Available online at SciVarse ScienceDirect

Chinese Herbal Medicines (CHM)

ISSN 1674-6384

Journal homepage: www.tiprpress.com E-mail: chm@tiprpress.com

Review

Systems Biology Application in Research on Sustainable Utilization of Chinese Materia Medica Resources

Sheng Wang[1], Hai-yu Xu[1], Lan-ping Guo[1], Lu-qi Huang[1], Chang-xiao Liu[2*]

1. State Key Laboratory Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China

2. Tianjin Binhai Research Center for Food and Drug Regulatory Science, Tianjin Center for New Drug Evaluation Research, Tianjin Institute of Pharmaceutical Research, Tianjin 300193, China

ARTICLE INFO

Article history

Received: February 5, 2015

Revised: May 21, 2015

Accepted: June 11, 2015

Available online:

DOI:

ABSTRACT

This paper reviews the progress of systems biology applied to research on sustainable utilization of Chinese materia medica (CMM) resources in the following aspects: identification and evaluation of CMM resources, analysis of biosynthesis and their regulation of active ingredients in medicinal plants, metabolic engineering and synthetic biology research of medicinal plants, and molecular breeding of medicinal plants. Development of systems biology is currently leading to extremely broad applications in the field of CMM resources, and systems biology will become a significant approach for the sustainable utilization of CMM resources.

Key words

biosynthesis; Chinese materia medica; design breeding; evaluation; identification; metabolic engineering; synthetic biology; systems biology

1. Introduction

Systems biology is a new field of biology that aims to develop a systems-level understanding of biological systems (Kitano, 2001). Systems biology involves the study of a variety of molecules with different structures and functions as well as their interactions at the levels of cells, tissues, organs and organism, and quantitatively describes and predicts biological function, phenotype and behavior through computational biology.

Kitano (2002) indicated that a systems-level understanding of a biological system can be derived from insight into four key properties (Figure 1): system structures, system dynamics, control methodology, and design methodology (Kitano, 2002). The ideas and methods in the systems biology field essentially include the integration of multi-information and the establishment of system models. Genomics, transcriptomics, proteomics, metabolomics, interactomics, and phenomics constitute the major technologies and platforms of systems biology (Aderem, 2005). However, systems biology is not a simple accumulation of "omics" data; Instead, it would extract some optional mathematical models from the "omics" data. These mathematical models not only could simulate the behavior of biological systems, but also can predict the future behavior of the system in case of stimulus and outside interference (Kitano, 2001).

Systems biology is characterized by integrity research. Traditional Chinese medicine (TCM) plays a role in the overall

*Corresponding author: Liu CX E-mail: liuchangxiao@163.com

Fund: National Natural Science Foundation of China (81130070, 81430096); Important National Science & Technology Specific Projects (2012BAI29B02, 2012BAI28B002)

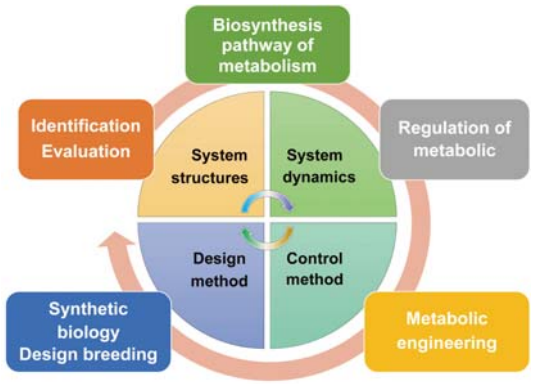

Figure 1 Four key properties of systems biology and their application in research on sustainable utilization of CMM resources (Kitano, 2002)

regulation of disease treatments and "from the overall concepts". TCM and systems biology are well matched in the understanding of overall concepts (Wang and Xu, 2014). In the modern research and development of CMM, systems biology methods can be used to study the pharmacological effects, mechanisms of action, drug safety and compatibility rules of CMM (Liu, 2006). In research on the sustainable utilization of CMM resources, the methods and techniques of systems biology are also very important. Methods and techniques of systems biology can provide scientific methods and bases for the identification and evaluation of CMM resources, analysis of metabolic pathways for effective compounds of CMM, systematic study of metabolic regulatory networks, synthetic biology research on medicinal ingredients as well as research in genetics, breeding and quality improvement of medicinal plants, to ensure the quantity and quality of CMM resources and achieve the biological diversity and ecological balance in the full utilization of CMM resources (Figure 1).

