Lu-qi Huang *Editor*

Molecular Pharmacognosy

Springer

Molecular Pharmacognosy

Lu-qi Huang
Editor

Molecular Pharmacognosy

Springer

Editor
Lu-qi Huang
Institute of Chinese Materia Medica
China Academy of Chinese Medical Sciences
Dongzhimeneii, Beijing, China

ISBN 978-94-007-4944-3 ISBN 978-94-007-4945-0 (eBook)
DOI 10.1007/978-94-007-4945-0
Springer Dordrecht Heidelberg New York London

Jointly published with Shanghai Scientific and Technical Publishers
ISBN: 978-7-5478-1379-9 Shanghai Scientific and Technical Publishers

Library of Congress Control Number: 2012948639

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

With its development over almost 200 years, pharmacognosy has existed as one applied science with comparably impeccable theory and technology. With the passage of time, it also faces various problems which cannot be solved by current technology and methodology, e.g., the exact identification of the "species" level of medicinal plants without flowers and fruits, sustainable uses of rare and endangered medicinal resources, and directional control of the qualities of medical materials. All of these greatly need to be settled by introducing new technology and methodology.

Ever since the discovery of DNA's double helix structure and its semiconservative replication, the molecular biology technology has almost permeated into all the fields of life science, thus generating quite a number of interdisciplinnary subjects. It can be clearly anticipated that molecular biology will still lead the development in life science in the twenty-first century. Without keeping out of the affair, pharmacognosy will combine with molecular biology and bring about new areas in its studies.

During the collision between pharmacognosy and molecular biology, how can we grasp and choose the binding point of the two? What is the theoretical basis for their binding? This is the problem I often pondered upon during my graduate studies. After continuous studies and thoughts as well as discussions with my teachers and classmates, I put forward the concept of "molecular pharmacognosy" in the article *Anticipation on the Application of Molecular Biological Technology in Pharmacognosy* published in 1995. I had never thought that its publication would arouse such strong resonance among so many scholars. They raised their own ideas, which encouraged me to do further researches. Along with 5-year research quest and practice, the rudiment of a new science, "molecular pharmacognosy," came into being with its own theory and technology.

In the past 15 years, owing to my research work and other researchers at home and abroad, the technology of molecular biology has been widely applied and practiced in relevant fields of pharmacognosy. Molecular pharmacognosy has been further developed theoretically and systematically. Up to now, the Chinese edition of molecular pharmacognosy has been copied twice with new research contents added in each edition, thus making the system of molecular pharmacognosy more abundant and more complete gradually. Besides, the course of molecular pharmacognosy has

been opened up in Beijing University of Chinese Medicine, West China University of Medical Sciences, Fudan University, Huazhong University of Science and Technology, and other institutions of high learning. The innovative teaching material of molecular pharmacognosy was published in 2008. All of these sped up the advancement of the cultivation of research talents in molecular pharmacognosy and the subject's development.

Along with deeper researches of molecular pharmacognosy and more perfection of the subject's theoretical system, there are more researches in the field worldwide. Internationalization of molecular pharmacognosy, a new subject, is on schedule. Under this circumstance, we compiled this internationalized edition of molecular pharmacognosy in cooperation with experts and scholars in this field from Mainland China, Hong Kong, Taiwan, South Korea, and Japan in the hope for furthering the internationalized development of molecular pharmacognosy.

Initialized in the middle of 2010, this book is divided into nine chapters. The first chapter, "Emerging molecular pharmacognosy," mainly introduces the historical background, concepts, and research contents of molecular pharmacognosy as well as its relations with other subjects. "Methodology," being the second chapter, from the perspective of molecular pharmacognosy, introduces some common methods of molecular pharmacognosy researches and part of the new research thinking and methods, e.g., "Ingredient difference phenotypic cloning." Furthermore, some practical problems solved by molecular pharmacognosy have been discussed from the aspect of methodological application. The third chapter is "Molecular Identification of traditional medicinal materials." It mainly discusses the common methods to identify the Chinese medical molecules, and then, there is some exploration of molecular identification of the family, genus, inter-species, and under-species of original Chinese medical plants. The fourth chapter is "The mechanism of formation of Dao-di Herbs," with the specialty of Chinese herb Dao-di and theoretical hypothesis that have been formed as the main contents. The fifth chapter is "Seeking for new members of origin materials (a new usage of plant species) for CMM," centered on the introduction of the theory and methodology of pharmaphylogeny. The sixth chapter is "Gene modification of pharmic plants germplasm resources" with the main contents of the concept, classification, collection, and identification of pharmic plants' germplasm resources. The eighth chapter is "Regulation of the 'active constituents' production of medicinal plants." It mainly talks about the regulation of the active constituents' production of medicinal plants in biotransformation, generic engineering, and other methods. The ninth chapter is "Molecular mechanism and regulation on biosynthesis of active ingredients of medicinal plants" with the main contents of research overview, significance, as well as the basic strategies and ways of the biosynthesis of active ingredients of medicinal plants.

The publication of this book is intended to explicate the concept, theory, and methodology of molecular pharmacognosy and encourage more people in the same pursuit to solve more new problems of pharmacognosy with fully utilizing the developing technology and methodology of biology within the framework of molecular pharmacognosy.

It should be noted that compilation of this book is the fruit of all the international experts in the field of pharmacognosy. Considering the molecular pharmacognosy is a new inter-discipline to be further improved, we sincerely hope that more professionals will raise more advice and support to improve the maturity and development of molecular pharmacognosy.

March, 2012 Lu-qi Huang

Contents

Contributors

Paul Pui-Hay But School of Life Sciences, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, P.R. China

Wen-yuan Gao Tianjin Key Laboratory for Modern Drug Delivery & High-Efficiency, School of Pharmaceutical Science and Technology, Tianjing University, Tianjin, China

Lan-ping Guo Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China

Lu-qi Huang Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China

Kenji Kondo Botanical Raw Material Research Dept. (Molecular identification of BRM), Tsumura & Co., Ibaraki, Japan

Ming Li State Key Laboratory of Phytochemistry and Plant Resources in West China (CUHK), Institute of Chinese Medicine, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, P.R. China

Tsai-yun Lin Department of Life Science, Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, Taiwan

Chao-yi Ma Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China

Kee-yoeup Paek Department of Horticultural Science, Research Center for the Development of Advanced Horticultural Technology (Horttech), Chungbuk National University, Cheongju, Republic of Korea

Pang-Chui Shaw State Key Laboratory of Phytochemistry and Plant Resources in West China (CUHK), Science Centre, Institute of Chinese Medicine, and School of Life Sciences, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, China

Xue-yong Wang College of Chinese Materia Medica, Beijing university of Chinese Medicine, Beijing, China

Rong-min Yu Biotechnological Institute of Chinese Materia Medica, Jinan University, Guangzhou, China

Wei Gao School of Traditional Chinese Medicine, Capital Medical University, Beijing, China

Chapter 1
Emerging Molecular Pharmacognosy

Lu-qi Huang

Abstract Molecular pharmacognosy is a science dealing with crude drugs, which refer to either a kind of new products that originate in plants, animals, and minerals or natural medicinal materials that are directly used for medical care or as raw materials for medicine after a simple processing. As a subject, pharmacognosy has gone through four stages of development: pharmacognosy in ancient times, pharmacognosy in early modern times, pharmacognosy in modern times, and period of natural pharmacognosy. Molecular pharmacognosy has been given the following main tasks: systematic assortment of varieties of Chinese herbs and study of quality standardization, conservation of medicinal plant and animal biodiversity and research of sustainable utilization of crude drugs resources, medicinal plant marker breeding and new variety cultivation, gene regulation of metabolic pathway and directional control of the quality of Chinese herbal medicines, the use of genetic engineering and tissue culture technique to achieve high-level expression and production of natural active ingredients or genetically modified ingredients, genetic engineering, and green pollution-free medicinal plant.

1.1 Historical Development of Pharmacognosy

1.1.1 Pharmacognosy in Ancient Times (Before the Nineteenth Century)

The medicines originated in Egypt and India. Medicines were recorded both in *papyrus* of Egypt about 1,500 BC and later in *Ajur veda* of India. In *papyrus*, crocus, dried ox-bile juice, castor oil, and so on were mentioned. In about 77 AD, Dioscorides,

Lu-qi Huang (✉)
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China
e-mail: huangluqi@263.net

Lu-qi Huang (ed.), *Molecular Pharmacognosy*,
DOI 10.1007/978-94-007-4945-0_1,

a Greek doctor, kept a record of about 600 kinds of crude drugs in his compiled book *De Materia Medica*, a book that had played an important role in pharmacology and botany by the fifteenth century. Ancient Rome also promoted its development. The *Historia*, written by Pliny (23–79 AD), gave a brief account of nearly 1,000 species of plants, most of which could be used for medicines. Wild herbs were used to treat illnesses in the Soviet Union before the eleventh century.

From ancient times to the middle of the nineteenth century, pharmacology had been in its traditional stage for all countries in the world. At that time, knowledge about medicines came mainly from senses and practical experiences, and the major contents for any medicinal book were mostly about medical effectiveness and partly about its name, origin, morphology, and identifying sensory characteristics. Due to underdeveloped science, people could hardly know all the details about medicines. Besides, it was hard for people to reach an agreement on them due to differences in location and personal experiences.

1.1.2 Pharmacognosy in Early Modern Times (1815–1930)

It was at the beginning of the nineteenth century that pharmacognosy came into its real being. In 1815, C.A. Seydler, a German who used the word "pharmakognosie" in his book named *Analecta Pharmacognostica*, was referred to as the father of pharmacognosy. Pharmakognosie etymologically means knowledge about crude drugs. In 1825, Martius, a German scholar, set up "pharmakognosie" as a subject to study in college, and then a new discipline named pharmakognosie emerged in natural science. According to Martius, pharmacognosy, as a part of merchandizing, was a study to research drug base taken from nature to test its purity and to check impurities or adulterants. From then on, German scholars published works, named pharmakognosie, on plant and animal drugs successively. In 1880, Gendō Oi, a Japanese scholar, translated pharmakognosie as pharmacognosy. Japanese early study of pharmacy was based on quoting and researching Chinese herbs.

In 1806, Serturner, a German, clarified the cell to be the basic unit of plant structure, and then microscopes were used to research the internal structure of crude drugs. In 1857, Schleiden published a book named *Grundriss der Pharmakognosie des Pflanzenreiches* (foundation of pharmacognosy of plants), in which he gave a detailed description of microstructure of a great many plant crude drugs. Later, Berg in 1865 and Vogl in 1887 published anatomical atlas of crude drugs successively, which furthered the development of identifying crude drugs by the use of microscopes, and then this method became the most important one to identify crude drugs. Meanwhile, the qualitative and quantitative methods of chemistry were used in crude drug identification. Fluorescence analysis and chromatography were used in sequence in the latter half of the nineteenth century and in the beginning of the twentieth century, both of which enriched the research field of pharmacognosy and promoted a greater development than the early method of identifying crude drugs by shape and smell.

All in all, the reason why pharmacognosy has become an independent subject is closely related to the development of international traffic and trade at that time. In the first half of the nineteenth century, the rapid progress in international trade gave rise to an increase in variety of medicines and resulted in enlarging scope of raw materials and medium products. Crude drugs were sold after being broken down or crushed into powder. In order to hunt for profits, some merchants took advantage of the difficulty in identifying powder and usually mixed drugs at low prices with those at high prices, and even taking fake products as fine products. Therefore, the problem of identifying the authenticity and quality of crude drugs arose. With advancement in bioscience and widespread application of microscope, pharmacognosy made great progress in the middle of the nineteenth century and finally became an independent subject. The early work of pharmacognosy was establishing quality standards for crude drugs in business.

1.1.3 Pharmacognosy in Modern Times (1930 to the Late 1990s)

Since the 1930s, development of biology and chemistry enriched methods and ways of studying goods pharmacognosy. The development of the bioassay of intensity of drug action (biological potency) advanced the study of active ingredients of crude drugs and strengthened quality evaluation of them. Chemical and physical methods, such as colorimetry, spectrophotometry, and fluorescence analysis, were all applied to identification of crude drugs gradually. When pharmacognosy developed in lines of morphology and chemistry, many new disciplines emerged. For example, with accumulation of chemical composition of plants in type and number, a new subdiscipline – plant chemotaxonomy, came into being through exploration of chemical composition of plants and their genetic relationship. The new discipline not only had taxonomy significance but also pushed for new sources of crude medicines.

1.1.4 Period of Natural Pharmacognosy (At the End of the Nineteenth Century to the Early Twenty-First Century)

Going through the first three stages of development, pharmacognosy became an established applied discipline with advanced technology and theory. In 1970s and 1980s, many universities canceled the "pharmacognosy" course from their curriculum. However, at the end of this century, with humans "returning to nature" and uprising of modern life science, pharmacognosy has presented a strong vitality and broad prospects. Progresses in separation of the chemical composition, structure determination, and quantitative technology made ^{1}H NMR, ^{13}C NMR, DNA

fingerprint identification, etc. possible to be used in the identification of crude drugs thus promoting its standardization and normalization.

Through constant exploration, a batch of new developing points was made sure. In 1995, Huang Luqi first mentioned the concept of molecular pharmacognosy in his paper *Prospect of Application of Molecular Biology Technology to Pharmacognosy*. This paper aroused a strong resonance in the circle. People with similar interests took whatever means to put forward their ideas and encouraged him to further the exploration and systematize the theory. Based on the encouragement he received, he worked with great effort for many years and published *Molecular Pharmacognosy* in Peking University Medical Press in June 2000. This book made the original pharmacognosy give rise to the birth of a new branch – molecular pharmacognosy.

1.2 Concept of Molecular Pharmacognosy and Its Development

1.2.1 Generation of Molecular Pharmacognosy

Watson and Crick's discovery of DNA structure in 1953, marking a new era for life science, changed life medicine and thinking ways of scholars in its relevant fields a lot, and since then people began to re-understand life's nature and laws from the level of biological macromolecules. Although DNA's discovery and determination bear no stamp on the development of pharmacognosy, they still had an immeasurable impact on the whole life science. Molecular biology developed rapidly and penetrated into the application fields of biomedicine, which brought into existence a large number of interdisciplinary sciences, frontier disciplines. Genetic engineering technology based on molecular cloning and reorganization rose, and the related tissue culture technology, especially molecular marker technology based on PCR, sprung up like mushrooms, all of which contributed to fast development of pharmacognosy and a full extension and enrichment of pharmacognosy research areas and methods. Confliction and mutual integration between pharmacognosy and molecular biology brought into forth a new interdisciplinary called molecular pharmacognosy.

