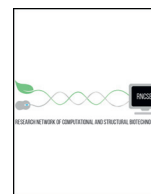




ELSEVIER



COMPUTATIONAL
AND STRUCTURAL
BIOTECHNOLOGY
JOURNAL

journal homepage: www.elsevier.com/locate/csbj

YaTCM: Yet another Traditional Chinese Medicine Database for Drug Discovery

Baiqing Li^a, Chunfeng Ma^{a,c}, Xiaoyong Zhao^a, Zhigang Hu^a, Tengfei Du^a, Xuanming Xu^a,
Zhonghua Wang^{b,*}, Jianping Lin^{a,b,c,*}

^a State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy and Tianjin Key Laboratory of Molecular Drug Research, Nankai University, Haihe Education Park, 38 Tongyan Road, Tianjin 300353, China

^b Biodesign Center, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, 32 West 7th Avenue, Tianjin Airport Economic Area, Tianjin 300308, China

^c Platform of Pharmaceutical Intelligence, Tianjin International Joint Academy of Biomedicine, Tianjin 300457, China

ARTICLE INFO

Article history:

Received 8 August 2018

Received in revised form 4 November 2018

Accepted 6 November 2018

Available online 23 November 2018

Keywords:

TCM

Systems pharmacology

Drug discovery

Pathway analysis

ABSTRACT

Traditional Chinese Medicine (TCM) has a long history of widespread clinical applications, especially in East Asia, and is becoming frequently used in Western countries. However, owing to extreme complicity in both chemical ingredients and mechanism of action, a deep understanding of TCM is still difficult. To accelerate the modernization and popularization of TCM, a single comprehensive database is required, containing a wealth of TCM-related information and equipped with complete analytical tools. Here we present YaTCM (Yet another Traditional Chinese Medicine database), a free web-based toolkit, which provides comprehensive TCM information and is furnished with analysis tools. YaTCM allows a user to (1) identify the potential ingredients that are crucial to TCM herbs through similarity search and substructure search, (2) investigate the mechanism of action for TCM or prescription through pathway analysis and network pharmacology analysis, (3) predict potential targets for TCM molecules by multi-voting chemical similarity ensemble approach, and (4) explore functionally similar herb pairs. All these functions can lead to one systematic network for visualization of TCM recipes, herbs, ingredients, definite or putative protein targets, pathways, and diseases. This web service would help in uncovering the mechanism of action of TCM, revealing the essence of TCM theory and then promoting the drug discovery process. YaTCM is freely available at <http://cadd.pharmacy.nankai.edu.cn/yatcm/home>.

© 2018 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

Throughout the history of drug development, natural products have often proved to be a critical starting point in drug design [1]. All herbal-based medicines are derived from natural compounds [2]. Moreover, for several thousand years, traditional Chinese medicine (TCM) has been widely used in the treatment and prevention of various diseases, playing a vital role in improving the health of Asian people. It is estimated that approximately one-third [3] of the top-selling drugs are derived from medicinal herbs. For example, Realgar-Indigo naturalis formula can treat acute promyelocytic leukemia [4], whereas artemisinin, the major compound of the herb *Artemisia carvifolia* can

treat malaria [5,6]. Additionally, several other well-known drugs are all derived from Chinese herbs. These include Taxol, a chemotherapy medication used to treat breast cancer [7]; Danshensu, a phenolic acid used for its cardio-protective effect [8]; and Salvicine, a diterpenoid quinone compound used for the treatment of several solid tumors [9]. Interestingly, recent studies have shown that some substances derived from TCM herbs may also have toxic effects [10,11]. For example, aristolochic acids, derived from the woody vines of *Aristolochia* plants, are strongly associated with liver cancer [12]. Therefore, great efforts have been made to identify, isolate, and investigate various compounds of TCMs to better understand their biological background and mechanism of action, including toxicity.

The human body is regarded as one of the most complex known systems, comprising metabolic and regulatory networks and pathways. Thus, a silver-bullet approach or mono-target [13,14] drug interventions are not effective strategies in every case owing to the complex pathologies of systemic diseases like cancers, cardiovascular diseases, and neurodegenerative disorders [15]. However, in such instances, network pharmacology approaches [16] combining multi-component, multi-target drug design are highly useful because of the ability of the

Abbreviations: TCM, Traditional Chinese medicine.; MV-SEA, Multi-voting similarity ensemble approach.; HHN, Herb-herb network.; AM, Astragalus Membranaceus.; RA, *Arctostylodes Macrocephala*.; SD, Saposhnikovia Divaricata.

* Corresponding authors at: State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy and Tianjin Key Laboratory of Molecular Drug Research, Nankai University, Haihe Education Park, 38 Tongyan Road, Tianjin 300353, China

E-mail addresses: wang_zh@tib.cas.cn (Z. Wang), jianpinglin@nankai.edu.cn (J. Lin).

<https://doi.org/10.1016/j.csbj.2018.11.002>

2001-0370/© 2018 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/>).

drugs to target multiple proteins or networks involved in a disease [17–19]. The accumulation of large amounts of biomedical data, along with the improvement of knowledge-intensive computational technologies, make the network pharmacology approach a viable option to uncover new drugs and their mechanisms of action [20]. Therefore, at the molecular level, TCM formulas are indeed multi-component and multi-target agents, which in essence is the combination therapy strategy of multi-component drugs. However, owing to the complex chemical components and mechanism of action, understanding TCM at the molecular level is still a challenging task.

During the last decade, many databases specializing in TCM have emerged. For example, Hou's CTM [21,22] database is a 3D structure repository of components extracted from TCM that have been experimentally investigated for characteristics such as drug-likeness and clinical effects [23,24]. Database@Taiwan [25] once contained data on the largest number of herbal compounds with 3D structures in the world (it contained nearly 24,000 pure compounds isolated from 453 Chinese herbs). Phytochemical database [26] includes 78 protein targets for 2,597 Chinese herb compounds. The HIT [27] database focuses on herbal compounds and related targets, but it only contains 586 compounds with 1,301 related targets. The latest version of the TCMID [28] database is one of the largest databases of oriental Chinese medicine, including 8,159 herbs, 43,413 ingredients, and 4,633 diseases. TCMSp [29], a database of systems pharmacology for drug discovery contains 499 Chinese herbs, 29,384 ingredients, and 3,311 targets. There are also other distinctive databases, such as CVDHD [30] (a cardiovascular disease herbal database), CEMTDD [31,32] (Chinese ethnic minority traditional drugs database), and TCMGeneDIT [33] (linking TCM and participating genes).

To keep pace with the ever-increasing number of published TCM compounds and provide analytical tools for TCM-based drug discovery, we developed Yet another Chinese Traditional Medicine database (YaTCM). Our database provides comprehensive biologically relevant information of isolated TCM compounds, including prescriptions, herbs, ingredients, definite or putative protein targets, pathways and diseases. The database incorporates an analysis toolkit that includes a variety of modules, such as similarity search and substructure search for potential structures; an in-house multi-voting similarity ensemble approach (MV-SEA) [34] for predicting protein targets; interaction network analysis; pathway analysis; and functional similarity of herb pairs. The output of the tools is rendered through a visualization network. To our knowledge, the following five key aspects make YaTCM unique, when compared to those of other related TCM databases: (1) Integration of large-scale structural data [47,696 natural compounds, 6,220 herbs, 18,697 targets (with 3,461 therapeutic targets included), 1,907 predicted targets, 390 pathways and 1,813 prescriptions] with manually curated information; (2) Incorporation of >50 key ADMET-related properties from diverse sources for active compound screening; (3) Establishment of prescription-herbs-compound-target-disease-pathway networks and herb-pairs networks for studying mechanism of action of TCM; (4) Mapping the TCM ingredients and targets to KEGG pathways [35] (<https://www.kegg.jp/>), which combines KEGG gene set enrichment analysis; and (5) Predicting potential targets for TCM molecules by an in-house multi-voting chemical ensemble approach [34] (MV-SEA). YaTCM could serve as a useful tool for uncovering the mechanism of action of TCMs at a systematic level, validating anecdotal evidence and theory. It also has the potential to contribute significantly to the drug discovery process and modernization and wider adoption of TCM.

2. Materials and Methods

2.1. Data Sources

To build a comprehensive TCM database, various resources were manually integrated, including published TCM databases (TCMID [28], Database@Taiwan [25], TCMSp [29]), Therapeutic Target Database

(TTD) [36], ChEMBL, the KEGG database [35], books [37–47] and scientific literature mined from PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>). The following steps were taken to manage the quality and updating of data: (1) manually extracted the TCM crude information from resources above; (2) collected prescription information mainly through text-mining of relevant literature (i.e., books and publications), Formulas of Chinese Medicine, and the latest release of Chinese Pharmacopoeia [48] (2015 edition); (3) integrated herb information, and removed duplicates from different resources with respect to the herb's Chinese name, English name or Latin name; (4) integrated compound information, removed duplicates with respect to English name, chemical compound name or structure, and established a network relationship between YaTCM structure and ChEMBL bioactivities; (5) integrated target and disease information and removed duplicates with respect to respective item name; (6) identified KEGG compounds and targets, and constructed relationship between YaTCM and KEGG entries by calculating Tanimoto similarity between KEGG and YaTCM compounds; (7) established the relationship between the database entries for prescription, herb, compound, target, disease and pathway by using in-house scripts; (8) updated the database bi-weekly by conducting literature text-mining and manual curation, in addition to repeating the procedures above. The data entries present in YaTCM have the following components: herbal recipe, the molecular structures of active ingredients, protein targets, KEGG pathways, and diseases. To visualize and graphically analyze the relationship of the different aspects of YaTCM data, we developed a visualization interface using ECharts (<http://echarts.baidu.com/>) and NetworkX (<http://networkx.github.io/>), allowing the information of the interactive network to be displayed within a webpage. Fig. 1 shows the main schema of YaTCM where the data is stored in a PostgreSQL database. The following subsections will describe how the data is compiled and organized.

