Strategy of Systems Biology for Visualizing the “Black Box” of Traditional Chinese Medicine

Yu Gu, Gao-Song Wu, Hou-Kai Li, Wei-Dong Zhang

Institute of Interdisciplinary Integrative Medicine Research, Shanghai University of Traditional Chinese Medicine, Shanghai, China; Department of Phytochemistry, College of Pharmacy, Second Military Medical University, Shanghai, China.

Abstract

Traditional Chinese medicine (TCM) has been practiced for thousands of years in China. TCM formula, usually composed of several or even dozens of herbal medicines, is the main form of TCM practicing, which is extremely complex due to multiple components and therapeutic targets, especially the characteristics of formula compatibility. Thus, it is an enormous challenge for the modernization of TCM. Systems biology is a strategy for investigating the complex interactions between genes, mRNA, proteins, and metabolites by using integrated omics approaches. In recent years, systems biology has been increasingly adopted in TCM study. This review comprehensively summarized status of syndrome and application of TCM formulae in clinical and preclinical studies and discussed the advances of systems biology in TCM research. Then, a “Disease-Syndrome-Formulae-Effect” strategy was proposed for TCM research. Combination of systems biology and “Disease-Syndrome-Formulae-Effect” strategy provided a novel approach to understand the complex interactions among biological systems, drugs, and complex diseases from a network perspective, thus facilitating the modernization of TCM. The objective of this manuscript is to provide comprehensive and up-to-date review on the application of systems biology in TCM research, as well as the perspective of TCM modernization with systems biology.

Keywords: Formula, omics, systems biology, traditional Chinese medicine

Introduction

“One target, one component” mode is classical mode for drug development for dozens of years, especially in Western medicine. However, this classical strategy encounters huge challenge nowadays because there are few diseases that are due to the abnormality of single gene.\(^1\)\(^-\)\(^4\) Recently, multitargeted agents have been used to treat disease with better efficacy and safety, shifting the therapeutic approach from a “one target, one drug” mode to a “network target, multiple component therapeutics” mode.\(^5\)\(^-\)\(^7\) For example, a combination of metformin and aspirin can produce a synergistic effect on pancreatic cancer by action at different targets.\(^8\) Sunitinib is a multitargeted kinase inhibitor used for renal cell carcinoma therapy.\(^9\) Thus, a network-targeted combination would be a better therapy for the treatment of some diseases. The human body is not only a complex and highly connected system but also dynamically adjusts within the boundary, that is, the so-called homeostasis.\(^10\) A certain physiological or pathological phenomenon should not be treated as a separate entity, but rather it should be treated and prevented from a holistic perspective, which is similar with the “holistic view” in TCM.\(^11\)

TCM has been developed and practiced for thousands of years in China, which has made great contribution for the

Address for correspondence: Prof. Wei-Dong Zhang, Institute of Interdisciplinary Integrative Medicine Research, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China; Department of Phytochemistry, College of Pharmacy, Second Military Medical University, Shanghai 200433, China.
E-mail: wdzhangy@hotmail.com
Prof. Hou-Kai Li, Institute of Interdisciplinary Integrative Medicine Research, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China.
E-mail: houkai1976@126.com

Received: 02-01-2020, Accepted: 01-05-2020, Published: 05-08-2020

How to cite this article: Gu Y, Wu GS, Li HK, Zhang WD. Strategy of systems biology for visualizing the “Black box” of traditional Chinese medicine. World J Tradit Chin Med 2020;6:260-70.
prevention and treatment of disease in China and even Asian countries. There are a variety of theories in TCM such as “Yin-Yang” and “Five Elements.” The practice of TCM is mainly through prescription of formulas, which usually compose two or more medicines (including herbs, animals, and minerals) and contain hundreds or even thousands of components. Actually, TCM formula is a “black box” in terms of its complex components and mechanisms. It is almost impossible to understand the mechanism of TCM with the reductionism-oriented approaches.

