CPMCP: a database of Chinese patent medicine and compound prescription

Chang Sun1,2,#, Jipeng Huang1,2,#, Rong Tang1,2,#, Minglei Li1,2, Haili Yuan3, Yuxiang Wang4, Jin-Mao Wei1,2,* and Jian Liu1,2,*

1College of Computer Science, Nankai University, No. 38 Tongyan Road, Tianjin 300071, China
2Centre for Bioinformatics and Intelligent Medicine, Nankai University, No. 38 Tongyan Road, Tianjin 300071, China
3Department of Acupuncture and Moxibustion, First Teaching Hospital of Tianjin University of Traditional Chinese Medicine and National Clinical Research Center for Chinese Medicine Acupuncture and Moxibustion, No. 88 Changle Road, Tianjin 300030, China
4Tianjin Key Laboratory of Epigenetics for Organ Development of Preterm Infants, Medical Research Centre, Tianjin Fifth Central Hospital, No. 41 Zhejiang Road, Tianjin 300071, China

*Corresponding author: Tel: +8613820022571; Email: weijm@nankai.edu.cn
Correspondence may also be addressed to Jian Liu. Tel: +8613820022571; Email: jianliu@nankai.edu.cn
#These authors contributed equally to this work.

Citation details: Sun, C., Huang, J., Tang, R. et al. CPMCP: a database of Chinese patent medicine and compound prescription. Database (2022) Vol. 2022: article ID baac073; DOI: https://doi.org/10.1093/database/baac073

Abstract

Although several traditional Chinese medicine (TCM)-related databases have emerged, they focus on researching single medicinal materials, which is far from sufficient for clinical research and application. In comparison, compound prescriptions are more informative and meaningful in TCM, for they embody the information on the compatibility of TCM besides the relatively isolated information about single medicinal materials. The compatibility information is essential in TCM because it conveys not only what components are involved to treat special diseases but also how to combine these single medical materials. We established a database of Chinese patent medicine and compound prescription (CPMCP). It demonstrates the prescription information of Chinese patent medicines (CPMs) and ancient Chinese medicine prescriptions (CMPs). CPMCP reports their comprehensive and standardized information such as the components, indications and contraindications. It is worth mentioning that we organized relevant experts and spent lots of time manually mapping the functions of compound prescriptions in ancient Chinese to the standardized TCM symptom vocabularies, obtaining a total of 71,414 associations between compound prescriptions and TCM symptoms. In this way, CPMCP established the associations between TCM and modern medicine (MM) according to the associations between TCM symptoms and MM symptoms. In addition, to further exhibit the compatibility mechanism of compound prescriptions, CPMCP summarizes a set of common drug combination principles by analyzing the existing prescriptions. We believe that CPMCP can promote the modernization of TCM and make greater contributions to MM.

Database URL: http://cpmcp.top

Key points

- Chinese patent medicine and compound prescription (CPMCP) is an online database for traditional Chinese medicine (TCM).
- CPMCP promotes the modernization of TCM.
- CPMCP helps to reveal the compatibility mechanism of compound prescriptions.

Introduction

Traditional Chinese medicine (TCM) is an empirical science summed up by ancient Chinese based on thousands of years of clinical experience (1–3). Today, TCM still plays an irreplaceable role in maintaining the daily health of Chinese people (4). The doctors of TCM treat diseases by using natural medicinal materials (i.e. natural herbs, animals or minerals) (5, 6). For example, Artemisia annua was widely used to treat malaria in ancient China. In 1972, Tu et al. successfully extracted artemisinin (C15H22O5), a colorless crystal, from A. annua (7). It was confirmed that artemisinin is the main active ingredient in A. annua for the treatment of malaria. This infers that TCM and modern medicine (MM) share the same theoretical basis at the molecular level. That is, compounds treat diseases by regulating the activity of molecular targets (8–10). However, treating disease with a single natural material may cause serious side effects due to the complex ingredients of that material (11). Therefore, Chinese medicine practitioners try to treat diseases by flexibly combining a...
variety of natural medicines to form compound prescriptions (8). For example, the alkaloids in Pinellia (a natural herb) can treat hiccups and belching. However, taking unprocessed Pinellia orally will inhibit the secretion and release of PGE2, a protective mediator of the gastric mucosa, thereby injuring the gastric mucosa (which can be seen as the side-effect of Pinellia). Ginger, another herb for treating hiccups and belching, can effectively antagonize the inhibition of Pinellia to protect the gastric mucosa with its active ingredient, 6-gingerol. Therefore, TCM practitioners try to combine Pinellia and ginger, which not only can enhance the curative effect of the herbs but also can alleviate the side effects caused by Pinellia (12, 13). The compound prescription has the characteristics of multicomponent, multi-approach and multi-target. Multiple components in the prescription can act synergistically on multiple targets and regulate different signal pathways, thereby treating complex diseases.