2.2. Prescription-Herb Relationships

Prescription-herb relationships were deduced mainly by combining TCMID (16) databases and text-mining output of two books—Formulas of Chinese Medicine [47], and Chinese Pharmacopoeia (2015 edition). TCMID consists of >46,000 prescriptions, however, most of them contain rather vague information on usage and dosage as it was accumulated over several thousand years of empiricism and folklore, recorded in ancient natural language. According to the books Formulas of Chinese Medicine and Chinese Pharmacopoeia (2015 edition), the 1813 most commonly used prescriptions were manually curated, and these contained additional properties like Chinese name, English name, phonetic name, related herbs and its ingredients, traditional explanation, traditional usage, traditional and modern application and their English description. (Table S1 in Supporting information).

2.3. Herb-Ingredients Relationship

Herb-ingredients relationships were collected from public TCM databases, TCMID, TCMSp and Database@Taiwan, in addition to a text-mining approach. The current version of TCMID contains 8,159 herbs,

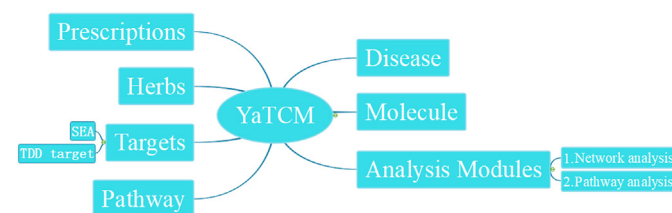


Fig. 1. The main database schema for YaTCM. Each record contains six data fields, for prescription, herbs, ingredients, diseases, targets, pathways. There are two analysis modules—network analysis and pathway analysis. It includes one target prediction function, similarity ensemble approach.

with 43,413 compounds, TCMSP contains 29,384 compounds, Database@Taiwan contains 453 herbs with >20,000 compounds. To enrich the herbs and compounds repertoire, 10 additional books [37–46] (TCM Series of Active Components) (the information of each book detailed in Table S2 in Supporting information) were manually curated. Furthermore, to avoid outdated information in our database, text-mining of PubChem abstracts was performed on articles from 2013 to-date. Some search queries used for retrieving articles were: “traditional Asian medicine,” “traditional Chinese medicine,” “TCM,” “herbal,” “Chinese,” “oriental,” and “combined” among others. The PubChem search retrieved >1000 articles. All data from the sources were manually curated and added to the compiled ingredients data, together with their additional details, including herb's information such as Chinese name, English name, medicinal part, herb images, category, effect and indication.

2.4. Ingredients-Targets-Disease Relationships

Ingredients-targets-disease relationships were compiled from entries in TCMID, TTD [36], and OMIM databases (<http://www.omim.org/>).

The data used for the ingredients category were standardized using the open source data analytics, reporting and integration platform, KNIME [49], in the following steps: (1) The canonical SMILES of compounds were generated and the duplicates were removed. (2) Physicochemical descriptors, including LogP, the number of H-bond donors/acceptors, the number of rotatable bonds, and molecular weight, were calculated for each ingredient compound using RDKit (<http://rdkit.org/>). In addition, 48 molecule-related ADMET descriptors (see Table S3 in Supporting information) were calculated by QikProp [50] module of Schrödinger 2018 and stored in YaTCM. Each ADMET parameter has its respective suitable druggability threshold. It is worth noting that ADMET parameters in a total of 39 molecules in the database are absent because of abnormal valence and could not be recognized by RDKit or Schrödinger 2018. (3) The Morgan fingerprint [51] together with the Tanimoto coefficient [52] were employed for mapping ingredients to ChEMBL database (<https://www.ebi.ac.uk/chembl/>) and KEGG (<http://www.genome.jp/kegg/>) metabolism pathway database. (4) Tao' classical drug-likeness prediction method [16] was added as the drug classifier by calculating the Tanimoto similarity between YaTCM structures and 8842 molecules from DrugBank (version 5.1.1) based on 4885 Dragon soft descriptors. In YaTCM, we considered molecules with drug-likeness >0.18 to be bioactive compounds [29]. Table 1 show the data statistics of YaTCM and a comparison with other TCM databases. It is evident from Table 1 that YaTCM contains the most ingredients, targets (predicted targets or not), and pathways, out of all presented databases.

2.5. Ingredient Category

We divided these ingredients into 10 categories based on structural characteristics of natural products. These included steroids, alkaloids, cardiac glycosides, flavonoids, phenylpropanoids, phytosterols, quinones, steroidal saponin, terpenoids, and triterpenoid. Each category contained a number of subclasses. There are 60 subclasses in total (See Table S4 in Supporting information for more details).

3. Utility and Discussion

3.1. Target Prediction

In silico target prediction is critical for drug discovery and development, on which there is a substantial body of research [53]. One such method—SEA—has proved to be a promising tool and has been successfully applied in drug repositioning [54–56] and natural products target prediction [57]. In our previous work, we proposed a multi-voting SEA model (MV-SEA) for small molecule target prediction [34]. By utilizing the MV-SEA model, all compounds in YaTCM can be systematically predicted and mapped to single protein targets from the ChEMBL database. Four different molecular fingerprints (MACCS, Morgan, Topological and Atom pair) were used to build MV-SEA models. Prediction results of MV-SEA were depicted in respective compound detailed pages. In addition, to increase the functionality of target prediction, the YaTCM built-in “Tools” section includes an in-house developed “Chemical Screening” module and external webserver links to “PharmMapper” and “ChemMapper,” which were implemented to identify potential drug targets of the query compounds. “Chemical Screening” allows users to submit a batch of molecules, which were also predicted and mapped in ChEMBL, and obtain the potential targets by applying the in-house MV-SEA scripts [34]. Users receive the computed results via email.

3.2. KEGG Pathway Mapping

The concept of metabolite-likeness, which is derived from the fact that drugs are more structurally similar to endogenous metabolites than library compounds, is a useful principle in drug design [58]. Kim et al. [4] reported that the degree of structural similarity between compounds derived from traditional oriental medicine and human metabolites is even higher than that between small drug molecules and human metabolites. They argue that mapping of TCM compounds to human metabolic pathway based on structural similarity may be useful in analyses of mechanisms of action [4]. Therefore, in YaTCM, all compounds were mapped to KEGG human metabolic pathway by calculating structural similarity of TCM-metabolite pairs. In addition, we further mapped therapeutic targets to KEGG human signal transduction pathways. To quantitatively elucidate the importance of these pathways, we further applied commonly used gene-ontology enrichment analysis [59] to rank the score of these pathways. Fig. 2 shows the pathway mapping result in prescription Yu Ping Feng. In Fig. 2, the corresponding compound and protein are outlined by a red circle and a blue rectangle, respectively. It is possible to click on these objects to acquire the corresponding relationship between YaTCM and KEGG.

3.3. Network Analysis

3.3.1. Herb-Herb Interaction

Hu et al. [60] apply a simple definition that two targets are related to each other if they share at least five active compounds. In our previous work, the PhIN [61] database defined target-target interaction pairs in terms of shared compounds or shared scaffolds. As an analogy of this methodology, in YaTCM, two herbs are defined as an interacting functionally similar pair if they share a certain number (user can define

Table 1
Data statistics and comparison with other TCM databases.

Database	Ingredients	Herbs	Prescriptions	Targets	Predicted targets	Pathways	Update date
YaTCM	47,696	6220	1813	18,697	1907	390	2018–10
TCMID	43,413	8159	46,914	17,521	0	0	2017
TCMSP	29,384	499	0	3311	0	0	2014–06
Database@Taiwan	24,033	453	0	0	0	0	2011