Although some advances have been made such as active component and therapeutic target identification of TCM, there is still a long way to go for elucidating the exact mechanisms underlying TCM practice. As a result, systems biology is highly valued for bridging TCM with modern science. Systems biology is focused on studying the interactions among molecules, cells, and organs and integrates different types of biological information (genomes, RNAs, proteins, metabolites, phenotypes, and so on) based on high-throughput platforms and computational and mathematical modeling. High-throughput techniques and information-rich assays can be used to decipher the pharmacodynamic basis and quality control of formula and shift the research paradigm from “one target, one component” mode to “network target, multicomponent” mode. In this review, we mainly discussed the characteristics of TCM theory and omics-based systems biology approaches. We also proposed a model of “Disease-Syndrome-Formulae-Effect” for TCM study.

**Characteristics of Traditional Chinese Medicine**

TCM is a systemic medicine and views human body as a wholeness [Figure 1]. There are two core characteristics of TCM: (1) Holism. TCM believes that the human body is an organic whole, which is composed of organs and tissues. Each tissue and organ have its unique functions to maintain the health of body; thus, TCM treatment emphasizes restoration of homeostasis of the whole person instead of curing the ill organs and (2) syndrome differentiation and treatment (meaning TCM treatment according to syndrome differentiation). Syndrome, also called “Zheng” in Chinese, is another core characteristic in TCM, which is different with the “syndrome of disease” in Western medicine. Syndrome or Zheng is a phenotypic summary during the development of disease, which usually reflects the general health status of the patients and varies accompanied with the progression of disease, that is, the same disease may have different syndromes and same syndrome in different diseases. Therefore, syndrome differentiation of disease lays the foundation for personalized disease therapy, in that different diseases could be treated with the same method on the basis of identical syndrome, and the same disease could be treated with different ways due to diversified syndromes present.

At present, most clinical used medicines are single target oriented, with inevitable side effects and limited efficacy. Actually, the etiology of most diseases is far more complicated than supposed to be single gene related. The principle of TCM formulae usually includes the four functional components: Jun (君), Chen (臣), Zuo (佐), and Shi (使). Medicines of Jun target the main cause of the disease, while those of Chen are supposed to increase the efficacy of Jun drug and relieve the secondary symptoms. Medicines of Zuo facilitate the therapeutic effects of Jun and Chen and reduce the toxic or side effects of these medicines. Shi usually promotes the absorption of all components and orchestrates the components to the targets. The formula is based on the compatibility of TCM and it is not the simple combination of medicines. TCM formula is usually better for reducing adverse effects and toxicity or improving therapeutic efficacy than the equivalent doses of individual ingredients by targeting multitargets of disease simultaneously, which is called synergetic effect. These properties indicate that there are multifaceted and mutual interactions existed among the complex components in the formulae with nonlinear and complex characteristics. The dynamic process of formulae in vivo is important for affecting TCM efficacy. As a result, studies on TCM formulae should take the static and dynamic processes of TCM formulae into consideration, that is, determining the chemical components of TCM formulae and the metabolized forms in vivo. However, it is still a great challenge to unveil the therapeutic mechanism of TCM formulae.