Molecular pharmacognosy had the following three theoretical bases. First, the development of molecular biology brought all branches related to biology into a molecular level. Pharmacognosy, with a major concern of plant and animal crude drugs, touched upon many biological theories and methods, so there is no exception to pharmacognosy. Secondly, crude drugs originated in animals and plants, whose cells contained DNA, the material basis to store, duplicate, and transmit genetic information. DNA was also the material base of molecular biology, so pharmacognosy had the combined material base – DNA with molecular biology. This made molecular biology theoretically and methodologically applied to pharmacognosy. Thirdly, the study level of crude drugs of animals and plants in pharmacognosy

developed from organism, tissue, organ, and cell into genetics. Therefore, the advancement of modernization of pharmacognosy had surely a close relation to molecular biology, thus inevitably promoting the study of pharmacognosy into a molecular level.

1.2.2 Concept of Molecular Pharmacognosy

Molecular pharmacognosy is a science dealing with study of classification, identification, cultivation, and protection of crude drugs and production of effective element at molecular level. Based on theories and methods of pharmacognosy and molecular biology, molecular pharmacognosy is a promising and prospective branch in pharmacognosy. It can be said that molecular pharmacognosy carries on traditional contents and mission of pharmacognosy and endows pharmacognosy with new tasks and challenges as well.

Discussed from sources of medicines, crude drugs include herbal, animal, and mineral drugs in the broad sense, while in the narrow sense mainly herbal and animal drugs. The study objects of molecular pharmacognosy are limited to the narrow sense. According to different organism levels and gradual combination relation, it can be divided into six major biological levels: gene, cell, organ, organism, population, and community. This biological mode based on multiple levels is called biological spectrum. Each unique level in the biological spectrum is discovered with a historical process. Generally speaking, developing the level of organism in micro- and macro-direction brings into gradual discovery of all levels, as it did in the case of the discovery of cells and genes. Study of the unique scientific questions at each level gives rise to independent branching in life science. The branches such as molecular biology and cytobiology could be mutually promotive but never could be substitutes for each other. These days, pharmacognosy shows a major concern for tissue, organ, organism, and population, upon which relatively mature and independent theory and methods, such as pharmacognosy, histology, and morphology, are brought forth. Molecular pharmacognosy deals with crude drugs at genetic level with a theoretical and methodological basis on gene-level branches – molecular biology.

1.2.3 Major Concern and Main Task for Molecular Pharmacognosy

With a theoretical basis on pharmacognosy, molecular pharmacognosy presents problems to the field of study in pharmacognosy whose major contents can be summarized by authenticity and excellence as presented as follows: (a) Discerning the false from the genuine so as to settle the problem of variety confusion. Due to rise in scope of use and dosage of medicine, plants, animals, and parts with similar appearance or homonym are taken as the same drugs to be used in different regions,

thus leading to a lot of confusion; therefore, it is necessary to discern the false from the genuine in terms of their origins, distribution areas. Only in this way can quality be guaranteed. (b) Quality assessment. A systematic study is conducted on crude drugs with multi-origin and genuine quality, including place of origin, harvesting, processing, storage, and the influence of transportation upon active ingredients to confirm high-quality variety and factors that may have effect on it. More than that, excellent varieties should be researched and cultured to achieve fast growth, high quality, and high yield in order to meet the needs of medication.

Scientific connotation of all specific information of authenticity and excellence mentioned above is related to difference in their DNA (except mineral medicines). The fake and the genuine may have different DNA composition due to difference in their origin of varieties; thus, DNA, the genetic material, differentiates the genuine from the fake. Therefore, the scientific connotation of studying molecular pharmacognosy is to research DNA and its relation of its expression of difference to authenticity and excellence of crude drugs.

Molecular pharmacognosy has been given the following main tasks:

1.2.3.1 Systematic Assortment of Varieties of Chinese Herbs and Study of Quality Standardization

Assortment of varieties of Chinese herbs, based on classical taxonomy, can be applied to systematic assortment, classification, and identification, but too many human factors are involved, especially for planted groups of Chinese herbal medicines, such as identification of authentic raw materials, which still remains a problem. Development of species biology and molecular systematics provides an effective weapon and basis for the study of system and evolution, classification and identification. Modern species concept has been widely accepted by taxonomists, and pharmacognosists can never ignore theory and fruits in species biology and molecular systematics. Penetration of molecular systematics and application of biological engineering technology provide crude drugs' classification and identification with an effective weapon and basis to test molecules. To establish a method and system to identify crude drugs based on species biology, molecular systematics, Chinese medicinal resources, and herbalism and to further the development of systematic assortment and standardization of quality of Chinese herbal medicines at the population, individual even genetic level, fall into one of the major concerns of molecular pharmacognosy.

1.2.3.2 Conservation of Medicinal Plant and Animal Biodiversity and Research of Sustainable Utilization of Crude Drugs Resources

DNA diversity is the essence of biodiversity. Molecular makers based on DNA polymorphism analysis and molecular systematics based on genome sequence analysis can directly test DNA variation patterns and determine the key units to

protect, so the study of molecular systematics of medicinal plants and animals may presume developing status and endangered degree of the population, thus rendering new operative methods for measurement of biodiversity and countermeasures taken to protect rare medicinal plant and animal resources. Moreover, the application of research findings in molecular systematics based on DNA polymorphism makes the work of hunting for and enlarging the scope of medicinal plant and animal resources more effective and efficient. Expounding the relevance among genetic relationship of DNA molecules, active ingredients, and efficacy in combination of chemical taxonomy, obtaining molecular genetic background of important chemical elements so as to identify whether unknown plants have the genes to produce specific chemical composition, are shortcuts to hunt for and enlarge the scope of crude drug materials by use of molecular systematics.

1.2.3.3 Medicinal Plant Marker Breeding and New Variety Cultivation

In the process of exploring molecular theory and practice, it can be said that molecular detection of genetic diversity and molecular systematics lays a basis for understanding and transforming nature. Exploring and harnessing molecular markers with important traits is the purpose for us to understand and transform nature. With support of cell engineering and genetic engineering, molecular pharmacognosy research becomes more practical. QTL method, which combines breeding technology to the key medicinal plant and animal genetic linkage map constructed by the use of molecular genetic markers, makes it possible to provide information about the mapping of the quantitative trait loci, such as the quantity of genes with target trait, the genetic effects, ways of interaction between genes, and the decomposition of the quantitative traits, which can never be provided by traditional quantitative genetics.

1.2.3.4 Gene Regulation of Metabolic Pathway and Directional Control of the Quality of Chinese Herbal Medicines

More and more attention should be paid to the basic research of secondary metabolite biosynthesis, especially that the research of the gene regulation of key enzymes will be particularly noticeable and then become one of the most challenging and promising direction in molecular pharmacognosy study, for the major source of active ingredients in Chinese herbs is secondary metabolites. The presence or absence of secondary metabolites and their amount decide the quality of Chinese herbal medicine, so carrying out the genetic engineering and improving the content of active ingredients of Chinese herbal medicines will help ease the pressure on the resources of Chinese herbal medicines.

1.2.3.5 The Use of Genetic Engineering and Tissue Culture Technique to Achieve High-Level Expression and Production of Natural Active Ingredients or Genetically Modified Ingredients

The use of genetically modified organisms as a bioreactor to produce exogenous gene-encoding goods is among the most attractive in genetic engineering, and thus it is called "new-generation pharmaceutical factories." It has many advantages: it can express complex natural protein in a natural state, it can be obtained continually from animal milk and blood and extracted from bodies of plants, and it can also pass through the digestive tract. In addition, for plant or animal protein with strong active ingredients but side effects as well, such as scorpion venom and trichosanthin, the sequence that decides on toxicity can be removed or inhibited in expression, thus strengthening the expression of the active parts.

Hairy root cultures and crown gall culture – new technologies in combination of plant genetic engineering and cell engineering in recent years – open up a new road for the research and development of production of active pharmaceutical ingredients. With the improvement in expression efficiency and the expanding of the scope of receptor plants, biotechnology will certainly bring a new impetus to the research of adding new genetic characteristics into traditional crude drugs; meanwhile, with the development of new bioreactor technology and the establishment and improvement of efficient cell culture, the commercialization and industrialization of the biotechnology of the natural medicines will speed up.

1.2.3.6 Genetic Engineering and Green Pollution-Free Medicinal Plant

The problem of pesticide pollution in medicinal plants has aroused public concern, for it damages environment, endangers the health, and limits the export of Chinese herbal medicines. Thus, implementing GAP production, advocating green pollution-free medicinal plants, and controlling pesticides without farm chemical in the process of growing medicinal plants become goals for people to achieve. How to improve the ability of medicinal plants against pests via genetic engineering is also one of the tasks for molecular pharmacognosy.

1.3 Relation of Molecular Pharmacognosy to Other Disciplines

As a new and pioneering branch in pharmacognosy being developed, molecular pharmacognosy combines pharmacognosy and molecular biology. With its birth, pharmacognosy develops into its intensive micro-study inevitably. Pharmacognosy itself is a basic multidisciplinary science for applications, and molecular biology is based on modern sciences like life medicine, so molecular pharmacognosy is sure

compatible and open cross-disciplinary, interdisciplinary, and multidisciplinary with a rich connotation and denotation.

Molecular pharmacognosy is closely related to the following fields of study:

Identification of Traditional Chinese Medicine – With a purpose to identify the genuine from the fake and the good from the bad and to guarantee safety and efficacy, molecular pharmacognosy differs from the traditional identification methods like original plant identification, character identification, microscopic identification, and physical and chemical identification in using DNA molecular genetic technology to directly analyze the polymorphism of the genetic materials so as to determine differences in intrinsic and external performance between different varieties of traditional Chinese medicine, thus providing identification of traditional Chinese medicine with a new convenient and accurate method.

Resource Science of Traditional Chinese Medicine – It is a science to study traditional Chinese medicine in varieties, quantities, geographical location, temporal and spatial variation, rational exploitation, and scientific management. Based on the study of distribution of traditional Chinese medicine resources, resource science of traditional Chinese medicine is aimed to use cost-effective optimization techniques to make a reasonable arrangement for traditional Chinese medicine resources to be harvested, processed, and utilized, so that the society, economy, and ecology can all achieve a coordinated and balanced program of development and thus provide sufficient high-quality raw materials to people's health care and pharmaceutical industries. Therefore, resource science of traditional Chinese medicine is a comprehensive natural science or an emerging multidisciplinary, cross-disciplinary, interdisciplinary science with a management nature. The concept originated in early 1980s and worked as an independent discipline in late 1980s. *Resource Science of Traditional Chinese Medicine*, the first book used for teaching, was compiled and published in May 1993 by Zhou Ronghan, a professor in College of Traditional Chinese Medicine of China Pharmaceutical University in Nanjing. Resource science and molecular pharmacognosy have a close relation for the latter lays a new theoretical basis for resource identification, germplasm diversity detection, searching and expanding new varieties, and new resources of traditional Chinese medicine, thus promoting the development of resource science of traditional Chinese medicine.

Pharmaceutical Botany – As a science that deals with study of pharmaceutical plants by use of morphology, structure, and taxonomic knowledge and methods in botany, pharmaceutical botany shows a major concern for systematic study of botanical knowledge to research the identification and classification of pharmaceutical plants, to investigate their resources, to sort out the types of Chinese herbal medicine, and to ensure accurate and effective medication. Molecular pharmacognosy will provide genetic evidence to pharmaceutical plants in their systemic evolution, the search for new drug source, and the cultivation of new varieties.

Chapter 2
Methodology

Lu-qi Huang, Xue-yong Wang, Wei Gao, and Kenji Kondo

Abstract This chapter mainly introduces the development of methods and techniques involved in pharmacognosy and molecular pharmacognosy. Firstly, ideas and principles of methodology of molecular pharmacognosy study are introduced in this chapter. Of them, some of new theories like "pharmaphylogeny" and new methods like "ingredient difference phenotypic cloning" are firstly introduced in integrity together. Secondly, study objects and problems involved in molecular pharmacognosy are prominently introduced in this chapter. Finally, technologies often used and examples of the study are also briefly introduced.

From this chapter, we could understand thoroughly the theories, study objects, and usual methods of molecular pharmacognosy.

Lu-qi Huang (✉)
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China
e-mail: huangluqi@263.net

Xue-yong Wang
College of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing, China
e-mail: wangxydoctor@yahoo.com.cn

Wei Gao
School of Traditional Chinese Medicine, Capital Medical University, Beijing, China
e-mail: weigao@ccmu.edu.cn

K. Kondo
Botanical Raw Material Research Department (Molecular Identification of BRM), Tsumura & Co., Ibaraki, Japan
e-mail: kondou_kenji@mail.tsumura.co.jp

Lu-qi Huang (ed.), *Molecular Pharmacognosy*,
DOI 10.1007/978-94-007-4945-0_2, © Shanghai Scientific and Technical Publishers and Springer Science+Business Media Dordrecht 2013

2.1 Introduction

Pharmacognosy is the study of medicines derived from the natural sources. Long time ago, people learned to use the herbal medicines originated from the natural sources, such as the roots and barks of medicinal plants. After longtime use of herbal medicine, a great deal of crude drugs was recorded.