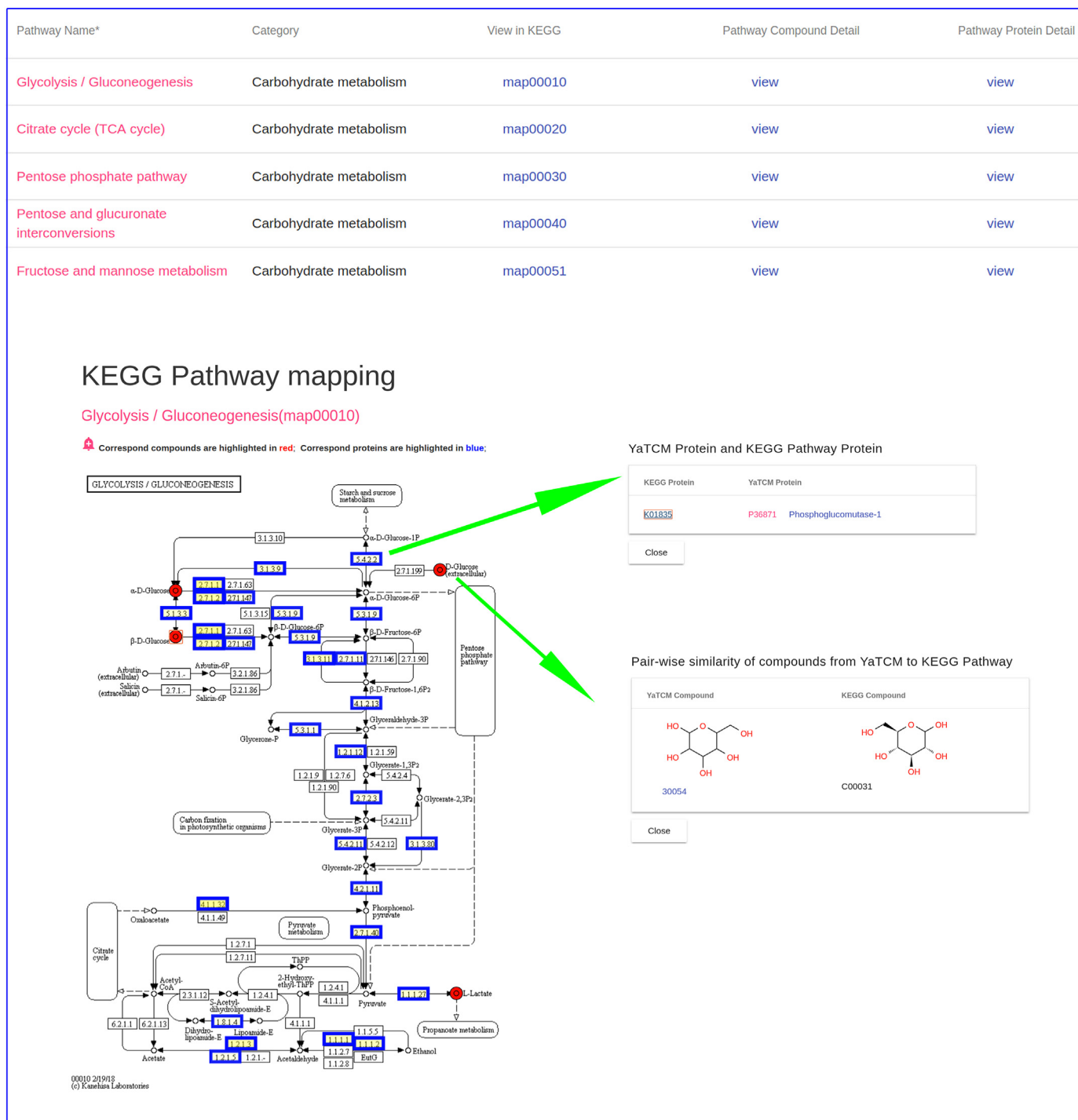


Fig. 2. Pathway analysis. To obtain the comprehensive pathway analysis of TCM, YaTCM maps related ingredients and TTD targets to one specific pathway. Corresponding compounds and protein targets were highlighted in red point and blue rectangle respectively. These can be clicked to reveal the corresponding relationship between YaTCM and KEGG data.

the cutoff number) of therapeutic targets or diseases. To give the user a global view of the herb-herb network (HHN) in YaTCM, we employ visualization of interaction networks to display the network in herbs' detail pages where user can specify the number of shared targets or diseases. Take one famous traditional Chinese herb *Glycyrrhiza Uralensis* for example (<http://cadd.pharmacy.nankai.edu.cn/yatcm/herb/5742/graph>): there are 288 functionally similar herb pairs that share the 135 common targets (Fig. 3A), and 511 herb pairs that share the 200 potential diseases (Fig. 3B) with *Glycyrrhiza Uralensis*.

3.3.2. Prescription Exploration

When treating diseases using TCM, a prescription or herbal formula containing various herbs is usually applied. To better understand the molecular mechanism of action of TCM prescription or herbal formulas, YaTCM can explore a TCM prescription or herbal formulas from a network and pathways perspective. Firstly, the common therapeutic targets (Fig. 4A) and diseases (Fig. 4C) between any two herbs are identified and rendered as an interactive network graph. Secondly, the herb-herb interacting pairs were constructed based on a certain number

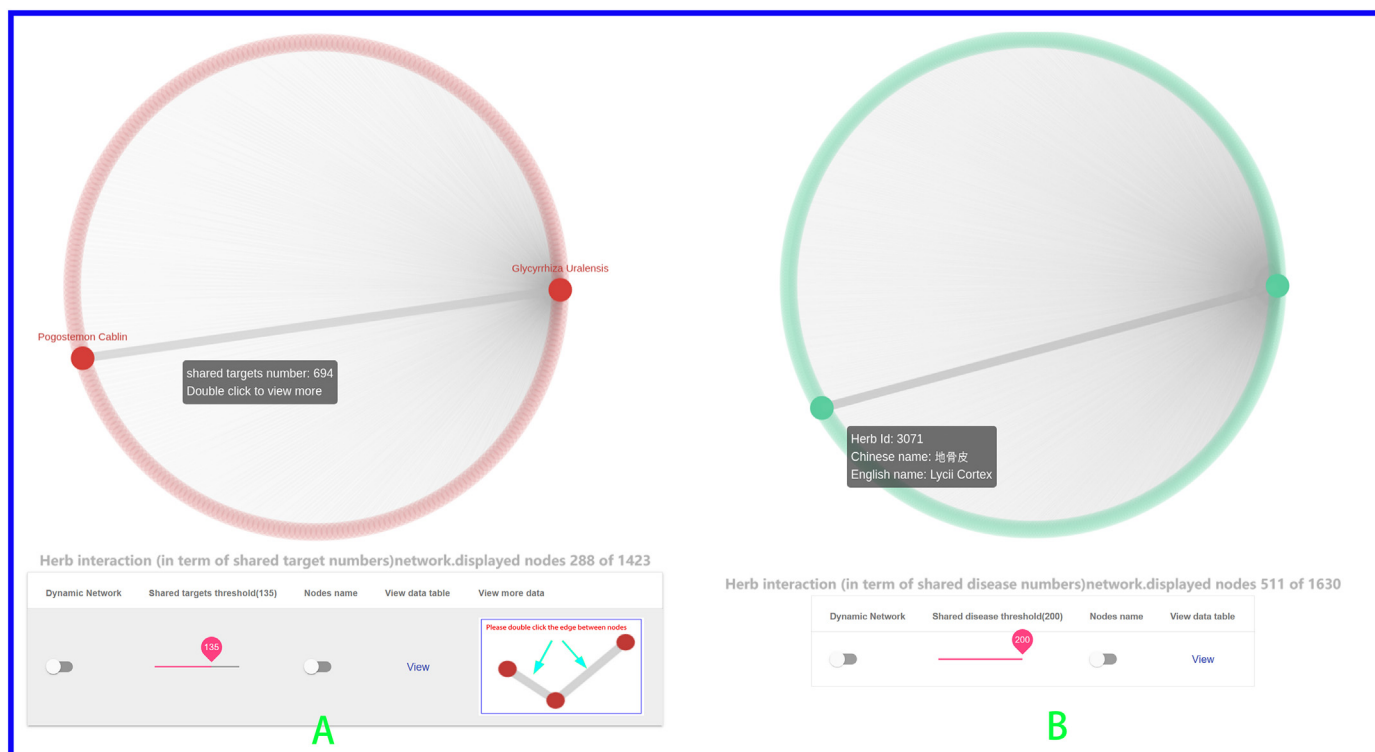


Fig. 3. Herb-herb network. User can obtain functionally similar herb pairs by adjusting the number of shared TTD target (A) or diseases (B). Double clicking the edge between two nodes will all lead to comprehensive pathway analysis and network analysis page between the two herbs.

of common therapeutic targets, which are either involved in the same KEGG pathway (Fig. 4D) or share a certain number of diseases (Fig. 4B). Clicking the edge between two nodes reveals three links, a list of common KEGG pathways (Fig. 4D,2), a list of shared therapeutic targets (Fig. 4D,1) and a list of therapeutic targets that appeared only in the common KEGG pathway (Fig. 4D,3). More detailed information is presented in the Case Study section.

4. Implementation and Application

4.1. Searching and Browsing

Firstly, YaTCM can be accessed through web browsers. Users can browse the entire database containing information about herbs, prescriptions, pathways, targets and diseases, and retrieve their preferred information from the server. Secondly, users are also allowed to retrieve category specific information by entering their preferred query. Thirdly, four analysis tools were incorporated in YaTCM, which are structure retrieval, MV-SEA, pathway analysis and network analysis.

In structure retrieval module, using JSME [62] (a JavaScript-based molecular editor) sketcher, users can build or import a molecular structure and perform a similarity or substructure search from YaTCM database. For instance, one can retrieve a set of compounds in database containing specific groups like aromatic rings, fused rings, heterocyclic or polycyclic rings and so on, combining with similarity degree. As shown in Fig. 5, seven commonly used query fields are supported, including structure, prescription, herb, molecule, target, pathway, disease. After drawing a molecular structure and setting other parameters, users can submit the task (Fig. 5A), and then the candidate molecules list will be loaded in the consequent page for inspection (Fig. 5B). Fig. 5C shows the base information page of Ephedrine, including English name, Chinese name, smiles, CID, CAS, category, compound MS, and related herbs.

