellent effects in curing diseases. Ephedrine extracted from the TCM herb Ma Huang (Herba ephedrae) has been used to treat asthma [9]. Artemisinin from the TCM herb QingHao (Herba Artemisiae annuae) is an anti-malaria drug and can reduce malaria mortality rates, and the discovery of artesiminin has been awarded the Nobel Prize in Physiology or Medicine in 2015 [10]. Although great achievements have been made, the disease treatment mechanisms of most TCM remain unclear, which greatly hindered its development and modernization.

To dig for treasures from TCM, the network pharmacology method was proposed, which analyzes the molecular associations between drugs and diseases in a system level through networks. In recent years, network pharmacology has been widely used in the discovery of active ingredients of drugs and TCM [11], the interpretation of treatment mechanisms of TCM [12], and the analysis of drug combinations and prescription compatibility [13]. As indicated in Web of Science, more than 1900 network pharmacology method-based works have been published since 2016. The applications of network pharmacology not only promoted the scientific interpretation of TCM [13], but also provided scientific and technological supports for drug development and rational clinical use of TCM [14]. Recently, the World Federation of Chinese Medicine Societies released the “Network Pharmacology Evaluation Method Guidance-Draft” [15], where the difficulties and challenges of network pharmacology were pointed out. One of the challenges
is that the data required by network pharmacology were scattered in different resources. Although a series of databases for network pharmacology have been developed, such as TCMSP (Traditional Chinese Medicine Systems Pharmacology) [16], BATMAN-TCM (a bioinformatics analysis tool for molecular mechanism of traditional Chinese medicine) [17], TCMD (traditional Chinese medicine integrative database) [18], and ETCM (The encyclopedia of traditional Chinese medicine) [19], the data collection and integration process are still time-consuming and laborious, which precludes the quick identification of ingredients and targets of TCM. In addition, most of the existing databases only focus on TCM or diseases rather than exploring the disease treatment mechanisms of TCM.

Considering the promising futures and challenges of network pharmacology, it’s necessary to develop a comprehensive and automatic platform for network pharmacology analysis. In the present work, by integrating the data from multiple resources, the intelligent network pharmacology platform unique for traditional Chinese medicine (INPUT) was developed. The current version of INPUT deposited 130 common diseases, 490 common herbs, 2158 active ingredients and 10,424 related genes, which can be freely accessed at https://cbcb.cdtcm.edu.cn/INPUT/. More importantly, the bioinformatics tools for mechanism analysis, protein–protein interaction (PPI) network analysis, screening of pivotal ingredients and hub genes were also integrated in INPUT, which can be directly used to perform network pharmacology analysis. An overview of data integration of the INPUT platform was shown in Fig. 1.

2. Materials and methods

INPUT integrates four kinds of data, namely herbs, active ingredients, targets and diseases, from various resources as summarized in Table 1.

![Table 1 Overview of the data in INPUT.](https://cbcb.cdtcm.edu.cn/INPUT/)

| Components                  | Data sources                                                                 | Number |
|-----------------------------|------------------------------------------------------------------------------|--------|
| Herbs                       | Pharmacopoeia of the People's Republic of China (2015 version)               | 490    |
| Active ingredients          | TCMSP (https://old.tcmsp-e.com/tcmsp.php)                                   | 2158   |
| TCM-related common diseases | TCM-related books                                                            | 130    |
| TCM-related targets         | TCMSP (https://old.tcmsp-e.com/tcmsp.php), BATMAN-TCM, SEA, Pharmmapper, Drugbank, GeneCards, CTD, OMIM, DisGeNET | 3208, 9639 |

2.1. The herb information

The INPUT database contains the general information of 490 commonly used traditional Chinese herbs, including the name, pinyin, English name, source and category of herbs. These information were derived from two resources, namely TCMSP and Pharmacopoea of the People's Republic of China (2015 version) [16].