2. Identification and evaluation of CMM resources

Identification and evaluation of CMM resources have a bearing on the safety and effectiveness of clinical drugs. Currently, identification and evaluation of CMM resources have been developed from overall macroscopic subjective judgments to microscopic accurate detection of a single component, through morphology studies at the organism, organ, cell, and DNA levels (Chen et al, 2014). However, there are some limitations with respect to evaluating CMM through single or partial index components. The ideas and methods of systems biology have been applied to evaluating the quality of CMM resources. Huang et al investigated the categories, differentiation, and quality differences of CMM at the molecular level and proposed bimolecular marking methods in the identification of CMM by analyzing DNA markers and metabolic markers (Huang et al, 2015). Using the methods of genomics and metabolomics, they integrated the data from different omics and completed the identification and evaluation of CMM resources. Taking as an example the CMM *Astragali Radix*, *Chinese Pharmacopoeia* (2010 version) describes the dry roots of *Astragalus membranaceus* (Fisch.) Bge. or *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao. as quality products. Due to the similarity in medicinal properties, it is hard to distinguish the original plants. Duan et al used the AFLP molecular marker technique combined with metabolomic studies of GC-TOF/MS to identify *A. membranaceus* and *A. membranaceus* var. *mongholicus* by which they amplified DNA sequence with 11 pairs of primers to obtain 100- to 500-bp fragments. It was found that two bands M40E41-5 and M33E41-2 were amplified only in *A. membranaceus*, and M38E35-3 only in *A. membranaceus*; As a result, the three bands were used as identification markers of these two species (Duan et al, 2012). PCA, OPLS-DA, and other analysis techniques were employed to distinguish 48 plant samples of *Astragalus* L. by GC-TOF/MS, and the analysis of partial correlation coefficients was conducted for 85 AFLP markers and 1193 metabolite peaks; And the results showed a significant correlation between the 21 AFLP markers and 122 metabolite peaks. Two original plants from *Astragalus* L. were successfully identified with three markers and eight metabolic markers. And it was found that with an increase in extensive cultivation areas, the content of soluble sugars, fatty acids, proline, and other compounds varied with the adaption of plants to different habitats, laying the foundation for the later identification and quality evaluation of *Astragali Radix*. (Duan et al, 2012).

In addition, Yuan et al performed the transcriptome analysis of honeysuckle (*Lonicerae Flos*) with the active ingredients as the guiding direction and an established chemical quality evaluation method for honeysuckle based on gene expression profiles. They found that different bathes of honeysuckle had the significant differences in quality, i.e., its active ingredients, because of different germplasm, origins, and floral development stages, could seriously threaten the drug safety. They used chlorogenic acid, luteolin, and other active ingredients of honeysuckle as the guiding directions and separately analyzed the expression profiles of honeysuckle bud, full bloom, and waste flower with the use of de novo transcriptome, constructed biosynthetic pathway networks for honeysuckle phenolic acids, flavonoids, terpenoids, and fatty acids, and screened the key enzyme genes closely related to active ingredient accumulation according to "gene expression-ingredients content" correlation analysis. In addition, they established the quality evaluation methods of honeysuckle based on its gene expression profiles, which could be used in the quality evaluation of the honeysuckle variant, *Lonicera japonica* var. *chinensis* (P. Watson) Baker (Yuan et al, 2012).

3. Analysis of biosynthesis pathways of active ingredients and their regulation of medicinal plants

Secondary metabolites are a class of small molecular compounds that are produced in the process of plant development as well as its adaptation to the environment

(Chen, 2006), many of which are active ingredients of medicinal plants. Analysis of biosynthetic pathways in plants is the core content of medicinal plant secondary metabolite research. The major secondary metabolic pathways have been known, such as shikimic acid pathways for phenols and isoprenoid diphosphate (IPP) and pathways for terpenes; Additionally, flavonoids, lignins, and other relatively conservative secondary metabolic pathways have been well studied (Kang et al, 2014; Chen 2006). However, the synthesis and accumulation processes of many secondary metabolites are also poorly understood. Currently, only a few active compounds, such as paclitaxel, artemisinin, and periwinkle indole alkaloids in medicinal plants, are deeply known in terms of their biosynthetic pathways, and most should be further elucidated in the biosynthesis pathways.

Most plant secondary metabolites are distinct in their distribution to a specific taxonomic group, which is different from the primary metabolites existed in all living plants. Thus they have recently been re-named as "specialized metabolites" (De Luca et al, 2012; Weng et al, 2012). Although the secondary metabolism in plants has species-specificity, there is little application of model plant mutant screening and other new genetic techniques in gene screening and function identification in the biosynthesis of secondary metabolism. The development and application of genomics, transcriptomics, metabolomics, and other omics techniques provide the effective ideas and methods for the comprehensive and systematic analysis of secondary metabolism pathways. Currently, due to a decrease in the sequencing cost, genome or transcriptome sequences of some medicinal plants is more accessible than before, and the reverse genetics method is becoming an important method in plant secondary metabolism studies (Wang et al, 2013).

With the ideas and methods of systems biology, Huang's research team obtained the systematic results in studying the biosynthetic pathway of secondary metabolites tanshinone of *Salvia miltiorrhiza* Bge (Danshen). The hairy roots of *S. miltiorrhiza* produced the phenotypic differences in the tanshinone content by elicitor stimulation. For multiple groups of *S. miltiorrhiza* hairy root with phenotypic differences, the analysis of data was conducted in metabolomics and proteomics, and also on the transcriptome, using gene chips (Cui et al, 2007a; 2007b). By multivariate analysis, a number of gene fragments that are closely related to tanshinone secondary metabolism were screened, and full-length cDNA was obtained. Researchers have cloned *SmCPS* (Gao et al, 2008), *SmKSL* (Gao et al, 2009), and the key P450 gene *CYP76AH1* for the catalytic generation of ferruginol (Guo et al, 2013). They found a new branch of the unique diterpenoid biosynthetic pathway for tanshinone and greatly promoted the analysis of the tanshinone biosynthetic pathway.