The related information of Ephedrine located at the bottom of the details page (Fig. 5D), includes related prescription (Fig. 5D,2), related herbs (Fig. 5D,7), protein targets (Fig. 5D,4), pathways (Fig. 5D,6),

diseases (Fig. 5D,3), MV-SEA predicted targets (Fig. 5D,5), related molecules (same or similar to ChEMBL molecules) (Fig. 5D,1). The network button on the result page can be used to render a 'Force' dynamic network graph or 'Circle' static network graph. Users can filter superabundant no-drug molecular nodes and links through setting four vital ADMET parameters and Tao's drug-likeness parameter [16]. The four ADMET parameters, include stars (default value is 0–4), QPlogBB (default is $-3.0-1.2$), RuleOfFive (default is 0–4), PercentHumanOralAbsorption (default is 0.0–100.0) (see Table S3 in Supporting information). The default value of drug-likeness parameter is >0.18 . Networks can be exported in Portable Network Graphic (PNG) format. (Fig. 5D,8).

4.2. Web API

The website was built with Django framework (V. 1.10.3) (<https://www.djangoproject.com/>) as back-end, Angular framework as front-end, and deployed using Nginx (<http://nginx.org>). YaTCM uses JSME [62] for structure sketch and it displays in a web browser. RDKit (<http://rdkit.org>) was used for small molecule manipulation (such as fingerprint generation and similarity calculation). EChart (<https://echarts.baidu.com/index.html>) and NetworkX (<https://networkx.github.io>) were used for rendering the network visualization.

5. Case Study

Yu Ping Feng is an ancient and popular Chinese prescription that has proven to be useful in the treatment of respiratory tract diseases, such as chronic bronchitis and pulmonary fibrosis [63,64]. There are three herbal medicines (*Astragalus Membranaceus*–AM, *Atractylodes Macrocephala*–RA, *Saposhnikovia Divaricata*–SD). According to Chinese medicine theory, each herb has a different role and AM is thought to play a pivotal role, and is sometimes named "King herb" or "principle medicine." [48] It is evident from the results in Table 2 that AM have potential effects on more pathways, targets, and diseases overall, which is

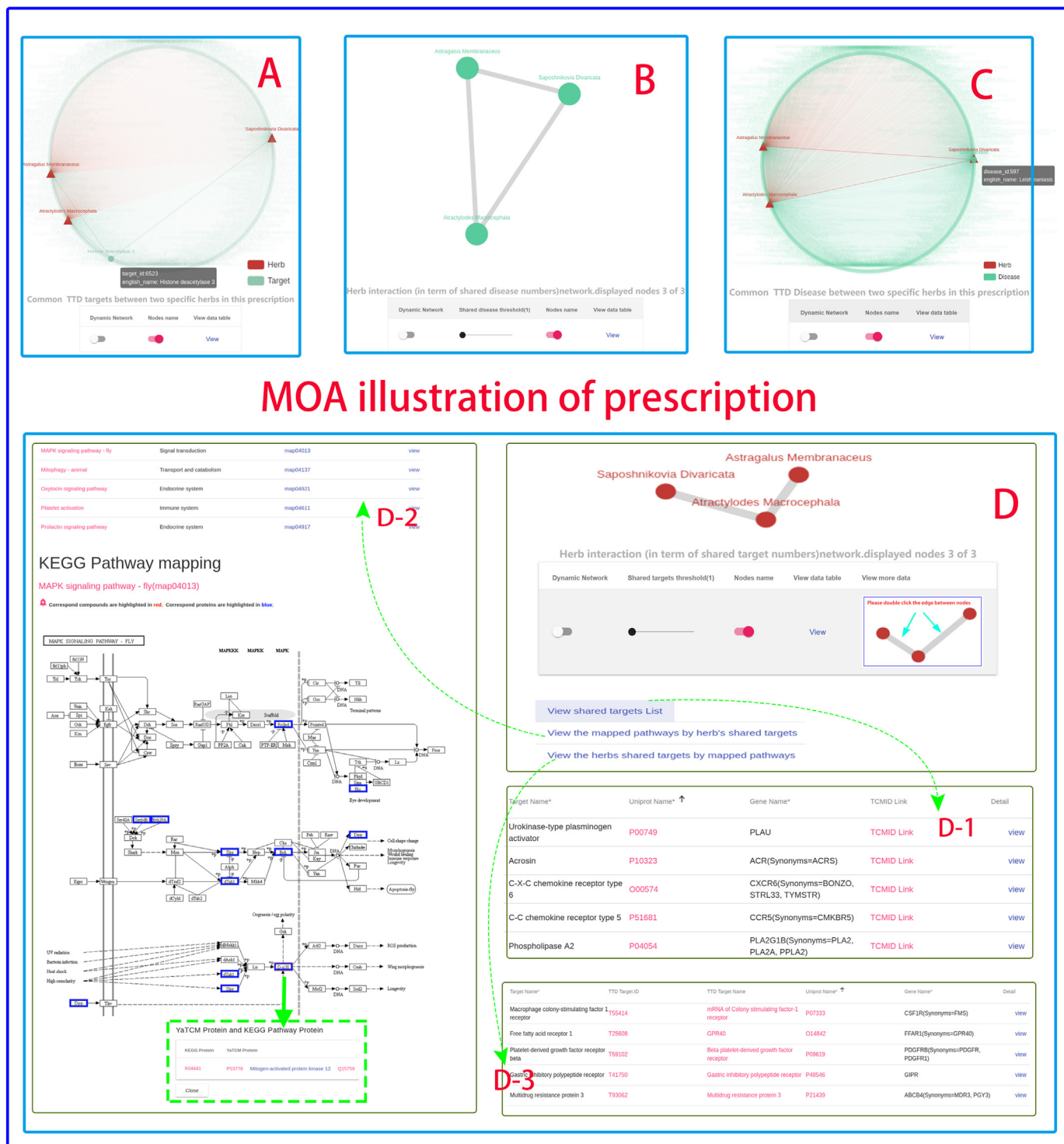


Fig. 4. Prescription exploration. Clicking the edge between any two herbs reveals three function links, a list of shared therapeutic targets (arrow D-1), a list of common KEGG pathways (arrow D-2), and a list of therapeutic targets that were present only in the common KEGG pathway (arrow D-3).

in strong agreement with the Chinese medicine theory. This information was obtained only from statistical analysis and evaluation and not through experimental validation. Users can obtain the detailed information in respective pages. However, highly complex chemical ingredients and mechanism of action makes deeper understanding of Yu Ping Feng challenging. Here, we show a step-by-step process to deep mining the biological basis of pharmacology of Yu Ping Feng.

By searching using a query “yupingfeng power,” the detailed information page of the prescription can be retrieved. Besides the description

of the prescription (Fig. 6), the page consists of seven fields, namely 3 herbs, 497 molecules, 5502 targets, 379 pathways, 1102 diseases, network, and graph. All molecules were mapped to KEGG human metabolic pathway by calculating structural similarity of TCM-metabolite pairs. Related therapeutic targets are also mapped to KEGG human signal transduction pathways in ‘Pathway’ toggle. A commonly used gene-ontology enrichment analysis [59] was applied to rank the score of mapped pathways. By setting the parameter p-value < 0.05, 183 of 379 pathways can be ranked by their statistical significance indicators of

YaTCM Prescription Herb Compound Target Pathway Disease **Search** Tools Help

Structure Prescription Herb Molecule **Target** Pathway Disease

Structure Type: Structure Substructure
Tanimoto threshold: 0.9

English Name: Ephedrine
Chinese Name: 麻黄碱
Smiles: CN[C@H](C)[C@H](O)C1CCCC1
CID: 9294
CAS: 6912-63-6 299-42-3 321-96-0

Enriched	English Name	Formula	Mol Weight	Align
	Ephedrine	C10H17NO	165.12	1.33
	Ephedrine	C10H17NO	165.12	1.33
		C10H17NO	201.09	1.75
		C10H17NO	201.09	-1.56

Highuang (Ephedra), Minutao (Ephedra Minuta), Sidajiao (Sida Acuta), Gerardiana (Ephedra Gerardiana), Monosperma (Ephedra Monosperma), Lepidosperma (Ephedra Lepidosperma), Equisetina (Ephedra Equisetina), Distachya (Ephedra Distachya), Sinica (Ephedra Sinica), Regelliana (Ephedra Regelliana), Likiangensis (Ephedra Likiangensis), Intermedia (Ephedra Intermedia), Pinellia Ternata (Pinellia Ternata), Procera (Ephedra Procera), Przewalskii (Ephedra Przewalskii), Tibetica (Ephedra Intermedia Var. Tibetica), Saxatilis (Ephedra Saxatilis)

Molecular Formula	Mol Weight	PSA	Align	HBD	HBA	HB
C10H17NO	165.12	32.26	1.33	2	2	3

Chinese Name: 麻黄白术天麻汤
English Name: banxia baizhu tianma decoction
Phytoic Name: ban via bai shu tian ma tang
Modern Application (english): This prescription is often used for otogenic vertigo, hypertension, neurological vertigo, epilepsy, facial nerve paralysis.
Traditional Application (english): To regulate the spleen to eliminate dampness and resolve cold phlegm.

Chinese Name: 半夏厚朴汤
English Name: banxia houpu decoction
Phytoic Name: ban via hou pu tang
Modern Application (english): This prescription is often used for hysteria, stomach nerve function, chronic pharyngitis, chronic bronchitis, esophageal spasm.
Traditional Application (english): To promote qi circulation, disperse accumulation of pathogen, relieve asthma.

