Molecular networks for the study of TCM Pharmacology

Jing Zhao, Peng Jiang and Weidong Zhang

Submitted: 30th September 2009; Received (in revised form): 26th November 2009

Abstract

To target complex, multi-factorial diseases more effectively, there has been an emerging trend of multi-target drug development based on network biology, as well as an increasing interest in traditional Chinese medicine (TCM) that applies a more holistic treatment to diseases. Thousands of years’ clinic practices in TCM have accumulated a considerable number of formulae that exhibit reliable in vivo efficacy and safety. However, the molecular mechanisms responsible for their therapeutic effectiveness are still unclear. The development of network-based systems biology has provided considerable support for the understanding of the holistic, complementary and synergic essence of TCM in the context of molecular networks. This review introduces available sources and methods that could be utilized for the network-based study of TCM pharmacology, proposes a workflow for network-based TCM pharmacology study, and presents two case studies on applying these sources and methods to understand the mode of action of TCM recipes.

Keywords: molecular networks; disease-associated networks; drug-associated networks; traditional Chinese medicine; pharmacology

INTRODUCTION

Traditional Chinese medicine (TCM) has a history of thousands of years. Considerable knowledge has been accumulated concerning in vivo efficacy and safety of TCM in targeting complex chronic diseases. Compared with the principles of western medicine, the TCM approach treats the function and dysfunction of living organisms in a more holistic way. In TCM theory, disease status is considered as the unbalance of the whole body system, and concoctions of natural products are formulated to regain the balance of the system. Currently, due to the emerging systems-based multi-target drug development paradigm [1–5], the drug discovery field is showing an increasing interest in TCM and considers it to be a source of inspiration [6–8]. However, a huge obstacle for the advancement of TCM is that, in most cases, the mode of action of TCM related to the therapeutic effectiveness is generally not known.

A TCM formula is a complex combination of many natural species such as plants, animals and minerals, each of which contains considerable numbers of chemical compounds. Its therapeutic effects mainly depend on the composition and content of effective substances. From the viewpoint of chemical structures, there is a high extent of overlap between TCM components and western drugs [9]. Therefore, at the molecular level, TCM formulae are multi-component and multi-target agents, essentially acting in the same way as the combination therapy of multi-component drugs [10]. It could be deduced that the therapeutic effectiveness of a TCM formula

Corresponding author. Dr Weidong Zhang, 1106 Yaoxue Building, 325 Guohe Road, Shanghai 200433, P. R. China. Tel: +86-21-81871244; Fax: +86-21-81871244; E-mail: wdzhangy@hotmail.com

Jing Zhao is visiting professor at Department of Natural Medicinal Chemistry, Second Military Medical University, and professor at Department of Mathematics, Logistical Engineering University. She has a PhD in bioinformatics. Her research interests include network biology and TCM pharmacology.

Peng Jiang is PhD student at Department of Natural Medicinal Chemistry, Second Military Medical University. His research interests include bioactive natural constituent isolation and TCM pharmacology.

Weidong Zhang is professor and chief director of Modern Research Center for Traditional Chinese Medicine, Second Military Medical University of China. He has a PhD in Natural Medicinal Chemistry. His research interests include bioactive natural constituent isolation and TCM pharmacology.

© The Author 2009 Published by Oxford University Press. For Permissions, please email journals.permissions@oxfordjournals.org
is achieved through collectively modulating the molecular network of the body system by its active ingredients.

During the past decade, the fast development in ‘omics’ technology and systems biology has facilitated systems-level understanding of biological processes concerning the interactions of genes, proteins and environmental factors, thus affording new possibilities for uncovering the molecular mechanisms related to the therapeutic efficacy of TCM from a systematic point of view [11]. Systems biology depicts the complex interactions at different levels as various networks and elucidates the underlying mechanisms of biological systems by studying these networks [12]. Applying network-based systems biology to the study of TCM pharmacology may open up the possibility to understand the explicit targets of TCM active ingredients and their interactions in the context of molecular networks. In this paper, we survey available sources and developments concerning molecular networks that could be applied in the study of TCM pharmacology. We then present two case studies on applying these sources and methods to understand the mode of action of TCM.