With the progress of genomic research, scientists have a new understanding of the regulation of plant secondary metabolites, namely, gene clustering. Gene clusters exist in the distribution of the plant's secondary metabolism genes. In contrast to the operon-like gene clusters of prokaryotes, the functional gene clusters of plants usually encode enzyme proteins in a specific metabolic pathway, which jointly regulates a certain metabolic process of plants (DellaPenna and O'Connor, 2012; Winzer et al, 2012). For example, in a section of 480-kb region for Chromosome 4 of rice (*Oryza sativa* L.), 14 terpene synthase genes are distributed (Qi et al, 2004); In the biosynthesis process of cucumber bitter principle, six key *CYP450* genes form a gene cluster in the 35 kb region of Chromosome 6 (Shang et al, 2014). In addition, the genes in a cluster would be coexpressed and coregulated.

While an understanding of genes and proteins continues to be important, the focus is on understanding the system structures and dynamics of secondary metabolism.

Biosynthesis of medicinal plant secondary metabolites usually occurs during a certain period of time and in specific organs or cells, and thus, it has a strong time-specificity and tissue-specificity. Because most of the secondary metabolism compounds are closely related to the adaptability of plants to the environment, the secondary metabolism is also vulnerable to induction of a variety of biological and non-biological environmental factors, which involves environment-specificity. Judging at the molecular level, the regulation of medicinal plant secondary metabolites can occur at the genomic level, transcriptional level, and post-transcriptional level, and thus, it is a complex process with multi-levels, multi-steps, and multi-way synergy. Systems biology aims to find a connection or specific interactions of life activities at different levels (e.g., gene level, transcription level, protein level, and metabolism level) or among molecules at the same level, thereby establishing complex system models that could display emergent behavior and emergent rules.

Currently, studies in synthesized signal regulation of medicinal plant secondary metabolites, transcription factor regulation, and post-transcriptional regulation of translation levels have made some progress, but most of the studies are still limited to functional studies (Patra et al, 2013; Zhao et al, 2012; Pino Del Carpio et al, 2014).

In recent years, some scholars have begun to focus on the identification of gene-regulatory logic and biochemical networks. Rischer et al combined genomewide transcript profiling by cDNA-amplified fragment-length polymorphisms with metabolic profiling of elicited *Catharanthus roseus* (L.) G. Don cell cultures, which yielded a collection of known and undescribed transcript tags and metabolites that are associated with terpenoid indole alkaloids. Then, they determined gene-to-gene and gene-to-metabolite networks by searching for correlations between the expression profiles of 417 gene tags and the accumulation profiles of 178 metabolite peaks. Additionally, these networks revealed that the different branches of terpenoid indole alkaloid biosynthesis and various other metabolic pathways are subject to differing hormonal regulation (Rischer et al, 2006).

Additionally, Gao et al combined metabolomics and transcriptomics to investigate the inducible biosynthesis of the bioactive diterpenoid tanshinones from *S. miltiorrhiza*. The network analysis indicated a biphasic response of Danshen terpenoid metabolism to elicitation, with early induction of sesqui- and tri-terpenoid biosynthesis, followed

by later and more sustained production of the diterpenoid tanshinones (Gao et al, 2014).

Once we have attained an understanding of the network structure, we will be able to investigate the network dynamics (Kitano, 2002). To unveil how regulatory behaviors on secondary metabolism reshape biological processes, Shi et al constructed and analyzed a dynamic regulatory network of secondary metabolic pathways in *Arabidopsis thaliana* (L.) Heynh. (Lv et al, 2014). They developed a method for constructing a dynamic regulatory network significantly in biological function by integrating regulatory interactions, large-scale microarray data, and evolutionary conservation of TFBSs. This dynamic network is efficient in systematically exploring regulatory rewiring (or crosstalk) on pathways to explain the mechanism of regulation.

Actually, the analysis of dynamics and structure on the basis of network dynamics is an overlapping process, and the dynamic regulatory network can largely improve the understanding of perplexing regulatory rewiring in secondary metabolism; In addition, dynamic analysis could yield the useful predictions of unknown interactions.

For dynamically analyzing a biological system, we must establish a model to obtain an in-depth understanding of the systems behavior or to predict complex behaviors in response to complex stimuli (Kitano, 2002). However, the current regulation research of secondary metabolism for medicinal plants is still focused on a single level, such as gene regulatory networks or transcription factor regulatory networks. Systematic study of dynamic regulatory networks for medicinal plant secondary metabolism has not been reported, and the construction of dynamic network regulatory models for medicinal plant secondary metabolism would still require long-term research.

4. Metabolic engineering and synthetic biology research on medicinal plants

Metabolic engineering mainly achieves its intended purpose by changing metabolic flows or extending metabolic pathways and building new metabolic pathways using the approach of genetic engineering. Metabolic engineering of medicinal plants involves genetic engineering to increase the content of the active ingredients in the medicinal plants, to solve the problem of medicinal plant resources deficiency. If the content can be increased by genetic engineering methods, then great economic and social benefits would be achieved, which would be conducive to the sustainable use of CMM resources, especially some rare and endangered medicinal plant resources.