**Systems Biology**

Systems biology is an integrated discipline, which studies all the elements in the system such as genes, mRNA, proteins, and small molecules and their dynamic alterations under specific conditions. The main advantage of systems biology is taking all the elements as a whole and integrating all levels from gene to cell, tissue, and even individual. On the other hand, it also reveals the dynamic change under different conditions and time periods. The goal of systems biology is to describe the biological functions, phenotypes,
and behaviors with various omics techniques including genomics, transcriptomics, proteomics, metabolomics, and metagenomics with computational and mathematical modeling. It aims to apply a holistic and “top-down” approach to study the whole system by measuring information as many as possible rather than the strategy of “down-up,” which divides the system into various parts. Systems biology is believed to be a bridge linking TCM with modern sciences. The integration of the data from multomics with bioinformatics is the main process for systems biology study, which is summarized in Table 1. These omics technologies not only observe the identification and quantification of the biomolecules at different levels but also analyze the complex interactions based on massive data. Genomics, transcriptomics, and proteomics decipher the alterations at DNA, RNA, or protein levels, respectively. In addition, epigenomics has been applied in the quality assessment and identification of “authentic” TCMs and investigating the mechanism of its formation. Some efforts have been made to investigate the TCM characteristics with epigenetics approach such as holism, yi-yang balance, and syndrome classification suggesting that the efficacy of TCM formulae is associated with regulation on epigenetics. In addition, phenomics was also applied in TCM research, which tried to explain the relationship between phenotype and genotype. Moreover, SymMap presents the symptom-herb knowledge, and the TCM symptoms map to the modern medicine symptoms, which would make the definitions of TCM symptoms and potential mechanism clearer. The transcriptome and proteome can not only be used as the approach for biomarker discovery and identification of therapeutic targets as well. Metabolites are components of the end products of gene expression and are produced by the action of metabolic enzymes, and therefore, metabolomics can sensitively reflect the changes in body. In addition, the gut microbiota is associated with the development of various diseases, especially metabolic diseases such as diabetes, obesity, and nonalcoholic fatty liver disease. Moreover, increasing evidence has revealed that gut microbiota plays an important role in mediating TCM activity. Microbiota may contribute to the transformation or production of metabolites facilitating the absorption of TCM components. TCM formulae are mainly administrated orally, and therefore unavoidably interact with gut microbiota. For example, the microbiota is capable of metabolizing daidzein, an isoflavone.

<table>
<thead>
<tr>
<th>Table 1: Summary of omics technologies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Omnics</strong></td>
</tr>
<tr>
<td>Genomics</td>
</tr>
<tr>
<td>Transcriptomics</td>
</tr>
<tr>
<td>Proteomics</td>
</tr>
<tr>
<td>Metabolomics</td>
</tr>
<tr>
<td>Metagenomics</td>
</tr>
</tbody>
</table>

phytoestrogen in soy, to equol and O-desmethylandolensin, which are more biologically active in treating breast and prostate cancers.\cite{53} Studies have shown that the efficacy of TCM treatment is associated with the modification on gut microbiota. For example, Gegen Qinlian Decoction (GGQLD) was used for the treatment of type 2 diabetes (T2D) in the clinic, in which 187 T2D patients were divided into high-, moderate-, low-dose GGQLD or the placebo groups. There were 17 and 9 GGQLD-enriched species negatively correlated with fasting blood glucose or glycated hemoglobin, respectively,\cite{54} suggesting the involvement of gut microbiota in its efficacy. Dahuang Mudan Decoction (DHMDD) alleviated the pathological changes of dextran sulfate sodium-induced ulcerative colitis. DHMDD significantly promoted *Firmicutes* and *Actinobacteria* abundance, while markedly reduced *Proteobacteria* and *Bacteroidetes* levels. Treatment with DHMDD increased the abundance of *Butyricicoccus halopraecaudorum*, a butyrate-producing bacterium. Meanwhile, the concentration of short-chain fatty acid (SCFA) restored by DHMDD in intestinal tract.\cite{55} These results suggested that the effect of TCM was associated with gut microbiota modulation. Biological molecules do not function separately; moreover, the molecules and the related interaction will form different types of network which are essential for acknowledging the function and mechanism of the complex system including gene interaction, gene–protein interaction, protein interaction, and metabolic and gut microbiota network. These complex networks construct the phenotype together [Figure 2]. It is anticipated that omics techniques will promote our understanding of the diagnosis and treatment of disease. Undoubtedly, a bulk of information generated by omics represents a huge challenge. To achieve this goal, bioinformatics approaches should be applied for strengthening the relationship among different omics and explaining the biologically relevant information. Different from other “omics,” network pharmacology aims to provide the joint analysis of the massive data from omics and even construct drug-target-disease networks, which shifts the research strategy from a “single component, single target” to “multicomponent, multitargets.”\cite{56} Then, it will be possible to have a more comprehensive understanding of the mechanisms underlying the multicomponent and multitarget effects of TCM with these omics tools. Databases for TCM information, software, biomolecular, and phenotype information could be a good resource for providing information to underlie the TCM. The detailed information of common databases and software is shown in Table 2.