**DISEASE-ASSOCIATED NETWORKS**

In cells, there are many interactions at different levels between genes and gene products. These interactions are deeply involved in the pathogenesis of diseases. Most diseases, especially complex chronic diseases, are not caused by changes in a single causal gene but by an unbalanced regulating network resulting from the dysfunctions of multiple genes or their products [13–16]. On the one hand, genes associated with the same disorder tend to share common functional features and be co-expressed in specific tissues, and their protein products have a tendency to interact with each other [16]. On the other hand, different disorders are related to each other through the functional networks or pathways shared by their disease genes [16–19]. Moreover, a complex disease as a networked system also exhibits redundancy and robustness [20], like other molecular networks [21–24]. Usually, blocking one target cannot change the phenotype [3]. Instead, alternative compensatory signaling routes can be activated to bypass the inhibition of a single target protein [20, 21], counteracting the drug’s efficacy and causing undesired side-effects. Thus it has been realized that, to treat these diseases, drugs should target a disease-associated network rather than a single target.

From a pharmacological perspective, genes and proteins suspected to be involved in a pathophysiological process can also be potential drug targets for intervening in that disease process. The Online Mendelian Inheritance in Man (OMIM) database [25, 26] contains information on all known Mendelian disorders and associated genes. It is a valuable source for finding drug targets. On the other hand, genes associated with some categories of diseases are less related to the treatment. Specifically, the network analysis of the relationship between drug targets and disease genes suggested that known targets for some categories of diseases, such as endocrine, hematological, cardiovascular and psychiatric disease, are preferentially associated with their disease genes, whereas targets for other disease categories, such as cancer, muscular, skeletal, gastrointestinal and dermatological disease are associated with fewer disease genes than average [27]. For the latter situations, targeting proteins interacting with the disease genes, or directly targeting the interactions could be other options [28].

In recent years, some efforts have been made to identify the biological process or molecular network underlying one specific disorder by the integrated analysis of heterogeneous data sources, including genetics, transcriptomics, proteomics and interactome data, combined with computational methodologies. Many specific disease-associated networks have been constructed, including those related to diabetes mellitus, cancers, asthma, Alzheimer’s disease, and cardiovascular diseases [29–39]. In addition, some cellular network or signaling pathway databases have systematically collected pathways associated with specific diseases reported in literature [40, 41]. For example, the KEGG database [41, 42] includes over two hundred pathways partitioned into five sections, in which the section of human diseases consists of pathways concerning cancers, immune disorders, neuro-degenerative diseases, metabolic disorders and infectious diseases, and the information is updated regularly.

We conducted a comprehensive literature search about the interactome of disease genes and proteins and found nearly 50 publications about disease networks concerning six classes of diseases (metabolic disorders, cancers, central neural system diseases, cardiovascular diseases, immune diseases and others). We list the disease-associated networks and
Molecular networks for TCM study

The references in details in the Supplementary Table 1. It can be seen that some disease networks in the table were constructed from gene expression-level data, but drugs usually act on proteins. This kind of network can also be useful in pharmacology study, because several studies have revealed the correlation between mRNA and protein expression levels [43, 44]. The table also shows that some diseases, such as type-2 diabetes mellitus [36, 45, 46], colon cancer [47, 48] and asthma [37, 49–51], have been studied by different groups of researchers, and thus several networks have been constructed for one specific disease. Naturally, different approaches and data sources could generate different networks, which help to explain the underlying mechanisms of the disease from various perspectives, while the common components of the networks may suggest the key factors involved in the disease. For instance, all the four overlapping genes (CCL11, IL13, IL4, IL9) in the four asthma-associated networks [37, 49–51] appear in the list of genes mentioned most often in asthma-related literature [50]. Earlier studies suggested that IL13, IL4 and IL9 are proinflammatory cytokines that activate the JAK-STAT pathway [52, 53], an important pathway to induce inflammation in asthma [54]. Although diseases with constructed networks are far from comprehensive compared with those in the OMIM database, the methodologies used to construct them could be applied to obtain the networks for other diseases. Refer to ref. [55] for a review on computational approaches for identifying disease-associated genes and protein networks.