In recent years, research in medicinal plant metabolic engineering has made great progress, and the effective content of medicinal plant secondary metabolites can be increased by plant transgenic engineering. The main strategies include the over-expression of key genes of the biosynthetic pathway, co-transformation of multi-genes, RNAi inhibiting competitive metabolic bypasses and the regulation of transcriptional factors. For example, Chinese researchers have accomplished *Artemisia annua* L. transgenic research in many aspects and increased the artemisinin content in transgenic *A. annua* by "source opening" and "flow regulating" measures, such as the over-expression of FPS genes in the artemisinin biosynthetic pathway (Wang et al, 2013), over-expression of *CYP71AV1* and *CPR* (Xiang et al, 2012), inhibiting the expression of the squalene synthase gene (Yang et al, 2008), inhibiting the expression of the caryophyllene synthase gene (Chen et al, 2011), and the over-expression of transcription factors in the artemisinin biosynthetic pathway (van der Fits and Memelink, 2000), or by regulating light-signaling pathways (Hong et al, 2009). In addition, Tang's group co-transformed belladonna with *PMT* (1,4-butanediamine-*N*-methyl transferase) genes and *H6H* (hyoscyamine 6-β-hydroxylase) genes, in such a way that the scopolamine content in the transgenic root of *Atropa belladonna* L. (411 mg/L) increased by nine times of the wild type (43 mg/L), greatly increasing the synthesis and accumulation of tropane alkaloids (Zhang et al, 2004).

In addition, the use of signaling pathways and transcription factors as regulatory targets is a new strategy in metabolic engineering, i.e., the modification of a transcription factor that could control multiple biosynthetic genes could more effectively regulate plant secondary metabolism, to improve the accumulation of specific compounds. For example, high expression of transcription factor ORCA3 with the AP2/ERF domain, which is involved in the biosynthetic pathway of periwinkle diterpenoid indole alkaloid, will lead to the over-expression of several genes that are related to diterpenoid indole alkaloid biosynthesis and the accumulation of diterpenoid indole alkaloids (van der Fits and Memelink, 2000). Tang et al produced the over-expression of *AaORA* in *A. annua* and regulated, positively and significantly, the expression levels of *ADS*, *CYP71AV1*, *DBR2*, and *AaERF1*, which resulted in a significant increase or decrease in artemisinin and dihydroartemisinic acid (Lu et al, 2013).

Synthetic biology is based on the intentional design of artificial biological systems using genetic engineering with systems biology and engineering methods, and synthetic biology involves artificial design and synthesis from gene fragments, DNA molecules, gene regulatory networks, and signaling pathways to cells. This approach could achieve artificial designs and build new biological systems with specific physiological functions by clarifying and simulating the basic rules of biosynthesis. In 2000, Kool proposed that synthetic biology is a form of genetic engineering based on systems biology (Kool et al, 2000).

With the successful application of synthetic biology techniques in research and in the production of paclitaxel, artemisinin, tanshinone, and ginsenosides, synthetic biology applications in research on the sustainable utilization of CMM resources have gradually attracted widespread attention. Based on the efficient and directional heterologous biosynthesis of complex and diverse Chinese medicinal active ingredients of CMM in synthetic biology, this approach will effectively solve many problems that are encountered in the research of CMM, and it will also provide a new strategy and

technology for the sustainable use of CMM resources (Huang et al, 2014).

The application of synthetic biology in the sustainable utilization of CMM mainly includes the following processes: first clone the genes in the biosynthetic pathway of the active ingredients from medicinal plant, identify functions one by one, and analyze the biosynthetic pathway of medicinal active ingredients through the study of gene functions; Then, design and integrate the heterologous biosynthesis pathways that refer to botanical approaches, load the artificially designed pathways into basal cell (e.g., *Escherichia coli*, yeast) genomes to build microbial cell factories, and finally, optimize the fermentation conditions to achieve the efficient fermentation and the production of medicinal active ingredients and intermediates (Huang et al, 2014).

Huang et al analyzed tanshinone biosynthetic pathways and obtained heterologous production in synthetic biology. Gao et al first cloned and identified two enzymes (*SmCPS* and *SmKSL*) of its precursor, miltiradiene, by functional genomics methods, and built metabolic pathways in *E. coli*, with a yield of up to 2.5 mg/L (Gao et al, 2009). Then, Zhou et al established module pathway engineering strategies and quickly assembled a miltiradiene biosynthesis pathway in yeast cells; Based on this work, the involved precursor supply, rate-limiting steps, substrate transport, metabolic flux distribution, and other issues were systemically considered through the designed module combination. The genes of encoded *SmCPS*, *SmKSL*, farnesyl pyrophosphate synthase, *GGPP* synthase, and mevalonate reductase were operated. The best engineering strains were cultured in 15 L fermenter, with a yield of miltiradiene up to 365 mg/L (Zhou et al, 2012). They successfully designed and developed a combination of functional modules for the regulation and control of *Saccharomyces cerevisiae* terpenes biosynthesis, with a yield of miltiradiene of up to 488 mg/L (Dai et al, 2012). Subsequently, Guo et al used comparative transcriptomics and RNA sequencing to identify 14 *CYP450* genes that were related to miltiradiene synthesis; They found that *CYP76AH1* could catalytically convert miltiradiene for the synthesis of ferruginol by establishing an *in vitro* catalytic reaction, and then, they used the module pathway engineering strategy to integrate *CYP76AH1* and *CYP* reductase *SmCPR1* into miltiradiene for the biosynthesis of yeast. Engineering the strains can synthesize 10.5 mg/L ferruginol, and the synthesis of ferruginol was found to have a negative correlation with the accumulation of miltiradiene, laying the foundation for further resolving the downstream tanshinone biosynthetic pathways and the microbial synthesis of tanshinone (Guo et al, 2013).