Since 2019, the novel coronavirus (severe acute respiratory syndrome coronavirus 2), which broke out in more than 200 countries, has caused the novel coronavirus disease...
Table 1. Number and sources of various entities in CPMCP

| Entities            | Number | Source                                      |
|---------------------|--------|---------------------------------------------|
| CPMs                | 1469   | Compendium of National Standards for Chinese Patent Medicines |
| CMPs                | 656    | Various Chinese Medicine Works              |
| Medicinal materials | 1557   | Chinese Pharmacopoeia                       |
| Ingredients         | 26,341 | TCMD, TCMP, and TCM-ID                      |
| Targets             | 20,965 | HIT and TCMP                               |
| Diseases            | 14,086 | OMIM and Orphanet                          |
| MM symptom          | 1148   | MeSH, SIDER, and Unified Medical Language System (UMLS) |
| TCM symptoms        | 2285   | Chinese Pharmacopoeia, Research on the Standardization of TCM Terminology, and Pathology Terminology Standardization |

(COVID-19), a highly contagious and pathogenic viral pneumonia (14). Up to now, there are still no related specific medicines developed in the world. However, several TCM compounds have been proven to be effective in relieving the symptoms and pain of patients with COVID-19. In the COVID-19 diagnosis and treatment plan (trial version 7) of China, Qingfei Paidu Decoction (QPD) is used in the clinical treatment of mild, routine and severe patients. Zhao et al. analyzed the mechanism of QPD in the treatment of COVID-19 from the perspective of network pharmacology. QPD can be used as an antiviral agent to inhibit the translation of viral messenger RNA by targeting ribosomal proteins which are necessary for viral replication (15). Xu et al. believe that the main active ingredients in QPD against COVID-19 are quercetin, luteolin, kaempferol, naringin and isorrine (16). These five components can regulate COVID-19-related targets such as Mitogen-activated protein kinase 1 (MAPK1) and interleukin-6 and signaling pathways such as tumor necrosis factor and NF-kappa-B through compound–target interactions. Through interaction, these components can inhibit the inflammatory response, regulate immune function, reduce lung damage and thus treat COVID-19. Yao et al. also found that the main active components of QPD are well binging to targets such as RNA-dependent RNA polymerase (RdRp) and Papain Like protease (PLpr), thereby regulating a variety of related signaling pathways to treat symptoms such as fever and cough caused by COVID-19 (17). These works suggest that TCM is an important and effective medical science that can provide an important reference for the development and improvement of MM.

In recent years, several online databases related to TCM have emerged, such as HIT (18), TCMID (10), SymMap (19) and HERB (20). They reported the active components and targets of natural medicine and try to associate them with modern diseases. However, these works focus on a single medicinal material, ignoring the more clinically valuable TCM compound prescriptions.

In light of the above facts and to further facilitate clinical research and application, we established a database...
of Chinese patent medicine and compound prescription (CPMCP). It reports the prescription information of Chinese patent medicines (CPMs) and ancient Chinese medicine prescriptions (CMPs), including comprehensive and standardized information such as the combination of medicinal materials, indications and contraindications (the detailed and innovative information of CPMCP is shown in Figure 1A and B). Due to the relevant descriptions of prescriptions being recorded in ancient Chinese, we have organized relevant experts to manually map and verify the functions and indications of prescriptions to 2285 standardized TCM symptom vocabulary (as shown in Figure 1C), obtaining a total of 71,414 associations between compound prescriptions and TCM symptoms. In this way, CPMCP establishes the associations between TCM and MM according to the associations between TCM symptoms and MM symptoms. Finally, CPMCP reports the relevant information of eight categories of entities collected from different data sources, including (I) CPMs, (ii) ancient CMPs, (iii) natural medicinal materials, (iv) chemical ingredients, (v) targets, (vi) diseases, (vii) TCM symptoms and (viii) MM symptoms. The associations between these entities are also reported in our database (as shown in Figure 1A). The number and sources of various entities in CPMCP are shown in Table 1.