The disease-associated networks have the promise of allowing for the identification of potential target sets for therapeutic intervention in the corresponding diseases. Studies in network biology have suggested a correlation between topology and function of molecular networks [12, 21, 24, 56]. Thus, it is important to consider the topology of disease-associated networks, as well as the network positions of proteins, when identifying potential target combinations. Several metrics that quantitatively measure the importance of nodes or edges in networks have been used to identify potential targets. Betweenness measures the degree to which a node is participating in communication between pairs of other nodes. A study on an asthma-associated protein network indicated that protein nodes with large degrees and large betweenness metrics could be putative targets for asthma [37]. Similarly, Hwang et al. suggested that bridging nodes, i.e. linkers of modular subregions of a network, are promising drug targets from the standpoints of high efficacy and low side effects [57]. Choke points in the metabolic network correspond to enzymes that either uniquely produce or consume a given metabolite. It was found that choke points in bacteria metabolic networks could be potential targets for antibiotics [58]. From the perspective of network regulation, in order to treat a disease efficiently while minimizing undesired side effects, a drug should act only on those overactive signaling pathways while preserving other normal cellular processes. Some mathematical models and algorithms have accordingly been set up to identify potential target combinations, such as the minimum knockout problem [59], the min-interference problem [60], the OPMET model [61], and the multiple target optimal intervention (MTOI) model [62]. Recently, a software Tilde (Target Identification) was developed to detect optimal inhibitor positions in disease-associated networks and pathways by simulating the effects of different modifications of reaction combinations [63].

**DRUG-ASSOCIATED NETWORKS**

Small-molecule drugs generally perform their therapeutic functions by binding to cavities of proteins, thereby influencing their biological activities. To understand the therapeutic mechanisms of a drug, it is critical to identify the biological processes its targets participate in, the drug–target interactions and target–target interactions.

The DrugBank database [64, 65], Therapeutic Target Database (TTD) [66, 67], SuperTarget [68, 69], Matador [68, 70], and Potential Drug Target Database (PDTD) [71, 72] have collected known information of drug targets. The search tool for interactions of chemicals (STITCH) database [73, 74] integrates information about interactions of chemicals and proteins from different types of databases. The information provided by each database has its own focus. Thus they could be complementary in application. For instance, we searched the targets of an anticholesteremic agent simvastatin in each database and got different results (Supplementary Table 2), with HMG-CoA reductase (HMGCR), the primary target of simvastatin, as their intersection. The PDTD database focuses on targets with known 3D-structures and provides a web server TarFisDock to predict the potential
binding targets of a drug \textit{in silico}. The TTD also provides target similarity and drug similarity search to enable a user to find similarity targets or drugs of an input protein sequence or drug structure. These tools provided by the PDPD and TTD could be applied to predict the putative targets of the active compound extracted from TCM recipe. Matador is a manually annotated subset of SuperTarget which provides additional binding information and indirect interactions. The Therapeutically Relevant Multiple-Pathways (TRMP) database [75, 76] integrates information on therapeutic targets and disease-associated signaling pathways. Once the targets of the main active compounds of TCM formula are identified by \textit{in-silico} predictive or experimental approaches, they can be mapped onto specific disease-associated networks or pathways and target databases of known western drugs to construct drug-associated networks of TCM compounds. See the two case studies in the following section for illustration.

Constructing the target protein network for a specific disease or drug could help us to understand the effects of drugs on diseases. Hopkins constructed a network between the literature reported 44 potential targets associated with asthma [4], in which each node denotes a drug target, and two nodes are linked if there is at least one drug targeting both of them. This network could be applied to explore combination therapy for asthma by multi-component drugs. Cases and Mestres collected a curated list of 214 cardiovascular targets by literature mining [77]. This target set could be utilized to construct a therapeutic network for cardiovascular diseases by mapping the proteins to the human protein interactome [78–82].

**NETWORK-BASED TCM PHARMACOLOGY**

The material sources of TCM are natural products, including plants, animals and minerals, each of which includes many chemical constituents. Although a TCM recipe usually contains hundreds even thousands of components, only a few bioactive compounds contribute to the therapeutic effects. On the other hand, compounds isolated from natural products have been important sources of new drugs or drug leads. As can be seen in Supplementary Table 3, many compounds identified from TCM materials are also drugs approved by the FDA. Therefore, identifying the effective bioactive compounds of TCM is very important for TCM pharmacology study, as well as modern drug discovery.

Once the active compounds of a TCM recipe are known, the remaining tasks are to identify the targets of each compound, and to study the targets in the context of disease networks and drug-associated networks. Here we propose a workflow for network-based TCM pharmacology study, as shown in Figure 1.

In this section, we survey major approaches for the identification of TCM effective active compounds and their targets, and then present two case studies that investigate the molecular mechanisms of TCM from a network-modulation point of view. A herbal drug and a TCM formula are studied respectively. Existing research results are surveyed and applied to construct drug-associated networks.

---

**Figure 1:** Workflow for network-based TCM pharmacology study.
Identification of TCM effective active compounds and their targets

To identify the bioactive compounds from the complex constituents of a TCM recipe, the conventional method is to extract and separate some components directly from the TCM recipe, and then conduct pharmacological evaluation on each component respectively. In this way, the compound Astragaloside IV (AGS-IV) was extracted from the TCM material Astragalus membranaceus and developed as a new drug. Astragalus membranaceus has long been used in TCM for the treatment of cardiovascular diseases but its bioactive components were still unknown. Our laboratory isolated AGS-IV from aqueous extract of Astragalus membranaceus, performed a series of in vivo and in vitro pharmacological experiments, and validated the cardioprotective effects of AGS-IV [83, 84]. AGS-IV was thus identified as active compound of Astragalus membranaceus. In Supplementary Table 4 we list the specific techniques usually used for isolating active components from TCM.

Since only a few compounds are responsible for the therapeutic effects of TCM, biochromatography, which is based on the biological interactions between bioactive compounds and immobilized proteins, enzymes and antibodies, has been applied to quickly eliminate the interference of non-viable components and to identify bioactive compounds from TCM. The bioactive components in Artemisia capillaris Thunb were thereby identified in such a way based on their affinity to human serum albumin (HSA) which binds with most synthetic drugs [85]; and those of Radix Angelica Sinensis were screened out by immobilized liposome chromatography (ILC) which mimics the filtering ability of a cell membrane system to drug molecules [86].

As most TCM is taken orally, only the components that eventually appear in blood could be considered to have the chance of exerting their effects. Some of the components may actually be metabolites of the original compounds. A serum pharmacological screen strategy was thus proposed to identify the main components absorbed in blood after administration of TCM [87]. Applying this methodology, we studied the absorbed components in rat plasma after oral administration of Shexiang Baoxin Pill (SBP), a Chinese traditional patent medicine for the treatment of cardiovascular diseases. Totally 21 components, including 17 components from SBP and 4 metabolites, were observed from a comprehensive analysis of the chromatography of SBP, controlled plasma and dosed plasma. Fourteen of the identified compounds, which were present in high concentration and reported to have effects on cardiovascular diseases, were identified as main active compounds [88]. Further study will be carried out to identify the targets, investigate the mode of action, and conduct comparative pharmacological evaluation on the active compound combinations and SBP itself. Along these lines, it may be possible to develop a new multi-component drug consisting of a rational combination of the SBP active compounds for the treatment of cardiovascular diseases in the future. We list some TCM recipes whose main bioactive ingredients have been identified in Supplementary Table 5.

Several databases have been constructed for providing information concerning constituent herbs, bioactive compounds and other aspects of TCM recipes. The TCM database includes information about Chinese medicinal plants and bioactive compounds [89]. The 3D structure database of components from Chinese traditional medicinal herbs provides the basic molecular properties and optimized 3D structure of herbal compounds [90]. TCMID database (Traditional Chinese Medicine Information Database) collects comprehensive information of TCM including prescriptions, constituent herbs, herbal ingredients, molecular structure and functional properties of active ingredients, therapeutic and side effects, clinical indication and application and related matters [91]. These databases could be applied for data mining of effective bioactive compounds of TCM.

Proteomic technologies could profile changes in protein expression in response to drug treatment and identify differentially expressed proteins, and have been proved effective for the identification of protein targets of TCM active compounds [92]. From a technological point of view, the current applicable tools are two-dimensional gel electrophoresis (2-DE) for separation of proteins in a proteome, and mass spectrometry (MS) for protein identification [93]. On the other hand, in silico virtual screening approaches could provide alternative ways for low-cost and rapid predictions of targets of TCM active compounds. The methodologies for target prediction can be roughly grouped into two classes: the first class predicts targets of new compounds from those with known targets only based on compound chemical information [94–98], while the second class
utilizes 3D information about both the compound and the target protein to perform ligand–protein docking [99–102]. Recently, considerable efforts have been made to infer unknown drug-target interactions by integrating more information about drugs and targets, such as drug chemical structure, side-effects, target protein sequence, and drug-target network topology [103–106]. These approaches could be complementary when being applied to predict targets of TCM active compounds in silico.

_Ganoderma lucidum_ is a medicinal mushroom used in TCM for the prevention or treatment of a variety of diseases including cancers [107, 108]. Triterpenes in _Ganoderma lucidum_ have been regarded as the main anti-cancer active ingredients due to their ability to inhibit growth, induce apoptosis and cause cell cycle arrest of cancer cells [109–111]. In a work by Yue _et al._, a proteomic approach was applied to investigate the possible targets of ganoderic acid D (GAD), a main compound of Ganoderma triterpenes, in cancer cells, and 21 differentially expressed proteins were identified [112]. These possible GAD-target related proteins were evaluated by the _in silico_ ligand-protein inverse docking software INVDOCK [101]. Totally 7 of the 21 proteins were found to bind with GAD by the software. The protein–protein interaction network between the 21 putative targets was constructed, and the enrichment of 14-3-3 proteins and their central localizations in this network indicated that they could be important targets of GAD in cancer cells.

Refer to ref. [113–115] and [92] for detailed reviews about approaches and strategies to screen bioactive compounds from TCM recipes, to predict protein targets of small molecules _in silico_, and to identify targets of natural compounds by proteomics, respectively.

**Case study 1: Antidepressant activity of St. John's Wort**

St. John’s Wort (SJW) is an extract from the plant _Hypericum perforatum_ L. Numerous clinical trials have shown that SJW had significant antidepressant efficacy and lower side effects than standard antidepressants [116–119]. In many countries, it has been widely used for the treatment of mild to moderate forms of depression. SJW has been included in the pharmacopoeias of Germany and the US.

The main active ingredients of SJW are hyperforin (HP), hypericin (HY), pseudohypericin (PH), amentoflavone (AF), and several flavonoids (FL) [120]. Experimental results have suggested that HP, HY, PH and AF are able to pass the blood-brain barrier [121–123]. Furthermore, the antidepressant activity of SJW is highly associated with these active compounds [120, 124–128].

We conducted a comprehensive literature search and collected the neurotransmitter receptors, transporter proteins, and ion channels on which the SJW active compounds show effects (Supplementary Table 6). By mapping these proteins onto KEGG pathways, it was found that SJW intervenes in mainly three pathways, neuroactive ligand–receptor interaction, the calcium signaling pathway, and the gap junction related pathway. In Figure 2 we show the effects of the SJW active compounds on the system of neuroactive ligand–receptor interaction. It can be seen that the SJW active compounds act on different receptors respectively so as to regulate the uptake and transport systems of neurotransmitters in a multi-target pattern. In this way, SJW blocks the reuptake of multiple neurotransmitters such as serotonin, norepinephrine and dopamine and stimulates the release of these neurotransmitters.

We then extracted all the FDA-approved antidepressants, i.e., the drugs whose first four ATC code (Anatomical Therapeutic Chemical code) is N06A, and their targets from the DrugBank database. Integrating these data with information in Supplementary Table 6, we constructed the drug-target network for FDA approved antidepressants and SJW compounds, as shown in Figure 3. This network shows that the active compounds of SJW share same targets with different types of antidepressants such as monoamine oxidase (MAO) inhibitors and monoamine reuptake inhibitors, respectively, suggesting that the effect of SJW is similar to that of a combination of different classes of antidepressants.

However, the inhibitory effects of the SJW active compounds on each of the targets are lower than individual therapeutic dosages, thus it is inadequate to explain the antidepressant effect of the herb only from the inhibition of any single target [120]. For instance, SJW inhibits MAO only in millimolar concentrations, which is much weaker than conventional antidepressant MAO inhibitors [128, 130]. Therefore, it is likely that the actions of multiple active compounds of SJW result in an additive or synergistic antidepressant efficacy [131, 132], making SJW realize the same antidepressant efficacy as normal monotherapy at much lower doses of separate compounds.
In fact, many potential targets for central nerve system (CNS) drugs participate in multiple signaling pathways that keep normal physiological functions of cells. Only in overactive or unbalanced conditions do they hurt nerve cells [133]. CNS drugs that work by specific and high-affinity binding to their targets could block all activity including normal cellular processes. Thus they usually result in intolerable side effects. Therefore, in the treatment of CNS diseases, low-affinity binding agents [133] and drug combination strategy have been proved useful in reinforcing efficacy, limiting side effects, and improving compliance [134]. Accordingly, the significant antidepressant efficacy and lower side effects of SJW could be attributable to the synergetic actions of the low-dose combination of multiple active compounds.

**Case study 2: The effect of Realgar-Indigo naturalis formula on acute promyelocytic leukemia**

Acute promyelocytic leukemia (APL) is a subtype of acute myeloid leukemia (AML) caused by a specific
chromosome translocation t(15;17). It is a malignancy of the bone marrow in which there is an excess of immature cells (called promyelocytes) and a deficiency of mature blood cells in the myeloid line of cells. APL can be effectively controlled by the differentiating agent all-trans-retinoic acid (ATRA), which activates the retinoid receptor RAR and induces the promyelocytes to differentiate toward mature granulocytes [135]. A TCM formula, Realgar-Indigo naturalis formula (RIF), has been applied in China to treat APL since the 1980s. Clinical trials showed that 60-day RIF treatment on APL patients resulted in a complete remission (CR) rate of 98.3% [136]; while a CR rate of 95% for relapsed APL [137], and 5-year survival rate of 86.88% [138] were achieved after RIF treatment.

RIF is a TCM formula consisting of four kinds of materials, realgar, Indigo naturalis, Salvia miltiorrhiza, and Radix pseudostellariae. In TCM theory, multiple agents contained in one formula must work synergistically. Realgar is regarded as the principal component of the formula RIF, and the other three are adjuvant components to assist the effect of realgar. Studies in recent years showed that the main active compounds of realgar, Indigo naturalis and Salvia miltiorrhiza are tetraarsenic tetrasulfide (As₄S₄, A) [139], indirubin (I) [140] and tanshinone IIA (T) [141], respectively. Applying approaches of modern biological research, a group of Chinese scientists investigated the multi-target, synergetic actions of the three active compounds in RIF and successfully illustrated the therapeutic mechanism of the TCM formula at

Figure 3: Drug-target network of FDA approved antidepressants and SJW compounds. A target protein node and a drug node are linked if the protein is targeted by the corresponding drug. This graph is drawn with the software Cytoscape [129].
Their in vivo experiments on a murine APL model showed that mono-therapy with A significantly prolonged the overall survival, while ATI combination exhibited the most potent therapeutic efficacy compared with mono- or bi-therapy of A, T, and I. In vitro experiments showed that A or T alone induced a certain degree of differentiation of APL cells, and ATI combination resulted in synergistic effects that caused APL cells to differentiate toward mature cell types. At the molecular level, ATI combination strengthened the regulation on APL associated proteins such as PML-RARα and C-Myc.

To understand the therapeutic mechanism of RIF in the context of network regulation, we collected the results of ref. [142] concerning the effects of A, T, I alone and their different combinations on APL associated proteins and listed them in Supplementary Table 7. We also searched the OMIM database and found six APL disease genes. We called the proteins in Supplementary Table 7 and those encoded by the six APL genes as RIF-associated proteins.

We first constructed a protein–protein interaction network for the human genome based on the HPRD [78] data and mapped the RIF-associated proteins onto this network. Then we adopted the Steiner minimal tree algorithm [143] to identify a minimum sub-network, which includes as many RIF-associated proteins and as few other proteins as possible, while each RIF-associated protein can interact with another through at most one bridge protein. We used the P-value [51] to quantitatively measure whether a network is more enriched with proteins of a specific Gene ontology (GO) term than what would be expected by chance. Given significance level α = 0.05, a P-value smaller than α demonstrates low probability that the proteins of same GO term appear in the network by chance. As can be seen in Figure 4A, the RIF-associated proteins are tightly connected together due to their direct interactions, while the network is significantly enriched with proteins whose GO terms are regulation of cell differentiation and cell proliferation \((P = 1.26 \times 10^{-6}, 1.09 \times 10^{-10})\), two biological processes highly associated with the progress of cancers. Specifically, the GO suggests that five of the proteins (CEBPA, CEBPB, PML, RB1 and NCOA6) are involved in the biological process of myeloid cell differentiation \((P = 1.72 \times 10^{-5})\). This protein–protein interaction network indicates a possible concerted functional mechanism of RIF on the APL associated proteins.