In the synthetic biology research of ginsenosides, Chinese scholars successfully constructed the proto-panaxadiol biosynthetic pathways in *S. cerevisiae* and found that squalene epoxidase plays a key role in the control of triterpenoids biosynthesis; Based on this finding, the yield of protopanaxadiol increased by 262 times by improving the activity of 3-hydroxy-3-methylglutaryl coenzyme A reductase, farnesene pyrophosphate synthase, squalene synthase, and squalene epoxidase; Finally, the yield of protopanaxadiol increased to 1189 mg/L by optimization with a biphasic fermentation process (Dai et al, 2012). Subsequently, Dai et al imported other plants on the basis of this yeast cell: β-amyrin synthase, oleanolic acid synthase, dammarenediol-II synthase, protopanaxadiol synthase, protopanaxatriol synthase, and NADPH-cytochrome P450 reductase. The obtained strains can produce three types of aglycons that are contained in ginseng cells, 17.2 mg/L protopanaxadiol, 15.9 mg/L protopanaxatriol, and 21.4 mg/L oleanolic acid (Dai et al, 2014). Those studies showed that the complete synthesis of aglycone ginsenosides can be achieved in yeast, indicating that in addition to be extracted from plants, ginsenosides can be obtained by the biosynthetic method in the future.

5. Design breeding for medicinal plants

Because human demand for herbal medicine is increasing, wild medicinal plants are sharply reduced in number or are endangered, and a large part of the CMM resources are being transferred from wild into cultivated. Because the cultivation of Chinese herbal medicines has special requirements in terms of the environment, climate, soil, and other factors, quantitative or qualitative changes occur in the quality of many types of Chinese herbal medicines after they are transformed from the wild to a cultivation environment. There is an urgent need for the development of medicinal plant breeding to meet the needs for national GAP base construction in China and the diverse quality needs of the medicinal plants. Plant molecular breeding includes genetic engineering breeding, marker-assisted breeding, and molecular design breeding. Molecular design breeding is a systemic study that is based on molecular marker-assisted breeding, combined with computer science, bioinformatics, crop genetics, biostatistics, and other sciences. In molecular design breeding, the "design" concept in systems biology is used for research on medicinal plant breeding, as an important aspect in the research of sustainable utilization of CMM resources. Molecular design breeding of medicinal plants can solve the deficiency, objective ambiguity, and other problems in the genetic breeding process and control the quality of the medicine from the source based on the systemic research in the metabolic pathways of active compounds under different external environment influences and related metabolic networks.

Molecular design breeding should be conducted in three steps (Podlich et al, 1999). First, to locate the quantitative trait locus (QTLs) that are related to agronomic traits, second, to evaluate the allelic variation of these sites, and third, to conduct the design breeding. Currently, the study of molecular design breeding is mainly used in crops, which has just started in medicinal plants. QTL mapping, gene fine mapping, and cloning research have been conducted for a few medicinal plants, such as *A. annua* and mangosteen, to obtain a yield of active ingredients. This approach has laid a good foundation for the implementation of design breeding for medicinal plants. In recent years, domestic and foreign scientists have also made good progress in molecular design breeding for medicinal plants.

Graham et al (2010) reported a genetic map of *A. annua* and identified the key loci that could improve agricultural yields, decrease production costs, ensure a steady global supply of the drug, and improve the grower confidence in the crop. They used the deep sequencing of the plant transcriptome to successfully identify genes and markers, which will facilitate the crossing of highly productive varieties. Then, they used a pure line to establish the first genetic linkage and QTL maps for the plant species, and they subsequently validated a positive QTL for the artemisinin yield.

On the other hand, Liu et al (2011) used "Yehong No.1" as the female parent and *Changtan* fruit as the male parent to construct 150 F1 filial generations as the mapping groups, and they used ISSR and SRAP techniques for the initial construction of a genetic linkage map of Luohanguo [*Siraitia grosvenorii* (Swingle) C. Jeffrey ex A. M. Lu & Zhi Y. Zhang], laying a solid foundation for further building high-density molecular marker linkage maps and performing research on gene cloning and molecular marker-assisted breeding. This study could contribute greatly to the theory and practice of the genetic breeding for Luohanguo, shorten the breeding process, improve breeding efficiency, and provide a valuable scientific basis for improving Luohanguo.

6. Discussion and conclusions

For the world's sustainable medicines and global health care, the medicinal plant resources are represented the foundation of primary health care. Contemporary harvesting methods for medicinal plants are severely depleting these critical indigenous resources. However, maintaining and enhancing the availability of quality medicinal agents on a sustainable basis is an unappreciated public health care concept. However, since biogenesis of medicinal plants is quite complex, the production and accumulation of the main active ingredients (secondary metabolites) are influenced by various biotic and abiotic factors either from gene or environments, the complexity and diversity may affect quality control of crude drugs and utilization of the active ingredients.