2.2 Ideas and Principles

2.2.1 Identification [1]

Identification is one of the important study contents of pharmacognosy. The main work of it is to discriminate the genuine and false crude drugs and to evaluate the excellent and inferior in quality of crude drugs. The crude drugs mainly originate from plants, and partly originate from creatures and minerals. It is very important to identify and evaluate the crude drugs because the crude drugs from various species will lead to the difference both in contents and categories of active ingredients. The identification is often based on the following characters of crude drugs: the morphology (character identification), the tissue structure (microscopic identification), the secondary metabolites (physical and chemical identification), the genetic information (molecular identification), and other aspects of characters. With the development of science and technology, the research area of the identification technology is constantly evolving. The methods of pharmacognosy study mainly include (1) identification of botanical origin, (2) trait identification, (3) microscopic identification, (4) physical and chemical identification, and (5) molecular identification. The current identification scope of pharmacognosy includes:

Related Species Identification: Many crude drugs and Chinese herbal medicines are multisource species. This part of crude drugs needs to be identified. At present, the closely related species identification and quality evaluation of crude drugs are mainly made by using the traditional identification methods. Because of the similarities in closely related species on morphology, tissue structure, and chemical composition, their identification is very difficult. As technology advances, the traditional identification methods combined with molecular biology techniques such as the popular DNA bar coding technology will be an effective means to identify the closely related species.

Identification for Analogs of Valuable Chinese Materia Medica: The rare and valuable Chinese materia medica (CMM) is one of treasures of traditional Chinese medicine; meanwhile, it is also the object of counterfeit medicines easily produced. The resources of many rare CMM such as Cordyceps sinensis, antler, agallochum, saffron, *Dendrobium candidum*, musk, etc. are limited, and they are more expensive. Identification of these herbs is often difficult. Take herbs of

agallochums for example. Most of them are imported from the Southeast Asian countries. Both inferior and counterfeit varieties are frequently found in the market. Because the shapes between the spurious and genuine are very similar, in addition that the spurious' 95% ethanol extract is higher than that of the genuine one, the common identification method is sometimes difficult to entirely ensure the identification results. Recently, with the DNA bar coding technology, E-bionic technology, and GC-MS technology widely used, the authenticity of identification for valuable CMM has been greatly improved.

Identification for the Genuineness of CMM: The genuineness refers to the study of Chinese materia medica which is produced from specific areas and recognized as the authentic and famous in clinic effects. Genuine herb is the medicine of the original species with long-term breeding in specific environmental conditions and specific production process and eventually becomes a certain crude drug in good quality. The formation of genuine medicine can be summarized as genetic model oriented, eco-oriented, and so on. From a species perspective, authentic herbs can originate from both single varieties and multi-varieties. But even if the single varieties of herbs are considered, their original genetic characteristics of species will continue to divide and form a local special variation of the base of gene. As a result, it is the formation of a genuine herb because of this genetic diversity and variability. Since authentic species of medicinal herbs and non-genuine ones are often very similar in morphology, properties, and chemical composition of crude drugs, their identification is often very difficult. As an important means of modern molecular biology techniques, the DNA molecular markers and DNA bar coding will play a key role in explaining the biological mechanism in the formation of the genuineness of CMM.

2.2.2 *Pharmaphylogeny [2]*

When it comes to the field of study in pharmacognosy, there is a very important part that aims at exploring the relations among the distribution of biologically active substances, efficacy and plant evolution. Xiao Pei-gen firstly put forward the theory of pharmaphylogeny which was concluded from the theory of certain links among plant genetic relationship, chemical composition, and efficacy. It provides a theoretical guidance for the development of medicinal plant resources. Pharmaphylogeny is the study of the correlations among genetic relationships, chemical components, and efficacies of medicinal plants. It is a new interdisciplinary subject in the field of medical plants study, which involves plant taxonomy, plant phylogenetics, phytochemistry, pharmacology, numerical taxonomy, genomics, information science, and other computer technology-related disciplines (Fig. 2.1).

Pharmaphylogeny is characterized by multidisciplinary and infiltration and the scope of their research must be the combination of multidisciplines. Therefore, the study also involves multidisciplinary content:

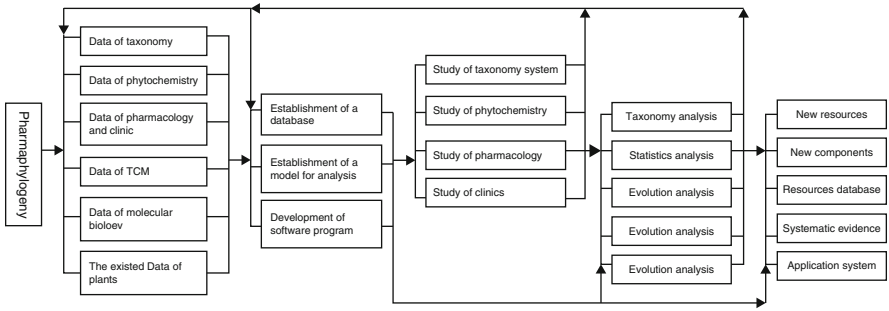

Fig. 2.1 The framework of pharmaphylogeny

2.2.2.1 Information Science and Intelligence Science Research

With thousands of years of experience in using traditional Chinese medicine, in combination with a large number of international research data on active ingredients, pharmacological effects, and clinical efficacy of medicinal plants, the key task that pharmaphylogeny tries to fulfill is how to use these valuable documents and resources of medicinal plants. Medicinal plants research information system is about the use of modern computer technology, information processing technology, and especially the "Knowledge Discovery in Database, KDD" technology, combination with the research results of chemistry, mathematics, biology, and Chinese medicine to establish and improve the database of medicinal plants, database of natural chemistry, and knowledge base of Chinese medicine. The goal is to dig the information of efficacy and related components in medicinal plants, combined with knowledge of plant systems to explore the law of distribution of active components in medicinal plants.

2.2.2.2 Distribution Law of Chemical Components in Plants

It is the basis to develop new drugs that find the distribution law of chemical components in plant system. Biological diversity leads to the variety of secondary metabolites, which makes it possible to provide human a variety of drugs. However, there is no theory to introduce the new drugs creation, which is bound to do less effectively. Therefore, the main task of pharmaphylogeny is to explore the distribution law of chemical components in plants and guide the screen and development of new drugs.

2.2.2.3 Chemosystematics

As the plant secondary metabolites of medicinal plants, active components in plants subject to gene regulation. It also has the genetic characteristics of the plants. The components' accumulation in plants is closely related to species of plants and their phylogenesis. Therefore, the chemical components of plant can be thought as a

strong evidence of plant systematics. So another task of pharmaphylogeny is to study the chemical evidence of plant systematics and enrich and improve the chemical taxonomy of plants.

2.2.3 *Ingredient Phenotype [3]*

As to the study of pharmacognosy, it is the focus of the quality formation in herbs. And the quality of herbs is based on the content of effective component (mainly the secondary metabolites) of herbs, whose production is closely related to secondary metabolism in medicinal plants. The secondary metabolism is controlled by enzymes in some related secondary metabolism pathway. Therefore, the core of the molecular pharmacognosy study is to disclose the relationship between genes encoding enzymes in the secondary metabolism pathway and secondary metabolites.

When it comes to the use of molecular biology techniques, such as gene cloning, transgenic technology to study the quality of tradition Chinese medicine, the greatest difficulty encountered, compared with the crop, is unknown or little known about the genetic information of medicinal plants. In another word, the unclear genetic information especially about genes in secondary pathway in medicinal plants has become a serious constraint of genetic engineering about secondary metabolism in medicinal plants. Because of the cloning of unknown functional genes in the unknown downstream of secondary pathway related to active components biosynthesis, it will become a difficult job but an important task that has to be completed. Although the structure of active components is clear, their biosynthesized secondary metabolic pathways in medicinal plants are not clear, which brings the most difficulties in the related functional gene cloning and genetic engineering study. Currently, the rice and Arabidopsis genome sequencing have been completed. As the same with the human genome plan coming into post-genomic era focused on functional genomics, the post-genomic era of plant is also focused on the function of genes. With the reference to Arabidopsis genetic information, many important plant functional genes will be found. However, because of the big difference between the active components of medicinal plants and the structure of secondary metabolites in Arabidopsis, added with the unknown secondary metabolic pathway, it is difficult to design and synthesize primers to clone functional genes related to the secondary metabolisms based on the reference genome information of Arabidopsis. Since most of secondary metabolic pathways for active ingredients biosynthesis in medicinal plant are not clear, it will be extremely difficult to clone genes according to similar information of sequence of genes from other plants.

The classic study of plant secondary metabolic pathways is isotopic tracer method. Because this method has some difficulties in operation, along with isotopes of radioactive pollution and other unforeseen factors, the isotopic tracer method is limited to be widely used to some extent. To this point, we introduce a new idea and method, the “ingredient difference phenotypic cloning,” for secondary metabolic pathway of active ingredient and its related functional genes cloning. This method does not require clear information of pathway for active ingredients

biosynthesis and sequence information of homologous genes. It has lots of advantages, such as high-speed and high-flux cloning, without clear information about gene sequence of secondary metabolism-related enzymes and the pathways. This will undoubtedly provide new ideas and effective means to solve the bottleneck problem of active related ingredients. Its concepts, mechanism, and methods are introduced as follows:

2.2.3.1 Concept

"Ingredient difference phenotypic cloning," a sort of phenotype cloning, is an effective strategy and method for the cloning of genes encoding key enzymes that regulate the secondary metabolisms. It is based on the phenotype of differences in ingredients (secondary metabolites) to clone functional genes that belong to some unknown secondary metabolic pathways by the technology of gene differential expression. The phenotypes of differences in ingredients include the difference in the content of them and the presence or absence of certain ingredients. This method has advantages of high flux, fast, and high efficiency in gene cloning, and the most predominance of it is to clone the unknown functional genes in unclear secondary metabolic pathway.

2.2.3.2 Postulate

Under external stimulation, such as a variety of environmental stresses and elicitors' stimulation, levels of gene expression in plant cells often abnormally increase, and it may lead to dramatic increase in secondary metabolites' accumulation. Since ingredients (secondary metabolites) in medical plants are easily tested, We can regard it as the phenotype of differential expression of genes and clone some objective genes in a certain secondary metabolic pathway by some technologies, such as differential display, suppression subtractive hybridization, and cDNA microarray.

2.2.3.3 Method

When it comes to the design of ingredient difference phenotypic cloning, the most important factor that has to be considered is how to make the biggest difference of active ingredients in plant cells between dealt groups of plants and the control ones. The operation in details is as follows: Firstly, add elicitors like heavy metals and other biological elicitors into the culture medium to make the biggest difference of active ingredients in the dealt groups of plant cells compared with that of control ones. After dealing with elicitors, the activity of secondary metabolism in plant cells will be stimulated and enhanced quickly, and the production of secondary

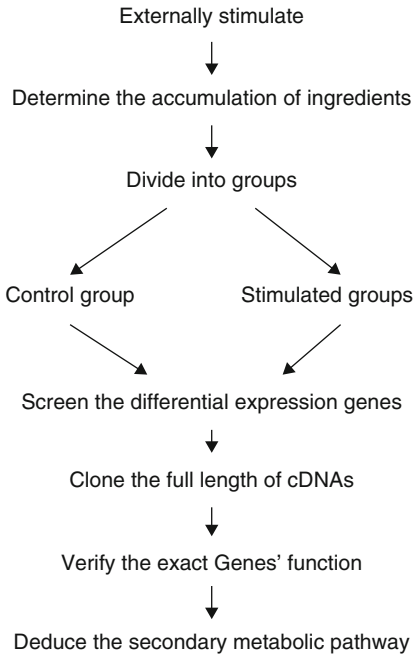

Fig. 2.2 Steps of ingredient difference phenotypic cloning

of metabolism, the active ingredients will be accumulated rapidly at the same time. Secondly, determine the content and categories of ingredients in different groups of cultured plant cells with HPLC or LC-MS and select pair groups with the biggest different content of ingredients for the study of "ingredient difference phenotypic cloning." Finally, screen the objective fragment of genes with differential expression, clone the full length of cDNA, analyze the sequence, and verify the function. The steps of the method can be shown in Fig. 2.2.

In this study, the first to do is to acquire the biggest variability of ingredients. Several methods can be used in ingredients analysis. For determination of the most active ingredients, HPLC method is usually used, but for some of volatile ingredients, GC method can be used, and for the unknown structure of ingredients, LC-MS is an ideal determination method. As to gene cloning method in the study of "ingredient difference phenotypic cloning," the mRNA differential display reverse transcription PCR (DDRT-PCR), the suppression subtractive hybridization (SSH), cDNA microarray, and so on will be good methods available.

Advantages: Firstly, it has the advantages of high flux, fast and high efficiency in gene cloning. Secondly, it will effectively clone the functional genes without necessity to know the exact homologous sequence information for multi-gene cloning and the ingredient-related biosynthesis pathway.

Table 2.1 Genes of key enzymes, precursors, and production of secondary metabolism in *Salvia miltiorrhiza*

GenBank code	The encoding enzymes	Precursors	Productions
–	4-Coumarate—CoA ligase (Sm4CL)	4-Coumarate	4-Coumarate—CoA
F635969	Acetoacetyl-CoA thiolase (Sm AACT)	Acetyl-CoA	Acetoacetyl-CoA
EF534309	2-Phospho −4-(cytidine-5′-diphospho)-2-C-methyl-D-erythritol kinase (Sm CMK)	2-Phospho-4(cytidine-5′-diphospho)-2-C-methyl-D-erythritol	4(Cytidine-5′-diphospho)-2-C-methyl-D-erythritol
EF635967	Isopentenyl-diphosphate delta-isomerase (SmIPP)	Dimethylallyl diphosphate	Isopentenyl diphosphate
EF635968	Farnesyl diphosphate synthase (SmFPPS)	Geranyl diphosphate (GPP)	Farnesyl diphosphate synthase (FPP)
EF635966	Entkaurene synthase-like (Sm KS L)	Copalyl diphosphate (CPP)	Pimar-15-en-8-yl+

2.2.3.4 Applications

Clone of Genes Encoding Key Enzymes of Secondary Metabolic Pathways

With the introduction of "ingredient difference phenotypic cloning" strategy, we have cloned six cDNA fragments of genes encoding key enzymes involved in secondary metabolism of *Salvia miltiorrhiza* hairy roots dealt with elicitors of yeast extract by cDNA microarray analysis (Table 2.1). Among these six functional genes, five of them (SmAACT, SmCMK, SmIPPI, SmFPPS, SmKSL) encoded the enzymes of tanshinone biosynthesis and one (Sm4CL) was involved in the biosynthesis of salvianolic acid by blastx analysis, which involved secondary metabolic pathway analysis with online software of KEGG. After the analysis, full-length cDNA of these 5 genes were cloned by 3′race-PCR and 5′race-PCR method. The five genes involved in tanshinone biosynthesis were registered in GenBank online database. Their GenBank code was shown to be in the order: F635969, EF534309, EF635967, EF635967, EF635968, EF635966.