Chinese Name: 半夏泻心汤
English Name: banxia xixin decoction
Phytoic Name: ban via xie xin tang
Modern Application (english): This prescription is often used in acute and chronic gastritis, chronic colitis, chronic hepatitis, early liver cirrhosis.
Traditional Application (english): To regulate cold and heat, relieve distension to eliminate abdominal mass.

Chinese Name: 保和丸
English Name: bahe pills
Phytoic Name: Bahe Wan
Modern Application (english): This prescription is often used in acute and chronic gastritis, acute and chronic enteritis, indigestion, infant diarrhea.
Traditional Application (english): To regulate the stomach and promote digestion.

Chinese Name: 藜芦煎丸
English Name: Carapax Trionychis Bolus
Phytoic Name: bie jia jian wan
Modern Application (english): This prescription is commonly used in cirrhosis. To promote qi circulation, promote hepatoprotection, liver cancer, uterine fibrosis, ovarian cysts.
Traditional Application (english): To promote circulation, eliminate dampness, resolve phlegm, resolve the hard lump.

Chronic obstructive pulmonary disease	480-482, 484-486	480-346, 347
Chronic breathing disorder	480-486	294, 297
Asthma disease	331	630
Alging oral Exacerbation of acute asthma	483	159139, 395
Acute asthma	1943, 1945, 1948, 2073, 2074, 2075, 2076, 2077, 2078, 2084	1781, 1783, 1784, 1785, 1786, 1787, 1788, 1789, 1790, 1791, 1792, 1793, 1794, 1795, 1796, 1797, 1798, 1799, 1800, 1801, 1802, 1803, 1804, 1805, 1806, 1807, 1808, 1809, 1810, 1811, 1812, 1813, 1814, 1815, 1816, 1817, 1818, 1819, 1820, 1821, 1822, 1823, 1824, 1825, 1826, 1827, 1828, 1829, 1830, 1831, 1832, 1833, 1834, 1835, 1836, 1837, 1838, 1839, 1840, 1841, 1842, 1843, 1844, 1845, 1846, 1847, 1848, 1849, 1850, 1851, 1852, 1853, 1854, 1855, 1856, 1857, 1858, 1859, 1860, 1861, 1862, 1863, 1864, 1865, 1866, 1867, 1868, 1869, 1870, 1871, 1872, 1873, 1874, 1875, 1876, 1877, 1878, 1879, 1880, 1881, 1882, 1883, 1884, 1885, 1886, 1887, 1888, 1889, 1890, 1891, 1892, 1893, 1894, 1895, 1896, 1897, 1898, 1899, 1900, 1901, 1902, 1903, 1904, 1905, 1906, 1907, 1908, 1909, 1910, 1911, 1912, 1913, 1914, 1915, 1916, 1917, 1918, 1919, 1920, 1921, 1922, 1923, 1924, 1925, 1926, 1927, 1928, 1929, 1930, 1931, 1932, 1933, 1934, 1935, 1936, 1937, 1938, 1939, 1940, 1941, 1942, 1943, 1944, 1945, 1946, 1947, 1948, 1949, 1950, 1951, 1952, 1953, 1954, 1955, 1956, 1957, 1958, 1959, 1960, 1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969, 1970, 1971, 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1979, 1980, 1981, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 2679, 2680, 2681, 2682, 2683, 2684, 2685, 2686, 2687, 2688, 2689, 2690, 2691, 2692, 2693, 2694, 2695, 2696, 2697, 2698, 2699, 2700, 2701, 2702, 2703, 2704, 2705, 2706, 2707, 2708, 2709, 2710, 2711, 2712, 2713, 2714, 2715, 2716, 2717, 2718, 2719, 2720, 2721, 2722, 2723, 2724, 2725, 2726, 2727, 2728, 2729, 2730, 2731, 2732, 2733, 2734, 2735, 2736, 2737, 2738, 2739, 2740, 2741, 2742, 2743, 2744, 2745, 2746, 2747, 2748, 2749, 2750, 2751, 2752, 2753, 2754, 2755, 2756, 2757, 2758, 2759, 2760, 2761, 2762, 2763, 2764, 2765, 2766, 2767, 2768, 2769, 2770, 2771, 2772, 2773, 2774, 2775, 2776, 2777, 2778, 2779, 2780, 2781, 2782, 2783, 2784, 2785, 2786, 2787, 2788, 2789, 2790, 2791, 2792, 2793, 2794, 2795, 2796, 2797, 2798, 2799, 2800, 2801, 2802, 2803, 2804, 2805, 2806, 2807, 2808, 2809, 2810, 2811, 2812, 2813, 2814, 2815, 2816, 2817, 2818, 2819, 2820, 2821, 2822, 2823, 2824, 2825, 2826, 2827, 2828, 2829, 2830, 2831, 2832, 2833, 2834, 2835, 2836, 2837, 2838, 2839, 2840, 2841, 2842, 2843, 2844, 2845, 2846, 2847, 2848, 2849, 2850, 2851, 2852, 2853, 2854, 2855, 2856, 2857, 2858, 2859, 2860, 2861, 2862, 2863, 2864, 2865, 2866, 2867, 2868, 2869, 2870, 2871, 2872, 2873, 2874, 2875, 2876, 2877, 2878, 2879, 2880, 2881, 2882, 2883, 2884, 2885, 2886, 2887, 2888, 2889, 2890, 2891, 2892, 2893, 2894, 2895, 2896, 2897, 2898, 2899, 2900, 2901, 2902, 2903, 2904, 2905, 2906, 2907, 2908, 2909, 2910, 2911, 2912, 2913, 2914, 2915, 2916, 2917, 2918, 2919, 2920, 2921, 2922, 2923, 2924, 2925, 2926, 2927, 2928, 2929, 2930, 2931, 2932, 2933, 2934, 2935, 2936, 2937, 2938, 2939, 2940, 2941, 2942, 2943, 2944, 2945, 2946, 2947, 2948, 2949, 2950, 2951, 2952, 2953, 2954, 2955, 2956, 2957, 2958, 2959, 2960, 2961, 2962, 2963, 2964, 2965, 2966, 2967, 2968, 2969, 2970, 2971, 2972, 2973, 2974, 2975, 2976, 2977, 2978, 2979, 2980, 2981, 2982, 2983, 2984, 2985, 2986, 2987, 2988, 2989, 2990, 2991, 2992, 2993, 2994, 2995, 2996, 2997, 2998, 2999, 3000, 3001, 3002, 3003, 3004, 3005, 3006, 3007, 3008, 3009, 3010, 3011, 3012, 3013, 3014, 3015, 3016, 3017, 3018, 3019, 3020, 3021, 3022, 3023, 3024, 3025, 3026, 3027, 3028, 3029, 3030, 3031, 3032, 3033, 3034, 3035, 3036, 3037, 3038, 3039, 3040, 3041, 3042, 3043, 3044, 3045, 3046, 3047, 3048, 3049, 3050, 3051, 3052, 3053, 3054, 3055, 3056, 3057, 3058, 3059, 3060, 3061, 3062, 3063, 3064, 3065, 3066, 3067, 3068, 3069, 3070, 3071, 3072, 3073, 3074, 3075, 3076, 3077, 3078, 3079, 3080, 3081, 3082, 3083, 3084, 3085, 3086, 3087, 3088, 3089, 3090, 3091, 3092, 3093, 3094, 3095, 3096, 3097, 3098, 3099, 3100, 3101, 3102, 3103, 3104, 3105, 3106, 3107, 3108, 3109, 3110, 3111, 3112, 3113, 3114, 3115, 3116, 3117, 3118, 3119, 3120, 3121, 3122, 3123, 3124, 3125, 3126, 3127, 3128, 3129, 3130, 3131, 3132, 3133, 3134, 3135, 3136, 3137, 3138, 3139, 3140, 3141, 3142, 3143, 3144, 3145, 3146, 3147, 3148, 3149, 3150, 3151, 3152, 3153, 3154, 3155, 3156, 3157, 3158, 3159, 3160, 3161, 3162, 3163, 3164, 3165, 3166, 3167, 3168, 3169, 3170, 3171, 3172, 3173, 3174, 3175, 3176, 3177, 3178, 3179, 3180, 3181, 3182, 3183, 3184, 3185, 3186, 3187, 3188, 3189, 3190, 3191, 3192, 3193, 3194, 3195, 3196, 3197, 3198, 3199, 3200, 3201, 3202, 3203, 3204, 3205, 3206, 3207, 3208, 3209, 3210, 3211, 3212, 3213, 3214, 3215, 3216, 3217, 3218, 3219, 3220, 3221, 3222, 3223, 3224, 3225, 3226, 3227, 3228, 3229, 3230, 3231, 3232, 3233, 3234, 3235, 3236, 3237, 3238, 3239, 3240, 3241, 3242, 3243, 3244, 3245, 3246, 3247, 3248, 3249, 3250, 3251, 3252, 3253, 3254, 3255, 3256, 3257, 3258, 3259, 3260, 3261, 3262, 3263, 3264, 3265, 3266, 3267, 3268, 3269, 3270, 3271, 3272, 3273, 3274, 3275, 3276, 3277, 3278, 3279, 3280, 3281, 3282, 3283, 3284, 3285, 3286, 3287, 3288, 3289, 3290, 3291, 3292, 3293, 3294, 3295, 3296, 3297, 3298, 3299, 3300, 3301, 3302, 3303, 3304, 3305, 3306, 3307, 3308, 3309, 3310, 3311, 3312, 3313, 3314, 3315, 3316, 3317, 3318, 3319, 3320, 3321, 3322, 3323, 3324, 3325, 3326, 3327, 3328, 3329, 3330, 3331, 3332, 3333, 3334, 3335, 3336, 3337, 3338, 3339, 3340, 3341, 3342, 3343, 3344, 3345, 3346, 3347, 3348, 3349, 3350, 3351, 3352, 3353, 3354, 3355, 3356, 3357, 3358, 3359, 3360, 3361, 3362, 3363, 3364, 3365, 3366, 3367, 3368, 3369, 3370, 3371, 3372, 3373, 3374, 3375, 3376, 3377, 3378, 3379, 3380, 3381, 3382, 3383, 3384, 3385, 3386, 3387, 3388, 3389, 3390, 3391, 3392, 3393, 3394, 3395, 3396, 3397, 3398, 3399, 3400, 3401, 3402, 3403, 3404, 3405, 3406, 3407, 3408, 3409, 3410, 3411, 3412, 3413, 3414, 3415, 3416, 3417, 3418, 3419, 3420, 3421, 3422, 3423, 3424, 3425, 3426, 3427, 3428, 3429, 3430, 3431, 3432, 3433, 3434, 3435, 3436, 3437, 3438, 3439, 3440, 3441, 3442, 3443, 3444, 3445, 3446, 3447, 3448, 3449, 3450, 3451, 3452, 3453, 3454, 3455, 3456, 3457, 3458, 3459, 3460, 3461, 3462, 3463, 3464, 3465, 3466, 3467, 3468, 3469, 3470, 3471, 3472, 3473, 3474, 3475, 3476, 3477, 3478, 3479, 3480, 3481, 3482, 3483, 3484, 3485, 3486, 3487, 3488, 3489, 3490, 3491, 3492, 3493, 3494, 3495, 3496, 3497, 3498, 3499, 3500, 3501, 3502, 3503, 3504, 3505, 3506, 3507, 3508, 3509, 3510, 3511, 3512, 3513, 3514, 3515, 3516, 3517, 3518, 3519, 3520, 3521, 3522, 3523, 3524, 3525, 3526, 3527, 3528, 3529, 3530, 3531, 3532, 3533, 3534, 3535, 3536, 3537, 3538, 3539, 3540, 3541, 3542, 3543, 3544, 3545, 3546, 3547, 3548, 3549, 3550, 3551, 3552, 3553, 3554, 3555, 3556, 3557,