Systems biology is the computational and mathematical modeling of complex biological systems. An emerging engineering approach applied to biomedical and biological scientific research, systems biology is a biology-based inter-disciplinary field of study that focuses on complex interactions within biological systems, using a holistic approach to biological and biomedical research. The systems biology approach often involves the development of mechanistic models such as the reconstruction of dynamic systems from the quantitative properties of their elementary building blocks. A cellular network can be modeled mathematically using methods coming from chemical kinetics and control theory. Due to the large number of parameters, variables and constraints in cellular networks, numerical and computational techniques are often used.

Update, the ideas and methods of systems biology have been widely used in the field of sustainable utilization of TCM, including the identification and comprehensive evaluation of CMM resources, biosynthetic pathways of active ingredients for medicinal plants and their metabolic regulation system models, biosynthesis of medicinal active ingredients, and molecular design breeding for medicinal plants. Systems biology is characterized by integrity research, which has much in common with the theories of TCM. The combination of these two approaches provides CMM opportunities and challenges to re-examine and develop its own characteristics.

Scientists have been attached great importance to the development of predictable metabolic engineering. Various existing databases and instrumental analysis methods have made such systems analysis possible to a certain extent. However, the rational engineering of complicated metabolic networks involved in the production of biologically active plant compounds has been greatly impeded by our poor understanding of the regulatory and metabolic pathways that underlie the biosynthesis of these compounds.

As shown in Figure 2, the sustainable utilization of CMM resources research based on systems biology is significant for the identification and evaluation of CMM resources, analysis of biosynthesis, their regulation of active ingredients, metabolic engineering and synthetic biology research, and molecular breeding of medicinal plants. Through CMM Resources based on systems biology, the formed CMM resource Data-bank will be beneficial to improve and control the quality of medicinal resources, but also conducive to the variation of active ingredients for CMM regulatory. Therefore, to carry out this research, not only conducive to the sustainable development and utilization of CMM resources, but also beneficial to ensure the safety and effectiveness of the CMM.

Identification and evaluation of CMM resources have developed from a macroscopic overall subjective judgment to a microscopic accurate detection of single component. The overall concept of systems biology would realize the integration of the medicinal plants at all levels from molecules (nucleic acids, proteins, and small molecules), to cells, tissues, and individuals. More specifically, the whole is greater than the sum of its parts, and thus, this concept is used in the establishment of a comprehensive system for the identification and evaluation of CMM resources.

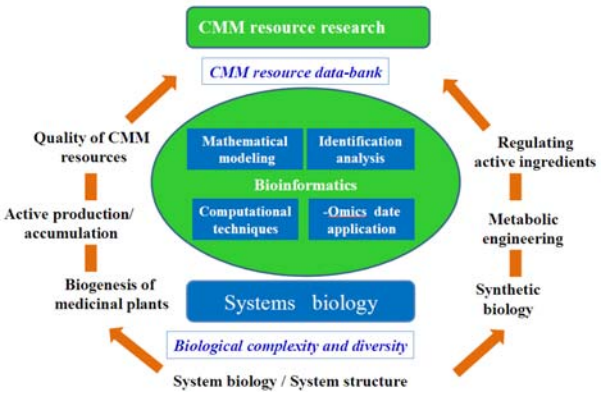

Figure 2 CMM resource research pathway based on systems biology

Methods of systems biology can be used to provide an overall understanding of the biological process of secondary metabolism of plants, to construct metabolic regulatory network systems for signal transduction regulatory levels and gene-regulatory levels of medicinal plant secondary metabolism, enzyme-regulatory levels, and metabolic flux regulatory levels, to fully reveal the biosynthetic pathways of secondary metabolites, to further operate any specific plant genes of the plant or interfere with the environment by building models, to accurately predict the functions of plants, to achieve the optimization of medicinal plant resources and their sustainable use.

In recent years, China has made great progress in the interpretation of biosynthetic pathways for tanshinone and ginsenoside as well as their biological research. In addition, unearthing the gene elements that are related to the active ingredients of CMM and the analysis of biosynthetic pathways will be the primary tasks of the current research in the synthetic biology of CMM. It can be expected that with the gradual maturity of synthetic biology and the deepening of its applications in traditional CMM area, synthetic biology will also greatly promote the development of the CMM area and become one of the important ways to develop the sustainable utilization of CMM resources.

As time is going on, systemic and dynamic concept of systems biology, and the regulation and design methods still have broad application prospects in the research field of CMM resources, with the significance in the effective protection of its diversity during rational development and utilization of CMM resources.

In summary, the production of Chinese herbal medicines is a major event in relation to people's livelihoods, and medicinal plant breeding is a prerequisite for the development of Chinese herbal medicines, while the booming development of systems biology has brought new opportunities for the medicinal plant breeding. This review recalled the progress of systems biology applied to research on sustainable utilization of CMM resources in the following aspects: identification and evaluation of CMM resources, analysis of biosynthesis and their regulation of active ingredients in medicinal plants, metabolic engineering and synthetic biology research of medicinal plants, and molecular breeding of medicinal plants. Development of systems biology is currently leading to extremely broad applications in the field of CMM resources, and systems biology will become a significant approach for the sustainable utilization of CMM resources.