Help to Systemically Reveal Between the Biological Network of External Stimulating Factor and Secondary Metabolism and Elucidate the Mechanism of Gene Expression Regulation of Secondary Metabolisms

Here we elucidate the effects of elicitors on secondary metabolism as an example; to begin with, the elicitors bind with membrane receptors and introduced changes in membranes of which lead to changes of membrane's permeability and internal ion distribution in membranes. At the same time, G-proteins may be coupled to recep-

tors and mediate elicitor-induced ion channel activation. Ion fluxes, especially Ca^{2+} influx, cause cytosolic free Ca^{2+} spiking which causes activation of protein kinases, peroxidases, NADPH oxidases, and phospholipases, which further generate other signaling messengers, such as reactive oxygen species, DAG, IP3, cAMP, lysoPC, JA, ethylene, NO, cADP ribose, and SA. All these messengers compose paralleling or cross-linking pathways to integrate these signals into regulation of transcription factors (TFs) [4]. Various transcription factors integrate these signals to activate gene expression by transcription machinery. Most genes for secondary metabolite synthesis are late response genes. The response genes' expression levels then affect their encoding enzymes in regulating secondary metabolite synthesis. That is the theory to explain the biological network relations between external stimulating factor and secondary metabolism.

With the development of functional genomics, proteomics, and metabolomics, many new and powerful tools could be applied to plant secondary metabolism study and improve overall understanding and practical manipulation of plant secondary metabolite production. The ingredient difference phenotypic cloning is a useful and powerful method mainly involved in transcript-profiling differential analysis and secondary metabolites related genes cloning in medical plants. This method helps to identify more genes involved in biosynthesis of secondary metabolites but also facilitates isolation of some possible signal transduction components such as transcription factor and other regulation genes.

Taking advantages of "ingredient difference phenotypic cloning" methods, we have at the same time acquired signal transduction proteins, sulfate transporter, DNA binding/transcription factor, and other secondary metabolism-related genes. The results and established methods lay a useful groundwork for future study of regulation mechanism of genes for ingredients biosynthesis (Table 2.2).

The ingredient difference phenotypic cloning method makes people's attention once again focus on the close relations between phenotype of secondary metabolites and genetic information in medicinal plants. To minimize the impact of external uncontrollable factors and make the maximum differential of secondary metabolites, the ideal strategy of the study system is to culture the tissue and cells of medicinal plants under the controlled condition and stimulate plant cells to produce the utmost differential phenotype by adding the elicitors into the culture medium. According to the phenotype difference, a great body of differential-expressed objective genes would be efficiently cloned with high-throughput techniques of microarray.

Since secondary metabolism of plant is often affected by external environmental factors, profiling a group of secondary metabolites from plants under various environmental conditions helps to understand metabolic flux and the related regulatory mechanisms. Because of the close relationship between geography-related environmental condition and biosynthesis of secondary metabolites for defense, plant cells have various strategies to control metabolic flow directions. This regulation is mainly controlled at enzyme and gene expression levels. Thus, the "ingredient difference phenotypic cloning" method can be widely used in the study on the formation mechanism of geo-authentic medical material at molecular level.

Table 2.2 The cloned partial regulation genes related to secondary metabolism of *Salvia miltiorrhiza* by "ingredient difference phenotypic cloning" method

		Ctrl sample signal intensity of hybridization		Treatment sample signal intensity of hybridization		
Code	Differential-expressed genes	1	2	1	2	Treatment/ctrl
chip44d06	Aquaporin (PIP2;8/PIP3B;)	358	614	690	862	1.61
chip39b01	CBL-interacting protein kinase	2,251	3,465	5,682	9,245	2.025
chip31h02	DNA binding/transcription factor	3,639	1,397	1,510	567	0.465
chip36h12	Major allergen Cor a 1	7,714	3,995	14,858	10,649	3.455
chip41b06	Sulfate transporter	2,345	1,868	4,580	3,100	2.515
chip45h02	Stress and pathogenesis-related protein	2,289	4,775	5,783	15,290	2.765
chip19g08	Pathogenesis-related protein 10	6,636	8,323	13,499	14,522	2.345
chip29h02	Metallothionein-like protein 1 (MT-1)	1,204	1,709	325	519	0.285

Combining the approaches of transcription profiling proteomics and secondary metabolite profiling, the method of "ingredient difference phenotypic cloning" offers the most powerful tool ever in studying all aspects of plant secondary metabolism as a whole.

2.2.4 Systems Biology

Secondary metabolite is a kind of micromolecular organic compound produced during the growth and development of plants along with the adaptation of outer environment. It has been estimated that the amount of secondary metabolites in plants is more than 100,000, including terpenoids, phenols, alkaloids, polyacetylenes, etc. [5]. So far, the studies on medicinal plant secondary metabolites have concentrated on such aspects as separation of chemical composition, structure determination, bioactivity, pharmacological actions, etc. However, the content of secondary metabolites in medicinal plants is relatively low, and the resources of natural medicinal plants are limited, which affect the quality control of medicinal plants and the exploitation of active ingredients. So it is obviously important to study the biological formation of secondary metabolites; to unearth genes of relevant enzymes, signal factors, or enviromental factors; and to systematically illustrate biosynthetic pathway, signal transduction pathway, and biological formation mechanism as well as the interaction of them.

Systems biology is a new field that aims at system-level understanding of biological systems, and it is the first time that we may be able to understand biological systems grounded in the molecular level as a consistent framework of knowledge after genomics, proteomics were put forward [6]. Different from molecular biology focusing on the individual ingredient, systems biology concentrates on constitution of all the compositions (such as gene, RNA, protein, etc.) in a biological system and the correlations of these compositions under specific condition [7]. It is a powerful tool to explore biology fully, and the thought and approach applying to secondary metabolites in medicinal plants is an effective way to fully proclaim the process from genes to secondary metabolites [8].

2.2.4.1 Biological Process of the Formation of Secondary Metabolites in Medicinal Plants

Biological process of the formation of secondary metabolites in medicinal plants, which are regulated by various biotic and abiotic factors either from gene or environments, is very complex. Now, there are various hypotheses existing in induced mechanism of the production and accumulation of secondary metabolites, including growth/differentiation balance (GDB), carbon/nutrient balance (CNB), optimum defense(OD), resource availability(RA), etc. [9–12]. From the angle of systems biology, the formation of secondary metabolites is a systematic biological process,

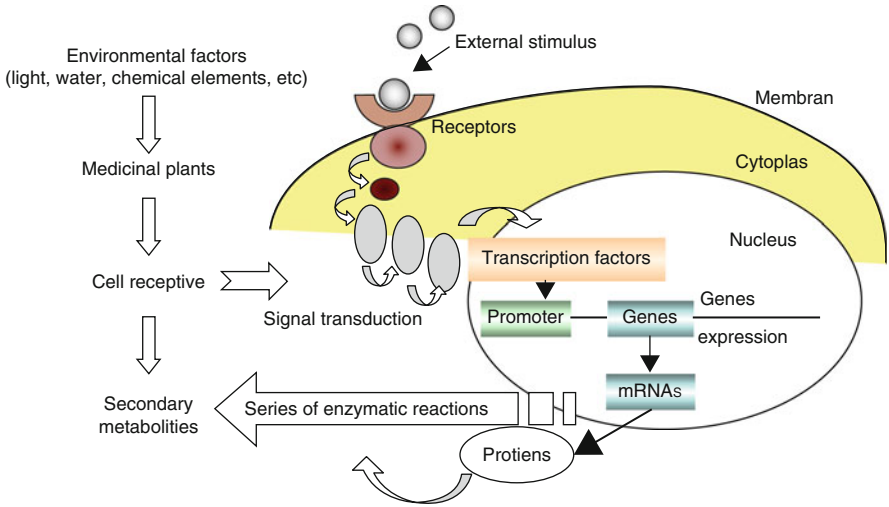

Fig. 2.3 The possible biological processes of secondary metabolites formation in medicinal plants

which consists of three portions, the stimulation of environmental factors (inner and outer environments), signal transduction, and the biological process catalyzed by gene expression and protease translation. The specific process is as follows: environmental factors stimulate receptors out of plant cells, making receptors activated. These receptors activate intracellular signaling cascade and then transcription factors to start the expression of specific genes. Afterward, genes are transcribed and translated into relevant proteases, in order to catalyze the production of secondary metabolites (Fig. 2.3).

2.2.4.2 Technology Platform and Basic Methods of Systems Biology

Compared with the method of reduction theory that molecular biology adopts, systems biology applies the method of systems science to quantitative studies on organisms as an entire system instead of isolated parts [13]. Classic molecular biology research is a vertical research to study a single gene or protein by various approaches. First, finding specific genes on the DNA level and then studying the function of genes by methods of gene mutation or gene knockout; on this basis, studying the space structures, modification of proteins, or protein-protein interaction. Genomics, proteomics, and the others are horizontal researches to study thousands of genes or proteins simultaneously by a single method. But the method of systems biology integrates both of them to be a three-dimensional study [14]. It can take full advantage of omics technologies to study the molecular difference among biosystems, infer the mechanism of environmental chemistry in biosystems, establish mathematical models to assess the modification or diversity of mRNA,

protein, illustrate the holistic biological effects, and describe biological functions, phenotypes, and behaviors.

The major technology and platform of systems biology consist of genomics, transcriptomics, proteomics, metabonomics, interactomics, and phenomics [15]. Genomics is about genome mapping (including genetic map, physical map, and transcription map), nucleotide sequences analysis, gene mapping, and gene function analysis to all genes of a species. And the common analysis methods of transcriptomics are differential display, gene chip, EST, MPSS, cDNA-AFLP, etc. [16, 17]. Recently, the German scientist Marc Sultan [18] utilized deep sequencing to get a brand-new view of the human transcriptome, and it is expected to apply to transcriptomics of other species. The major approaches of proteomics analysis are DEP, MS, etc. Metabonomics is a very important way to study medicinal plants and achieve modernization of traditional Chinese medicine [19–21], with common analysis methods of NMR, GC-MS, LC-MS, FTMS, CE-MS, etc. Genomics, transcriptomics, proteomics, and metabonomics detect and identify various molecules to study their functions on the level of DNA, mRNA, protein, and metabolin, respectively, forming multiple levels of biological information transfer. Interactomics studies interaction of molecules to discover and identify molecular pathways and networks, and to draw interaction maps systematically. Phenomics is regarded as a link between genotype and phenotype.

2.2.4.3 The Application of Systems Biology in Secondary Metabolites Study

Biosynthetic Genes and Pathways of Secondary Metabolites

Biosynthetic pathway is the core of study on secondary metabolites in medicinal plants, an extremely complicated process from the genes to biological phenotypes (secondary metabolites). People have had a basic understanding of the main part of secondary metabolic pathway through long-term studies, such as shikimic acid pathway of phenols and IPP pathway of terpenoids [5]. Because of a variety of metabolites, end products are often generated by structural modification after the formation of basic framework. Currently, except taxol, arteannuin, indole alkaloids in *Catharanthus roseus*, etc., the majority of secondary metabolic pathways are not yet fully elucidated, waiting to be further illustrated.

Our research group has adopted the thought and approach of systems biology to acquire systemic studies results on biosynthetic pathway of tanshinone as diterpene secondary metabolite in *S. miltiorrhiza* [22–25] (Fig. 2.4). Using elicitors to stimulate *S. miltiorrhiza* hairy root can lead to diverse phenotypes of the tanshinone content. And we analyze the metabolome, proteome, transcriptome of multiple materials with diverse phenotypes. The next step is to obtain full-length genes from the screened gene fragments which are closely related to tanshinone secondary metabolites, by multivariable analysis. SmCPS from clone is the first (+)-CPP synthetase in angiosperm; SmKSL is identified as a new diterpene synthetase, and it

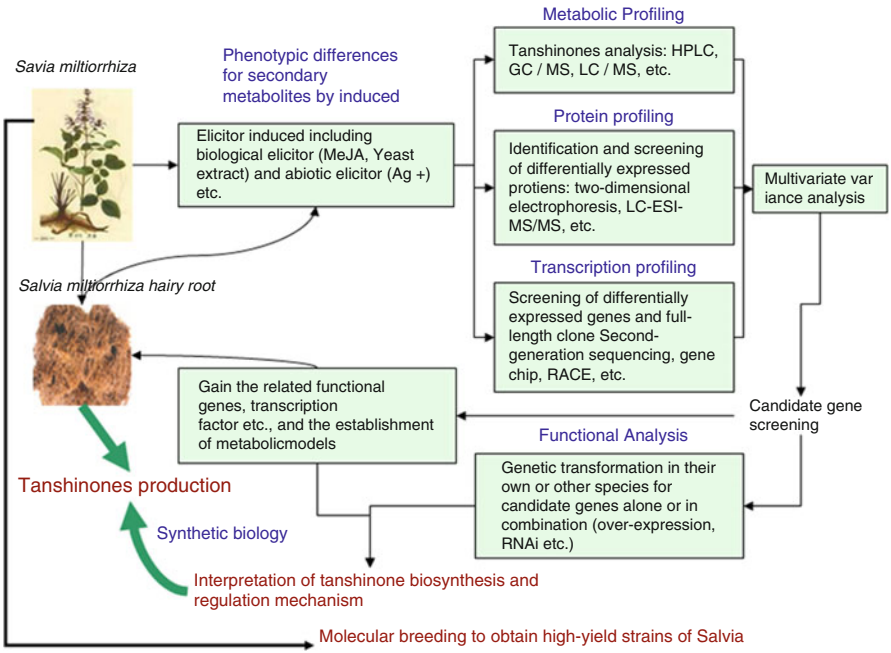

Fig. 2.4 Systems biology approach to explore tanshinone biosynthesis

can catalyze (+)-CPP into miltiradiene. This is a new and specific branch of tanshinone diterpene biosynthetic pathway, which puts tanshinone diterpene biosynthetic pathway two steps forward.

