Progress and Prospects of Research Ideas and Methods in the Network Pharmacology of Traditional Chinese Medicine

Zhongzhu Yuan1, Yingying Pan2, Tian Leng1, Yu Chu1, Haojie Zhang1, Jingru Ma1, Xiaoju Ma1

1School of Public Health, Chengdu University of Traditional Chinese Medicine, Chengdu, China; 2School of Basic Medicine, Chengdu University of Traditional Chinese Medicine, Chengdu, China.

Corresponding author: Xiaoju Ma, School of Public Health, Chengdu University of Traditional Chinese Medicine, Wenjiang campus, 1166 Liutai Avenue, Wenjiang District, Chengdu, Sichuan, China 611137; email: maxiao@cdutcm.edu.cn; ORCID: 0000-0001-7747-X

Received, May 24, 2022; Revised, June 5, 2022; Accepted, June 6, 2022; Published, June 19, 2022

ABSTRACT -- In recent years, the emerging network pharmacology has been extensively applied to the field of traditional Chinese medicine and has made great contributions to the modernization of TCM. Therefore, this paper provides an overview of the progress of research ideas and methods in the network pharmacology in the last few years in the field of traditional Chinese medicine and presents insights into future research methods and ideas in the network pharmacology. Problems with the current network pharmacology are discussed and prospects of its future development are put forward.

INTRODUCTION

In recent decades, traditional Chinese medicine (TCM) has attracted worldwide attention for its significant efficacy, relatively low toxicity, and cost (1-3). In addition, there is ample evidence that many widely used modern pharmaceuticals are derived from TCM (4, 5) and that TCM itself is characterized by “multi-component, multi-target and multi-pathway” (6, 7). The difficulty of modernizing traditional Chinese medicine lies in explaining its mechanism of treatment, which is also a problem that needs to be solved in the modernization of TCM (8). The thinking behind TCM is guided by the “holistic concept” and “dialectical argumentation,” emphasizing the integrity, unity, and relevance of things themselves (9, 10).

The recently emerging network pharmacology is a comprehensive discipline that integrates the theory of systematic biology (11), pharmacology, information network and computer science, using computer simulations and various databases to screen drug molecular targets and disease-related targets. It uses high-throughput screening, network visualization, and network analysis to reveal the complex network connection between drugs, targets, and diseases, analyzes and predicts the mechanisms of drugs. If necessary, it also further validates the effects of the drugs through corresponding experiments (12).

This paper concentrates on the progress of the network pharmacology in the field of TCM for the treatment of diseases and the general ideas of its research, as well as its problems in recent years, in order to provide reference for the application of the network pharmacology in the treatment of diseases, the research, and development of TCM.

THE DEVELOPMENT HISTORY OF NETWORK PHARMACOLOGY

With the generation of large amounts of data in the medical and life sciences, the emergence of cross-disciplines such as systems biology, bioinformatics, and computational biology has been promoted. Moreover, the network pharmacology as a discipline has started and developed steadily and rapidly (Figure 1).

In 1999, Chinese scholar Li proposed that there is a link between TCM and biomolecular networks (13). In 2000, Melita et al. also suggested the importance of the internet in the development of pharmacology in Europe and the need to set additional standards for modern pharmacological information from the internet (14).

In 2002, Li et al. suggested that TCM prescriptions may have an overall effect on gene networks of complicated diseases (15). In 2007, Yildirim et al. combined biology and networks to study drugs and found that most pharmaceuticals do not directly perturb target proteins (16). In the same year, Hopkins proposed network pharmacology for the first time describing it as “the next paradigm of pharmaceuticals research.” The network pharmacology suggests that pharmaceuticals act on multiple targets and produce therapeutic effects...
The development history of the network pharmacology

Figure 1. The development history of the network pharmacology through their interactions (17, 18).

In 2009, Pan proposed a new model for developing pharmaceuticals based on the network pharmacology (19). Shao Li proposed a "phenotypic-biological-TCM network" model for the study of TCM syndromes and prescriptions (20). And in 2010, Liu et al. focused on the network pharmacology and its progress, while expounding on its directions of development and prospects of application. They also elaborated on its limitations and shortcomings (21).