References

Aderem A, 2005. Systems biology: Its practice and challenges. *Cell* 121(4): 511-513.

Chen JL, Fang HM, Ji YP, Pu GB, Guo YW, Huang LL, Du ZG, Liu BY, Ye HC, Li GF, Wang H, 2011. Artemisinin biosynthesis enhancement in transgenic *Artemisia annua* plants by downregulation of the beta-caryophyllene synthase gene. *Planta Med* 77(15): 1759-1765.

Chen SL, Pang X, Song JY, Shi L, Yao H, Han J, Leon C, 2014. A renaissance in herbal medicine identification: From morphology to DNA. *Biotechnol Adv* 32(7): 1237-1244.

Chen XY, 2006. Plant secondary metabolism. *World Sci-Tech R&D* 28(5): 1-4.

Cui G, Huang L, Qu D, Yuan Y, Fu G, 2007a. Functional genomics studies of *Salvia miltiorrhiza*. Gene expression profiling of different stage of hairy root. *China J Chin Mater Med* 32(13): 1267-1232.

Cui G, Huang L, Tang X, Qu D, Wang X, Fu G, 2007b. Functional genomics studies of *Salvia miltiorrhiza*. Establish cDNA microarray of *S. imiltiorrhiza*. *China J Chin Mater Med* 32(12): 1137-1141.

Dai Z, Liu Y, Huang L, Zhang X, 2012. Production of miltiradiene by metabolically engineered *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 109(11): 2845-2853.

Dai Z, Wang B, Liu Y, Shi M, Wang D, Zhang X, Liu T, Huang L, Zhang X, 2014. Producing aglycons of ginsenosides in bakers' yeast. *Sci Rep* 4: 3698.

De Luca V, Salim V, Atsumi SM, Yu F, 2012. Mining the biodiversity of plants: A revolution in the making. *Science* 336(6089): 1658-1661.

DellaPenna D, O'Connor SE, 2012. Plant science. Plant gene clusters and opiates. *Science* 336(6089): 1648-1649.

Duan LX, Chen TL, Li M, Chen M, Zhou YQ, Cui GH, Zhao AH, Jia W, Huang LQ, Qi X, 2012. Use of the metabolomics approach to characterize Chinese medicinal material Huangqi. *Molecular Plant*. 5(2): 376–386.

Gao W, Cui G, Kong J, Chen K, Wang W, Yang Y, Huang L, 2008. Optimizing expression and purification of recombinant *Salvia miltiorrhiza* copalyldiphosphate synthase protein in *E. coli* and preparation of rabbit antiserum against SmCPS. *Acta Pharm Sin* 43(7): 766-772.

Gao W, Hillwig ML, Huang L, Cui G, Wang X, Kong J, Yang B, Peters RJ, 2009. A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. *Org Lett* 11(22): 5170-5173.

Gao W, Sun HX, Xiao H, Cui G, Hillwig ML, Jackson A, Wang X, Shen Y, Zhao N, Zhang L, Wang XJ, Peters RJ, Huang L, 2014. Combining metabolomics and transcriptomics to characterize tanshinone biosynthesis in *Salvia miltiorrhiza*. *BMC Genomics* 15: 73.

Graham IA, Besser K, Blumer S, Branigan CA, Czechowski T, Elias L, Guterman I, Harvey D, Isaac PG, Khan AM, Larson TR, Li Y, Pawson T, Penfield T, Rae AM, Rathbone DA, Reid S, Ross J, Smallwood MF, Segura V, Townsend T, Vyas D, Winzer T, Bowles D, 2010. The genetic map of *Artemisia annua* L. identifies loci affecting yield of the antimalarial drug artemisinin. *Science* 327(5963): 328-331.

Guo J, Zhou YJ, Hillwig ML, Shen Y, Yang L, Wang Y, Zhang X, Liu W, Peters RJ, Chen X, Zhao ZK, Huang L, 2013. CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. *Proc Natl Acad Sci USA* 110(29): 12108-12113.

Hong GJ, Hu WL, Li JX, Chen XY, Wang LJ, 2009. Increased accumulation of artemisinin and anthocyanins in *Artemisia annua* expressing the Arabidopsis blue light receptor CRY1. *Plant Mol Biol Rep* 27(3): 334-341.

Huang L, Gao W, Zhou Y, 2014. Application of synthetic biology to sustainable utilization of Chinese materia medica resources. *Acta Pharm Sin* 49(1): 37-43.

Huang L, Qian D, Deng C, 2015. Hypothesis and application of bimolecular marking methods in Chinese materia medica. *China J Chin Mater Med* 40(1): 15-18.

Kang YL, Pei J, Cai WL, Liu W, Luo J, Wu QH, 2014. Research progress on flavonoid metabolic synthesis pathway and related

function genes in medicinal plants. *Chin Tradit Herb Drugs* 45(9): 1336-1341.

Kitano H, 2001. *Foundations of systems biology*. MIT press, Cambridge.

Kitano H, 2002. Systems biology: A brief overview. *Science* 295(5560): 1662-1664.

Kool ET, Morales JC, Guckian KM, 2000. Mimicking the structure and function of DNA: Insights into DNA stability and replication. *Angew Chem Int Edit* 39(6): 990-1009.