Excavation of Signal Factors and Study of Signal Transduction Pathways

Cells for the intercellular communication and the process of cellular compression reactions receive signals from outside. And these signals are converted into intracellular signals or cascades. The typical signals include hormone, pheromone, heat, cold, light, osmotic pressure, and some other materials, like the appearance or the concentration changes of glucose, potassium ion, calcium ion, and cAMP. And the significant difference between signal transduction and metabolic process is metabolic process provides the transmission of quality, determined by a series of catalyze reactions but signal transduction undertakes information processing and transferring.

The biosynthesis of secondary metabolites in plants is a series of complex biochemical reactions controlled by correlative intracellular genes. But the environmental factors do not participate in secondary metabolites directly like outside stimulating factors, so there must be relevant intracellular signal molecules and corresponding signal transduction mechanisms to receive and conduct stimuli of external factors. Study has discussed that the correlative signal molecules and signal

transduction mechanisms, regulated by secondary metabolites biosynthesis in plant molecules, can contribute to the understanding of the regulation laws of secondary metabolites biosynthesis in molecules, as well as can provide a rationale for improving secondary metabolites output of cultured cells in production practice [26]. Since there have been in-depth signal transduction researches on plant disease resistance and defense responses, studies on the mechanism of secondary metabolism signal transduction in medicinal plants are still in the initial exploration stage. Xu Maojun research group has made progress on the study on the signal factors and signal transduction mechanism of mediating Forsythia suspense seeds into hypericin and accumulating Ginkgo flavonoid glycoside in Ginkgo cells [27]. Additionally, at the foundation of optimizing the cell culture conditions, it finds that UV-B is an external environment stress factor, which can induce the synthesis and accumulation of flavonoids in the cells from 5 to 30 h, using the established *Hypericum chinense* cells as materials, then, it can explore the signal transduction mechanism deeply induced by UV-B. It also considers that this process is influenced by NO and H_2O_2 signal molecules, and these molecules will have synergetic effects on the process, which is regarded as a new signal interaction phenomenon. Furthermore, the study deems NO mediating UV-B to induce the synthesis and accumulation of flavonoids and is related to the CHS genes expression, but H_2O_2 is not [28]. The most feasible application of systems biology is conducting detailed models of cell regulation and focusing on specific signal transductions and molecules at all levels, in order to have a deep understanding of drug discoveries on the basis of mechanism [29]. Obviously, these are beneficial explorations via the thought and methods of systems biology to study the secondary metabolite signal transduction pathway. And secondary metabolite signal regulation in plant cells is a quite complex system. Recently, there have been some definite progresses in the relevant researches, but now it is still far from completely knowing the mechanism of secondary metabolite signal transduction pathway. By finding signal molecules that can make the phenotype of secondary metabolites different, the application of systems biology methods is ultimate to make high-throughput isolation of mutants related to plant secondary metabolites, to clone the relevant genes of secondary metabolism regulation with studying the functions as well as to discuss about activating transcription factors to start specific genes expressing signal transduction pathways so that it can establish the interaction and network of signal factors, genes, and metabolites to illustrate signal transduction pathways based on the formation and accumulation of secondary metabolites.

Ecology of Medicinal Plant Secondary Metabolism

Plants in the growth progress will be influenced or even coerced by various environmental factors, including abiotic factors (such as light, temperature, soil, moisture, atmosphere, etc.) and biotic factors (damage by diseases and insects, herbivores, microorganisms, artificial inferences, etc.). Plants make adaptive responses to these factors at the foundation of morphological structure, physiology,

biochemistry, and gene expression. And secondary metabolite is one of significant biochemical regulators, for example, concentration of flavonoids, terpenes, alkaloids, and organic acids in plant tissues will definitely rise under water deficit conditions [30–33]. As a result of coupling with environment for a long period, plant secondary metabolism is playing an important role in enhancing plants self-protection ability and recording more environmental information than primary metabolism. Some scholars [34] proposed the concept of ecology of plant secondary metabolism. Compared with chemical ecology and plant physiological ecology, ecology of plant secondary metabolism pays much attention to both secondary metabolism itself and the way environmental factors induce the generation of these compounds. Therefore, it is an important task for ecology of plant secondary metabolism to illustrate how to induce activation of relevant acceptors, express genes, and conduct secondary metabolism.

Owing to the fact that secondary metabolism is complicated and plants are usually affected by different environmental factors at the same time, like drought and high temperature often coexist, the work of studying the relationship between ecological factors and secondary metabolism is full of challenge. Recently, there have been more and more related researches on the way to quantitative researches, because qualitative description cannot state the relationship fundamentally.

In terms of signal transduction, gene expression, or metabolites, we can adopt the methods of controlled experiment and systems biology to reveal how different outside stimuli induce and regulate secondary metabolism by receptors and mechanism of intracellular signal transduction. We can also establish a network of enviromental factors-genes-metabolites, and try to find out the pathway of enviromental factors activate related receptors, and the receptors initiate gene expression and regulation for producing the metabolites, which cloud clarify the physiological mechanism of the generation and accumulation of secondary metabolites in medical plants as well. Here we could utilize two strategies to achieve the goal: initially, using controlled experiments which control investigated environmental factors (like temperature factor) strictly, tracking and analyzing the key enzymatic genes expression in secondary metabolism and protein (enzyme) synthesis, detecting the content variation of secondary metabolites, thus positively proclaiming the relationship between environmental factors and secondary metabolism; then, analyzing the content of secondary metabolites with different phenotypes and the environmental factors that may affect the accumulation of secondary metabolites and detecting relevant genes expression. Therefore, through repeated confirmation, we read the relationship between plants and environment via secondary metabolism, cognizing plant secondary metabolism via ecology.

Metabolic Engineering of Medicinal Plant Secondary Metabolites

Metabolic engineering is mainly about altering metabolic flow, expanding metabolic pathway, or establishing new metabolic pathway to reach the expected target by genetic engineering. And the study has made great development. The group of

Professor Kexuan Tang converts genes of PMT (rate-limiting upstream enzyme putrescine *N*-methyltransferase) and H6H (the downstream enzyme hyoscyamine 6-hydroxylase) into henbane and produces 411 mg/L scopolamine in transgenic henbane hairy root which is over nine times more than that in the wild type (43 mg/L), improving the accumulation of tropane alkaloids greatly [35]. American Professor Jay D. Keasling et al. produced the antimalarial drug precursor artemisinic acid in engineered yeast by a series of methods of gene regulation [36]. This introduced single, multiple target genes or an integrated metabolic pathway to produce new target materials or increase the content of target metabolites in organisms.

Moreover, antisense RNA and technologies like RNA interference can reduce the expression level of target genes and thereby restrict competitive metabolism pathway, alter metabolic flow, and increase the content of target materials. The study of Allen et al. [37] shows that blocking the metabolism pathway of morphine production in opium poppy will lead to the accumulation of reticuline and its methylated derivatives.

A new strategy of metabolic engineering is treating signal pathway and transcription factor as regulating targets [5]. Modifying the transcription factors that control multiple biosynthesis genes will regulate plant secondary metabolism effectively and improve the accumulation of specific compounds. For example, in the biosynthetic pathway of diterpenoid indole alkaloids in *Catharanthus roseus*, high expression of the transcription factor ORCA3 with AP2/ERF functional domain will result in the overexpression of several relevant genes and accumulation of diterpenoid indole alkaloids [38].

Medicinal plant metabolic engineering aims at improving the content of some important secondary metabolite and its precursor to solve the problem of medicine sources. If the content can be enhanced by the method of gene engineering, there will be enormous economic and social benefits. So far, scientists have paid much attention to developing predictive metabolic engineering, and it utilizes the way of systems biology to integrate the data from metabolomics, proteomics, and transcriptomics and then to carry out repeated systematic simulation on the level of metabolic network, finally to get the result that is closer to true state. The existing database and instrumental analytical methods have made such system analysis possible to some extent.

2.2.4.4 Prospect

Systems biology is a collaborative study on the interaction of components in cellular network and other components in biosystem, the application of high-throughput genome-wide experimental techniques, and the integration of calculation methods and experiment results. Study on the production of secondary metabolites in medicinal plants by the thought and approach adopted in systems biology includes the matriculate way of secondary metabolites and signal transduction of signal factors. The most obvious feature of interactive relationship between generation and accumulation of metabolites and external environment is the holistic study from the reductionism perspective, which can

adequately explore genes, transcription factors, signal factors, and environmental factors related to secondary metabolite biosynthesis. Establishing the system model of genes expression and regulation in secondary metabolite biosynthesis provides rationale for fully interpreting molecular mechanism of the production of secondary and the metabolic engineering. And it is of great significance to explain the cause of active ingredients in traditional Chinese medicine, the formation mechanism of famous-region drug, or the reasonable development and utilization of medicinal plant resources systematically.

2.3 Study Objects

2.3.1 "Dao di" Medicinal Materials

Generally speaking, "dao di" medicinal materials are the genuine articles, the "real McCoy," and the "geo-authentic" medicinal materials with Chinese characteristic and good quality in clinical effect. The concept of "dao di" has a long history of significance in Chinese medicine, and it is essential to preserve this rich, clinically relevant information. A good diagnosis is nearly useless if one uses poor quality herbs. The biological meaning of "*Dao di*" medicinal materials refers to the same species of medicinal plants from differences areas. That is to say, a certain subunit of a species of medicinal plants with similar ecological structure character from different region forms the populations. Among the populations, if one of them yields good quality of medicinal material with superior clinical effects, we call them as "*dao di*" herbs or "*dao di*" medicinal materials and the correct geographic region is called as "*dao di*" (geo-authentic) region. Therefore, "*dao*" means the botanical species of a particular population. It should be the formation of the interaction between genotype and environmental factors and can be expressed as: phenotype = genotype + environmental modification.

An investigation found that when a species has a wider distribution area, its subunit of population from different regions often display different genotypes, or called local specialized genotype, and these genotypes are due to different ecological or geographical conditions shaped by long-term section. That is the genetic nature of "dao." So to speak, the connotation of molecular pharmacognosy is to study the difference of genetic material of DNA in similar medicinal materials and identify the genuine and false term, the superior and inferior term of them [1].

2.3.2 Tissue Culture

Use of genetically modified organisms (microorganisms, plant tissue culture) as a bioreactor to produce the exogenous gene-encoding products is one of the most attractive areas in genetic engineering, which is called "new generation of

pharmaceutical factory." It also has challenges and difficulties for the study of molecular pharmacognosy to produce drugs by application of transgenic plants. Hairy root culture technology developed in the last century has opened up a new way for the study of secondary metabolites biosynthesis, functional gene cloning, and regulation of secondary metabolism and R&D of new active pharmaceutical ingredients. So it has become a hot spot for pharmaceutical ingredients production through enhancement of secondary metabolites' accumulation by genetically modifying and regulating methods [1].

2.3.3 Animal Medicinal Materials

A great majority of Chinese medicinal materials originate from either plant or animal sources. Animal medicinal materials mainly include the animal's fur, skin, horns, and bones and insects' body. The challenge of correct identification for animal medicinal materials is compounded by substitutions and unscrupulous adulterations. The traditional phenotypic identification often encounters difficulties in animal materials. The recently developed DNA analysis becomes an important tool to complement organoleptic, morphological, anatomical, and chemical parameters. With the development of molecular markers, like the wide array of sequences and patterns in the genomic, chloroplast, and mitochondrial DNA, especially for the COI gene widely used in animal material marking, the animal material identification problem will be more and more easy to be solved [39]. The molecular markers for snake identification written into Chinese State Pharmacopeia (2010 edition) is thought to be a milestone in the development history of identification of pharmacognosy. Therefore, it is the major advantage of molecular pharmacognosy to solve the problem of animal identification.

2.4 Technology

The goal of molecular pharmacognosy is to solve the problems about medicinal materials to ensure safety, efficacy, and quality of traditional medicines. There are a number of methods involved in molecular pharmacognosy study.

Identification is the first task of molecular pharmacognosy. In the past, identification of Chinese medicinal material was mainly based on morphological features. The identification of morphological features is simple and inexpensive but heavily depends on the experience and judgment of the inspector. Then the method of microscopic features identification has been developed. This method is based on microscopic features, such as texture, tissue arrangement, and cell components. Other identification methods also have been developed by physical and chemical ways. These methods provide a more objective, standard, and accurate way for Chinese herbs' identification than the subjective judgment of inspector based on morphological features of medicinal materials.

One of the most reliable methods for identification of Chinese medicinal materials is by analyzing DNA. In terms of the mechanisms involved, DNA methods can be classified into three types, namely, polymerase chain reaction (PCR)-based, hybridization-based, and sequencing-based methods.

2.4.1 PCR-Based Method

PCR-based methods use amplification of the region(s) of interest in the genome; subsequent gel electrophoresis is performed to size and/or score the amplification products. PCR-based methods include PCR-restriction fragment length polymorphism (PCR-RFLP), random-primed PCR (RP-PCR), direct amplification of length polymorphism (DALP), inter-simple sequence repeat (ISSR), amplified fragment length polymorphism (AFLP), and directed amplification of minisatellite-region DNA (DAMD). Except PCR-RFLP and DAMD, these methods are suitable for Chinese medicinal materials which lack DNA sequence information, as they do not require prior sequence knowledge [40].

2.4.2 Hybridization-Based Methods

Nucleic acid hybridization is a process in which two complementary single-stranded nucleic acids anneal into a double-stranded nucleic acid through the formation of hydrogen bonds. The most obvious advantage is that if the probes are oligonucleotides shorter than 100 bases, hybridization is possible even after a considerable level of DNA degradation [41]. However, a relatively large amount of DNA is required and the process is time-consuming (because the hybridization step typically requires overnight incubation) [41].

2.4.3 Sequencing-Based Methods

DNA sequences can be used for studying phylogenetic relationships among different species [42]. Another advantage of using sequencing for species identification is that the identities of adulterants can be identified by performing sequence searches on public sequence databases such as GenBank. However, prior sequence knowledge is required for designing primers for amplification of the region of interest [41].

2.5 Case Study (I)

Candidate genes involved in tanshinone and salvianolic acid biosynthesis acquired by "ingredient difference phenotypic cloning" method.