Systems biology brings amounts of data for the diagnosis and treatment for diseases; meanwhile, the explanation for biological meaning has been a great challenge. However, data quality and reproducibility are demanding. The error rate has an attendance to increase when combining the multiple omics data due to the incompatibility and the quality of the public database. In spite of the current situation, more software, database, and algorithm are being explored.\cite{96}

**“Disease-Syndrome-Formula-Effect” Network of Traditional Chinese Medicine**

### Disease-related network

In biological system, there are various interactions at different levels among genes, proteins, metabolites, gut microbiota, and so on, which are closely related to the occurrence of diseases. There are very few cases of diseases that are caused by the alteration of single factor such as the SNP of single gene or malfunction of single protein.\cite{87} Taking stroke for example, it resulted from the abnormality of a variety of genetic and environmental factors.\cite{88} Previous publication summarized that there were a dozen of genes as the monogenetic cause of ischemic stroke such as *NOTCH3, HTRA1, TREX1, HBB, CBS, GLA, ABCC6, COL3A1, FBNI, tRNA Leu, APP, CST3*, and *COL4A1.*\cite{89} In addition, some proteins are considered as the potential therapeutic targets for stroke including 70-kDa heat shock protein,\cite{90} C1q/TNF-related protein 9,\cite{91} and thioredoxin-interacting protein.\cite{92} Protein carbonyl derivatives is related to oxidation stress and formation of amino acid such as histidine, lysine, cysteine, proline, arginine, and threonine,\cite{91-93} and some amino acids were reduced in ischemic stroke patients.\cite{94} In addition to genomics and transcriptomics evidence, metagenomics provided novel view for understanding the etiology of heart disease. For example, increased SCFA-producing bacteria including *Odoribacter, Akkermansia, Ruminococcaceae UCG005*, and *Victivallis* have been detected in cerebral ischemic stroke patients, suggesting that gut microbiota is involved in development of stroke.\cite{95} Therefore, elucidation of disease-related network from genes to metabolites, as well as gut microbiota, is necessary for seeking effective therapeutic options in clinic.

### Syndrome-related network

TCM syndrome is descriptive for the pathophysiological status of patients, which is composed of a set of abnormal symptoms in the human body such as the appearance of tongue,
### Table 2: Network pharmacology-related public databases and software for traditional Chinese medicine study