2.2. The active ingredients information

The active ingredients of the herbs were collected from TCMSP. With the name of herbs as keywords and the thresholds of oral bioavailability (OB) ≥ 30% and drug likeness (DL) ≥ 0.18 [20], 2158 active ingredients were obtained. In order to allow users to

![Fig. 1. An overview of data integration of INPUT platform.](https://cbcb.cdtcm.edu.cn/INPUT/)
get more information about the active ingredients, their molecular name, pubchem CID, molecular formula, molecular weight, isomeric smiles, canonical smiles, InChI, InChI key, IUPAC name and structure were collected from the Pubchem database (https://pubchem.ncbi.nlm.nih.gov) [21].

### 2.3. The diseases information

By consulting TCM-related books, we screened out 200 common diseases treated by TCM. In order to make the diseases more reliable for network pharmacological analysis, we queried the 200 diseases in the Drugbank [22], GeneCards [23], CTD (Comparative Toxicogenomics Database) [24], OMIM (Online Mendelian Inheritance in Man) [25] and DisGeNET [26] databases, and removed 70 diseases based on the rule as described in Supplementary Fig. S1. Finally, 130 diseases were retained in the INPUT (Supplementary Table S1).

### 2.4. The targets information

#### 2.4.1. TCM-related targets

The targets of TCM active ingredients were from four well-known databases, namely TCMSp (https://tcmspw.com/), BATMAN-TCM (https://biocat.ncpsb.org/batman-tcm/), SEA (https://sea.bkslab.org/) and Pharmmapper (https://www.lilab-ecust.cn/pharmmapper/). Considering that the targets in TCMSp were from Drugbank, all targets collected from TCMSp were retained without setting cutoffs [27]. The targets from BATMAN-TCM and SEA were predicted based on a similarity-based method [17,28]. According to previous experiences, the parameters of BATMAN-TCM were set as score ≥ 20 and p-value ≤ 0.05 [29], and the parameter of SEA was set as Tanimoto score ≥ 0.8 [19]. Pharmmapper can be used to find the best mapping posture of all targets in PharmTargetDB according to the uploaded molecular structure, and outputs potential drug targets [30]. The parameter of Pharmmapper were set as score ≥ 0.9 [31]. Finally, the UniProt (Universal Protein Resource, https://www.uniprot.org/) was used to standardize the name of the targets [32].

#### 2.4.2. Disease-related targets

Disease-related targets were harvested from Drugbank (https://go.drugbank.com) [22], GeneCards (https://www.genecards.org) [23], CTD (https://ctdbase.org) [24], OMIM (https://omim.org) [25], and DisGeNET (https://www.disgenet.org) [26]. Although these databases provided scores for disease-related targets to indicate its reliability, their scores were in different ranges, which hindered the direct integration of data from different resources. Therefore, a normalization procedure was performed to set the scores in the range of 0 to 1 (Supplementary Material). Accordingly, the targets with the normalized score no less than 0.5 were retained in INPUT.

#### 2.4.3. Protein-Protein interaction (PPI)

The STRING (https://string-db.org/) database [33] is a platform for searching known and predicted protein–protein interactions. The interaction between proteins includes direct physical interactions and indirect functional correlations [33]. In the present work, we choose human protein–protein interactions in the STRING database as the background data for the PPI functional analysis.

### 3. Results

The INPUT platform was developed by using the python-based Django framework, and the web page is written in html. All programs in INPUT were written in Python (Version 3.8) and R (Version 4.0.3). At present, INPUT provides four mainly functional modules, namely Network Pharmacology, Enrichment Analysis, Data Query and Browse (Fig. 2).

#### 3.1. Network pharmacology analysis

In this module, by providing the name of herbs or active ingredients and diseases, the associations between herbs or active ingredients and diseases could be analyzed. Besides using the deposited data in INPUT, users can also submit their own data in this module.

By taking the intersection between the targets of active ingredients and diseases, their common targets were obtained. According to the relationships among herbs, active ingredients, common targets and diseases, the herbs-active ingredients-common targets-disease network can be obtained (Fig. 3A). In this network, the ingredients were sorted according to their degree values. Accordingly, the pivotal ingredients for treating diseases will be screened out based their degree values in the network.