This example is desired to concretely introduce how to use the new method or strategy for cloning of functional genes involved in secondary metabolism. The advantages of high speed, high flux, and without need for clear genetic information of the method in gene cloning will be shown in details in this chapter.

2.5.1 Materials and Methods

2.5.1.1 Plant Material and Preparation of mRNA

The *S. miltiorrhiza* used for cDNA library construction was collected from Shangluo, Shanxi province. Hairy roots were grown in MS medium, and on day 18, postinoculation was dealt with YE and Ag^+ according to the reference [43] and harvested after 24 h. For RNA extraction, 5–10 g of hairy root material was frozen and stored in liquid nitrogen immediately after harvest.

2.5.1.2 Microarray Production

Microarray production was performed as described previously [22]. Briefly, the source of the clones arrayed was an *S. miltiorrhiza* root tissue cDNA library. The library was constructed in ZAP Express vector (Stratagene, La Jolla, CA). Eight thousand seven hundred and thirty-six individual phage clones were picked randomly and amplified by polymerase chain reaction (PCR) using the M13 and BK universal primers with the GeneAmp PCR system 9700 (Perkin Elmer, Foster City, CA). PCR products were purified using MultiScreen filter plates (Millipore) and eluted in 100 μL of 0.1 × TE (pH 8.0). After analysis by agarose gel electrophoresis, 4,354 samples were dried to completion, resuspended in 15 μL 50% DMSO (approximately 1 g/L), and then transferred to a 384-format plate to be subsequently used for spotting. Amplified cDNAs were spotted in duplicate onto silylated microscope slides (CEL Associates, Houston, TX) using a 16-pin print head and a custom-built arraying robot. After arraying, the slides were air-dried and stored in the dark. Each of the microarray experiments was performed in duplicate with the dyes reversed.

2.5.1.3 Preparation of Fluorescent Dye-Labeled DNA, Hybridization, Scanning, and Data Acquisition

cDNA labeled with a cyanine fluorescent dye (Cy5 or Cy3-dCTP) was produced by Eberwine's linear RNA amplification method and subsequent enzymatic reaction [44]. Specifically, double-stranded cDNAs (containing the T7 RNA polymerase promoter sequence) were synthesized from 10 lg total RNA using the cDNA synthesis system according to the manufacturer's protocol (Takara). A T7-oligo (dT) primer

(50-AAACGACGGCCAGTGATTGTAATACACTCACTATAGGCGCTTT-TTTTTTTTTTTTTT-30) was used instead of the polyT primer provided in the kit. After completion of double-stranded cDNA synthesis, cDNA products were purified using a PCR purification kit (Qiagen) and eluted with 60 μL elution buffer. One-half of the eluted double-stranded cDNA products was vacuum evaporated to 8 μL and used as a template in 20 μL in vitro transcription reactions at 37 °C for 3 h using the T7 RiboMAX Express large-scale RNA production system (Promega). The amplified RNA was purified using an RNeasy mini kit (Qiagen). Klenow enzyme labeling strategy was adopted after reverse transcription. Briefly, 2 mg amplified RNA was mixed with 2 lg random hexamers, denatured at 70 °C for 5 min, and cooled on ice. Then, 4 μL of first-strand buffer, 2 μL of 0.1 M DTT, 1 μL 10 mM dNTP, and 1.5 μL SuperScript II (Invitrogen) were added. The mixtures were incubated at 25 °C for 10 min, then at 42 °C for 60 min. The cDNA products were purified using a PCR purification kit (Qiagen) and vacuum evaporated to 10 μL. The cDNA was mixed with 2 lg random hexamers, heated to 95 °C for 3 min, and snap cooled on ice. Then, 10 μL buffer, dNTP, and Cy5-dCTP or Cy3-dCTP (Amersham Pharmacia Biotech) were added to final concentrations of 120 mM dATP, 120 mM dGTP, 120 mM dTTP, 60 mM dCTP, and 40 mM Cy-dye. Klenow enzyme (1 μL; Takara) was then added, and the reaction was performed at 37 °C for 60 min. Labeled cDNA was purified using a PCR purification kit (Qiagen) and resuspended in elution buffer. Labeled controls and test samples were quantitatively adjusted based on the efficiency of Cy-dye incorporation and mixed with 30 μL hybridization solution (50% formamide, 19 hybridization buffer; Amersham Biosciences). DNA in the hybridization solution was denatured at 95 °C for 3 min prior to loading onto a microarray. Arrays were hybridized at 42 °C overnight and then washed twice (0.2% SDS, 29 SSC at 42 °C for 5 min, then 0.29 SSC for 5 min at room temperature). Microarray data were analyzed using GenePix Pro 5.0 (Axon Instruments, Union City, CA). The scanned data were normalized by the global normalization method [45], which normalizes the image data between Cy3 and Cy5 channels by adjusting the total signal intensities of two images and removing unreliable spots. The unreliable spots were discarded based on the following screening. Spots containing clones that had poorly amplified or multiple bands, as well as those that were flagged because of a false intensity caused by dust or background on the array, were removed. Spots with <65 % of the spot intensity at >1.5-fold that of the background in both channels were ignored. Clones in one sample that had an average induction greater than twofold in another were determined as up- or downregulated. Data management and analyses were carried out using Microsoft Excel and Microsoft Access database. After normalization, we calculated the means and coefficients of variation for the observed signal intensities in each channel and the ratio of signals from two replicates.

2.5.1.4 Sequence Analysis

cDNA clones with different expression in the microarray experiments were sequenced using the Applied Biosystems dye terminator cycle sequencing Ready Reaction kit and a 3130 DNA sequencer. Vectors or imprecise nucleotides, such as

polyT and polyA, and double peaks were removed. EST assembly was performed to obtain unigenes using Staden Package (gap4) software (http://staden.sourceforge.net). The sequences were first compared with the GenBank database using BLASTX (http://www.ncbi.nlm.nih.gov/BLAST/). Genes with score values higher than 80 and identity values higher than 35 % were designated as significant homologous genes. Unigenes with E-values [5] were designated as unknown. The undesignated genes were then identified by comparison with sequences in the nonredundant nucleotide and EST databases of GenBank using the BLASTn algorithm. Gene ontology (GO) (http://www.geneontology.org) was used to describe gene and gene product attributes. All genes were classified in terms of biological process, cell component, and molecular function. Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/keg) [46] was used to identify biochemical pathways associated with hairy root development stages.

2.5.2 Results

2.5.2.1 Microarray Experiments

Microarrays were used to examine gene expressions quantitatively after hairy root dealt with YE + Ag^+. Previous research has shown that after dealt with YE + Ag^+, the secondary metabolite of tanshinone's accumulation increased dramatically [47]. In total, 4,354 unsequenced ESTs were picked from a cDNA library constructed from *S. miltiorrhiza* root tissue, amplified by PCR, and arrayed in duplicate on chemically modified microscope slides by using a robotic printing device. Experiments of comparing gene expression difference in two groups of hairy root YE + Ag^+ and control were performed. In each experiment, one mRNA population (target) was labeled with Cy3 and the other with Cy5. The labeled targets were then mixed and hybridized simultaneously to a microarray. To exclude artifacts, the researchers performed a reciprocal labeling experiment with each pair of targets, using the same techniques used in the first experiment except that the labels were exchanged. Statistically, only genes for which we had 8 data points (two duplicates per slide, two replications, and dye-swap experiments) were considered, and approximately 201 cDNA clones were selected for further analyses.

2.5.2.2 Detection of Differentially Expressed Genes

Analysis of the microarray data revealed significant changes in transcript levels of those genes for which the expression varied by more than twofold were considered to exhibit significant changes in expression. A total of 201 genes (significant at single test, $P<0.05$) were differentially expressed in hairy roots after 24 h' dealt with YE + Ag^+ compared with that of control. After sequencing and correction for redundancy (performed by sequence alignment), 196 unique differentially expressed cDNA clones were identified. Sequence alignment of the cDNA clones identified as differentially

Table 2.3 Results of KO and pathway analysis

Contig no.	KO definition and BLASTX	Pathway
chip46h02	Dimethylallyltranstransferase Geranyltranstransferase Farnesyl diphosphate synthase	Terpenoid biosynthesis
chip30h10	Ent-kaurene synthase	Diterpenoid biosynthesis
chip18f03	Isopentenyl-diphosphate delta-isomerase	Terpenoid biosynthesis
chip26g03	4-Coumarate—CoA ligase	Stilbene, coumarine, and lignin biosynthesis
chip39b01	CBL-interacting protein kinase	Benzoate degradation via CoA ligation
chip30h04	Acetyl-CoA C-acetyltransferase	Synthesis and degradation of ketone bodies, terpenoid biosynthesis
chip28b11	ATP:ADP antiporter, AAA family	Other ion-coupled transporters
chip23f06	4-Diphosphocytidyl-2-C-methyl-D-erythritol kinase	Terpenoid biosynthesis
chip16g09	Copalyl diphosphate synthase	Diterpenoid biosynthesis
chip20h03	HMG-CoA synthase	Terpenoid biosynthesis

expressed by microarray analysis revealed that several shared a high degree of sequence similarity. A total of 181 cDNA clones were successfully sequenced and acquired 130 unique genes. Among the 130 unique differentially expressed genes, 107 were categorized as genes of known function, 3 were homologues with low similarity.

2.5.2.3 Expression Profile of Genes Involved in Tanshinone and Salvianolic Acid Biosynthesis

The genes differentially expressed after dealt with YE + Ag^+ were then further analyzed using GO and KEGG pathways. On the basis of GO and KEGG analysis, candidate differentially expressed gene out of 130 unique genes were analyzed and notated. The researchers identified five genes involved in tanshinone biosynthesis: acetoacetyl-CoA thiolase (SmAACT, GenBank accession no. EF635969), 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (SmCMK, GenBank accession no. EF534309), isopentenyl diphosphate isomerase 2 (SmIPPI, GenBank accession no. EF635967), farnesyl diphosphate synthase (SmFPS, GenBank accession no. EF635968), and ent-kaurene synthase (SmKSL, GenBank accession no. EF635966); one gene involved in salvianolic acid biosynthesis. These genes were upregulated after dealt with YE + Ag^+ (Table 2.3).

2.5.3 Discussion

The "ingredient difference phenotypic cloning" being a useful and powerful method mainly involved transcript-profiling analysis of *S. miltiorrhiza* hairy root under the YE + Ag^+ treatment could display differential regulation of secondary

metabolism-related genes in *S. miltiorrhiza*. This method helps to identify more genes involved in biosynthesis of secondary metabolites of tanshinone and salvianolic acid if conditions satisfied. At the same time, the power of microarrays as a useful tool for novel gene discovery in "ingredient difference phenotypic cloning" method has been demonstrated in this study. In addition, because the cDNA clones were obtained from a recombinant cDNA library originating from the root of *S. miltiorrhiza*, the arrays are not representative of the entire transcriptome. Nevertheless, the data obtained provide an important and novel description of the expression of a large number of *S. miltiorrhiza* genes.

2.6 Case Study (II)

A new molecular identification method: anchored primer amplification polymorphism DNA

Since the inception of PCR technology, researches on the molecular identification of *Panax ginseng* and *P. quinquefolius* have attracted particular concern. In 1994, AP-PCR was used for the identification of *P. ginseng* and *P. quinquefolius* for the first time [48], and there have been 24 related reports presently, including RAPD [49–53], DNA sequence analysis [54–56], PCR-RFLP [50, 57], AFLP [58], SCAR [59], MARMS [60], repetitive sequence, DALP, minisatellite, and so on [56, 61]. Constant innovation of these methods lies in the gradual understanding of genomic information of *P. ginseng* and *P. quinquefolius* in which, MARMS is highly specific, fast, and accurate, but the primer design must be built on the basis of a large number of known sequences. RAPD is the most widely used method because there is no need to predict genome sequence, and the operation is simple and quick, but RAPD has defects including poor reproducibility, vulnerable to origin, and storage time of medicinal materials, thus restricting its application in the field of molecular identification. Therefore, it is important and difficult to explore molecular marker methods which are simple and easy to operate as well as have a good stability and strong operability in the molecular identification of Chinese materia medica. A new method reported in this chapter is based on RAPD method. And innovations were conducted on its two main factors including primers and annealing temperature. The method was named as anchored primer amplification polymorphism DNA (APAPD). First, APAPD method was established taking *P. ginseng* and *P. quinquefolius* as examples. Then, a wide range of review was conducted on the stability of its reaction system, and the stability of amplification results of different material. Meanwhile, validation and comparison were conducted combining with MARMS method reported in the literature. On this basis, APAPD method was applied to the identification of Tianhua-fen (Trichosanthes Radix) and Bai-zhi (Angelica Radix), achieving desired results. This indicates that APAPD method is a very promising new method for molecular identification of Chinese materia medica.

2.6.1 Materials and Methods

Thirty-four samples of *P. ginseng*, *P. quinquefolius*, and their adulterants; 28 samples of Tian-hua-fen; and 8 samples of Bai-zhi were collected from different areas of China. All samples were identified by the researcher Huang Luqi et al. in the Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, and they were placed in the herbarium of the institute.

Total DNA of Chinese materia medica was extracted using the modified CTAB method [62].

Primers were designed according to the existing ITS sequences of *P. ginseng*, *P. quinquefolius* (GenBank accession number: AJ786235, AY548192, U41680, U41689, U41688, U41687), Tian-hua-fen [63], and Bai-zhi [64]. Primers were about 20 bp in length. Following the general principles of primer design, areas which had large differences with adulterant sequences to be identified should be selected. MARMS primers used for the identification of *P. ginseng* by Shu Zhu et al. were applied for the verification of *P. ginseng* and *P. quinquefolius* [60]. Primers were synthesized by Sangon Biological (Shanghai) Co., Ltd.