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>URL</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Herb-related databases</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCM-ID</td>
<td>1197 prescriptions, 1102 herbs, and 12,117 compounds</td>
<td><a href="http://bidd.nus.edu.sg/group/TCMsite/Default.aspx">http://bidd.nus.edu.sg/group/TCMsite/Default.aspx</a></td>
<td>[57]</td>
</tr>
<tr>
<td>TCMID</td>
<td>46,929 prescriptions, 8159 herbs, 43,413 total ingredients, 8812 drugs, 4633 diseases, 1045 prescription ingredients, 788 herbal mass spectra and 3895 mass spectrometry of ingredients</td>
<td><a href="http://www.megabionet.org/tcmid/">http://www.megabionet.org/tcmid/</a></td>
<td>[58]</td>
</tr>
<tr>
<td>TCM Database@Taiwan</td>
<td>453 herbs, and more than 20,000 compounds</td>
<td><a href="http://tcm.cmu.edu.tw/">http://tcm.cmu.edu.tw/</a></td>
<td>[59]</td>
</tr>
<tr>
<td>TCMGeneDit</td>
<td>Association information about TCMs, genes, diseases, effects and ingredients</td>
<td><a href="http://tcmlifescience.ntu.edu.tw">http://tcmlifescience.ntu.edu.tw</a></td>
<td>[60]</td>
</tr>
<tr>
<td>TCMSP</td>
<td>499 herbs, 12,144 compounds, 3311 targets, and 837 associated diseases</td>
<td><a href="http://isp.nwu.edu.cn/tcmsp.php">http://isp.nwu.edu.cn/tcmsp.php</a></td>
<td>[61]</td>
</tr>
<tr>
<td>TM-MC</td>
<td>536 medicinal materials, 14,000 compounds, and provide links between medicinal materials and compounds</td>
<td><a href="https://omictools.com/tm-mc-tool">https://omictools.com/tm-mc-tool</a></td>
<td>[62]</td>
</tr>
<tr>
<td>CancerHSP</td>
<td>2439 anticancer herbs, 3575 anticancer ingredients and their molecular structures, and ADME parameters, anticancer activities on 492 different cancer cell lines</td>
<td><a href="http://isp.nwsuf.edu.cn/CancerHSP.php">http://isp.nwsuf.edu.cn/CancerHSP.php</a></td>
<td>[63]</td>
</tr>
<tr>
<td>CEMTDD</td>
<td>621 herbs, 4060 compounds, 2163 targets, and 210 diseases</td>
<td><a href="http://www.cemtdd.com/index.html">http://www.cemtdd.com/index.html</a></td>
<td>[64]</td>
</tr>
<tr>
<td>ETCM</td>
<td>403 TCM herbs, 3,979 TCM formulas, 7274 herbal ingredients, 2266 targets, and 4541 related diseases</td>
<td><a href="http://www.nrc.ac.cn:9090/ETCM/">http://www.nrc.ac.cn:9090/ETCM/</a></td>
<td>[65]</td>
</tr>
<tr>
<td>HIT</td>
<td>Provide integrative information between medicinal herbs, herb active compounds and the protein targets</td>
<td><a href="http://lifecenter.sgst.cn/hit/">http://lifecenter.sgst.cn/hit/</a></td>
<td>[66]</td>
</tr>
<tr>
<td>TCM-MESH</td>
<td>6235 herbs, 383,840 compounds, 14,298 genes, 6204 diseases, 144,723 gene-disease associations, 163,221 side effect, and 71 toxicity</td>
<td><a href="http://mesh.tcm.microbioinformatics.org/">http://mesh.tcm.microbioinformatics.org/</a></td>
<td>[67]</td>
</tr>
<tr>
<td>Biomolecular network resources</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DrugBank</td>
<td>A unique bioinformatics and cheminformatics resource that combines detailed drug data with comprehensive drug target information</td>
<td><a href="https://www.drugbank.ca/">https://www.drugbank.ca/</a></td>
<td>[68]</td>
</tr>
<tr>
<td>STITCH</td>
<td>A database of known and predicted interactions between chemicals and proteins</td>
<td><a href="http://stitch.embl.de/">http://stitch.embl.de/</a></td>
<td>[69]</td>
</tr>
<tr>
<td>ChEMBL</td>
<td>A manually curated database of bioactive molecules with drug-like properties</td>
