Integrating transcriptional profiling with network-based methodologies for revealing molecular mechanisms of traditional Chinese medicine

L Huang¹, L Jiang¹, Q Kuang¹, M Li¹, Z Wen¹,²*, L He³*

Abstract

Introduction

Traditional Chinese medicine (TCM), which originated in ancient China and has been developed for thousands of years, is widely used in China and the west nowadays as a part of complementary and alternative medicine. In order to advance the modernization of TCM, more evidence of the efficacy and pharmacological mechanisms at molecular level is still necessary for better understanding of TCM. The microarray technology can provide deep insights into the mechanisms of action at transcriptional level and is successfully applied to drug discovery and development. Based on gene expression data, a number of genes/targets, for which the expression levels significantly changed between the biological states before and after treatment by TCM in vivo or in vitro, can be identified by various statistical algorithms and used to interpret complex therapeutic effects of TCM. Moreover, the network-based approaches have recently been proposed to systematically investigate the molecular mechanisms of TCM from a systematic point of view. This paper reviewed the TCM researches based on the microarray data as well as the network-based system biology. The discussions on the perspective of integrating the microarray technology with the network-based approaches were also given in this review.

Conclusion

TCM uses holistic and synergistic approaches for disease treatment and keeping the yin-yang balance of body energy. Microarray technology can monitor TCM-induced changes of whole genomics and be used to identify the differentially expressed genes for further interpretation of TCM efficacy. Further research is needed to improve our methods and understanding of TCM.

Introduction

Traditional Chinese medicine (TCM) has been widely used in China for over thousands of years as the ancient prevention and medical treatment of various diseases. Under the guidance of TCM theory, the herbal combinations so-called TCM formulae were applied to the treatment of cold and hot syndromes according to the four natures of the herbs, namely hot, warm, cool and cold, for the purpose of restoring the yin-yang balance of the body energy. Although TCM is considered as part of complementary and alternative medicine nowadays in China and the west and there are many successful examples in clinical applications, the usage of TCM formulae is only based on the empiric clinical experience and the mechanisms of action of TCM are still unclear. Therefore, more molecular evidence of the therapeutic efficacy and pharmacological mechanism is necessary for better understanding of TCM and advancing its modernization.

Unlike western medicine, TCM uses the holistic and synergistic approaches for disease treatment and a TCM formula may target multiple molecules/pathways at the same time. In order to interpret the complex therapeutic effects of TCM, the methodologies of genomics and system biology are required for providing the system-level insight into TCM pharmacology. As a powerful tool in genomic researches, the microarray technology can monitor the expression levels of tens of thousands of genes simultaneously and give a global view of biological systems. The mechanisms of action of TCM can be reflected by the changes of gene expression profiles before and after treatment by TCM in vivo or in vitro. The target genes and the corresponding signalling pathways are subsequently identified for further biological analysis. In recent years, the network-based approaches in the researches of system biology have been proposed to investigate the drug–target interactions from a systematic point of view, especially for the multi-target drug discovery. In classical drug discovery, a drug is considered to act on an individual drug target. However, because of the existing interactions of the proteins and genes in a biological system, the drugs will act on multiple targets (proteins) instead of a single one. Likewise, the active components in a TCM formula will also target multiple proteins and cause complicated interactions within a biological system. The network-based system...
Review

Discussion

The authors have referenced some of their own studies in this review. The protocols of these studies have been approved by the relevant ethics committees related to the institution in which they were performed.

Identifying the genes/pathways impacted by TCM based on microarray data

In clinical application, the TCM formulae generally consist of several or dozens of herbs. Based on the theory of TCM, the pharmacological efficacy of Chinese herb medicines comes from the holistic and synergistic effects of all the herbs. The multiple active components in a TCM formula will cause very complex phenotypic changes within a biological system. For interpreting the complex therapeutic effects of TCM at the molecular level, identifying the genes and gene products impacted by TCM is one of the crucial steps in the procedure of investigating the TCM-induced changes in the biological system. The microarray technology can provide a global view of the biological system by monitoring tens of thousands of genes simultaneously. The phenotypic changes induced by TCM can also be reflected by the changes of gene expression profiles before and after treatment by TCM in vivo or in vitro. According to the identified differentially expressed genes and the corresponding signalling pathways, further biological analysis is conducted to reveal the mechanisms of pharmacological efficacy of TCM. Recently the microarray-based gene expression profiling has been used for exploring the molecular mechanisms of TCM formulae.