Table 2
Data statistic of Yu Ping Feng in YaTCM.

English name	Chinese name	Molecules	Pathways	Targets	Disease
<i>Astragalus membranaceus</i>	黄芪	199	376	4921	1081
<i>Atractylodes macrocephala</i>	白术	100	314	1068	738
<i>Saposhnikovia divaricata</i>	防风	220	322	1748	907
Total	3	497	379	5502	1102

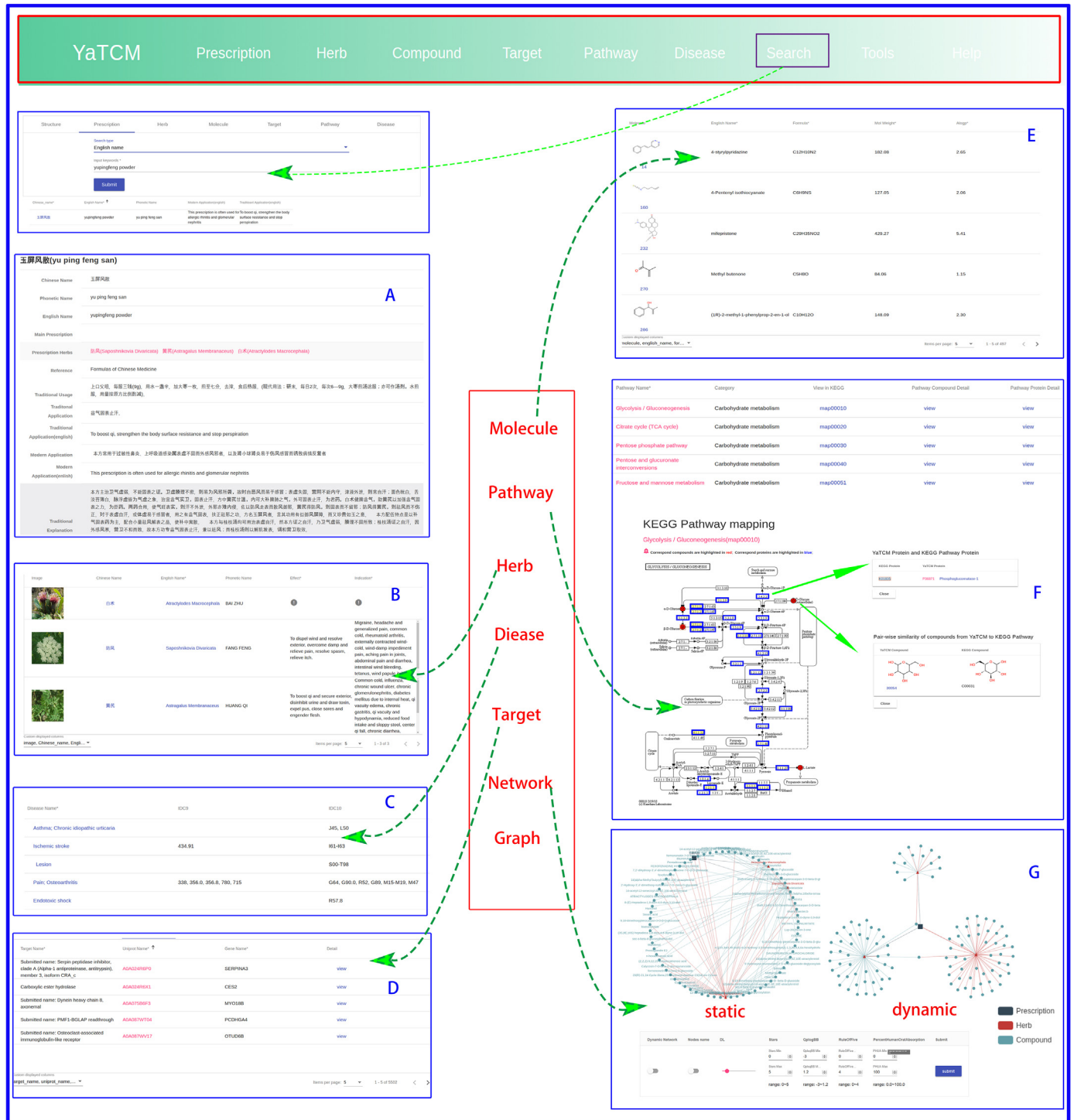


Fig. 6. The search page of prescription category as seen by submitting different names. When Yu Ping Feng is queried, a variety of information is presented. Users can select their preferred information to display and further explore the mechanism of action through variable tools, such as pathway analysis and network pharmacology analysis.

enrichment information including p-value, q-value, bg-ratio, gene-ratio and p-adjust. However, the remaining 196 pathways could not be ranked by this analysis owing to their low statistical significant enrichment information. These two complementary parts were listed on the web page <http://cadd.pharmacy.nankai.edu.cn/yatcm/prescription/71/prescription-pathway>. For a better view of all layers, prescription-herb-molecule-target-pathway-disease network was created. The network visualizations can be downloaded as PNG format.

To investigate effect of three different herbal medicines, and explore the commonalities among herbs, we obtained four interactive networks using 'Graph' toggle switch (Supplementary Figs. S1–S4). Users can retrieve the shared targets, pathways or diseases between any two or any more herbs. For instance, when submitting the *Saposhnikovia Divaricata* and *Atractylodes Macrocephala*, we can get

425 shared targets (Supplementary Fig. S1) (297 vital targets located in pathways Supplementary Fig. S2), 258 shared pathways, 733 shared diseases (Supplementary Fig. S3, Supplementary Fig. S4), which are critical for understanding the pharmacological mechanism of Yu Ping Feng. For example, TGF- β 1 (Transforming Growth Factor β 1) is one of the cytokines involved in regulating inflammation, and tissue fibrosis [65], which plays a vital role in the propagation of lung fibroblasts. Furthermore, ERK [a member of Mitogen-activated protein kinase (MAPK)] can regulate the expression of inflammatory mediators and cytokines (including TGF- β 1) [66]. Baicalin (obtained from AM) can down-regulate TGF- β 1-induced ERK signaling pathway [67]. The mapping relationships generated from YaTCM (Fig. 7) are consistent with the experimental results and provide novel insights into the mechanism of pulmonary fibrosis and the underlying mechanism may become the theoretical foundation for clinical use of Yu Ping Feng.