<td><a href="https://www.ebi.ac.uk/chembl/">https://www.ebi.ac.uk/chembl/</a></td>
<td>[70]</td>
</tr>
<tr>
<td>PubChem</td>
<td>A key chemical information resource for the biomedical research community</td>
<td><a href="https://pubchem.ncbi.nlm.nih.gov/">https://pubchem.ncbi.nlm.nih.gov/</a></td>
<td>[71]</td>
</tr>
<tr>
<td>HPRD</td>
<td>A database of curated proteomic information pertaining to human proteins</td>
<td><a href="http://hprd.org/">http://hprd.org/</a></td>
<td>[72]</td>
</tr>
<tr>
<td>MINT</td>
<td>Verify protein-protein interactions mined from the scientific literature</td>
<td><a href="https://mint.bio.uniroma2.it/">https://mint.bio.uniroma2.it/</a></td>
<td>[73]</td>
</tr>
<tr>
<td>STRING</td>
<td>A database of known and predicted protein-protein interactions</td>
<td><a href="http://string-db.org">http://string-db.org</a></td>
<td>[74]</td>
</tr>
<tr>
<td>DIP</td>
<td>Experimental data between proteins</td>
<td><a href="http://dip.doe-mbi.ucla.edu">http://dip.doe-mbi.ucla.edu</a></td>
<td>[75]</td>
</tr>
<tr>
<td>BioGRID</td>
<td>An open access database dedicated to the curation and archival storage of protein, genetic and chemical interactions for all major model organism species and humans</td>
<td><a href="http://thebiogrid.org/">http://thebiogrid.org/</a></td>
<td>[76]</td>
</tr>
<tr>
<td>Reactome</td>
<td>An open-source, open access, manually curated and peer-reviewed pathway database</td>
<td><a href="https://reactome.org/">https://reactome.org/</a></td>
<td>[77]</td>
</tr>
<tr>
<td>IntAct</td>
<td>A freely available, open source database system and analysis tools for molecular interaction data</td>
<td><a href="https://www.ebi.ac.uk/intact/">https://www.ebi.ac.uk/intact/</a></td>
<td>[78]</td>
</tr>
<tr>
<td>HAPPI</td>
<td>A comprehensive online pathway data resource guidebook</td>
<td><a href="http://discoveryinformatics.uab.edu/HAPPI/">http://discoveryinformatics.uab.edu/HAPPI/</a></td>
<td>[79]</td>
</tr>
<tr>
<td>OMIM</td>
<td>A comprehensive, authoritative compendium of human genes and genetic phenotypes</td>
<td><a href="http://www.omim.org">http://www.omim.org</a></td>
<td>[80]</td>
</tr>
<tr>
<td>UMLS</td>
<td>Integrates more than 2 million names for some 900,000 concepts biomedical vocabularies, as well as 12 million relations among these concepts</td>
<td><a href="http://www.nlm.nih.gov/research/umls/">http://www.nlm.nih.gov/research/umls/</a></td>
<td>[81]</td>
</tr>
<tr>
<td>HPO</td>
<td>Contain phenotype vocabulary, disease-phenotype annotations and the algorithms</td>
<td><a href="http://www.human-phenotype-ontology.org/">http://www.human-phenotype-ontology.org/</a></td>
<td>[82]</td>
</tr>
<tr>
<td>Software</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytoscape</td>
<td>An open source software platform for visualizing complex networks and integrating these with any type of attribute data</td>
<td><a href="https://cytoscape.org/">https://cytoscape.org/</a></td>
<td>[83]</td>
</tr>
<tr>
<td>Pajek</td>
<td>A tool for complex network analysis and visualization</td>
<td><a href="http://mrvar.fdv.uni-lj.si/pajek">http://mrvar.fdv.uni-lj.si/pajek</a></td>
<td>[84]</td>
</tr>
<tr>
<td>iSMART</td>
<td>An integrated cloud computing web server for TCM for online virtual screening</td>
<td><a href="http://ismart.cmu.edu.tw/">http://ismart.cmu.edu.tw/</a></td>
<td>[85]</td>
</tr>
</tbody>
</table>