In 2011, Li put forward the concept of "network targets" and proposed an algorithm to predict collaborative drug combination based on network targets (22). Since then, the network pharmacology has grown swiftly, and its research methods have gradually matured. We summarized the research methods and ideas of the network pharmacology in recent years (Table 1).

In 2012, Li et al collected the active ingredients of Ligusticum Chuanxiong, Dalbergia Odorifera and Corydalis Yanhusuo. The 3D structures of the active components and the targets associated with cardiovascular diseases. Then the disease-related targets and active ingredients were used as reverse molecules docked to obtain the therapeutic targets, and finally pathway enrichment analysis was performed, including renin-angiotensin-aldosterone (RAAS) pathway and vascular endothelial growth factor (VEGF) pathway, to possibly elaborate on the three key mechanisms of TCM to cure cardiovascular diseases (23).

In 2014, Li et al. also obtained the anti-T2DM components and effective targets in the Ge-Gen-Qin-Lian decoction (GGQLD) formula through the reverse molecular docking approach, and the effective targets were analyzed through a GO enrichment tool, demonstrating that GGQLD could treat T2DM by modulating the complex network associated with multiple pathological processes of T2DM. Finally, 4- Hydroxymephenytoin was selected for in vitro experiments to find out its effects on insulin secretion and resistance, which revealed its anti-diabetic potential (24).

In 2016, Tang et al. explored the active ingredients of XuanHuSuo powder (XHSP) and their corresponding targets to screen for targets related to osteoarthritis (OA). They drew the Protein-Protein Interaction (PPI) topology of "drug ingredient target – disease target – other human genes", considering the indirect interaction between genes. 41 major nodes were selected to go through GO enrichment analysis and were found related with regulation of nitric oxide (NO) biosynthetic process, response to interleukin-1 (IL-1), regulation of interleukin-1β (IL-1β) production, response to cytokine stimuli, and response to estrogen stimuli (25).

In 2016, Tang et al. found the active ingredients and corresponding targets of Xiaoyao powder (XYP) and searched the Therapeutic Target Database TTD for Infertility-related targets. A PPI network of "drug ingredient targets - disease targets - other human genes" was constructed. Then the pivotal genes were selected for GO and KEGG enrichment analyses, and possible mechanisms of XYP to cure Infertility was obtained (26).

In 2018, Xu et al. predicted the relevant targets of Baicalin and explored known relevant targets of Ischemic Stroke. The intersection of the two targets was used to construct a PPI network. Key targets were selected using the Cytoscape plugin, and the targets were subject to GO and KEGG enrichment analysis to present possible mechanisms of Baicalin to cure Ischemic Stroke (27).

In 2019, Li et al. constructed a PPI network by taking the intersection of the active targets of Gypenosides and Thyroid-Associated Ophthalmopathy-related targets to select the key targets for GO and KEGG enrichment analysis. And they finally verified the efficacy of the drug ingredients by molecular docking. The potential mechanisms of Gypenosides for the treatment of
The development course of the network pharmacology research ideas.

<table>
<thead>
<tr>
<th>No.</th>
<th>Year</th>
<th>Author</th>
<th>Title of study</th>
<th>Methods of study</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>March 2016</td>
<td>Tang, et al.</td>
<td>A Network Pharmacology Approach to Uncover the Pharmacological Mechanism of XuanHuaSuo Powder on Osteoarthritis.</td>
<td>Target of union set, Construction of PPI network, GO enrichment analysis</td>
</tr>
<tr>
<td>5</td>
<td>April 2018</td>
<td>Xu, et al.</td>
<td>Systematic Understanding of the Mechanism of Baiacalin against Ischemic Stroke through a Network Pharmacology Approach.</td>
<td>The target of intersection, Construction of PPI, GO and KEGG enrichment analysis</td>
</tr>
<tr>
<td>6</td>
<td>July 2019</td>
<td>Li, et al.</td>
<td>Research on the Potential Mechanism of Gypenosides on Treating Thyroid-Associated Ophthalmopathy Based on Network Pharmacology.</td>
<td>The target of intersection, Construction of PPI, GO and KEGG enrichment analysis, verification about Molecular docking</td>
</tr>
</tbody>
</table>

Thyroid-Associated Ophthalmopathy was also searched (28).

Thus far, the number of related studies in China and abroad has gradually increased, and plenty of articles related to the network pharmacology have emerged one after another, attracting more and more attention from the academic community (29). This has brought unprecedented opportunities for TCM research.