In order to explore the mechanism of TCM or active ingredients in the treatment of diseases, the Gene Ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analysis on common targets were performed by using the clusterProfiler package [34]. The results of KEGG analysis (Fig. 3B) and GO analysis (Fig. 3C–3E) will facilitate researchers to decipher the mechanism of TCM in treating disease.

In addition, the PPI for common targets (Fig. 3F) could also be obtained by setting different confidence scores, such as low confidence (0.15), medium confidence (0.4), custom value (0.5), high confidence (0.7) and highest confidence (0.9) [33]. In the PPI network, the greater the degree value of the node, the more important the node is [35]. Accordingly, the hub genes can be screened out according to the degree value of the nodes in the network (Fig. 3G).

#### 3.2. Enrichment analysis

To facilitate researchers to analyze their own in-house data, the GO and KEGG pathway analysis function were also integrated in the ‘Enrichment Analysis’ module (Fig. 4). By submitting the gene symbols or the query file, the results of KEGG and GO functional enrichment analysis, PPI networks and hub genes of the query genes could be obtained (Fig. 3B–3G).

#### 3.3. Data query

In order to help users to collect the targets of herbs, ingredients and diseases, INPUT platform provides the ‘Data Query’ module. In this module, by selecting keyword and submitting the name of herbs, the name of diseases, or the name/Pubchem CID/InChI Key of ingredients, users can get the corresponding targets and the related information (Fig. 5). Both the internal and external links (such as TCMSp, Pubchem and NCBI) were provided as well. All the query results can be freely downloaded.

#### 3.4. Browse

In the “Browse” module, the data of the INPUT platform were classified into 4 categories, namely herb, ingredient, gene, and disease (Fig. 6A), and sorted in alphabetical order, respectively (Fig. 6B). Users can get the desired information by clicking on the name. Take clicking on the name of the herb Baibu as an example (Fig. 6C), its herb name, Chinese pinyin, English name, source plant, category, herb-related ingredients, and herb-related genes will be shown in a new page (Fig. 6D).
Fig. 2. The main modules of INPUT. INPUT includes four main modules, namely Network Pharmacology, Enrichment Analysis, Data Query and Browse.

Fig. 3. The results of network pharmacology analysis. (A) The herb-active ingredient-common targets-disease network. Nodes of different colors indicate different categories, blue is for herb, red for common targets, yellow for disease, and green for active ingredients. The size of the node is related to its importance. (B) The enriched KEGG pathways. The X-axis represents GeneRatio of the KEGG term. The Y-axis represents the name of KEGG terms. The color of the circle indicates the adjusted p value; the size of the circle is related to the number of genes in the term. (C-E) The enriched GO terms, namely in biological processes (C), molecular functions (D) and cellular components (E). The X-axis represents the number of enriched genes. The Y-axis is the related terms. (F) The PPI network for common targets. (G) Hub genes obtained from PPI network. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
3.5. Case study

Recently, Liu et al. have deciphered the antidepressant mechanism of Xiaoyaosan (XYS) by integrating metabolic experiments and network pharmacology [36]. To demonstrate the reliability and effectiveness of INPUT, it was applied to illustrate the antidepressant mechanism of XYS. The formula of XYS consists of Radix Angelicae Sinensis (Danggui), Radix Bupleuri (Chaihu), Radix Paeoniae Alba (Baishao), Poria (Fuling), Rhizoma Atractylodis Macrocephala (Baizhu), Radix Glycyrrhizae (Gancao), Herba Menthae Haplocalycis (Bohe) and Rhizoma Zingiberis Recens (Shengjiang).

By entering the name of these herbs and the disease ('depression') in the Network Pharmacology module, the herb-active ingredient-common targets-disease network (Supplementary Fig. 2 and Table S2), the degree value of active ingredients (Table S3), GO and KEGG enrichment analysis (Supplementary Fig. 3 and Table S4), PPI network (Supplementary Fig. 4A and Table S5) and hub targets (Supplementary Fig. 4B) were obtained. All these results were automatically yielded by the INPUT platform.