For the high-intensity noise-induced audiogenic seizure (AGS), Li et al. identified 134 differentially expressed genes from the gene expression profiling of the inferior colliculus of mice and discovered that the TCM Qingyangshenylcosides could effectively prevent many of the AGS-induced gene expression changes, which indicated the pharmacological efficacy of Qingyangshenylcosides on the prevention of audiogenic seizure. Yang et al. selected the genes impacted by a TCM formula Yan-Gan-Wan and revealed that these genes were associated with TNFα expression and NF-κB activation, which indicated the protective effects of Yan-Gan-Wan on carbon tetrachloride hepatotoxicity. Pan-Hammarström et al. investigated the underlying mechanisms of a TCM formula V1-28 used in the treatment of chronic human immunodeficiency virus and hepatitis B virus infections based on the gene expression data of human lymphocytes. Zhang et al. elected 483 cancer-related genes from the gene expression profiling of lung cancer cell lines according to the changes of expression levels before and after treatment by a Zilongjin tablet and interpreted its potential anti-cancer mechanisms.

In our previous studies, we explored the phytoestrogenic activity and the Nrf2-inducing activity of a TCM formula Si-Wu-Tang, which is widely used in China for the treatment of women’s diseases. The gene expression profiles of MCF-7 cells induced by Si-Wu-Tang at medium and high concentrations showed an excellent match to the CMAP gene expression profile of oestradiol-treated MCF-7 cells, suggesting a widely claimed usage of Si-Wu-Tang as an effective TCM for women’s diseases with a possible phytoestrogenic effect. Furthermore, a number of Nrf2-regulated genes were also significantly impacted by Si-Wu-Tang, but not by oestradiol. The Nrf2 (nuclear factor erythroid 2-related factor 2) is considered to be a key molecular target for chemopreventive agents. Therefore, this result indicated that Si-Wu-Tang and its components could impact multiple components of the carcinogenic process by affecting the Nrf2-regulated genes.

Network-based approaches revealing the TCM-associated molecular interactions

In real biological systems, e.g. the cells, there are various interactions among the genes and proteins. These interactions existing in the systems can be well characterized by the approaches of network-based system biology, which represents the relationships between the molecules as a network, e.g. protein–protein interaction network. The molecules in the biological system are considered as nodes and the direct (actually physical interactions) or indirect (e.g. associated in same signalling pathways) between the molecules are considered as edges. Recently, the network-based approaches have been widely used for biomarker identification, disease-related candidate gene prioritization and multi-target drug discovery.

In conventional drug discovery, a drug is designed for specifically binding to a drug target with maximal selectivity and affinity, which is known as the ‘one drug, one target’ approach. However, complex diseases, e.g. cancers, are usually caused by multiple interactions between the genes and other molecules instead of the change of a single gene. For a biological network, a drug should be designed to target multiple genes or proteins, which have crucial functions in the system and show important features in this complex network, e.g. having the maximal number of edges.

Licensee OA Publishing London 2013. Creative Commons Attribution License (CC-BY)