**CURRENT RESEARCH IDEAS OF NETWORK PHARMACOLOGY IN TCM**

With compound (multiple potential active ingredients) are studied along the same lines as single component drugs. The today’s development of the network pharmacology in TCM concentrates mainly on compound drugs owing to their abundance as compared to single component drugs.

**General Network Research Steps**

The first step is to get the active ingredients and the corresponding targets of the compound in the drug. The data come from two sources, one is directly from the database; another is after detection of the main active ingredients using analytical techniques such as liquid chromatography and mass spectrometry (30) and finding the corresponding targets from the database. The main databases include TCMSP (31), TCMIP (32), and BATMAN-TCM (33). As the algorithms used in each database are different, we believe that data from two or more databases may be subject to a second screening through third-party databases, such as using the SwissADMA database in combination with the PubChem database to identify the final active ingredients and its corresponding targets.

In the steps to obtain disease-related targets, we conclude that the data sources mainly come from various databases by searching keywords. There are two categories of databases. One includes Online Mendelian Inheritance in Man (OMIM) database (34), Genecards database (35), Uniprot database (36), and Therapeutic Target Database (TTD) (33) etc., which retrieve disease-related targets directly. The other retrieves clinical samples and their corresponding matrix data, such as the gene expression omnibus database (GEO) (37) and The Cancer Genome Atlas (TCGA) database (38).

Intersect a drug target that corresponds to a certain disease and use that target to develop treatment with TCM prescription. In order to find the
key targets, the interactions between the intersected targets will be clarified by the STRING database (39) to set up a PPI network.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses can be used to study the mechanism of compound drugs to cure diseases. GO analysis includes Cellular Component (BP), Cellular Component (CC) and Molecular Function (MF). KEGG analyzes the pathways for the treatments of diseases. To better demonstrate the treating mechanisms of compound drugs, we can map an “active ingredients-targets-pathways” network by Cytoscape.

To further validate the effectiveness of the crucial drug components. Molecular docking is performed to validate the core target and corresponding ingredients. The active ingredient and targets’ protein structures are obtained at RSCB PDB (Protein Structure Database) (https://www.rcsb.org) (40), molecular docking is performed using AutoDock Vina (41). Lastly, outcomes of molecular docking with higher activity will be visualized by PyMol.

Finally, in vitro, or in vivo experiments are performed for validation.

Existing Research Ideas of TCM Prescriptions

The purpose of the current network pharmacology is mostly to study the mechanisms of action of prescriptions. TCM prescriptions are divided into categories of clinical and no clinical applications. The co clinical application include studies involving studies such as detection of responsible genes. We found that the first one can be used to study prescriptions’ mechanisms to treat diseases by the general network pharmacology research approach. Zeng et al. elucidated the mechanism of Qinghuang Powder to treat Myeloid Leukemia by the general network pharmacology research ideas (42). Yang et al. described the mechanism of Xinbao pill in curing myocardial ischemia-reperfusion injury based on the general network pharmacology research ideas. The idea of Yang’s research is to evaluate the clinical effectiveness of the prescriptions by meta-analysis, and then apply those prescriptions with good efficacy to explore their mechanisms in disease treatments using the general network pharmacology research ideas (43). Wang et al. evaluated the clinical efficacy of Sanzi Yangqin decoction (SZYQD) in the treatment of chronic obstructive pulmonary disease (COPD) and analyzed the mechanism through the network pharmacology. They chose luteolin for in vivo validation and found that it significantly inhibits COPD-related targets (44). Geng et al. evaluated the clinical efficacy of the Zhishe Tongluo capsule in the treatment of cerebral infarction by combining meta-analysis and the network pharmacology and explored its mechanism. The core targets include ALB, AKT-1, IL-6, and the key active ingredient is Quercetin.