Liu C, 2006. Systems biology and modern research of traditional Chinese medicines. *J Tianjin Univ Trad Chin Med* 25(3): 115-118.

Liu L, Ma X, Wei J, Qin J, Mo C, 2011. The first genetic linkage map of Luohanguo (*Siraitia grosvenorii*) based on ISSR and SRAP markers. *Genome* 54(1): 19-25.

Lu X, Zhang L, Zhang F, Jiang W, Shen Q, Zhang L, Lv Z, Wang G, Tang K, 2013. AaORA, a trichome-specific AP2/ERF transcription factor of *Artemisia annua*, is a positive regulator in the artemisinin biosynthetic pathway and in disease resistance to *Botrytis cinerea*. *New Phytol* 198(4): 1191-1202.

Lv Q, Cheng R, Shi T, 2014. Regulatory network rewiring for secondary metabolism in *Arabidopsis thaliana* under various conditions. *BMC Plant Biol* 14: 180.

Patra B, Schluttenhofer C, Wu YM, Pattanaik S, Yuan L, 2013. Transcriptional regulation of secondary metabolite biosynthesis in plants. *BBA-Gene Regul Mech* 1829(11): 1236-1247.

Pino Del Carpio D, Basnet RK, Arends D, Lin K, De Vos RC, Muth D, Kodde J, Boutilier K, Bucher J, Wang X, Jansen R, Bonnema G, 2014. Regulatory network of secondary metabolism in *Brassica rapa*: Insight into the glucosinolate pathway. *PLoS One* 9(9): e107123.

Podlich D, Cooper M, Basford K, Geiger H, 1999. Computer simulation of a selection strategy to accommodate genotype environment interactions in a wheat recurrent selection programme. *Plant Breeding* 118(1): 17-28.

Qi X, Bakht S, Leggett M, Maxwell C, Melton R, Osbourn A, 2004. A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. *Proc Natl Acad Sci USA* 101(21): 8233-8238.

Rischer H, Oresic M, Seppanen-Laakso T, Katajamaa M, Lammertyn F, Ardiles-Diaz W, Van Montagu MC, Inze D, Oksman-Caldentey KM, Goossens A, 2006. Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells. *Proc Natl Acad Sci USA* 103(14): 5614-5619.

Shang Y, Ma Y, Zhou Y, Zhang H, Duan L, Chen H, Zeng J, Zhou Q, Wang S, Gu W, Liu M, Ren J, Gu X, Zhang S, Wang Y, Yasukawa K, Bouwmeester, HJ, Qi X, Zhang Z, Lucas WJ, Huang S, 2014. Plant science. Biosynthesis, regulation, and domestication of bitterness in cucumber. *Science* 346(6213): 1084-1088.

van der Fits L, Memelink J, 2000. ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Science* 289(5477): 295-297.

Wang L, Fang X, Yang C, Li J, Chen X, 2013. Biosynthesis and regulation of secondary terpenoid metabolism in plants. *Sci Sin Vitae* 43(12): 1030.

Wang Y, Xu A, 2014. *Zheng*: A systems biology approach to diagnosis and treatments. *Science* 346(6216 suppl): S13-S15.

Weng JK, Philippe RN, Noel JP, 2012. The rise of chemodiversity in plants. *Science* 336(6089): 1667-1670.

Winzer T, Gazda V, He Z, Kaminski F, Kern M, Larson TR, Li Y, Meade F, Teodor R, Vaistij FE, Walker C, Bowser TA, Graham IA, 2012. A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine. *Science* 336(6089): 1704-1708.

Xiang L, Zeng LX, Yuan Y, Chen M, Wang F, Liu XQ, Zeng LJ, Lan XZ, Liao ZH, 2012. Enhancement of artemisinin biosynthesis by overexpressing dxr, cyp71av1 and cpr in the plants of *Artemisia annua* L. *Plant Omics* 5(6): 503-507.

Yuan Y, Song L, Li M, Liu G, Chu Y, Ma L, Zhou Y, Wang X, Gao W, Qin S, Yu J, Wang X, Huang L, 2012. Genetic variation and metabolic pathway intricacy govern the active compound content and quality of the Chinese medicinal plant *Lonicera japonica* Thunb. *BMC Genomics* 13: 195.

Zhang L, Ding R, Chai Y, Bonfill M, Moyanno E, Okasman-Caldentey KM, Xu T, Pi Y, Wang Z, Zhang H, Kai G, Liao Z, Sun X, Tang K, 2004. Engineering tropane biosynthetic pathway in Hyoscyamus niger hairy root cultures. *Proc Natl Acad Sci USA*, 101(17):6786-6791.

Zhao HW, Ge F, Sun Y, Liu DQ, Chen CY, 2012. Transcription factors involved in plant terpenoid biosynthesis and their application prospect. *Chin Tradit Herb Drugs* 43(12):2512-2518.

Zhou YJ, Gao W, Rong Q, Jin G, Chu H, Liu W, Yang W, Zhu Z, Li G, Zhu G, 2012. Modular pathway engineering of diterpenoid synthases and the mevalonic acid pathway for miltiradiene production. *J Am Chem Soc* 134(6): 3234-3241.