**TCMs:** Traditional Chinese medicines
pulse, symptoms, and signs. Physicians of TCM can make diagnosis based on these abnormal syndrome appearances. [96,97] Interestingly, a particular TCM syndrome can be present in different diseases (same syndrome for different diseases), which is the basis for the TCM theory of “different diseases with identical therapy.” On the other hand, different syndromes can be observed on the same disease (same disease with different syndromes), which is the basis for TCM theory of “same disease with different therapy.” However, the syndrome differentiation is dependent on the clinical observation and the physicians’ experience via four diagnostic ways of TCM: looking, listening and smelling, asking, and touching, in TCM practice. However, these traditional practices are rather subjective and personal experience oriented, which makes the diagnosis be inconsistent or even contradictory among different TCM physicians on the same patient. Thus, it is a huge challenge for realization of TCM practice standardization.

Chronic hepatitis B (CHB) is a typical virus-infected disease by HBV. In TCM, CHB patients are usually stratified into three different syndromes including liver–gallbladder dampness–heat syndrome, liver depression and spleen-deficiency syndrome, and liver–kidney yin deficiency syndrome. The diagnosis of these different syndromes for CHB patients is necessary for TCM physicians performing different therapies with TCM theory. Meanwhile, lots of biological markers including genes, miRNA, proteins, and metabolites corresponding to the relevant syndromes have also been identified using omics approaches, which are supposed to be either potential markers for CHB diagnosis or the biological basis of TCM syndromes. [12,98–100] Importantly, the dynamic changes of TCM syndrome is another problem in the syndrome differentiation study. For instance, miRNA could contribute to TCM syndrome classification such as the progression from excessive to deficient syndromes in CHB [99] as well as DNBs PLG and F12. [101] In addition, the fact of different diseases with the same syndrome is another important characteristic of TCM. For example, either coronary heart disease (CHD) or chronic renal failure (CRF) is characterized with dampness syndrome in TCM. Metabolomics study showed that patients of CHD or CRF with dampness syndrome also had common metabolic features in serum metabolic profiles compared to healthy subjects, [25] suggesting the metabolic basis for the rationale of different diseases with identical syndrome in TCM.

Therefore, the differentiation of TCM syndromes of diseases will provide practical information of disease therapy, in particular personalized therapy. Nevertheless, more omics-based studies are needed to uncover the biological basis of TCM syndromes.

**Syndrome-based traditional Chinese medicine therapy with formulae**

Based on the above discussion, the relationship between TCM syndrome and disease is the rationale for pursuing a syndrome-based TCM therapy in addition to disease-based therapy. There are usually two strategies for syndrome-based TCM therapy, i.e., same disease with different therapies and same therapy for different diseases. Taking Huanglian Jiedu decoction (HLJDD) for example, HLJDD is a classical TCM formula consisting of *Coptidis Rhizoma, Scutellariae Radix, Phellodendri Chinesis Cortex*, and *Gardeniae Fructus*. The main efficacy of HLJDD is heat-clearing and detoxifying to improve heat and blood-astasis syndrome (HBSS). HBSS is characterized with abnormal changes in blood rheology, microcirculation, hemodynamic changes, and tissue hyperplasia in patients. [11] HLJDD has been implicated in therapy for cardiovascular- and cerebrovascular-related diseases. Earlier studies suggested that HLJDD could change gene expression involved in regulating smooth muscle contraction, Ca(2+) homeostasis, and NO pathway with reduced hypertension. [102] Further studies showed that it exerted neuroprotective effects, regulated glutamate/GABA-glutamine, enhanced cholinergic neurons function, reduced oxidative stress and inflammatory- induced protective autophagy through regulating mitogen-activated kinase signals after cerebral ischemia. [103–106] In addition, HLJDD is also used for the treatment of diseases that are characterized with HBSS such as hyperglycemia, T2D diseases, and Alzheimer’s disease by attenuating inflammation and restoring gut dysbiosis. [107–109] Thus, the characteristic of TCM syndrome is the basis for disease therapy with formulae including same disease with different therapies and same therapy for different diseases as well.

Moreover, some researchers have used integrated omics to comprehensively analyze the mechanism of formula treatment. For example, transcriptomic and proteomic combination was utilized to understand the antifibrotic effects of Fuzheng Huayu (FZHY) in CCl4-induced liver fibrosis in rats. After FZHY treatment, ten core genes/proteins were found, of which Ugt2a3, Cyp2b1, and Cyp3a18 involved in retinol metabolism. [110] Similarly, antifibrosis mechanism of gypenoside, a saponin extract derived from *Gynostemma pentaphyllum* (Thunb.) Makino., was evaluated by integrated proteomics and metabolomics. Gypenoside may altered glycolysis metabolism and protected against the damage of aldehydes and lipid peroxidation via the upregulation of ALDH. [111] Besides, TCM syndrome differentiation and formula treatment may be based on the molecular subtyping such as CYP1A2-G2964A locus for syndrome-based FZHY efficacy in HBC. [112]