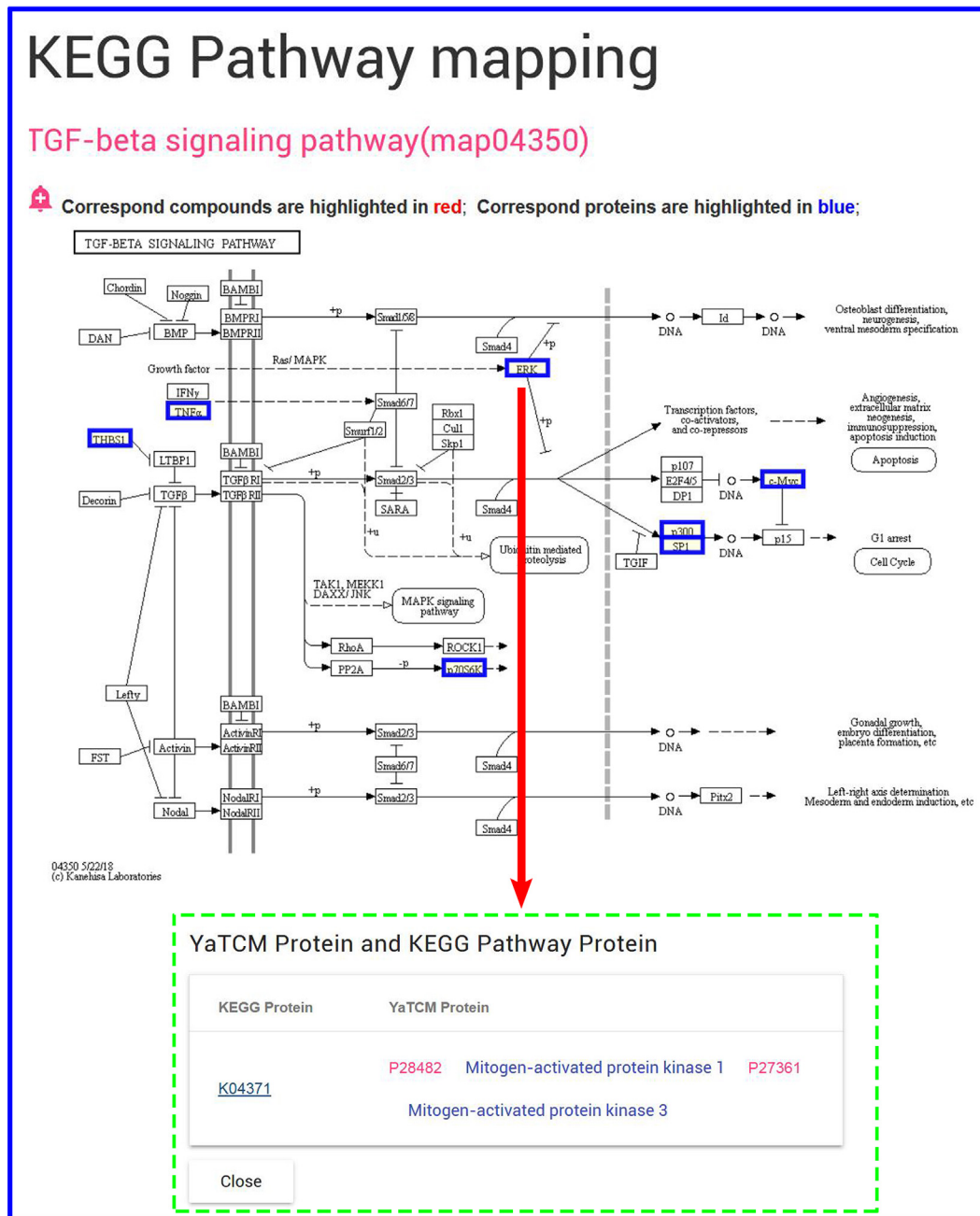


Fig. 7. The mechanism of Yu Ping Feng. The mapping relationships generated are consistent with the experimental results and provide novel insights into the mechanism of pulmonary fibrosis.

6. Conclusion

The YaTCM database is a freely available resource that provides a comprehensive relationship network of TCM components, including prescriptions, herbs, ingredients, definite or putative protein targets, pathways and diseases. In addition, it offers an analysis toolkit that allows the user to perform similarity and substructure searches for potential structures, MV-SEA for predicting protein targets, network analysis, pathway analysis, and identification of functionally similar herb pairs. YaTCM can potentially contribute to unravel the mechanism of action of TCM at systematic level, validate TCM anecdotal evidence, facilitate drug discovery and design processes, and last but not least, accelerate the modernization and popularization of TCM. However, much work remains to be done; for example, refining ingredient structures with correct chirality. Particularly, we are planning to implement deep-learning based method to provide a reasonable depiction of the drug-likeness of TCM ingredients.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2018.11.002>.

Availability of data and material

YaTCM is freely accessible at <http://cadd.pharmacy.nankai.edu.cn/yatcm/home>. It will be updated continually.

Authors' contributions

JPL, ZHW conceived the article. BQL constructed the database back-end, drafted the manuscript, designed herb-herb network and prescription mining function, wrote the user manual of database and description of the website. ZHW designed MV-SEA protein prediction function. ZGH, CFM and ZFZ constructed the database front-end. XYZ, XMZ, TFD participated in data collecting and processing. All authors read and agreed to the final manuscript.

Acknowledgements

This study was supported by the National Key R&D Program of China (No. 2017YFC1104400).