Since TCM prescriptions involve the use of a variety of medicinal materials, the compatibility mechanism between medicinal materials is important (21). CPMCP summarizes a set of common drug combination principles by analyzing the combination of medicinal materials in existing prescriptions. We believe that CPMCP can promote the modernization of TCM and make greater contributions to MM.

Materials and Methods

Tool implementation and usage

CPMCP is built on Python 3.9 and its framework Flask 2.0 and front-end framework React 17.0. Front-end pages are developed based on Material UI 5.0 which follows Google’s Material Design. State management is designed based on MobX 6.3. As for routing, we use React-Router 5.3. Front-end and back-end interactions are handled by axios 0.24. All front-end data’s visualization is rendered using ECharts 5.3. All data are stored and managed by MariaDB 10.7.3, an open source relational database, and the object-relational mapping middleware is Flask-SQLAlchemy 2.5.1. The operating system is based on Ubuntu 20.04 LTS version, and the web server is deployed using Nginx 1.21. All services are managed using Docker 20.10.16. The database has been tested in multiple operating systems, browsers and network environments to ensure its stability.

Data sources of CPMCP

The data sources of CPMCP can be divided into two categories: online databases and existing publications. For publications, CPMCP obtained 2202 TCM compound prescriptions, which consist of 1484 CPMs and 718 CMPs. The information on CPMs was extracted from the “Compendium of National Standards for Chinese Patent Medicines” (2002 edition), which is approved by the State Drug Administration of China. The details of CMPs were collected from various classical TCM-related works. CPMCP shows the names, components, usages, applications and contraindications of these prescriptions. According to the components and the applications of these prescriptions, CPMCP extracted 1575 natural medicinal materials and 2285 TCM symptoms from the Chinese Pharmacopoeia (2015 edition). The names of TCM symptoms were standardized according to an authoritative TCM publication named “Research on the Standardization of TCM Terminology” (2016 Edition) and a published TCM terminology integration platform. According to “Pathology Terminology Standardization” (2015 edition), the CPMCP obtained information on the definitions, loci and attributes of these symptoms.

In terms of database, CPMCP collected MM symptom terms from MeSH database (2015 version) (22) and SIDER database (2016 version) (23). According to the annotations in the SymMap database (19), the TCM symptoms were mapped to MM symptoms. After that, we extracted the chemical ingredient information of natural medicinal materials from TCMD (10), TCMSP (24), and TCM-ID databases (25) and obtained the target genes of these ingredients from HIT (18) and TCMSP databases. Diseases in CPMCP were gotten from OMIM (26) and Orphanet databases (27). The disease-gene associations were obtained from HPO (28), DrugBank (29) and National Center for Biotechnology Information (NCBI) databases (30). The number and sources of various entities in CPMCP are shown in Table 1.

Results

Search for information that interests you

On the home page (Figure 2), you can search for the content you are interested in by selecting different filters, including CPMs, CMPs, medicinal materials, targets, diseases, TCM symptoms and MM symptoms (in this step, fuzzy search is supported). The specific items contained in each type of entity

Figure 3. Associations between a CPM and other entities.
are given in Figure 1B. For each type of entity, we have provided some samples below the search box for users’ reference. If you want to browse all entries for a certain type of entity, you can click the search button directly by specifying the filter type.

The associations between entities are visualized

According to the different types of associations among the entities, CPMCP constructed and visualized the heterogeneous graph for each entity. Taking the CPM, Egui Yangxue Koufu, as an example (as shown in Figure 3), its components (medicinal materials and ingredients) and related TCM symptom are demonstrated intuitively.

It is worth noting that the functions and indications of CPMs and CMPs are described in ancient Chinese, which brings difficulties to practitioners of TCM in the process of using them. We organized relevant experts to manually map the functions of compound prescriptions in ancient Chinese to the standardized TCM symptom vocabularies (as shown in Figure 1C), obtaining a total of 71 414 associations between compound prescriptions and TCM symptoms. In this way, CPMCP established the association between compound prescription and disease through the meta-path: compound prescriptions → TCM symptoms → MM symptoms → diseases. This function of CPMCP contributes to the modernization of TCM. Moreover, CPMCP established another meta-path: compound prescriptions → medical materials → ingredients → targets → diseases. This meta-path can explain how a prescription treats a specific disease from the perspective of MM. It provides much valuable information for the interpretability of TCM.

CPMCP summarized the common compatibility of medicinal materials

The compatibility information is essential in TCM because it conveys not only what components are involved to treat special diseases but also how to combine these single medical materials. Therefore, when TCM practitioners construct a TCM compound prescription, they need to check the compatibility mechanism of the medicinal materials involved in the prescription. CPMCP summarizes a set of common drug combination principles by analyzing the combination of medicinal materials in existing prescriptions.

For each medicinal material, CPMCP built a doughnut chart to demonstrate the nine most common compatibilities (Figure 4). Complete compatibility information was given in the form of a table sorted by their frequency (Figure 5). For example, in Figure 5, we can intuitively see that Baishao has appeared together with Danggui 63 times. It can be inferred from this that Baishao and Danggui are often used together. In this way, CPMCP revealed the habits and taboos of collocation among medicinal materials.

Figure 4. Nine most common compatibilities of a medicinal material named Baishao.
### Conclusion

In this work, a TCM-related database, CPMCP, was proposed. CPMCP reports the comprehensive and standardized information of CPMs and ancient CMPs, such as their components, indications and contraindications. To promote the modernization of TCM, CPMCP mapped the instructions of TCM compound prescriptions to the standardized TCM symptom vocabularies manually. In this way, CPMCP established the associations between the TCM compound prescriptions and diseases. Since the compatibility information is essential in TCM, CPMCP summarized the common compatibility of medicinal materials by analyzing the existing TCM compound prescriptions. This function is helpful for revealing the habits and taboos of collocation among medicinal materials. We believe that CPMCP can promote the modernization of TCM and make greater contributions to MM.

### Funding

National Key R&D Programs of China (2018YFC1603800, 2018YFC1603802, 2020YFA0908700, 2020YFA0908702); National Natural Science Foundation of China (61772288, 61872115); Natural Science Foundation of Tianjin City (18JCZDJC30900).

### Conflict of interest

The authors have declared no competing interests.

### Authors’ contributions

C.S. wrote the manuscript, R.T. collected and processed the data, J.H. developed the website, C.S., J.H. and M.L. assisted...
data processing, H.Y and Y.W. verified the data and J.-M.W. and J.L. conceived the study.

References
1. Cheung,F. (2011) TCM: made in China. Nature, 480, 882–883. 10.1038/480882a.
2. Wenli,S., Shahrajabian,M.H. and Qi,C. (2019) Therapeutic roles of goji berry and ginseng in traditional Chinese. J. Nutri. Food Sec., 4, 293–305.
3. Zhang,Y.-W. and Cheng,Y.-C. (2019) Challenge and prospect of traditional Chinese medicine in depression treatment. Front. Neurosci., 13, 190. 10.3389/fnins.2019.00190.
4. Liu,Y., Yang,Z., Cheng,J. et al. (2016) Barriers and countermeasures in developing traditional Chinese medicine in Europe. Front Med., 10, 360–376. 10.1007/s11684-016-0455-x.
5. Cheng,X., Chen,X., Su,X. et al. (2014) DNA extraction protocol for biological ingredient analysis of Liuwei Dihuang Wan. Genom. Proteom. Bioinform., 12, 137–143. 10.1016/j.gpb.2014.03.002.
6. Zhao,J., Jiang,P. and Zhang,W. (2010) Molecular networks for the study of TCM Pharmacology. Brief. Bioinform., 11, 417–430. 10.1093/bib/bbp063.
7. Tu,Y. (2016) Artemisinin—a gift from traditional Chinese medicine to the world (Nobel lecture). Angew. Chem. Int. Ed. Engl., 55, 10210–10226. 10.1002/anie.201601967.
8. Hai-Yu,X., Yan-Qiong,Z., Zhen-Ming,L. et al. (2018) ETGM: an encyclopedia of traditional Chinese medicine. Nucleic Acids Res., D1. 10.1093/nar/gky987.
9. Qiao,L., Huang,W., Zhang,X. et al. (2021) Evaluation of the immunomodulatory effects of anti-COVID-19 TCM formulae by multiple virus-related pathways. Sig. Transduct. Target. Ther., 6, 1. 10.1038/s41392-021-00475-w.
10. Ruichao,X., Zhao,F., Meixia,Z. et al. (2013) TCMID: traditional Chinese medicine integrative database for herb molecular mechanism analysis. Nucleic Acids Res., 41, D1089–D1095. 10.1093/nar/gkt1100.
11. Zhao,Z., Li,Y., Zhou,L. et al. (2020) Prevention and treatment of COVID-19 using Traditional Chinese Medicine: a review. Phytomedicine, 85, 153308. 10.1016/j.phymed.2020.153308.
12. Fueki,T., Tanaka,K., Obara,K. et al. (2020) The acrid raphides in tuberous root of Pinellia ternata have lipophilic character and are specifically denatured by ginger extract. J. Nat. Med., 74, 722–731. 10.1007/s11168-020-00475-w.
13. Wu,H., Tan,X., Cai,B. et al. (1996) Effect of ginger-processing on l-ephedrine contents in rhizoma Pinelliae. China J. Chinese Materia Medica, 21, 157–158.
14. Xu,Z., Li,Y., Lei,Q. et al. (2021) COVID-ONE-hi: the one-stop database for COVID-19-specific humoral immunity and clinical parameters. Gen. Proteom. Bioinform., 19, 669–678. 10.1016/j.gpb.2021.09.006.
15. Zhao,D., Cai,H., Luo,W. et al. (2020) One case of COVID-19 in Changchun treated by integrated Traditional Chinese and Western Medicine. Jilin J. Chin. Med.
16. Xu,D., Xu,Y., Wang,Z. et al. (2020) Mechanism of Qingfeipaidu decoction on COVID-19 based on network pharmacology. Trad. Chin. Med. Pharmacol. Clinic.
17. Yao,W., Zhai,Y., Lin,L. et al. (2020) Preliminary study on active ingredients of qingfei oral liquid combining “composition-target-common pathway” network and molecular docking technology against new coronavirus pneumonia. Nanjing J. Univ. Trad. Chin. Med.
18. Yan,D., Zheng,G., Wang,C. et al. (2021) HIT 2.0: an enhanced platform for herbal ingredients’ targets. Nucleic Acids Res., 50, D1238–D1243. 10.1093/nar/gkaa1011.
19. Yang,W., Feilong,Z., Kuo,S. et al. (2018) SymMap: an integrative database of traditional Chinese medicine enhanced by symptom mapping. Nucleic Acids Res., 47, D1110–D1117. 10.1093/nar/gky1021.
20. Fang,S.S., Dong,L., Liu,L. et al. (2020) HERB: a high-throughput experiment- and reference-guided database of traditional Chinese medicine. Nucleic Acids Res., 49, D1197–D1206. 10.1093/nar/gkaa1063.
21. Wang,S., Hu,Y., Tan,W. et al. (2012) Compatibility art of traditional Chinese medicine: from the perspective of herb pairs. J. Ethnopharmacol., 143, 412–423. 10.1016/j.jep.2012.07.033.
22. Minguet,F., Salgado,T.M., Boogerd,L. et al. (2015) Quality of pharmacy-specific Medical Subject Headings (MeSH) assignment in pharmacy journals indexed in MEDLINE. Res. Soc. Adm. Phar., 11, 686–695. 10.1016/j.sapharm.2014.11.004.
23. Michael,K., Ivica,L., Juhl,J.L. et al. (2016) The SIDER database of drugs and side effects. Nucleic Acids Res., 44, D1075–D1079. 10.1093/nar/gkv1075.
24. Ru,J., Li,P., Wang,J. et al. (2014) TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. J. Cheminform., 6, 13–19. 10.1186/1758-2946-6-13.
25. Chen,X., Zhou,H., Liu,Y. et al. (2006) Database of traditional Chinese medicine and its application to studies of mechanism and to prescription validation. Br. J. Pharmacol., 149, 1092–1103. 10.1038/sj.bjp.0706945.
26. Amberger,J.S., Bocchini,C.A., Schiettecatte,F. et al. (2015) OMIM.org: Online Mendelian Inheritance in Man (OMIM), an online catalog of human genes and genetic disorders. Nucleic Acids Res., 43, D789–D798. 10.1093/nar/gku1205.
27. Weinreich,S.S., Mangon,R., Sikkens,J.J. et al. (2008) Orphanet: a European database for rare diseases. Ned. Tijdshr. Geneskd., 152, 518–519.
28. Köhler,S., Vasilyevsky,N.A., Engelstad,M. et al. (2017) The human phenotype ontology in 2017. Nucleic Acids Res., 45, D865–D876. 10.1093/nar/gkw1039.
29. Wishart,D.S., Feunang,Y.D., Guo,A.C. et al. (2018) DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res., 46, D1074–D1082. 10.1093/nar/gky1037.
30. Brown,G.R., Hem,Y., Katz,K.S. et al. (2015) Gene: a gene-centered information resource at NCBI. Nucleic Acids Res., 43, D36–D42. 10.1093/nar/gku1055.