**“Disease-Syndrome-Formulae-Effect”**

Based on the previous studies, we proposed a new strategy “Disease-Syndrome-Formula-Effect” network for TCM study [Figure 3]. In this model, the key basis is the “same disease with different syndromes” according to TCM theory. Actually, most diseases are characterized with different syndromes at different stages, which is the basis for adopting different therapies by TCM physicians on patients with identical disease. In Western medicine, the standardized clinical treatment guidelines are the main principles for disease therapy on patients with the same disease, which usually neglecting the fact of personalized difference in
were significantly altered in QDBSS model.\textsuperscript{[110]} Meanwhile, QSBSS is another main syndrome of CHD, in which about 104 long non-coding RNAs, 2 circular RNAs, and 697 genes were significantly altered.\textsuperscript{[126]} Proteomics revealed that JAK-STAT pathway-related and immune-related proteins were differently expressed in QSBS rats.\textsuperscript{[121]} The comparison between QDBSS and QSBSS in CHD indicated that intracellular adhesion molecule-1 (ICAM-1) was higher in QDBSS than QSBSS.\textsuperscript{[122]} Evidence showed that BYHWD was effective in treating QDBSS via inhibiting C-reactive protein and regulating endothelium-derived vasoactive factors.\textsuperscript{[123]} The cardioprotective effects of BYHWD included targeting angiogenesis via Cav-1/VEGF signaling pathway with expression upregulation of Cav-1, VEGF, VEGFR2, and p-ERK, as well as inactivation of Tgf-β/Smads and MAPKs signaling triggered fibrosis.\textsuperscript{[124,125]} A clinical trial has shown that XFZYD could treat QSBS and its related diseases such as CHD.\textsuperscript{[126]} Metabolomics analysis indicated that the mechanism of XFZYD might be involved in energy metabolism, lipid metabolism (e.g. phospholipid and polyunsaturated fatty acid), amino acid metabolism, and bile acid metabolism pathways.\textsuperscript{[127-129]} These studies provided proof-of-concept evidence of TCM therapies on the basis of syndrome differentiation, which highlight the significance of “Disease-Syndrome-Formulae-Effect” strategy.

Conclusions and Perspectives

TCM is a “black box” in terms of its complicated components and therapeutic mechanisms. Currently, very few evidences have been acquired for understanding this “Black box” as a whole. Systems biology is increasingly applied in TCM research due to their common principle of holism. The fast advances of omics approaches enhance the capacity for visualizing the “black box” of TCM with systems biology strategy. The proposed “Disease-Syndrome-Formula-Effect” network is a promising model for TCM research, and further, more investigations are needed for elucidating the therapeutic mechanisms underlying TCM formulae and theory. Systems biology could elucidate the mechanism of disease and syndrome from multiple dimensions and form a network of all factors for analysis from the aspects of genes, proteins, metabolites, and microorganisms, rather than a single target (Figure 4).

In this paper, we proposed the model of “Disease-Syndrome-Formulae-Effect” network for TCM study. It is necessary to choose classical TCM formulae (included in the Chinese Pharmacopoeia) with long-term clinical application to interpret the modern scientific connotation using multidisciplinary methods. In addition to the determination of active components in TCM formulae, the endogenous metabolites alteration should be included because many metabolites are functional, in addition to as endpoints of metabolism under disease or therapy. On the one hand, it should be noted that the study on TCM formulae should be under the guidance of TCM theory. On the other hand,
systems biology is practical for bridging the theory of TCM with modern science. In fact, there are many animal oriented or even in vitro experiments on TCM, which definitely provide experiment-based evidence for understanding TCM theory or mechanism underlying TCM formulae. However, the TCM research should particularly emphasize the significance of clinic-based studies by incorporating holistic and reductive approaches together. Many studies on TCM have been carried out by adopting omics approaches individually or in combination; we should pay more attention to integrating the knowledge from molecule, cell, tissue, organ, to system levels to obtain the system dynamic changes. As a result, we proposed the strategy of constructing the “Disease-Syndrome-Formulae-Effect” network with systems biology approaches, which is practical for visualizing the “black box” of TCM from molecular to holistic perspective.

Financial support and sponsorship
This work was funded by the National Key Research and Development Program of China (2017YFC1700200 and 2020YFC0845400), Professor of Chang Jiang Scholars Program (No. 81520108030), and Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning from Shanghai Municipal Education Commission.

Conflicts of interest
There are no conflicts of interest.

REFERENCES
Strategy of systems biology for visualizing the TCM


38. Kell DB, Oliver SG. Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. Bioessays 2004;26:99-105.