The PCR reaction system (25 μL) was as follows: 10 mmol · L^{-1} Tris–HCl (pH 9.0), 50 mmol · L^{-1} KCl, Mg^{2+} 1.5 mmol · L^{-1}, dNTP 0.15 mmol · L^{-1}, Taq E 1U (Invitrogen, Promega, etc.), primer 0.15 μmol · L^{-1}, and template DNA 50–200 ng. PCR amplification was conducted on AB I9700 amplification instrument. Primer screening and optimization of PCR conditions were conducted on all APAPD primers first using typical materia medica (usually four samples) to be identified. Preliminary screening was conducted following RAPD general procedures: predenaturation at 94 °C for 5 min, followed by 40 cycles: denaturation at 94 °C for 45 s, annealing at 37 °C for 1 min, and extension at 72 °C for 1 min 30 s, with a final extension at 72 °C for 5 min after 40 cycles. Annealing temperature was gradually increased in primers which could produce polymorphic bands in quality products and adulterants to eliminate nonspecific bands to determine the optimal reaction PCR parameters. Amplification products were electrophoresed on a 2.0% agarose gel containing EB in 1 × TAE buffer. And observation and photographing were conducted under SYNGENE gel imaging system.

2.6.2 Results

2.6.2.1 Establishment of APAPD Method of *P. ginseng* and *P. quinquefolius*

When PCR annealing temperature was 37 °C, the primer Pg-q36F showed good effects on the amplification of *P. ginseng* and *P. quinquefolius*, presenting clear bands and significant polymorphic bands. Gradual increase in annealing temperature displayed that primers could amplify at 37–60 °C, but when at 40–50 °C, PCR amplification results were stable, single 849-bp band was amplified from *P. ginseng*,

and 864-bp and 792-bp bands were amplified from *P. quinquefolius*; the bands gradually blurred at 55–60 °C. To ensure that all sources of materia medica and adulterants could be effectively amplified, PCR conditions were determined as follows: predenaturation at 94 °C for 5 min, followed by 40 cycles: 94 °C 45 s, 40 °C 1 min, and 72 °C 1.5 min, with a final extension at 72 °C for 5 min after 40 cycles.

2.6.2.2 Study on Accuracy of APAPD Method

To test the accuracy of primer Pg-q36F identifying *P. ginseng* and *P. quinquefolius*, 11 kinds of adulterants which have been presented on the market were amplified using primer Pg-q36F, respectively. Meanwhile, all *P. ginseng* and *P. quinquefolius* samples were verified using MARMS primers PgjqtK1966R, PqtK896F, PgS481F, and P-S712R. Results showed that in primer Pg-q36F, there were only 849-bp band amplified from *P. ginseng*, 864-bp and 792-bp bands amplified from *P. quinquefolius*, while no corresponding band presented in all adulterants. In MARMS primers, all *P. ginseng* presented 649-bp and 249-bp bands, and all *P. quinquefolius* presented 649-bp band. It indicated that the identification result of primer Pg-q36F was consistent with that of primers in the literature, and quality products could be distinguished from all kinds of adulterants, indicating that primer Pg-q36F could be used as identification primer of *P. ginseng* and *P. quinquefolius*.

2.6.2.3 Study on Stability of APAPD Method

In the PCR reaction system, the quality of Taq enzyme was the main factor to affect identification results. In the MARMS identification of *P. ginseng* and *P. quinquefolius*, 249-bp band was amplified from both *P. ginseng* and *P. quinquefolius* using ordinary Taq polymerase, so they could not be identified, and correct results could be obtained only using the high-fidelity Taq polymerase. Using primer Pg-q36F, ordinary Taq polymerase of Invitrogen, Promega, and five domestic companies were selected for amplification respectively, and the results obtained from all Taq polymerase were consistent. It indicated that the primer is undemanding in PCR reaction system and ordinary Taq polymerase could meet the requirements, being easy to be reproduced in laboratories.

In the long-term cultivation process of *P. ginseng*, different farm species such as Da-maya, Er-maya, Huangguo, and changbo were presented [63]. *P. quinquefolius* is native to the USA and Canada. Since the successful introduction into China, large-scale cultivation has been started in many areas. The prices of *P. quinquefolius* showed great differences according to its different qualities, for example, 3.80 yuan/g, 1.80 yuan/g, and 0.98 yuan/g of *P. quinquefolius* were sold in Tong Ren Tang Pharmacy. In addition, a lot of *P. quinquefolius* were processed into decoction pieces, thereby increasing the difficulty in the identification of *P. ginseng* and *P. quinquefolius*. Therefore, the correct identification of different sources of herbs, such as different

areas, different prices, different processing methods, and different storage time, is the first step to ensure the safety of clinical pharmacy. Therefore, in this chapter, four farm species of *P. ginseng*, medicinal materials, samples and powder of *P. ginseng* sold in different pharmacies as well as medicinal materials, decoction pieces, and samples of different areas and different prices of *P. quinquefolius* were selected as experimental materials, with broad representation. PCR amplification was conducted on all samples using primer Pg-q36F. The results showed *P. ginseng* of different sources steadily amplified 849-bp band, and *P. quinquefolius* of different sources steadily amplified 864-bp and 792-bp bands. It indicated that *P. ginseng* and *P. quinquefolius* could be identified steadily using the primer.

2.6.2.4 Study on Applicability of APAPD Method in Chinese Material Medica of Tian-hua-fen and Bai-zhi

Among Tian-hua-fen primers TkS1-64 F, TkS2-112 F, and TkS2-130R, TkS1-64 F showed the best amplification effect, manifesting that the polymorphism of quality products and adulterants was obvious, so quality products and adulterants could be accurately identified. PCR cycles were identified as follows: predenaturation at 94 °C for 5 min, followed by 40 cycles: 94 °C 30 s, 50 °C 45 s, and 72 °C 1 min, with a final extension at 72 °C for 5 min after 40 cycles. By detecting 19 batches of Chinese material medica of Tian-hua-fen of different sources, the 560-bp and 960-bp bands were determined as characteristic identification bands of Tian-hua-fen, while other bands such as 1,930-bp, 1,400-bp, 839-bp, and 715-bp bands could be used as secondary identification bands because they could not steadily reproduce among different PCR reaction systems or material medica from different areas. Characteristic identification bands of each adulterant were the following: Trichosanthes hupehensis 686 bp, 800 bp, 938 bp, 1,260 bp; Trichosanthes laueri-bractea Hayata 686 bp, 800 bp, 938 bp, 1,260 bp; Guizhou Trichosanthes 760 bp, 1,259 bp; Trichosanthes pedata Merr. et Chun 900 bp; Trichosanthes truncata C. B. Clarke 760 bp; Momordica cochinchinensis 770 bp, 1,373 bp; Melothria hetero-phylla (Lour.) Cogn. 673 bp, 786 bp, 919bp, 1,189 bp; Trichosanthes cucumeroides Maxim 865 bp, 1,296, 2,118 bp, 2,669 bp; and Trichosanthes lepiniana (Naud.) Cogn had no amplification bands.

In Bai-zhi primers AfS1-100 F and AfS1-120R, AfS1-100 F showed obvious amplification polymorphism, Bai-zhi could be clearly distinguished from Angelica porphyrocaulis Nakai et Kitagawa, Angelica dahurica (Fisch. ex. Hoffm.) Benth. ex. Franch. dt. Sav, and Angelica amurensis Schischk. PCR cycles were determined as: predenaturation at 94 °C for 5 min, followed by 40 cycles: 94 °C 30 s, 40 °C 45 s, and 72 °C 1 min, with a final extension at 72 °C for 5 min after 40 cycles. By detecting 17 batches of Chinese material medica of Bai-zhi of different sources, 740-bp band was determined as characteristic identification bands of Bai-zhi; 740-bp, 917-bp, and 1 032-bp bands as those of Angelica porphyrocaulis Nakai et Kitagawa; 740-bp and 1 032-bp bands as those of Angelica dahurica (Fisch. ex. Hoffm.) Benth. ex. Franch. et. Sav; and 500-bp and 1 032-bp bands as those of

Angelica amurensis Schischk. Being different from polymorphic bands of *P. ginseng*, *P. quinquefolius,* and Tian-hua-fen, characteristic bands of quality products of Bai-zhi also presented in *Angelica porphyrocaulis* Nakai et Kitagawa and *A. dahurica* (Fisch. ex. Hoffm.) Benth. ex. Franch. dt. Sav, but *A. porphyrocaulis* Nakai et Kitagawa increased 740-bp and 917-bp bands compared with Bai-zhi; *A. dahurica* (Fisch. ex. Hoffm.) Benth. ex. Franch. dt. Sav increased 740-bp bands compared with Bai-zhi; the combination of their bands formed characteristic fingerprints that could be used for accurate identification.

2.6.3 Discussion

Tian-hua-fen is an important class of Chinese materia medica. Trichosanthin which can terminate pregnancy and has anti-HIV activity is extracted from the root of Trichosanthes kirilowii Maxim. Studies on Trichosanthes serving as a composite species are known as "the most intractable taxonomic problem in eastern Asia Cucurbitaceae center." There are up to 28 kinds of commercial herbs of Tian-hua-fen, including 19 species of congeneric plants; some of them are highly toxic [63]. This study group has identified Tian-hua-fen, Trichosanthes hupehensis, and Momordica cochinchinensis using RAPD technique and protein immunoassay technique [65], but the results using RAPD method are affected by storage time. Certain difficulties exist in the identification of Tian-hua-fen by applying the method [66], while primer TkS1-64F can amplify samples from different sources, and DNA fingerprint can be obtained from quality products and eight kinds of adulterants of Tian-hua-fen only using one primer, providing a guarantee for the application of materia medica of Tian-hua-fen.

Existing commercial herbs of Bai-zhi are cultivated and divided into *A. anomala* (Radix angelicae dahuricae) and *A. dahurica* (Radix angelicae dahuricae). Because the original source of wild plants has not been really figured out, the specific name of traditional Chinese medicine Bai-zhi has been changed for many times in identification, and there has been no unified final conclusion. This study group conducted exhaustive research on germplasm resources of Bai-zhi from morphology, chemical composition, ITS sequence analysis, and RAPD to prove that the source of wild germplasm of the traditional Chinese medicine Bai-zhi (including Angelica anomala, Radix angelicae dahuricae, Angelica dahurica, and Radix angelicae dahuricae) is Angelica formosana H. Boiss. only distributed in the southeast region of China (Taiwan Province based) currently. While Angelica formosana H. Boiss., *A. dahurica* (Fisch. ex. Hoffm.) Benth. ex. Franch. et. Sav, and *A. porphyrocaulis* show a close genetic relationship. Angelica am urensis Schischk is an outgroup in the study. Results of the polymorphism of primer AfS1-100F used in this chapter also support the above conclusion; characteristic identification bands of Angelica form osana de Bioss are consistent with those of 4 kinds of commercial Bai-zhi, both *A. dahurica* (Fisch. ex. Hoffm.) Benth. ex. Franch. et. Sav and *A. porphyrocaulis* contain identification bands of Bai-zhi but increase 1–2 bands compared with Bai-zhi,

while Angelica am urensis Schischk has no characteristic bands which Bai-zhi, *A. dahurica* (Fisch. ex. Hoffm.) Benth. ex. Franch. dt. Sav and *A. porphyrocaulis* commonly have, suggesting APAPD method can also be used for the study of genetic relationship among plants.

The identification of *P. ginseng*, *P. quinquefolius*, Tian-hua-fen, and Bai-zhi indicates that APAPD method has the following advantages:

① Simple and easy to operate. Although the primer design is more difficult, ideal identification primers can be obtained by designing 2 or 3 primers, thus avoiding the trouble of screening a large number of random primers

② Good stability and reproducibility. Due to the increased primer length and specificity, there are only 1–5 amplification bands generally; the origin and storage time of medicinal materials have no effect on PCR results; moreover, APAPD primers are undemanding in PCR reaction conditions, so they are easy to be promoted and reproduced in laboratories

③ Large amount of information provided. Both quality products and most of adulterants can be amplified using the method; therefore, standard identification electrophoretogram of quality products and adulterants can be established, respectively, to achieve an accurate identification of quality products and adulterants. The increasing of APAPD primers will provide standard DNA identification fingerprint for more materia medica and provide a powerful tool for the quality control of Chinese materia medica.

References

1. Huang, L. Q. (2000). *Molecular pharmacognosy*. Beijing: Beijing Medical University Press.
2. Chen, S. B., Peng, Y., Chen, S. L., et al. (2005). Introduction of pharmaphylogeny. *World Science and Technology Modernization Traditional Chinese Medicine, 7*(6), 97–103.
3. Wang, X. Y., Cui, G. H., & Gao, W. (2009). New method of "ingredient difference phonetypical cloning" for functional gene cloning from medicinal plants. *China Journal of Chinese Materia Medica, 34*(1), 40–44.
4. Zhao, J., Davis, L. C., & Verpoorte, R. (2005). Elicitor signal transduction leading to production of plant secondary metabolites. *Biotechnology Advances, 23*, 283–333.
5. Chen, X. Y. (2006). Plant secondary metabolism. *World Sci-Tech R&D, 28*(5), 1–4.
6. Kitano, H. (2002). *Foundations of systems biology, chapter systems biology: Toward system-level understanding of biological systems*. Cambridge, MA: MIT Press.
7. Hood, L., Heath, J. R., Phelps, M. E., et al. (2004). Systems biology and new technologies enable predictive and preventative medicine. *Science, 306*(5696), 640–643.
8. Oksman-Caldentey, Kirsi-Marja, Inze, Dirk, & Oresic, Matej. (2004). Connecting genes to metabolites by a systems biology approach. *Proceeding of the National Academy of Sciences of the United States of America, 101*(27), 9949–9950.
9. Kong, C. H., Xu, T., Hu, F., et al. (2000). A helopathy under environmental stress and its induced mechanism. *Acta Ecologica Sinica, 20*(5), 849–854.
10. Herms, D. A., & Mattson, W. J. (1992). The dilemma of plants: To grow or to defend. *Quarterly Review Biology, 67*, 283–335.
11. Rosenthal, G. A., & Janzen, D. H. (1979). *Variable plants and herbivores in natural and managed systems*. New York: Academic.