A new strategy, the so-called network pharmacology, which identifies a set of drug targets for any disease by network analysis and investigates the effects of the drugs binding to these targets by chemical biology approaches, has been developed for modern drug discovery. Considering the holistic and synergistic effects of herb medicines, TCM uses the formulae including several or dozens of herbs for disease treatment. For the purpose of depicting the complex mechanisms of action of TCM formulae, the network pharmacology is also proposed in TCM researches. In clinical application, the Compound Danshen Formula (CDF) is commonly used for the treatment of cardiovascular diseases. To investigate the underlying mechanisms of action of CDF, Li et al. screened the candidate active compounds from TCM chemical databases and their potential targets for constructing three networks, namely compound-pathway network, compound-target network and target-disease network. By analysing the properties of these networks, e.g. topology of the network, a novel compound-target network for CDF was subsequently constructed to identify the key potential targets of active compounds of CDF and to illustrate the mechanisms of action of CDF on cardiovascular diseases. Tao et al. identified 58 bioactive components from a TCM formula Radix Curcumae, which was often used for the prevention of cardiovascular diseases, and predicted 32 potential targets related to cardiovascular diseases by constructing and analysing a well-characterized candidate compound–candidate target network. The results showed the synergistic actions of active compounds in Radix Curcumae for the prevention of cardiovascular diseases. In addition, the network pharmacology combined with the pharmacokinetic analysis was also proposed by Pei et al. as a strategy for revealing the material basis of Chinese herbal formulae and used to discover four bioactive compounds from Bushen Zhuanggu formula.

Perspective of integrating the microarray technology with network-based approaches for TCM researches

Although both the microarrays and network-based approaches can investigate a biological system from a systematic point of view, the limitations still exist when applied to TCM researches. For the microarray technology, all the genomic changes of cells detected by microarrays are induced not only by the active components of TCM, but also by the other unrelated components in the herbs. It will cause a high false positive in identifying TCM-induced differentially expressed genes and mislead the biological interpretation. How to accurately screen the differentially expressed genes induced by TCM active components from microarray gene expression profiling is an important issue for TCM pharmacology. The network-based approaches are useful for prioritization of disease-related genes by associating the candidate genes with the seed genes, which are known as the pathogenic genes, through ‘guilt by association’ methods across various knowledge-based data sources. However, when none or only a few of the seed genes are known beforehand for a new disease, this kind of method becomes in applicable or ineffective.

To improve the accuracy and applicability of the methodologies in discovering the mechanisms of action of TCM, integrating the microarray technology with network-based approaches was proposed in several researches. Nitsch et al. replaced the prior knowledge about the biological process by the microarray gene expression data and constructed a global distance network deduced from functional protein association network by a kernel method. Using the disease-induced gene expression profiling instead of the prior knowledge of a disease for the construction of network can overcome the limitations of previous network-based approaches to some extent. The results showed that the pathogenic genes were successfully prioritized from four monogenic diseases by the global distance network. Chen et al. characterized the hot and cold syndromes of rheumatoid arthritis by the microarray gene expression profiles in CD4+ T cells and explored the mechanisms of rheumatoid arthritis by combining the gene expression data with the protein–protein interaction network. Li et al. established a network balance model with the prior-known genes and the differentially expressed genes identified by microarrays for chronic superficial gastritis (CAG), for which the patients were classified into hot and cold syndromes based on the TCM theory. When the hot or the cold syndrome of CAG occurred, the model became imbalanced. Then, the potential biomarkers of hot/cold syndrome were identified for the treatment and the mechanism discovery of CAG.

Conclusion

As a part of complementary and alternative medicine in China and the west, TCM uses the holistic and synergistic approaches for disease treatment and keeps the yin-yang balance of the body energy. It is an important issue to reveal the pharmacological efficacy and mechanisms of action of TCM. Microarray technology can monitor TCM-induced changes of the whole genomics and be used to identify the differentially expressed genes for further interpretation of TCM efficacy. Network pharmacology depicted the complex interactions of the molecules in the biological system before and after treatment by TCM and was successfully used to discover the multiple genes and other molecules targeted by TCM and illustrate the mechanisms of action of TCM efficacy. Network-based approaches to some extent.
TCM from a systematic point of view. Compared with the microarray technology and network pharmacology, the strategy of combining microarray gene expression profiling with network-based approaches is more applicable for investigating the molecular mechanisms of TCM. In general, the continuous development of the techniques in genomics and system biology will make the pharmacological mechanisms of TCM clearer and allow the use of TCM to become safer and more effective, which will promote the modernization of TCM.

Acknowledgement

This work was supported by the National Nature Science Foundation of China (No. 21205085).

References

32. Pei L, Bao Y, Liu S, Zheng J, Chen X. Material basis of Chinese herbal formulas explored by combining pharmacokinetics...