New Formulae. New formulae can be obtained by data mining, and there are two ways of data mining now: one is from prescription dictionaries, Chu et al. acquired the data on hypoglycemic prescriptions from Pharmacopoeia of the People's Republic of China 2020 and Drug Standard Query database. They analyzed the frequency and linking strength by Spass Modeler. Combined with the top 10 edible traditional Chinese medicines in the list of health-food for diabetes, four Chinese herbal medicines were selected by Spass Modeler. Then, the mechanism for the treatment of type II diabetes was explored by the network pharmacology. Another way of data mining is to conduct literature mining based on meta-analysis and then calculate high-frequency Chinese herbal medicines to construct a core of effective prescriptions (45). Zhu et al. evaluated the efficacy and safety of Chinese herbal medicines in the treatment of metastatic colorectal cancer (mCRC) by meta-analysis. 24 herbal medicines with high frequency of being prescribed were selected to make effective prescriptions. Their active ingredients and potential targets were explored by the network pharmacology approach (46). Among two data mining methods, the former one has an abundance of basic data however, prescriptions obtained may not always be clinically validated. The latter one is able to select clinically validated prescriptions by changing the screening criteria in meta-analysis. To sum up, results from the latter are more reliable.

Obtaining Disease-related Genes. Disease targets are mostly obtained directly from database. Clinical samples can be obtained from clinical databases. The comparative analysis of genomes between control groups and disease groups can be used to obtain differential genes i.e., disease targets. In the screening step of disease-related targets, Weighted gene co-expression network analysis (WGCNA) can be used to analyze the targets to make the data more credible and thus help to present the mechanisms of formulas in treating the disease.

WGCNA, a systems biology approach used to analyze genes’ correlations between different
specimens (46), is able to identify gene modules with frequent synergistic expressions. Then, validate the core genes or therapeutic targets according to the inter-connectivity of gene modules and the correlations between phenomes (47). WGCNA identifies gene modules that are clinically related from numerous genes and, lastly finds out the key genes related to the diseases through inter-connectivity of gene modules and the importance of genes for further validation (48).

Based on WGCNA, Liu et al. conducted a comprehensive analysis of the GSE51674 and GSE63446 datasets in GEO database, screened the differentially expressed miRNA has-miR-574 and obtained 18 corresponding targets, including CYP19A1, HTR2A, F7, ATP1A1, and others. They took the intersection of the core targets from PPI analysis with the corresponding targets of has-miR-574 and then performed molecular docking to validate the mechanism of Liu Wei Di Huang Wan in the treatment of osteoporosis associated with diabetic nephropathy (49).

Huang et al., based on WGCNA, analyzed 23518 gene expression profiles related to HCC from GEO database, the generated positive and negative correlation modules were matched respectively with elevated and down-regulated genes from differential genes. Then, the results were intersected with corresponding targets of Huangqin to obtain 21 therapeutic targets, including CETP, CYP26A1, TYMS, NEK2, etc. Finally, AURKB, CHEK1, and NEK2 were identified as the first three potential targets of Huangqin for the treatment of HCC by PPI, GO, and other analyses (50).

Wu et al. combined survival analysis and differential gene analysis based on WGCNA to obtain genes associated with pancreatic cancer from TCGA database. Then, the gene targets associated with pancreatic cancer from the TTD database were combined together with duplicate values being removed to obtain disease-related targets. The components and corresponding targets of the compound bitter ginseng injection were intersected with the disease-related targets. The PPI network of PC-CKI was therefore constructed to obtain 15 key potential targets, including AKT1, MAPK1, CCNB1, MAPK3, etc. (51).

According to the analysis results of WGCNA analysis, genes within the same module may have synergistic regulation, functional correlation, or be in the same pathway (52). It is a useful source for identifying core genes or therapeutic targets.

### FUTURE NETWORK PHARMACOLOGY

#### RESEARCH IDEAS

**Rote Learning: LASSO Regression**

In recent years, the network pharmacology has flourished, and its research ideas have matured. Hence, there has been a trend of deeper interactions between network pharmacology and computations, clinical studies, and experiments, as well as the interdisciplinary development with information science, biology, and medicine. Based on the current trend of network pharmacology, we suggest that the data of disease-related targets should come from more specialized databases. For example, a clinical data matrix could be chosen in terms of access to disease targets. This leads to the introduction of a new way data processing, LASSO regression in rote learning.

Least absolute shrinkage and selection operator (Lasso) regression was first proposed by Robert Tibshirani. It is an algorithm with penalty function, which can compress the number of subsets and is suitable to analyze the matrix data set correlated with each other between each data. And the disease-related data sets extracted from databases such as TCGA and GEO just meet the applicable conditions. Currently, Lasso regression has been successfully applied to Genome-Wide Association Study (GWAS) (53). It can also select gene loci in candidate gene studies (54) and detect gene-to-gene interactions (55).