References

- [1] Banerjee P, Erehman J, Gohlke B-O, Wilhelm T, Preissner R, Dunkel M. Super natural II—a database of natural products. *Nucleic Acids Res* 2015;43:D935–9. <https://doi.org/10.1093/nar/gku886>.
- [2] Bent S. Herbal medicine in the United States: review of efficacy, safety, and regulation. *J Gen Intern Med* 2008;23:854–9. <https://doi.org/10.1007/s11606-008-0632-y>.
- [3] Strohl WR. The role of natural products in a modern drug discovery program. *Drug Discov Today* 2000;5:39–41. [https://doi.org/10.1016/S1359-6446\(99\)01443-9](https://doi.org/10.1016/S1359-6446(99)01443-9).
- [4] Kim HU, Ryu JY, Lee JO, Lee SY. A systems approach to traditional oriental medicine. *Nat Biotechnol* 2015;33:264–8. <https://doi.org/10.1038/nbt.3167>.
- [5] Miller LH, Su X. Artemisinin: discovery from the chinese herbal garden. *Cell* 2011;146:855–8. <https://doi.org/10.1016/j.cell.2011.08.024>.
- [6] Tu Y. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. *Nat Med* 2011;17:1217–20. <https://doi.org/10.1038/nm.2471>.
- [7] Liu W, Furuta E, Shindo K, Watabe M, Xing F, Pandey PR, et al. Cacalol, a natural sesquiterpene, induces apoptosis in breast cancer cells by modulating Akt-SREBP-FAS signaling pathway. *Breast Cancer Res Treat* 2011;128:57–68. <https://doi.org/10.1007/s10549-010-1076-8>.
- [8] Li H, Xie Y-H, Yang Q, Wang S-W, Zhang B-L, Wang J-B, et al. Cardioprotective effect of Paeonol and Danshensu combination on isoproterenol-induced myocardial injury in rats. *PLoS One* 2012;7:e48872. <https://doi.org/10.1371/journal.pone.0048872>.
- [9] Wang M-W, Hao X, Chen K. Biological screening of natural products and drug innovation in China. *Philos Trans R Soc B Biol Sci* 2007;362:1093–105. <https://doi.org/10.1098/rstb.2007.2036>.
- [10] Teschke R, Wolff A, Frenzel C, Schulze J. Review article: herbal hepatotoxicity - an update on traditional Chinese medicine preparations. *Aliment Pharmacol Ther* 2014;40:32–50. <https://doi.org/10.1111/apt.12798>.
- [11] Lv W, Piao J-H, Jiang J-G. Typical toxic components in traditional Chinese medicine. *Expert Opin Drug Saf* 2012;11:985–1002. <https://doi.org/10.1517/14740338.2012.726610>.
- [12] Ng AWT, Poon SL, Huang MN, Lim JQ, Boot A, Yu W, et al. Aristolochic acids and their derivatives are widely implicated in liver cancers in Taiwan and throughout Asia. *Sci Transl Med* 2017;9. <https://doi.org/10.1126/scitranslmed.aan6446>.
- [13] Dar AC, Das TK, Shokat KM, Cagan RL. Chemical genetic discovery of targets and anti-targets for cancer polypharmacology. *Nature* 2012;486:80–4. <https://doi.org/10.1038/nature11127>.
- [14] Gujral TS, Peshkin L, Kirschner MW. Exploiting polypharmacology for drug target deconvolution. *Proc Natl Acad Sci* 2014;111:5048–53. <https://doi.org/10.1073/pnas.1403080111>.
- [15] Zimmermann GR, Lehár J, Keith CT. Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug Discov Today* 2007;12:34–42. <https://doi.org/10.1016/j.drudis.2006.11.008>.
- [16] Tao W, Xu X, Wang X, Li B, Wang Y, Li Y, et al. 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;145:1–10. <https://doi.org/10.1016/j.jep.2012.09.051>.
- [17] van der Greef J. Perspective: all systems go. *Nature* 2011;480. <https://doi.org/10.1038/480587a>.
- [18] Xu Z. Modernization: one step at a time. *Nature* 2011;480:S90–2. <https://doi.org/10.1038/480S90a>.
- [19] Schmidt BM, Ribnick DM, Lipsky PE, Raskin I. Revisiting the ancient concept of botanical therapeutics. *Nat Chem Biol* 2007;3:360–6. <https://doi.org/10.1038/nchembio0707-360>.
- [20] Cohen J. Bioinformatics—an introduction for computer scientists. *ACM Comput Surv* 2004;36:122–58. <https://doi.org/10.1145/1031120.1031122>.
- [21] Hou T, Qiao X, Xu X. Research and development of 3D molecular structure database of traditional chinese drugs. *Coll Chem Mol Eng Univ Beijing* 2001;59:1788–92.
- [22] Qiao X, Hou T, Zhang W, Guo S, Xu X. A 3D structure database of components from chinese traditional medicinal herbs; 2002. <https://doi.org/10.1021/ci010113h481-9>.
- [23] Tian S, Wang J, Li Y, Xu X, Hou T. Drug-likeness analysis of traditional Chinese medicines: prediction of drug-likeness using machine learning approaches. *Mol Pharm* 2012;9:2875–86. <https://doi.org/10.1021/mp300198d>.
- [24] Shen M, Tian S, Li Y, Li Q, Xu X, Wang J, et al. Drug-likeness analysis of traditional Chinese medicines: 1. property distributions of compounds and natural compounds from traditional Chinese medicines; 2012; 1–13.
- [25] Chen CY-C. TCM Database@Taiwan: the world's largest traditional Chinese medicine database for drug screening in silico. *PLoS One* 2011;6:e15939. <https://doi.org/10.1371/journal.pone.0015939>.
- [26] Ehrman TM, Barlow DJ, Hylands PJ. Phytochemical databases of Chinese herbal constituents and bioactive plant compounds with known target specificities. *J Chem Inf Model* 2007;47:254–63. <https://doi.org/10.1021/ci600288m>.
- [27] Ye H, Ye L, Kang H, Zhang D, Tao L, Tang K, et al. HIT: linking herbal active ingredients to targets. *Nucleic Acids Res* 2011;39:D1055–9. <https://doi.org/10.1093/nar/gkq1165>.
- [28] Xue R, Fang Z, Zhang M, Yi Z, Wen C, Shi T. TCMID: traditional Chinese medicine integrative database for herb molecular mechanism analysis. *Nucleic Acids Res* 2012;41:D1089–95. <https://doi.org/10.1093/nar/gks1100>.
- [29] Ru J, Li P, Wang J, Zhou W, Li B, Huang C, et al. TCMSPP: a database of systems pharmacology for drug discovery from herbal medicines. *J Chem* 2014;6:13. <https://doi.org/10.1186/1758-2946-6-13>.
- [30] Gu J, Gui Y, Chen L, Yuan G, Xu X. CVDHD: a cardiovascular disease herbal database for drug discovery and network pharmacology. *J Chem* 2013;5:51. <https://doi.org/10.1186/1758-2946-5-51>.
- [31] Huang J, Wang J. CEMTDD: Chinese ethnic minority traditional drug database. *Apoptosis* 2014;19:1419–20. <https://doi.org/10.1007/s10495-014-1011-2>.
- [32] Huang J, Zheng Y, Wu W, Xie T, Yao H, Pang X, et al. CEMTDD: The database for elucidating the relationships among herbs, compounds, targets and related diseases for Chinese ethnic minority traditional drugs. *Oncotarget* 2015;6. <https://doi.org/10.18632/oncotarget.3789> 1419–20.
- [33] Fang Y, Huang H, Chen H, Juan H-F. TCMGeneDIT: a database for associated traditional Chinese medicine, gene and disease information using text mining. *BMC Complement Altern Med* 2008;8:1–11. <https://doi.org/10.1186/1472-6882-8-58>.
- [34] Wang Z, Liang L, Yin Z, Lin J. Improving chemical similarity ensemble approach in target prediction. *J Chem* 2016;8:20. <https://doi.org/10.1186/s13321-016-0130-x>.
- [35] Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 2017;45:D353–61. <https://doi.org/10.1093/nar/gkw1092>.
- [36] Chen X. TTD: therapeutic target database. *Nucleic Acids Res* 2002;30:412–5. <https://doi.org/10.1093/nar/30.1.412>.
- [37] Zhou J. Active Components on Digestive and Respiratory Systems in TCM. *Sci Press*; 2012.
- [38] Zhou J. Anti-Inflammatory Active Components in TCM. *Sci Press*; 2012.
- [39] Zhou J. Antioxidant and Antisenescence Active Components in TCM. *Sci Press*; 2012.
- [40] Zhou J. Active Components against Microbial Infections in TCM. *Sci Press*; 2012.
- [41] Zhou J. Active Components against Parasiti in TCM. *Sci Press*; 2012.
- [42] Zhou J. Active Components on Cardiocerebro-vascular System in TCM. *Sci Press*; 2012.
- [43] Zhou J. Multitargets Active Components in TCM. *Sci Press*; 2012.
- [44] Zhou J. Enzyme Inhibitors in TCM. *Sci Press*; 2012.
- [45] Zhou J. Anticancer Active Components in TCM. *Sci Press*; 2012.
- [46] Zhou J. Active Components on Nervous System in TCM. *Sci Press*; 2012.
- [47] Jia DZ. Formulas of Chinese Medicine. *Med China Press Tradit Chinese*; 2010; 1–357.
- [48] Ed. CPC. The Chinese Pharmacopoeia IV. *Beijing Chem Ind Press*; 2015.
- [49] Berthold MR, Chebron N, Dill F, Gabriel TR, Kötter T, Meini T, et al. KNIME: the Konstanz information miner. *Data Anal Mach Learn Appl* 2008;319–26. https://doi.org/10.1007/978-3-540-78246-9_38.

- [50] QikProp. Version 2.8. New York, NY: Schrödinger, LLC; 2018.
- [51] Rogers D, Hahn M. Extended-connectivity fingerprints. *J Chem Inf Model* 2010;50:742–54. <https://doi.org/10.1021/ci100050t>.
- [52] Delaney JS. Assessing the ability of chemical similarity measures to discriminate between active and inactive compounds. *Mol Divers* 1996;1:217–22. <https://doi.org/10.1007/BF01715525>.
- [53] Peón A, Dang CC, Ballester PJ. How reliable are ligand-centric methods for target fishing? *Front Chem* 2016;4:1–10. <https://doi.org/10.3389/fchem.2016.00015>.
- [54] Keiser MJ, Setola V, Irwin JJ, Laggner C, Abbas AI, Hufeisen SJ, et al. Predicting new molecular targets for known drugs. *Nature* 2009;462:175–81. <https://doi.org/10.1038/nature08506>.
- [55] Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. Relating protein pharmacology by ligand chemistry. *Nat Biotechnol* 2007;25:197–206. <https://doi.org/10.1038/nbt1284>.
- [56] Cameron RT, Coleman RG, Day JP, Yalla KC, Houslay MD, Adams DR, et al. Chemical informatics uncovers a new role for moexipril as a novel inhibitor of cAMP phosphodiesterase-4 (PDE4). *Biochem Pharmacol* 2013;85:1297–305. <https://doi.org/10.1016/j.bcp.2013.02.026>.
- [57] Sá MS, de Menezes MN, Krettli AU, Ribeiro IM, Tomassini TCB, Ribeiro dos Santos R, et al. Antimalarial activity of physalins B, D, F, and G. *J Nat Prod* 2011;74:2269–72. <https://doi.org/10.1021/np200260f>.
- [58] Dobson PD, Patel Y, Kell DB. 'Metabolite-likeness' as a criterion in the design and selection of pharmaceutical drug libraries. *Drug Discov Today* 2009;14:31–40. <https://doi.org/10.1016/j.drudis.2008.10.011>.
- [59] Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2009;37:1–13. <https://doi.org/10.1093/nar/gkn923>.
- [60] Hu Y, Bajorath J. Polypharmacology directed compound data mining: identification of promiscuous chemotypes with different activity profiles and comparison to approved drugs. *J Chem Inf Model* 2010;50:2112–8. <https://doi.org/10.1021/ci1003637>.
- [61] Wang Z, Li J, Dang R, Liang L, Lin J. PhIN: a protein pharmacology interaction network database. *CPT Pharmacometrics Syst Pharmacol* 2015;4:160–6. <https://doi.org/10.1002/psp4.25>.
- [62] Bienfait B, Ertl P. JSME: a free molecule editor in JavaScript. *J Chem* 2013;5:24. <https://doi.org/10.1186/1758-2946-5-24>.
- [63] Li L, Li D, Xu L, Zhao P, Deng Z, Mo X, et al. Total extract of Yupingfeng attenuates bleomycin-induced pulmonary fibrosis in rats. *Phytomedicine* 2015;22:111–9. <https://doi.org/10.1016/j.phymed.2014.10.011>.
- [64] Song J, Li J, Zheng S, Jin Y, Huang Y. Anti-inflammatory and immunoregulatory effects of Yupingfeng (玉屏风) powder on chronic bronchitis rats. *Chin J Integr Med* 2013;19:353–9. <https://doi.org/10.1007/s11655-013-1442-6>.
- [65] Tahira Y, Fukuda N, Endo M, Suzuki R, Ikeda Y, Takagi H, et al. Transforming growth factor-beta expression in cardiovascular organs in stroke-prone spontaneously hypertensive rats with the development of hypertension. *Hypertens Res* 2002;25:911–8.
- [66] Peng S, Zhang Y, Zhang J, Wang H, Ren B. ERK in learning and memory: a review of recent research. *Int J Mol Sci* 2010;11:222–32. <https://doi.org/10.3390/ijms11010222>.
- [67] Huang X, He Y, Chen Y, Wu P, Gui D, Cai H, et al. Baicalin attenuates bleomycin-induced pulmonary fibrosis via adenosine A2a receptor related TGF-β1-induced ERK1/2 signaling pathway. *BMC Pulm Med* 2016;16:132. <https://doi.org/10.1186/s12890-016-0294-1>.