According to the methods in section 2.2, the screened out pivotal ingredients and hub targets of XYS for treating depression were summarized in Table 2, where all the pivotal ingredients and hub targets reported in Liu et al. were also provided. As shown in Table 2, all of the 7 pivotal ingredients and three of the 4 hub targets could be screened out by using INPUT. Most of the identified ingredients and hub targets are all top ranked.

By comparing the results of KEGG enrichment analysis, it was found that the three significant KEGG pathways (AGE-RAGE signaling pathway in diabetic complications, Neuroactive ligand-receptor interaction and Hepatitis B) reported in Liu et al.’s work were also obtained by INPUT (Supplementary Fig. 3 and Table S4). The GO analysis results (Table S4) demonstrated that INPUT could find the 12 BP terms, 12 MF terms and 10 CC terms of those reported in Liu et al.'s work. Although TCMSP, BATMAN-TCM, SEA and Pharmmapper are the famous databases of network pharmacology, none of them were able to finish the procedures for decoding the antidepressant mechanisms of XYS. These results demonstrate the reliability and usefulness of INPUT for identifying ingredients and hub targets.

4. Discussion

Network pharmacology plays key roles in exploring the disease treatment mechanisms of TCM. However, the data required by the network pharmacology analysis are scattered in different resources. In the present work, by integrating the data from multiple resources, we proposed an automatic platform, called INPUT, for network pharmacology analysis. To the best of our knowledge, INPUT is the first platform able to explore the mechanisms of TCM in the treatment of diseases based on network pharmacology. The efficiency and performance of INPUT in practical applications have been demonstrated by analyzing the antidepressant mechanism of XYS.

Although the data deposited in INPUT platform is incomplete and only the common diseases related to TCM were integrated, it can be used to perform network pharmacology analysis for the in-house data by using the integrated bioinformatics tools. In order to obtain the desired results, users only need to provide the targets-ingredients association, targets-ingredients-herbs association, and targets-disease association information.

Taken together, the present work proposed an automatic platform for network pharmacology analysis. The platform will be continuously updated and improved. We believe that INPUT will vigorously promote the scientific interpretation of TCM and drug discovery.
Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to express their appreciation to Dr. Juanjuan Kang for her constructive suggestions in developing the platform. This work was supported by National Nature Science Foundation of China (No. 31771471 and 61872268).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2022.03.006.

References

[1] Xu Q, Bauer R, Hendry BM, Fan TP, Zhao Z, Duez P, et al. The quest for modernisaton of traditional Chinese medicine. BMC Complement Altern Med 2013;13:132.
[2] Wang L, Ding T, Gong WL, Yang CH, Liu F. Effective components of traditional Chinese medicine for regulating TGF-Beta1/Smads signaling pathway in hepatic fibrosis. Zhongguo Zhong Yao Za Zhi 2019;44(4):666–74.
[3] Xu Q, Guo Q, Wang CX, Zhang S, Wen CB, Sun T, et al. Network differentiation: A computational method of pathogenesis diagnosis in traditional Chinese medicine based on systems science. Artif Intell Med 2021;118:102134.
[4] Sham TT, Chan CO, Wang YH, Yang JM, Mok DK, Chan SW. A review on the traditional Chinese medicinal herbs and formulae with hypolipidemic effect. Biomed Res Int 2014;2014:925302.
[5] Xiao LJ, Tao R. Traditional Chinese Medicine (TCM) Therapy. Adv Exp Med Biol 2017;1010:261–80.
[6] Xu W, Towers AD, Li P, Collet JP. Traditional Chinese medicine in cancer care: perspectives and experiences of patients and professionals in China. Eur J Cancer Care (Engl) 2006;15(4):397–403.
[7] Yen Y, So S, Rose M, Saif MW, Chu E, Liu SH, et al. Phase I/II study of PHY906/capetabine in advanced hepatocellular carcinoma. Anticancer Res 2009;29(12):4883–92.
[8] Hsiao WL, Liu L. The role of traditional Chinese herbal medicines in cancer therapy–from TCM theory to mechanistic insights. Planta Med 2010;76(11):1118–31.