Liu et al. analyzed 155 differential lncRNAs associated with breast cancer from TCGA database by a combination of LASSO regression, SVM-RFE algorithm and Cox regression to select 7 lncRNAs and build prognostic models. Under the guidance of co-expression analysis, they predicted that LINC01215 is a central immune-associated lncRNA, highly related to multiple immune pathways. It showed that 7 lncRNAs were related to survival outcomes of patients with breast cancer, immune infiltration degree and even tumor mutation load scores in breast cancer (56). Xu et al. analyzed 1149 m6A-related lncRNAs from TCGA database by LASSO-Cox regression and constructed a risk model with 12 m6A-related lncRNA, this provided clues for the prognosis prediction of LUAD patients and may contribute to explaining the mechanisms and processes of m6A-related lncRNAs (57).

Genome-wide association studies can identify risk genes that influence diseases. A large number of gene loci are often involved in GWAS studies. Such genetic studies in relation to a multitude of variables is often subject to the difficulty of reproducing the
results (58). Using Lasso regression, the number of SNPs can be appropriately reduced and genes that are consistently associated with the outcome variables can be screened to build reproducible models. There are three main advantages of using Lasso regression models (59-63), (i) able to deal with the multi-dimensional problems of the genomes; (ii) able to deal with multi-collinearity caused by LD; (iii) able to deal with multiple comparisons.

In contrast to WGCNA, the data of LASSO regression are derived from clinical databases like TCGA and GEO, etc. with complete cases or tissue samples. Data from these clinical databases generally require to be pre-processed in order to remove unimportant data and be used to analyze differentially expressed genes. Both of them can also be combined to screen for disease-related targets. We believe that TCM in network pharmacology may be able to incorporate more bioinformatics-related analysis methods in the future to increase the credibility of data.

The Disease as the Starting Point

The aims of most of the current research in the network pharmacology is to explore the mechanisms of Chinese herbal medicines that are already practiced clinically, i.e., the starting point is the TCM. However, we believe that since the ultimate goal is to control diseases, it may be possible to take the diseases as the starting point, collect the targets related to certain diseases, and screen out the core therapeutic targets, pathways, and biological processes through bioinformatics analysis methods. Then choose the appropriate drugs, and finally, drug groups for the disease can be confirmed through a systemic and holistic view.

CONCLUSION

Network pharmacology interprets the mechanisms of drugs in the treatment of diseases from the proposition of complex biological networks which coincides with the need to explore the true effectiveness of TCM. Most of the existing TCM prescriptions are derived from experiences without scientific explanations, but with their efficacy being based on empirical clinical practice. The application of the network pharmacology in TCM prescriptions is expected to achieve the transformation from empirical medicine to evidence-based medicine (64).

Despite its promising future, there are still some limitations in the network pharmacology. Firstly, the current network pharmacology in TCM mostly focuses on static theoretical analysis (64), while the metabolism in living organism is a dynamic living process. In addition, as the network pharmacology continues to develop, various database problems have gradually been found, one of which is that they have algorithms that produce different results. Thus, it is necessary to choose the appropriate and effective algorithm. Another is that some databases are incomplete. Theory needs to be combined with practice, and nowadays in vivo or in vitro experiments are generally required to make the analyses more credible. Finally, we believe that the network pharmacology is limited to a certain extent by the fact that it mines existing databases for the study of pathways and biological processes, which restricts the discovery of new targets, pathways, and biological processes. Nevertheless, the network pharmacology provides a strong impetus to interpret the principles of TCM based on modern science and technology.

REFERENCES


13. Li S. Possible correlation between TMC syndromes and regulatory mechanisms of molecular networks; proceedings of the Academic annual meeting of china association for science and technology, Hangzhou, zhejiang, China, F, 1999


30. Chen Q, Zhao Y, Li M, et al. HPLC-MS and


46. Chu Y, Chen H, Lv H, et al. Mechanism study on that treatment of type 2 diabetes mellitus with medicinal and edible Chinese medicine base on network pharmacology and molecular docking Chinese Traditional Patent Medicine, 1-6,


49. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis


