Modern bioinformatics meets traditional Chinese medicine

Peiqin Gu and Huajun Chen

Submitted: 15th May 2013; Received (in revised form): 28th August 2013

Abstract

Motivation: Traditional Chinese medicine (TCM) is gaining increasing attention with the emergence of integrative medicine and personalized medicine, characterized by pattern differentiation on individual variance and treatments based on natural herbal synergism. Investigating the effectiveness and safety of the potential mechanisms of TCM and the combination principles of drug therapies will bridge the cultural gap with Western medicine and improve the development of integrative medicine.

Dealing with rapidly growing amounts of biomedical data and their heterogeneous nature are two important tasks among modern biomedical communities. Bioinformatics, as an emerging interdisciplinary field of computer science and biology, has become a useful tool for easing the data deluge pressure by automating the computation processes with informatics methods. Using these methods to retrieve, store and analyze the biomedical data can effectively reveal the associated knowledge hidden in the data, and thus promote the discovery of integrated information.

Recently, these techniques of bioinformatics have been used for facilitating the interactional effects of both Western medicine and TCM. The analysis of TCM data using computational technologies provides biological evidence for the basic understanding of TCM mechanisms, safety and efficacy of TCM treatments. At the same time, the carrier and targets associated with TCM remedies can inspire the rethinking of modern drug development.

This review summarizes the significant achievements of applying bioinformatics techniques to many aspects of the research in TCM, such as analysis of TCM-related '-omics' data and techniques for analyzing biological processes and pharmaceutical mechanisms of TCM, which have shown certain potential of bringing new thoughts to both sides.

Keywords: bioinformatics; traditional Chinese medicine; systems biology; herbal synergism; linked life data

INTRODUCTION

Traditional Chinese medicine (TCM) is a unique Chinese health care system of ancient medical practice that covers a broad range of medical theories and practices that are based on ancient Chinese philosophy, including Yin–Yang theory, five-phase theory, herbal medicine and acupuncture [1]. It has played an important role in health maintenance for the people of Asia for thousands of years, and is becoming frequently used in Western countries [2]. TCM is also a part of the pantheon of traditional systems of medicine used in Asia that includes Siddha and Ayurveda, which are mainly practiced in India. Its therapies and treatment have been proven to have better effects on some fatal diseases such as malaria [3], benefiting the global health care with auxiliary therapies. In Eastern countries, TCM has a good reputation for improving the health of individuals and preventing or healing diseases, not to mention the great advantages in early intervention, combination therapies and personalized medicine for chronic diseases. In Western countries, the efficacy
of herbal medicine and acupuncture has been recognized [4]. The uniqueness of the TCM system is derived from the philosophical logic behind the daily practices, which accumulates thousands of years of empirical studies and provides a unique view of the relationships between the human body and the universe. Compared with modern medicine, the underlying knowledge of TCM is rather vague because its major corpus was written in ancient natural language; however, academic and industrial agents [5] all around the world, including the Chinese government [6], have been trying to find scientific evidence that could support a deeper understanding of the mysterious Eastern medicine using modern technologies [7].

The fundamental components of TCM are different from modern medicine, and those functional units form a microecosystem of balancing forces. Thus, the conventional reduction approach in biology research cannot effectively draw useful conclusions out of the holism of TCM [8]. The main challenge for researchers is how to discover the physical existence of these units and explain the mechanisms of coherently related parts as a whole system. A systematic model of TCM components and their interrelations would be helpful to explain the physical foundation of TCM theories, just like the anatomy in Western medicine (WM), yet the model has not been proposed. Take Qi for an example; it is currently impossible to detect the physical existence of the driving forces like Qi, and therefore the understanding of the mechanisms is difficult. As the biological model of the human body is fairly clear, discovering the association relationships between modern biology (e.g. genes, proteins, molecules) and TCM entities (e.g. Qi, syndrome, herb) may uncover the biological foundation and effectiveness of TCM.

Meanwhile, from the perspective of modern medicine, understanding the functional mechanisms of the whole organism has become more popular, with increasing attention on integrative medicine [9] and personalized medicine [10]. Taking TCM as one of their parts, integrative medicine and personalized medicine emphasize wellness and healing of the entire person and deliver personalized therapeutic solutions to every individual patient rather than disease-target therapies. Systems biology—a biology-based interdisciplinary field that has emerged as an alternative and promising method for exploring the complex systems of life using a more holistic perspective (instead of traditional reductionism)—proposes novel solutions for integrative medicine and personalized medicine involving TCM [11, 12].

### PROBLEMS AND CHALLENGES

TCM and WM are obviously different medical systems, both of which provide support for disease diagnosis, treatment and health prevention. Research on possible relevance between these two large medical systems is likely to solve the following problems:

- **Understanding the biological foundation of TCM**

  As we mentioned, explaining the biological foundation and core mechanisms of TCM based on the discovery of genes, proteins and molecules is gaining increasing interest among academic communities. The basic idea is to find the genetic, proteomic, cellular, molecular targets or associated biological pathways related to the three important parts of TCM: basic theories, diagnosis and treatment principles and herbal drugs. The biological findings can help researchers understand the mechanisms of TCM.

- **Modern drug development inspired by TCM**

  Different from Western herbal medicine, in which herbs are often delivered singly or combined into small formulas of herbs with the same function, Chinese herbalists usually prescribe combined drug therapies of multiple herbs to treat a disease. For some fatal diseases, Chinese herbal medicine has slower, but better, effects. These diseases are assumed to have complex causes and multiple targets. Investigating TCM-inspired multicomponent herbal drugs can be a future direction of multitarget drug development for effective personalized medicine.

The accumulation of a massive amount of biomedical data makes data- and knowledge-intensive computation technologies more important. Bioinformatics techniques [13], including computer science, mathematical and engineering methods, are broadly used to generate useful biological knowledge and to discover association relationships in biological data, e.g. inferring gene association networks [14]. Those methods include pattern recognition,
Combining systems biology with bioinformatics to probe the mystery of TCM potentially bridges the methodological gap between TCM and modern medicine, providing the scientific evidence to support simplification of the complexity of TCM, which can verify the value of TCM in modern medicine. These techniques serve as a driving force for the translation of TCM formulas and pathological mechanisms into practice.

In this review, we survey the available sources and bioinformatics methods that have been applied to these two concerned problems and subsequent subproblems (as shown in Figure 1), and present a vision regarding the use of semantically linked biomedical data in the future.

**TCM MEETS SYSTEMS BIOLOGY**

Pathological disorders can be regarded as a reaction of the human body to changes of the environment, including virus, bacteria, physical damage and internal pathological changes such as cancer. The cause and the reaction usually involve the structure and dynamics of cellular and organismal function. To understand biology at the system level, systems biology [16], an emerging field that focuses on complex interactions within biological systems, is characterized by using a holistic approach to model biological systems. It is gaining interest from academic communities and enables us to collect comprehensive data and gain system-level understanding of biological systems. A systems biological approach typically includes observing (measuring) data from different subcategories, e.g. ‘-omics’, physical parameters, developmental parameters, along with the various chemometric methods to extract information such as correlations and similarities between biological entities (e.g. genes, proteins, molecules).

The goals of the systems biology-based medicine potentially overlap with the holistic and integrative view of TCM, as they both aim toward personalized medicine. In the past decade, the question of...
whether systems biology can unify TCM and modern medicine has become an issue of growing concern. For example, to understand the Chinese medicinal concept of **Qi** in the perspective of systems biology, research suggested that the transmission of **Qi** along the meridians is based on molecules that travel via an intercellular communication system [17], in which nitric oxide is considered as a prime candidate for such a signaling molecule in the meridian system. The mechanotransduction properties of nitric oxide elicitation [18], recently modeled comprehensively, may provide important insights in this research.

Although TCM and modern medicine understand the living systems based on different criteria, there are similar disease patterns (e.g., insomnia), and the metabolites of herb-perturbed bodies can be studied and analyzed using screening to figure out the chemical compounds of the herbs and their possible targets, as demonstrated in Figure 2. Advances like this example in systems biology have enabled the discovery of biomarkers and provided the basis for the development of new targeted drugs, potentially facilitating ‘the right therapy for the right patient’, which fits the goal of TCM. A broad range of biological technologies [19] and computer models [20], e.g., genomics, proteomics and metabolomics, that come along with systems biology, can identify biomarkers of a particular disease or herbal drugs within the set of genes, proteins and metabolites of a given organism.

This diagram illustrates the following points: (i) Syndrome refers to a pattern of disharmony that is characterized by a collection of symptoms, which can be regarded as a certain profile along with one or more diseases. (ii) Identifying the mutant genotypes responsible for certain genetic diseases can create the links between genes and TCM entities, e.g., syndromes, symptoms and drugs. (iii) With the development of proteomics and metabolomics, the contributing compounds in Chinese herbs can be filtered and studied.

**TCM meets genomics**

Genetic information [21] is likely to elucidate the mechanisms underlying the fundamental biological processes perturbed in human diseases and regulated pathways affected by herb intervention. The combined genome–wide expression analysis with methods of systems biology can identify the functional gene networks for the sets of clinical symptoms that comprise the major information for pattern classification related to pathological and treatment patterns in TCM [22]. With the accumulation of biological data, data about genetic regulated diseases and the corresponding herbal therapies are growing. In recent years, many herbal medicines have been reported to be associated with various symptoms or diseases and may exhibit a variety of effects through regulation of a wide range of gene expressions or protein activities.

**Discovering TCM–gene relationships from corpus learning**

Hidden knowledge about the association relationships between genes and TCM entities implicitly exist in varieties of textual corpus. Those associations come from isolated small biological experiments all over the world. Through corpus analysis, pieces of association information can be gathered and integrated to support high-level data analysis.

The most popular TCM information database, TCMGeneDIT [23], collected information from public databases, including TCM herb database HULU, TCM-ID, NCBI Entrez Gene and medical subject headings vocabulary, to form a unique database that offers diverse association information related to TCM and participating genes. The collected terms of TCM names, gene names, medical subject headings vocabulary diseases and TCM ingredients were used to annotate the literature corpus obtained from the literature database PubMed using entity annotation. Useful co-occurrence and corresponding confidence values can then be observed through association discovery and rule-based extraction during corpus analysis. The final results contain association groups as (TCM, gene), (TCM, disease), (TCM, gene, disease), (TCM, ingredient), (TCM, effect) and (gene, ingredient). The creation of such a database could facilitate the understanding of therapeutic mechanisms involving TCM and gene interactions, which would further be used to inspire modern clinical research.

Wu et al. [24] also applied a text mining approach to uncover the functional relationships from PubMed literature. They used a bootstrapping method to extract disease names from TCM-related articles, and then extracted the relationship of symptom complex (SC) and disease based on disease names and SC terminological database. Here, SC refers to a holistic concept reflecting the dynamic,
functional, temporal and spatial morbid status of the human body. The term co-occurrence is used to identify the disease–gene and SC–gene relationships within the articles. It is suggested that the related genes of the same SC will have some functional interactions. In all, 1100 SC–disease relationships were collected, e.g. 72 related genes of kidney YangXu SC were filtered, such as chronic renal failure. In TCM, syndromes play an important role in disease diagnosis; it is reasonable to guess that the related genes of the same syndrome may also have some biological functional relationships. Similar work has been done to find syndrome–gene relations [25].

**Novel representation of drug–gene–disease patterns**

Genes share notable patterns with respect to diseases, syndromes and even drugs. Zhao et al. [26] proposed an alternative approach to define drug–gene–disease relationships, called ‘co-module’, to characterize closely related drugs, diseases and genes. A co-module consists of a gene module and those drugs and diseases associated with it. They implemented a Bayesian partition method to identify drug–gene–disease co-modules underlying the gene closeness data. In a gene closeness profile, a gene whose products are more highly interconnected in the network with drug targets or disease–gene products, receives a higher closeness score with respect to that drug or disease. Given these gene closeness data, one can find important genes interconnected with drug targets or disease-related genes. Though not directly mentioned in the article, the approach of discovering co-modules should be useful or highlight the genetic basis of TCM diseases and drugs.

**TCM patterns of clinical manifestations based on gene profiles**

Genome-wide expression analysis in *in vivo* studies can also be used with methods of systems biology to identify the functional gene networks for clinical patterns in TCM. In TCM, cold and hot patterns...
are two opposite properties that are used to identify and categorize the universe of things, such as food, diseases, etc., classify symptoms and conduct treatment. The clinical manifestations of a specific disease can be clustered into these two patterns before blood samples, and the genes significantly related to hot and cold patterns can be identified to study the mechanism of pathogenesis and diagnosis criteria of TCM. This approach has recently been implemented in a study of rheumatoid arthritis [27], a systemic autoimmune disease with unknown cause and various treatment methods. Through factor analysis and correlation analysis, the genes were used to search protein interaction information from protein interaction databases and literature data, and consequently several significant pathways were found to be related to TCM patterns.

From the aforementioned work, we have found that the associations between genome and TCM entities including diseases, syndromes, drugs and patterns have already existed in a broad range of literature, databases and patient samples, which are available for textural analysis and association discovery. The associations discovered can be analyzed to infer advanced relations, such as protein pathways related to diseases, and the molecular basis of TCM clinical patterns. In all, we proposed a basic workflow of such data processing jobs, as shown in Figure 3. Recently, a new scientific field of toxicogenomics has been proposed to investigate the toxicity and reactions based on the integration of genomics and toxicology, which is another important concern regarding the safety of the natural products used in TCM herbal medicine [28].

**TCM meets proteomics**

As proteins are vital parts of living organisms, proteomics gives us a better understanding of an organism by studying the structures and functions of proteins. In a model organism, proteomics [29] aims to identify and quantify the cellular levels of each protein that is encoded by the genome, through studying
protein biomarkers, identifying gene-modified proteins and establishing protein–protein interactions. One aim of systems biology is to model the whole cell by integrating such protein–protein interactions [30], which can likely be aided by using the whole physiome approaches of ancient systems of medicine such as TCM. It can be used to link genomics and metabolism, leading to an integrated understanding of the entire organism. In addition, the ability to detect expressed proteins has enormous potential for early diagnostics and intervention at curable stages of disease and multitarget drug development.

With the broad use of systems biology technologies, the adoption of the proteomics technologies can potentially leverage the study of pharmacological effects and their action mechanisms of TCM [31]. For example, using proteomics approaches, it was found that the Siwu decoction, which is an ancient TCM therapy for enriching blood tonic, could regulate the protein expression of the bone marrow of blood-deficient mice [32]. Fundamentally, proteomics can benefit the research in TCM in two ways: (i) Proteomics could provide a platform for studying the cellular basis of diagnostic principles of TCM. In modern therapeutics, individual variance concerning expression profiles of proteins is of great importance in TCM syndrome variance. (ii) Proteomics technologies can be used to screen the target molecules of the action of TCM, isolate and characterize new active components, as well as analyze toxic substances from TCM. Using proteomic approaches to identify therapeutic targets, to evaluate the effects of new drugs and to explore the functional mechanism of the effects of new drugs may meet the shortcomings of the conventional methodology being applied in the current studies.

Identifying functional community of syndrome-associated proteins

Syndrome (‘Zheng’ in Chinese) is an essential part of TCM theory, which refers to a characteristic profile of all clinical manifestations (e.g. symptoms) that can be identified by TCM practitioners as an important criterion during diagnosis. All these characteristics of TCM syndromes might be related to various proteins with different contents, functions, structures and interactions of proteins expressed for different syndrome behaviors, and the changes of proteins before and after treatment with single herb or formulas could be analyzed. Understanding of the characteristic changes in proteomics associated with a specific syndrome will facilitate syndrome identification and novel diagnostic approaches that will potentially lead to personalized health care strategies [33].

A practical way to investigate the proteome basis of syndrome is to study syndrome in the context of the expressed proteins of a specific disease, for example, coronary heart disease (CHD). According to experimental survey [34], blood stasis syndrome in TCM plays an important role in CHD patients. In clinical analysis, WM specifications and diagnosis criterion of the syndrome were both studied given blood samples of patient profiles. The case group and the control group were analyzed using traditional bioinformatics methods to obtain significant proteins related to the specific disease. As there are no standardized models of disease–syndrome, one way to study the associations between certain syndromes and proteomes is through pattern discovery over prior data of related proteins and symptoms. Mutual information between each pair of variables (e.g. proteins, symptoms) can be studied, and the set that has the high correlative value of variables will be considered significant. Thus, this approach can generate a bunch of identified patterns of closely related proteins and symptoms, and the validation results showed that discovering a protein–symptom pattern helps understand the molecular basis of syndrome.

Finding potential therapeutic targets of Chinese herbs

The components of Chinese herbal medicine, either a single herb or formulas, are complex chemical systems, and each component binds to its own target receptor. Typical proteomics technologies such as 2D electrophoresis with mass spectrometry can reveal statistically significant changes in the intensity of proteins, which will provide clues for studying the cellular mechanisms of Chinese herb medicine.

Salvia miltiorrhiza (SAME) (‘Dan Shen’ in Chinese) has been widely used for the treatment of cardiovascular and cerebrovascular diseases in Chinese herbal medicine, yet little is known about its cellular mechanisms except for the effect of relieving oxidative stress. Thus, in 2010, a research group from Taiwan [35] investigated the protein–protein interaction network related to the herb SAME using pathway analysis, to identify transcriptional factors in its cellular mechanisms. To evaluate the effect of salvianolic acid B (the main compound of SAME) on A10 cells, differential protein expression in Hcy-induced A10 cells with or without SAME treatment was uploaded into MetaCore bioinformatics software
for network analysis. Basically, the analysis algorithm attempted to deduce scoring processes regulated by differentially expressed proteins, and to find the shortest path in a protein network of the smallest possible number of direct interactions between differentially expressed proteins. The protein interaction network results indicated that proteins expressed after SAME treatment were primarily involved in regulating oxidation, the apoptotic process and cytoskeletal rearrangement. Other than that, SAME was related to the inhibition of reactive oxygen species levels, as well as modulation of several transcriptional factors such as c-Fos, C-Myc and p53. The study provided proteomic and pathway evidence that SAME exhibits a protective effect by inhibiting oxidative stress-induced damage.

Figure 4 demonstrates the common workflow of applying these computational methods to the analysis of symptom, syndrome, disease and related proteins. The principal task of today’s drug discovery process is to identify the biotargets that are related to diseases, clinical profiles, herbs or drugs. Proteomics appreciates the individual differences that are characterized by distinct profiles during disease development and drug response, and the interaction network or pathways among proteins show great value for exploring the potential value of TCM, including disease patterns and herbal medicine. The future of proteomics combined with TCM will enhance the development of drug discovery and personalized therapy.

**TCM meets metabolomics**

Metabolites are the products that specific cellular processes leave behind. Metabolomics [36] is a novel discipline that studies the metabolic profiles, which contribute to the identification of metabolic features corresponding to medical treatment or drug therapies. As a systemic approach, metabolomics adopts a ‘top-down’ strategy to reflect the function of organisms from the end products, and thus to understand the metabolic changes of a living system, usually through urine or plasma samples, which can further lead to the complete understanding of the body reaction.

The ability to understand the dynamic responses to the changes toward integrated living systems can make things easier in the study of complex systems such as TCM [37], with the observation of both endogenous and exogenous factors. Motivated by the intervention effects of Chinese formulas and herb medicine, the characters and changes of the metabolomics profiling can be analyzed with a range of statistical and machine-learning algorithms. The technologies of metabolomics have been successively applied to the key scientific issues in TCM by the international and domestic scholars with metabolomic technology, including syndrome, differential treatment of the individual, synthetic effect of Chinese herbs, molecular pharmacology [38] and quality control of Chinese medicine [39]. At the same time, metabolomics approaches may potentially
benefit the combination between TCM and WM [40] by facilitating the evaluation of the therapeutic effects of TCM formula and understanding of the possible action mechanisms of herbal medicine.

**Biomarker metabolites associated with certain Chinese herb**

Chinese herbal medicine provide modern medicine with a collection of complementary remedies for disease treatment and health maintenance. Some of the natural herbs are known to have good effect on the diseases that are not fully understood regarding their molecular mechanisms. Metabolomics can study the biomarker metabolites and perturbed pathways associated with these herbs, providing insights for drug action mechanisms.

A famous Chinese herbal remedy, ‘SuanZaoRen decoction’, has been used in TCM to treat insomnia for a long time, yet the molecular mechanism and key metabolites underlying it are not fully clear. A research institute in China presented a method with regard to the natural product extracted from SuanZaoRen decoction, ‘Jujuboside B (JuB)’ [41], which is considered to be the major pharmacological active compound responsible for insomnia treatment. Metabolic data of both control and case groups were analyzed to detect the enriched clusters. Multivariate projection approaches such as principal component analysis (PCA) and partial least-squares-discriminant analysis were applied to raw spectrometric data to identify various metabolites as potential biomarkers. Typically, the metabolic profiles of disease cases and controls are compared to identify spectral features and discriminatory variables. The supervised orthogonal projection to latent structures-discriminant analysis was implemented to separate significant metabolites from the two groups, which can be viewed as potential biomarkers, suggesting the metabolic profiles have significantly changed as a result of sleep deprivation. The intermediate data were subjected to computational systems analysis to further investigate the effects of JuB on the insomnia Drosophila metabolite profiles, the parallel PCA score trajectory plots, hierarchical clustering analysis and heatmap visualization were used to distinguish perturbed metabolomics models, which showed that JuB exhibits preventive efficacy by adjusting target metabolic pathways to their normal state. Once the influenced pathways are targeted and analyzed, new drugs can be designed for insomnia.

**Metabolomics study on the toxicity on Chinese herbs**

Although herbal remedies are available to treat various illnesses, little or no actual scientific basis of their toxicity and safety was presented. Under this circumstance, doctors cannot provide enough experimental evidence while guiding their patients regarding proper use or potential toxicity. Metabolites before or after herbal treatment provide measureable objects for studying the pharmacological activity and potential toxicity of herbs.

Aconite root (‘Fuzi’ in Chinese), a herb in TCM, has been popularly used in herbal medicine in Asia for thousands of years, commonly applied for various diseases, i.e. rheumatic fever, painful joints, etc. However, the original plant can be toxic and fatal and, to date, the toxicological risk of its usage is not clear. As a TCM herbal processing approach, Pao Zhi is supposed to reduce the toxicity of aconite root and exert maximal therapeutic efficacy. Studying the metabolomics profiling changes of processing can help define the scientific parameters for evaluating safety and toxicity of drugs [42].

To compare the metabolic profiles of crude herb and processed products, the resultant data matrices were processed for pattern recognition analysis of significant metabolic patterns. The data were imported to EZinfo 2.0 software for PCA, partial least-squares-discriminant analysis and orthogonal projection to latent structures analysis. To maximize cluster discrimination, the data were further analyzed using the orthogonal projection to latent structures-discriminant analysis method. Same as identifying targeting metabolites, variables in the computation results that contributed significantly to discrimination between groups were considered as potential biomarkers. The results indicated that NaK-ATP activity was increased and Lactate dehydrogenase (LDH) and Aspartate Transaminase (AST) values were decreased in the processed group compared with the crude group; this proved that herbal processing can actually achieve weak toxicity of toxic herbs.

Metabolomics approaches, as a strong link among genomics, proteomics and molecular biology, have shown increasing effect on providing biochemical evidence of the mechanisms of the metabolomics network in the study of TCM drug development and herbal medicine. Until now, using metabolomics in combination with TCM pharmacology is a convincing way for the discovery of novel biological active compounds as well as for proving herbal safety. No doubt that metabolomics will
gradually promote Chinese medicine research, and bioinformatics methods such as pattern recognition algorithms will speed up the process.

**TCM HERBAL MEDICINE MEETS MODERN PHARMACEUTICS**

In TCM, medical treatment is determined by a holistic characterization of the patient’s health status (i.e. disease, syndrome), and is thus followed by a prescription that comprises a group of herbs specially tailored for the patient. Most of the Chinese medicinal prescriptions have relatively complex combination of herbs and complex targeting preparation, which usually lead to one or more pharmacological mechanisms [43]. The active components have been isolated from those TCM herbs for chemical compound analysis, and some findings are used in modern pharmaceutical drugs [44]. As multicomponent drug development is gaining rising interest among numerous drug companies, investigating the effective components of TCM herbal remedies and medicinal plants [45] becomes a novel trend to invent new drugs [46].

The core philosophy of modern drug development is the biochemical decomposing process of herbal medicine and tagging the targeting biomarkers. For instance, an herb that we mentioned before—Fu zì, the cardiotonic active substance [44]—has been isolated and identified as higenamine, a beta adrenergic agonist with an isoquinoline structure related to catecholamines. The alternative way of studying the active components of the herbs is through targeted agents affected by herbs. The targeted agents refer to the drug carrier that selectively concentrates on the target site through local delivery or whole-body blood circulation, such as liposomes, nanoparticles and emulsion. [47]. Research on targeted agents can help profoundly understand the structural characteristics and mechanism of drug carrier and achieve clinical targeting drug delivery system.

**TCM Herb databases**

TCM database@Taiwan [48] is claimed to be the world’s largest 3D molecular structure database for TCM drug screening, download and search. The database contains >20,000 pure compounds isolated from 453 TCM ingredients in the form of 2D and 3D structures. The database supports simple and advanced web-based query options that can specify search clauses, such as molecular properties, substructures, TCM ingredients and 2D/3D visualization of molecules.

An integrated knowledge portal TCMOnline (http://www.cintcm.com) [49] is developed and maintained by China Academy of Traditional Chinese Medicine. Currently, TCMOnline database system is the largest TCM data collection in the world, integrating 17 branches in China and >50 TCM-related databases, including chemical ingredients of Chinese herbal medicines constituting Chinese medicinal formula. A comprehensive e-Science architecture [50] was built above these databases to support large-scale database integration and knowledge-intensive applications.

**Exploring functional herbal network**

Because of the common occurrence of compound therapies, studying functional herbal network of TCM herbs reflects the associations and the combinational patterns among various herbs, which can provide insights for studying the molecular mechanisms of herbs and improving drug design.

In the information science, mutual information (MI) of two variables refers to a quantity that measures the mutual dependence of the two variables, which, in this case, are two herbs. The MI between a pair of variables X and Y can be defined as: 

$I(X, Y) = H(X) + H(Y) - H(X, Y)$, where $H(X, Y) = -\sum_{x,y} p(x,y) \log p(x,y)$ is the joint entropy of X and Y.

A herb network can be constructed by studying the MI among herbs from the numerous herbal formulas. Li et al. [51] proposed a distance-based mutual information model to identify combination mechanisms among herbs. The formula data were collected from SIRC-TCM Herbal Formula database and transformed into a numeral matrix $A = (a_{ij})_{m \times n}$, where $a_{ij}$ = 0 means herb j is absent in formula i to indicate the relative position of the herbs in a formula. They provided an integrated scoring system combining the MI entropy characteristics and the ‘between-herb-distance’ d between herbs: $score(x, y) = MI(x, y)/d(x, y)$, which describes the tendency of herbs x and y to form a herb pair. The efficacy of distance-based mutual information model method to uncover the combination rules of herbs has been successfully validated through a frequently used herbal formula Liu Wei Di Huang by co-module analysis.

Sometimes to remove the noisiness of the entropy data, random permutation test is used to set a sound threshold for MI process. Random permutation test estimates the approximate sampling distribution
based on the large number of random permutation samples. The network construction of treatment cases based on this method applied in tumor diseases [52] indicated the complex entropy network is suitable for the analysis of interaction effects among multifactors through the analysis of examples. Similarly, the entropy network method can also be applied to construct subnetworks such as ‘herb–herb’, ‘herb–syndrome’, ‘herb–symptom’ and ‘syndrome–symptom’, and assess the correlation of multifactors by the statistical information of networks. Network-based analysis can be useful in predicting therapeutic mechanisms of specific herb on specific disease, such as SAME and atherosclerosis [53].

Common biological networks shared by different diseases and various herbal formulas might have meaningful evidence corresponding to the commonly existing functional networks. For Rheumatoid Arthritis and CHD diseases that sometimes are treated with similar therapies, literature data were extracted from PubMed and SinoMed, and analyzed with data slicing algorithm [54] to generate commonly existing biological networks.

**Identifying multicomponent targets**

The multicomponent synergy in Chinese herbs is of great significance for understanding TCM herbal medicine, and most importantly, of great value nowadays for the novel drug discovery. Usually the action mechanism of the multiple components is studied separately; however, a novel concept, ‘network target’ [55], has been proposed to consider simultaneously the disease mechanisms and drug actions on a network basis. A network target comes from the single target-based and multiple target-based drug studies, referring to a therapeutic target that is derived from systematic interventions of the biology network underlying a disease or pathological process. The mechanisms of herbal synergism [56] can be investigated by *in vivo* studies using drug combination analysis methods and network-based analysis of the molecular interaction profiles and pathways regulatory actions of active ingredients.

Li et al. [57] proposed a network target-based identification of multicomponent synergy (NIMS) method to access the synergistic strength of multicomponent therapeutics. They measured synergistic agent combinations by introducing and integrating two parameters, named topology score and agent score. Then, NIMS was applied to prioritize synergistic combinations from the agents including herbs or herb compounds. Two graph-based measures, betweenness and closeness, were used to capture the associations among variables, and the other measure, PageRank, to verify the node importance. Interestingly, when NIMS is applied to the angio-genesis network, two synergistic agent pairs, ‘Sinomenine and Matrine’ and ‘Sinomenine and Honokiol’, are found to be the main constituents of TCM herbal formulas Qing Luo Yin and Tou Gu Zhen Feng. The preliminary results demonstrated that NIMS, as a software tool, has the potential for screening synergistic combinations of Chinese herbal formulas, as well as modern drugs.

Wang et al. [58] developed a quantitative structure–activity relationship (QSAR) to study the multioject optimization of component combination in drug design. The basic idea of optimal combination of drugs is to obtain best biological activities and minimal side effects. In QSAR, a multicomponent combination of different doses of *n* components $C_i$ $(i = 1, 2, \ldots, n)$ was represented within a certain combination vector $[T_{R1}, T_{R2}, \ldots, T_{Rn}]$. They introduced a mathematical function $Y_R = f(X_R)$ to quantify the multicomponent combination of chemical composition and biological activities, where $X_R = C_i * T_{Ri} \quad (i = 1, 2, \ldots, n)$ stands for the composition of $R$. This algorithm-based knowledge discovery method first quantitatively interprets the relationship between combinations and their activities into a model, and then uses the proposed model to simulate the real-life drug formula combinations. It was found to be valuable in providing potential association information for studying the design of Chinese formula Shenmai.

The biological activities of the Chinese herbal medicine are unclear because of their complex nature and the possible interaction among the combined therapies. With the DNA microarray data, gene expression profiles were analyzed to compare disease-altered genes and drug-altered genes, and transcriptomic analysis was used to perform pathway analysis and gene expression similarities in different formulas. This method [59] has been applied to study the top 15 most used Chinese herbal formulas and the comparison results indicated that the TCM formula treatments might be related to metabolic, cardiovascular, neuroskeletal and hepatic diseases.

Network-based computational technologies based on systems biology strategies have been widely used in modern drug discovery, mostly in herbal synergy
Distinguishing the herbal properties

To improve the research of TCM pharmaceutics, studying the herb properties (HPs) of the herb recipes and keeping track of herb-related entities can be helpful in modulating the pharmacological and toxicological effects of the chemical ingredients of the constituent herbs. Here, HP refers to the fundamental parameters (such as characters, tastes, toxic states, meridians) in TCM herbs.

Ung et al. [64] divided the TCM-HPs into four classes: character (C), taste (T), meridian (M) and toxicity level (Tox). These classes can be further divided into 5, 5, 12 and 2 subclasses, respectively. Each herb can be represented with a vector \( h = (C, T, M, Tox) \) of the TCM-HPs with 39 features. The value of HP is 1 if the herb possesses the corresponding property, and it is 0 if the herb does not possess the property. For an herb pair composed of herbs A and B, two separate vectors \( h_{AB} = (h_A, h_B) \) and \( h_{BA} = (h_B, h_A) \) of dimension 78 can be formed, both of which were used to represent the herb pair. Given TCM herb pairs and randomly generated non-TCM herb pairs, artificial intelligence methods were applied to attempt to distinguish TCM herb pairs from non-TCM herb pairs. Mainly, probabilistic neural network method is used to classify the HPs, and k nearest neighbor is used to predict the class of unclassified vector (new herb pair). Known TCM herb pairs and non-TCM herb pairs in a training set were adjusted by using a separate testing set of TCM herb pairs and non-TCM herb pairs. By projecting the feature vector of new multitherb pairs, support vector machine (SVM) method can be used to determine whether it is a valid TCM herb pair based on its location with respect to the hyperplane. The classification results suggested that TCM-HPs of TCM herb pairs contain distinguishable features that contribute to the synergistic combinations.

Novel drug discovery

Although the success of the antimalarial drug artemisinin was inspired by Chinese herbal therapies, the pharmacological effectiveness of Chinese material medica has not led to the successful development of new drugs. Systems biology technologies and bioinformatics approaches [65, 66] are making promising progress in opening up the opportunities of developing new drugs [67] from herb–drug interactions [68].

It is recognized that Chinese herbs are usually screened for compounds that may be active against certain targets. In modern biomedicine, enzymes and receptors represent the most common drug targets. Target-based drug discovery is an important strategy for developing new agents. Thus, a new concept of drug discovery is a reverse approach to find and separate the active compounds from TCM by using virtual screening, immobilized enzymes [69], polyclonal antibodies and molecularly imprinted polymers. Databases of Chinese herbal constituents are being created [70], providing the community with more detailed information of Chinese herbs.

Moreover, intelligent computing algorithms such as data mining approaches [71] can be applied in TCM-inspired drug discovery. Zhou et al. [72] proposed a structure-activity relationship method to identify potential active ingredients in natural products. An SVM model was trained to learn the structural activity relationship of bioactive NP ingredients. Decision tree algorithms and rule set algorithms were applied to pick out important descriptors in bioactive ingredient prediction. The method should find evidence about the therapeutic mechanisms and synergies of natural products, whose ingredients can potentially compose drugs.

There are so many active fields in the integration of TCM herbal medicine and modern pharmaceutics, and drug discovery based on effective Chinese herbs seems promising with the help of various computational analysis methods and emerging data resources. Except for manual curation, knowledge about herbal information can also be extracted from textual data [73], with the assistance of domain-specific ontology models. In addition, currently in the field of TCM herbal medicine, a novel concept ‘herbogenomics’ is defined as the analysis of the biological effect of the target objects of a
particular herbal medicine through a profiling of the affected genomic and proteomic changes. Integrating genomic and proteomic profile changes into the research of the efficacy and toxicity of herbal medicine [74] will provide novel insights into all the mechanism and action studies of TCM herbal medicine.

**TCM MEETS EVIDENCE-BASED MEDICINE**

In Chinese medicine, diagnosis is highly related to contextual information, including not only a detailed profile of the patient at the moment but also the external factors such as climate, emotional states, joint problems, types of pain, fever, etc. Practitioners collect these symptoms through four typical methods: inspection, auscultation–olfaction, inquiry and palpation. Doctors will deliver diagnosis and treatment based on a holistic understanding of the patient’s condition. This diagnostic process in TCM is rather experience-based than science-based, so explicit evidence is missing in explaining the mechanisms, efficacy and safety of the diagnosis.

Recently, the novel discipline called ‘evidence-based medicine’ [75] aims to deliver optimized health care by informing clinical decision making in diagnosis and systems biology-supported personalized health strategies [76] with the use of computation estimate of the benefits and the harm. It can potentially enable the researchers from both TCM and WM in assessing the strength of the evidence found in various data sets and analysis and benefit the integrated treatments. Until now, evidence-based approaches in TCM [77] have been focusing on major fields such as improvement of poor treatment, reduction of severe adverse effects, unwanted interactions of standard therapy with herbal medicines and efficacy and safety of TCM treatments. Moreover, there are a lot of successful case studies [78] that showed finding evidence in TCM helps improve new drug discovery and functional food therapies in modern drug development.

**Understanding TCM concepts**

Here, the TCM concepts refer to the core entities in the basic theoretical systems of TCM that play an important role in mapping and shaping the universe, making treatment decisions and delivering personalized therapies. One of the understanding obstacles between WM and TCM is that they disagree with each other on the concept level. WM is based on modern mathematics and TCM is based on ancient understanding of the nature.

**Understanding ZHENG**

‘ZHENG’ (syndrome) differentiation is often used as a guideline in TCM disease diagnosis, and has been recently incorporated with biomedical diagnosis [79]. There is no equivalent concept in WM, although there could be a biomedical path that may explain the molecular and cellular mechanisms of the ‘ZHENG’ concept.

One of the explanations of the molecular basis of ZHENG is the context of neuro-endocrine-immune (NEI) network system [80]. In modern medicine, NEI system acts as a pivot in modulating host homeostasis and adjusting health through complex communication among hormones, cytokines and neuro-transmitters. NEI network and ZHENG network can be extracted from PubMed articles through text mining, and by topological comparison and the pathway analysis of networks of Hot ZHENG and Cold ZHENG, it was found that hormones are predominant in the Cold ZHENG network, immune factor is predominant in the Hot ZHENG network and these two networks are connected by neurotransmitters. Here, Hot ZHENG and Cold ZHENG are widely applied in the diagnosis and treatment of patients in TCM as two different patterns. The results demonstrated the distinguishable differences between Hot ZHENG and Cold ZHENG in TCM in terms of different targets and different receptors in a biological communication network.

Specifically, Ma et al. [81] focused on the implicit stratification of Cold Syndrome by surveying 4575 cases of Cold Syndrome patients and examining gene expression information of a typical Cold Syndrome pedigree by microarray. The symptom patterns of Cold Syndrome can be standardized using latent tree model. The Cold Syndrome was represented as quantitative scale that is composed of 20 factors, such as the 10 fixed cold body parts from head to feet. Each cold factor can have value 0, 1, 2 and 3, standing for none, mild, moderate and severe level, respectively. The latent variables derived from the latent tree model of Cold Syndrome are three categories of elements, namely, cold adaption, cold behaviors and cold areas. On combining the multifactor measurements with NEI network, results indicated that Cold Syndrome-related genes play an essential role in energy metabolism.
For more specific studies of syndromes, latent tree models have been applied in the validation of Kidney-Yin deficiency syndrome and Kidney-Yang deficiency syndrome in elderly women with menopausal symptoms. In the study of multitarget molecular pharmacology of Qi-deficiency and Blood-stasis syndrome (QDBS), high-throughput gene microarrays before and after treatment with Fuzheng Huayu Capsule were analyzed. Patients with QDBS suffer from energy deficiency and blood stasis. It is related with a bunch of diseases such as diabetes mellitus, dyslipidemia, hypertension, hepatitis and liver cirrhosis. Fuzheng Huayu Capsule is a recipe for treating liver fibrosis with QDBS. A pathway-based similarity comparison method was proposed based on a microarray database ‘Connectivity Map’, which collects microarrays corresponding to treatment of different small molecules in different human cell lines. A gene set repository Sigpathway was referenced to sort out the meaningful gene sets and pathway information according to the microarray data. The expression pattern similarity between the microarrays and in the Connectivity Map Database in every selected pathway was calculated using KS-test, whose result will be either positive or negative.

Understanding Yin–Yang

It is a well-known concept that refers to contrary forces in the natural world, and how they give rise to each other. The basic theories of TCM fundamentally rely on the understanding of Yin–Yang balance in the diagnosis and treatment of diseases. However, the Yin–Yang balance has not been studied by modern scientific means. One of the convincing explanations of Yin–Yang balance is antioxidation–oxidation with Yin representing antioxidation and Yang oxidation. In their article, they found the fact that the Yin-tonic traditional Chinese herbs have about 6 times more antioxidant activity and polyphenolic contents than the Yang-tonic herbs on average.

Advanced medical diagnosis

Medical diagnosis in TCM is a complex reasoning process based on the identification and differentiation of symptoms, syndromes and diseases. Modern diseases can sometimes find corresponding description, symptoms and treatments in TCM, e.g. Parkinson’s disease. However, there are several challenging problems that need to be solved to transform the reasoning process to an interpretable model of diagnostic rules. First, no regularized architecture of these entities exists because of the relatively non-standardized naming in ancient descriptions. Second, although the diagnosis in TCM is based on individuals’ experience, the clinical records of thousands of years of accumulation contain larger amount of information than well-structured data of prescriptions extracted manually from TCM literature. These data can be an important source for discovering useful patterns and regularities in TCM. Last, owing to different experience and background of TCM doctors, the same concept might be described in several different terms, which will cause inconsistency and duplicate issues.

Wang et al. addressed the problem of automatic symptom name normalization by measuring the similarity between the clinical symptom name to be normalized and all possible standard forms. The background knowledge is that most of the symptoms listed in clinical records have >1 synonym. Based on the experiment evaluation, after all kinds of similarity distance calculation within the features of dynamic programming, vector space model, it was found that the symptoms can be normalized into reasonable and accurate standard ones among three similarity metrics—literal similarity metrics, remedy-based similarity metrics and hybrid similarity metrics. This automatic symptom name normalization method is important in advancing a unified system of medical theories and diagnosis. Apart from similarity metric normalization, MI among symptoms provides an alternative way to study the relations between symptoms based on clustering of associated symptoms with Bayesian network construction.

In the process of TCM diagnosis, not only the normalization of symptom names matters but also the differentiation of syndromes matters. Mapping into an abstract model, syndrome differentiation is basically a classifier that classifies patients into different classes based on their symptoms. For example, Kidney-Yang deficiency is a TCM syndrome concluded from the observation of manifestations, from symptoms such as ‘cold limbs’. The relationships between symptoms and syndromes can be regarded as a model of latent trees, in which syndrome is represented as a latent variable and symptoms as manifest variables. Abstractly speaking, the task of TCM diagnosis is to classify the input to latent variables that have been defined, based on the differentiation of manifest variables. Formally, a latent tree
model can be represented as $M = (G, \theta)$, where $G$ stands for the model structure plus cardinalities of variables, and $\theta$ stands for the vector of probability parameters. In a case study of a subdomain of TCM diagnosis—kidney deficiency, a bunch of symptom variables were selected in data collection phase and analyzed with the kidney data using learning algorithm, and consequently generate models of latent trees as natural clusters. The results showed that there exist natural clusters in data that correspond to TCM syndrome types.

More interesting attempts tried to associate the basic theories of TCM diagnosis and treatment principles with a multilateral system of mathematical representations and relationships [90], namely, two kinds of opposite relations and one kind of equivalence relation, to formalize the treatment principles in mathematical way, and the reasoning of the mathematical assertions proves that the principles are true.

In modern TCM research, it is possible to integrate syndrome differentiation [91] with orthodox medical diagnosis leading to new scientific findings in overall medical diagnosis and treatment. The key challenge in TCM syndrome differentiation research is how to standardize the diagnostic procedure for syndrome. Modern technologies are applied into the syndrome differentiation in diagnosis, clinical research, pharmacological research, new drug discovery and medical equipment development. The way toward to identify evidence of TCM syndrome differentiation in TCM practice requires multidisciplinary collaborations from biomedical, bioinformatics, medical, pharmaceutical and TCM disciplines. Pattern differentiation in TCM will help identify a subset of patients who are more likely to respond to combination biomedical therapy treatment [92]. In addition, specific diagnostic examples such as apoptosis pathways in cancer related with Chinese medicine [93] and the QSAR study of the bioactivity of the TCM compounds in stroke treatment [94] were studied to understand the mechanisms of diseases and the novel therapeutics.

NEXT-GENERATION PLATFORM FOR BIOMEDICAL INTEGRATION

The rise of Semantic Web technologies in biomedicine

A massive amount of biological data has been generated, which has led to the adoption of tools for automated analysis of the biological knowledge and the full integration of heterogeneous data. Because of the data diversity and semantic heterogeneity of biological data [95], it is essential for the whole biomedical community to share a standardized data and knowledge representation paradigm. In the past decade, the amount of biological ontologies [96] has been growing, within the definition that ontology is a technology in Semantic Web used to represent and share knowledge about a domain by modeling the things in that domain and the relationships between those things.

Specifically, ontology is becoming a technology advocated by computer scientists and bioinformaticians to overcome problems of heterogeneity and to form a linked web of structured knowledge instead of sparse data distributions. In Semantic Web, the basic idea of representing knowledge is resource description framework (RDF) [97] and Web Ontology Language [98], in which the resources (e.g. specific genes, gene-encoded proteins) are formatted and linked by binary associations. The resources form a directed knowledge graph with each node assigned with unique web identifiers, which can enable the contents to be stored, managed and queried within any persistent systems or web environment. Through the abstractions of biomedical semantics [99], the Semantic Web offers an ideal platform for representing and linking biomedical information among different researchers. Community effort have already been made, including the growing repository Bio2RDF [100] that contains ~5 billion triples, including omics information [101] such as genes [102] and proteins [103] and other resources such as pathways [104], diseases [105] and drugs [106]. With the strong power of expressivity and connectivity of ontology resources, the knowledge-supported biomedical data should be able to import rules like ‘If Gene X is implicated in Disease D, and its Protein Product Y is a functional component of only Pathway P then Disease D directly perturbs Pathway P’, and answer complex cross-discipline questions that are associated with customized requests such as drug discovery [107, 108].

In today’s cyberinfrastructure of biological information processing [109], Semantic Web technologies are playing a more important role in greatly reducing the amount of coordination needed among participants, which will potentially accelerate the development of biomedical information communication.
Use case: modern biomedical integration

Standardized representation and computational analysis of biomedical data have become increasingly important, and this situation provides us with a speedy way of facilitating the communication of TCM and systems biology from accumulated high-throughput data. With the increasing amount of biomedical ontologies [110], we believe that the integration of biomedical data on the semantic level might bring hope to the research of TCM and modern medicine in two steps: (i) connect various kinds of information from published trustworthy data sources [111] and make the resources recognizable and linkable and (ii) introduce associated knowledge into the linked data and make the data available for association discovery and comprehensive knowledge inference. The basic idea of the connections is demonstrated in Figure 5.

By reconstructing the networks of biological entities such as gene, transcription factors, compounds and other regulatory molecules [112], and integrating significant findings in various facets of TCM mentioned in this article, a broader biomedical linked cloud can be built up and may be able to derive novel cross-discipline evidence based on logic-based reasoning, association discovery and network analysis. As an example, the idea of openly available RDF/Web Ontology Language data sets has been applied to integrate heterogeneous data of traditional medicine and modern pharmaceutical research [113], aiming to find evidence for pharmaceutical compounds from Chinese medicine that may treat depressive disorders or serve as lead compounds for the future pharmaceutical drug development. As a typical example of assessing the drug-target associations in semantically linked data [114], a wide range of databases that include compound-gene, drug–drug, protein–protein interactions, and side effects of drugs can be annotated with domain-specific semantic information, and the strength of association can be calculated based on the topology and semantics of the neighborhood in the drug linked data.

Cheung et al. [115] proposed a small specific example of the integration of WM and TCM on Alzheimer’s disease based on semantic ontology representation and association discovery mechanisms. In this example, _Huperzia serrata_ (HS) is a Chinese herb that strengths the kidney, mainly for curing aging disorders, whereas in some western biological experiments, it was found that a compound of the herb HS acts on the brain and can serve as a potential therapy for the Alzheimer’s disease. This knowledge was represented as several linked graphs of resources, and after linking the knowledge with newly found statements, they could assert that HS targets the brain (WM) instead of the kidney (TCM). For more information of drug discovery from Chinese medicine, researchers can refer to [116].

Another promising technology support for the biomedical domain is cloud computing of large-scale data [117], which is now becoming more important in the data-deluge era. It provides a group of hardware, software and computing frameworks to deal with the data effectively. Combining Semantic
Web technologies and cloud computing together [118] can allow biomedical researchers to perform the integration of vast amount of data and knowledge in a comparably reasonable time.

CONCLUSION
With the development of integrative medicine, the integration of TCM and modern informatics technologies is increasing rapidly. Bioinformatics technologies, such as text mining, relational database construction, similarity calculation and latent clustering, have great potential in building and discovering the explicit and implicit associations between TCM and modern biomedicine. Currently, the major focus of research is divided into two directions, understanding the mechanisms of TCM from the systems biology perspective and facilitating novel drug design based on the analysis of TCM herbal medicine [119]. The applications are still limited and massive amounts of research data all over the world are private and isolated, which will restrict the communication of the community and weaken the power of integrative medicine. Advances in Semantic Web, especially the effort on linking open life science data together, open a gate for worldwide access and integration of the knowledge discipline. On the other hand, the capability of consuming large scale of data today will allow researchers focus more on the data itself.

There are two basic steps for us to enrich the life science data integration with Chinese medicine knowledge. First, as there exists a wide range of biomedical ontologies in WM but little in TCM, it is an urgent need to transform the knowledge of herbal medicine, basic theories and disease diagnosis in TCM into structured models of ontologies. Second, find the associations between the knowledge of genes, protein, drugs and the knowledge formatted, or import the existing associations into the big linked biomedical data, which can be further used to generate meaningful connections, especially implicit facts as novel discoveries.

Key Points
- Combining systems biology with modern bioinformatics technologies to probe the mystery of TCM will potentially help bridge the methodological gap between TCM and modern medicine, providing the scientific evidence to support simplification of the complexity of TCM, which can verify the value of TCM in modern medicine.
- Decomposing the biochemical components from TCM herbs and tagging the targeting biomarkers using MI and network-target identification, the active components of herbal medicine can be investigated to invent modern multicomponent drugs.
- The way toward to identify evidence of TCM syndrome differentiation in TCM practice requires multidisciplinary collaborations from biomedical, bioinformatics, medical, pharmaceutical and TCM disciplines, e.g. explaining the molecular basis of syndromes through NEI network and identifying the patterns during syndrome differentiation through latent tree models.

FUNDING
China’s Natural Science Foundation Project (NSFC61070156/NSFC60525202); NSF of Zhejiang (LY13F020005); China National Cloud Initiative.

References


Modern bioinformatics meets traditional Chinese medicine