We also mapped the RIF-associated proteins onto KEGG pathways and generated a bipartite graph of protein-pathway association, in which a protein and a pathway were linked if the protein appeared in the pathway. Figure 4B shows that the RIF-targeted proteins are involved in a series of cancer pathways, five of which participate in the acute myeloid leukemia (AML) pathway, suggesting that the pathway is the key pathway modulated by RIF. In Figure 5 we show the targets of RIF on the

**Figure 4:** Functional networks of APL disease gene-encoded proteins and RIF-targeted proteins. (A) Protein interaction network. (B) Protein-pathway association network. [D]: Gene Ontology (GO) of the protein: regulation of cell differentiation; [P]: GO: regulation of cell proliferation; [B]: GO: regulation of cell differentiation, and regulation of cell proliferation. This graph is drawn with the software Cytoscape [129].
AML pathway and the effects of RIF on them. It can be seen that, on the one hand, by up-regulating C/EBPa and PU.1 proteins and down-regulating PML-RARa oncoprotein, RIF stimulates APL cell to differentiate; on the other hand, by inhibiting PML-RARa and c-Myc, RIF deters the promyelocytes from proliferating. In conclusion, RIF intervenes in the AML pathway by targeting multiple proteins localized at its two distinct but associated branches, hence resulting in a synergetic anticancer action on APL.

Figure 4B shows that RIF also targets on multiple proteins at the pathways of chronic myeloid leukemia (CML) and small cell lung cancer (SCLC), indicating that it is probably efficacious against these cancers. More research deserves being done in this direction.

**PERSPECTIVES**

Network-based TCM pharmacology seeks to develop a systematic understanding of the actions of TCM by considering their targets in the context of molecular networks. The sources and methods of molecular networks introduced here may facilitate the network-based study of TCM pharmacology. The examples in this paper suggest that by integrating information from different sources, network-based TCM pharmacology provides a perspective for better understanding of the holistic, complementary and synergetic essence of TCM at a molecular level. TCM, in essence, is combination therapy by multiple active compounds. Rich experience in the combinatorial use of natural products has been accumulated in TCM to achieve a synergetic therapeutic efficacy and reduced side-effects. By a combination of multiple chemical ingredients, TCM remedies elicit their beneficial effects by tinkering with different proteins in networks in a gentle way, achieving the same therapeutic efficacy of normal mono-ingredient agents at much lower doses of separate compounds. Thus the side effects of TCM are usually weaker than the monotherapy of western medicine. A great value of TCM is in its application for thousands of years and considerable knowledge accumulated concerning in vivo efficacy and safety.
two of the confounding problems facing new designed drugs. Thus drug discovery starting with well-validated TCM remedies is promising in developing new multi-target agents, or potent drug combinations that are individually less therapeutic but efficacious in combination. This approach also has the advantage of controlling the pharmacokinetics and drug–drug interactions of multiple components. We expect that, along this reverse drug discovery path, it is possible to develop new-entity drugs or efficient drug combinations at a lower cost of time and money.

Key Points
- At the molecular level, TCM recipes are multi-component and multi-target agents, essentially acting in a similar way as combination therapy using multi-component drugs.
- Network-based systems biology provides new tools and perspectives for the understanding of the mode of action of TCM.
- Identifying the effective bioactive compounds from the complex constituents of TCM is the foundation for TCM pharmacology.
- Disease-associated network and drug-associated network are proper context networks for elucidating the holistic, complementary and synergic essence of TCM from molecular level.

SUPPLEMENTARY DATA
Supplementary data are available online at http://bib.oxfordjournals.org/.

Acknowledgements
The authors would like to thank Dr Mikael Huss, Dr Yiczhu Shan and the anonymous reviewers for their helps to improve the manuscript.

FUNDING
The work was supported by program NCET Foundation, National Natural Science Foundation of China (10971227, 30725045), the Special Program for New Drug Innovation of the Ministry of Science and Technology, China (2009ZX09311-001, 2008ZX09101-Z-029), Shanghai Leading Academic Discipline Project (B906).

References