12. Coley, P. D., Brant, J. P., & Chapin, F. S. (1985). Resource availability and plant antiherbivore defense. *Science, 230*, 895–899.
13. Liu, H. L. (2009). Systems biology: Trending to holism. *Journal of Systems Sciences, 17*(1), 46–92.
14. Liu, C. X. (2006). Systems biology and modern research of traditional Chinese medicines. *Journal of Tianjin University of Traditional Chinese Medicine, 25*(3), 115–118.
15. Aderem, A. (2005). Systems biology: Its practice and challenges. *Cell, 121*(4), 511–513.
16. Rensink, W. A., & Buell, C. R. (2005). Microarray expression profiling resources for plant genomics. *Trends in Plant Science, 10*(12), 603–609.
17. Marnik, V., Johan, D. P., Michiel, J. T., & van Eijk, M. J. (2007). AFLP-based transcript profiling (cDNA-AFLP) for genome-wide expression analysis. *Nature Protocols, 2*(6), 1399–1413.
18. Sultan, M., Schulz, M. H., Richard, H., et al. (2008). A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science, 321*(5891), 956–960.
19. Jia, W., Jiang, J., Liu, P., et al. (2006). Application of metabonomics in complicated theory system research of traditional Chinese medicine. *China Journal of Chinese Materia Medica, 31*(8), 621–624.
20. Qi, L. W., Li, P., & Zhao, J. (2006). Metabonomics and modernization of TCM research. *World Science and Technology, 8*(6), 79–86.
21. Sumner, L. W., Mendes, P., & Dixon, R. A. (2003). Plant metabolomics: Large-scale phytochemistry in the functional genomics era. *Phytochemistry, 62*(6), 817–836.
22. Cui, G. H., Huang, L. Q., Tang, X. J., et al. (2007). Functional genomics studies of *Salvia miltiorrhiza* establish cDNA microarray of *S. miltiorrhiza*. *China Journal of Chinese Materia Medica, 32*(12), 1137–1141.
23. Gao, W., Cui, G. H., Kong, J. Q., et al. (2008). Optimizing expression and purification of recombinant *Salvia miltiorrhiza* copalyl diphosphate synthase protein in *E. coli* and preparation of rabbit antiserum against SmCPS. *Acta Pharmaceutica Sinica, 43*(7), 766–772.
24. Wang, X. Y., Cui, G. H., Huang, L. Q., et al. (2008). A full length cDNA of 4-(cytidine 5′-diphospho)-2-*C*-methyl-*D*-erythritol kinase cloning and analysis of introduced gene expression in *Salvia miltiorrhiza*. *Acta Pharmaceutica Sinica, 43*(12), 1251–1257.
25. Gao, W., Hillwig, M. L., Huang, L. Q., et al. (2009). A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. *Organic Letters, 11*(22), 5170–5173.
26. Xu, M. J. (2007). Nitric oxide: The possible key factor of secondary metabolism signal transduction network of plant cells. *Progress in Natural Science, 17*(12), 1622–1630.
27. Xu, M., Dong, J., Wang, H., et al. (2009). Complementary action of jasmonic acid on salicylic acid in mediating fungal elicitor-induced flavonol glycoside accumulation of Ginkgo biloba cells[J]. *Plant, Cell & Environment, 32*(8), 960–967.
28. Xia, L. J. (2008). *Interaction between nitric oxide and hydrogen peroxide signaling in UV-B-induced flavonoid accumulation of Hypericum perforatum cells*. Dissertation, Zhejiang Gongshang University.
29. Hiroaki, K. (2002). Systems biology: A brief overview. *Science, 295*, 1662–1664.
30. Jeyaramraja, P. R., Pius, P. K., Raj Kumar, R., et al. (2003). Soil moisture stress induced alterations in bioconstituents determining tea quality. *Journal of the Science of Food and Agriculture, 83*(12), 1187–1191.
31. Clark, R., & Menary, R. (1980). The effect of irrigation and nutrient on the yield and composition of peppermint oil. *Australian Journal of Agricultural Research, 31*, 489–498.
32. Li, X., Wang, Y., & Yan, X. F. (2007). Effects of water stress on berberine, jatrorrhizine and palmatine contents in amur cork tree seedlings. *Acta Ecologica Sinica, 27*(1), 58–64.
33. Zheng, Q. S., Liu, L., Liu, Y. L., et al. (2003). Effects of salt and water stresses on osmotic adjustment and osmotica accumulation in *Aloe vera* seedlings. *Journal of Plant Physiology and Molecular Biology, 29*(6), 585–588.
34. Yan, X. F. (2001). Ecology of plant secondary metabolism. *Acta Phytoecology Sinica, 25*(5), 639–640.

35. Zhang, L., Ding, R., & Chai, Y. E. T. A. L. (2004). Engineering tropane biosynthetic pathway in Hyoscyamus niger hairy root cultures. *Proceeding of the National Academy of Sciences of the United States of America, 101*(17), 6786–6791.
36. Dae-Kyun, R., Eric, M. P., Mario, O., et al. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature, 440*(7086), 940–943.
37. Allen, R. S., et al. (2004). RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy. *Nature Biotechnology, 22*(12), 1559–1566.
38. Dixon, R. A. (2005). Engineering of plant natural product pathways. *Current Opinion in Plant Biology, 8*, 329–336.
39. Huang, L. Q. (2011). Research strategy on molecular identification of animal medical material. *China Journal of Chinese Materia Medica, 36*(3), 234–236.
40. Yip, P. Y., Chau, C. F., Mak, C. Y., et al. (2007). DNA methods for identification of Chinese medicinal materials. *Chinese Medicine, 2*, 9.
41. Lockley, A. K., & Bardsley, R. G. (2000). DNA-based methods for food authentication. *Trends in Food Science and Technology, 11*, 67–77.
42. Dong, T. T., Ma, X. Q., Clarke, C., et al. (2003). Phylogeny of *Astragalus* in China: Molecular evidence from the DNA sequences of 5 S rRNA spacer, ITS, and 18S Rrna. *Journal of Agricultural and Food Chemistry, 51*, 6709–6714.
43. Ge, X. C., & Wu, J. Y. (2005). Tanshinone production and isoprenoid pathways in *Salvia miltiorrhiza* hairy roots induced by Ag^{+} and yeast elicitor. *Plant Science, 168*, 487–491.
44. Shi, Y. H., Zhu, S. W., Mao, X. Z., et al. (2006). Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. *The Plant Cell, 18*(3), 651–664.
45. Hihara, Y., Kamei, A., Kanehisa, M., et al. (2001). DNA microarray analysis of cyanobacterial gene expression during acclimation to high light. *The Plant Cell, 13*(4), 793–806.
46. Kanehisa, M., Goto, S., Hattori, M., et al. (2006). From genomics to chemical genomics: New developments in KEGG. *Nucleic Acid Research, 34*, D354–D357.
47. Wang, X. Y., Cui, G. H., Huang, L. Q., et al. (2007). Effect of elicitor on tanshinones' accumulation in hairy root of *Salvia miltiorrhiza*. *China Journal of Chinese Materia Medica, 32*(10), 976–978.
48. Cheung, K. S., Kwan, H. S., But, P. P. H., et al. (1994). Pharmacognostical identification of American and Oriental ginseng roots by genomic fingerprinting using arbitrarily primed polymerase chain reaction (AP-PCR). *Journal of Ethnopharmacology, 42*, 67–69.
49. Artiukova, E. V., Kozyrenko, M. M., Reunova, G. D., et al. (2000). Analysis of genomic variability of planted *Panax ginseng* by RAPD. *Molecular Biology (Mosk), 34*, 339–344.
50. Um, J. Y., Chung, H. K., Kim, M. S., et al. (2001). Molecular authentication of *Panax ginseng* species by RAPD analysis and PCR-RFLP. *Biological and Pharmaceutical Bulletin, 24*, 872–875.
51. Yukiko, T. K., Isao, A., & Ichio, I. A. (2001). Random amplified polymorphic DNA (RAPD) primer to assist the identification of a selected strain, aizu K-111 of *Panax ginseng* and the sequence amplified. *Biological and Pharmaceutical Bulletin, 24*, 1210–1213.
52. Cui, X. M., Lo, C. K., Yip, K. L., et al. (2003). Authentication of *Panax notoginseng* by 5S-rRNA spacer domain and random amplified polymorphic DNA (RAPD) analysis. *Planta Medica, 69*, 584–586.
53. Shim, Y. H., Choi, J. H., Park, C. D., et al. (2003). Molecular differentiation of *Panax* species by RAPD analysis. *Archives of Pharmacal Research, 26*, 601–605.
54. Hong, G. K., Yong, S., & Jae, E. C. (1997). The physical map of the chloroplast DNA from Korean ginseng (*Panax ginseng* C.A. Meyer). *Molecules and Cells, 7*, 136–139.
55. Andrew, C. C., Michael, K. B., Patricia, A. M., et al. (2004). Phylogeny of *Panax* using chloroplast trnC-trnD intergenic region and the utility of trnC-trnD in interspecific studies of plants. *Molecular Phylogenetics and Evolution, 31*, 894–903.
56. Shim, Y. H., Park, C. D., Kimdo, H., et al. (2005). Identification of *Panax* species in the herbal medicine preparations using gradient PCR method. *Biological and Pharmaceutical Bulletin, 28*, 671–676.

57. Fushimi, H., Komatsu, K., Isobe, M., et al. (1997). Application of PCR-RFLP and MASA analyses on 18S ribosomal RNA gene sequence for the identification of three Ginseng drugs. *Biological and Pharmaceutical Bulletin, 20*, 765–769.
58. Ha, W. Y., Shaw, P. C., Liu, J., et al. (2002). Authentication of *Panax ginseng* and *Panax quinquefolius* using amplified fragment length polymorphism (AFLP) and directed amplification of minisatellite region DNA (DAMD). *Journal of Agricultural and Food Chemistry, 50*, 1871–1875.
59. Wang, J., Ha, W. Y., Ngan, F. N., et al. (2001). Application of sequence characterized amplified region (SCAR) analysis to authenticate Panax species and their adulterants. *Planta Medica, 67*, 781–783.
60. Zhu, S., Fushimi, H., Cai, S. Q., et al. (2004). Species identification from Ginseng drugs by multiplex amplification refractory mutation system (MARMS). *Planta Medica, 70*, 189–192.
61. Leung, O. C., Ho, I. S. H., & Leung, F. C. C. (2002). Isolation and characterization of repetitive DNA sequences from Panax ginseng. *Molecular Genetics and Genomics, 266*, 951–961.
62. Cui, G. H., Tang, X. J., & Huang, L. Q. (2006). DNA isolation from starch and polysaccharide-rich plants. *China Journal of Chinese Materia Medica, 3*, 1365–1367.
63. Huang, L. Q., Yue, Z. X., Yang, B., et al. (1999). Systematic study of Trichosanthes L. *Journal of Jiangxi University of Traditional Chinese Medicine, 11*, 75–78.
64. Huang, L. Q. (2004). Systemic study on the germplasm resources of Dahurian angelica. *Journal of Jiangxi University of Traditional Chinese Medicine, 16*, 5–7.
65. Feng, C. Q., Huang, L. Q., Liu, C., et al. (2005). Study on Radax Trichosanthis identification using protein immunization detection technology. *Chinese Pharmaceutical Journal, 8*, 574–577.
66. Huang, L. Q., Wang, M., Yang, B., et al. (1999). Authentication of the Chinese drug tian-hua-en (RadixTrichosanthes) and its adulterants and substitutes using random amplified polymorphicDNA (RAPD). *Chinese Journal of Pharmaceutical Analysis, 1*9, 233–238.

Chapter 3
Molecular Identification of Traditional Medicinal Materials

Ming Li, Paul Pui-Hay But, and Pang-Chui Shaw

Abstract Traditional medicines are consumed by 80% of the population in the world for health maintenance and disease treatment. Adulteration and substitution of the source materials have been life-threatening problems growing along with their popularity. Consequently, a reliable identification method is important for safety and quality assurance of the traditional medicinal materials. Molecular techniques provide alternative means to conventional organoleptic and chemical authentication methods and are often more superior in accuracy, sensitivity, resolution and reproducibility. Since the early 1990s, a number of molecular techniques have been developed to identify traditional medicinal materials based on DNA fingerprinting (RFLP, AP-PCR, RAPD, AFLP, DALP, ISSR, PCR-RFLP, SCAR and isothermal amplification), DNA microarray and DNA sequencing (DNA barcoding and FINS). These techniques are capable of differentiating traditional medicinal materials and their adulterants and substitutes, and in some cases, distinguishing closely related species, subspecies, varieties, cultivars and species from different localities. This chapter introduces the major molecular identification techniques and reviews their applications in the identification of animal and botanical medicinal materials.

M. Li
State Key Laboratory of Phytochemistry and Plant Resources
in West China (CUHK), Institute of Chinese Medicine,
The Chinese University of Hong Kong, Shatin, NT, Hong Kong, P.R. China
e-mail: liming@cuhk.edu.hk

P.P.-H. But
School of Life Sciences, The Chinese University of Hong Kong,
Shatin, NT, Hong Kong, P.R. China
e-mail: paulbut@hotmail.com

Pang-Chui Shaw (✉)
State Key Laboratory of Phytochemistry and Plant Resources in West China (CUHK),
Institute of Chinese Medicine, and School of Life Sciences, Rm. 180, Science Centre,
The Chinese University of Hong Kong, Shatin, NT, Hong Kong, P.R. China
e-mail: pcshaw@cuhk.edu.hk

Lu-qi Huang (ed.), *Molecular Pharmacognosy*,
DOI 10.1007/978-94-007-4945-0_3,

3.1 Introduction

Traditional medicines have been used worldwide for centuries, and over 80% of the world population use traditional medicines to maintain health and cure diseases [1]. A fundamental prerequisite for the proper delivery of healthcare with traditional medicines is the use of authentic herbal materials. When adulterants or erroneous substitutes are dispensed instead, they could compromise treatments or even cause adverse reactions. In the early 1990s in Belgium, rapidly progressive interstitial fibrosis and end-stage renal failure were observed in some 80 women taking a slimming regimen of herbal medicinal product made from various herbs including Fangji (Stephaniae Tetrandrae Radix) and Houpo (Magnoliae Officinalis Cortex) [2]. It was later revealed that the herb Fangji was adulterated by another herb Guangfangji derived from *Aristolochia fangchi* which contains the carcinogenic aristolochic acids [3]. Many more cases of aristolochic acid nephropathy have subsequently been reported in many Western and Asian countries [4–8]. In addition, substitution of the traditional medicinal herb Lingxiaohua (Campsis Flos) derived from *Campsis grandiflora* by a toxic herb Yangjinhua (Daturae Flos) derived from *Datura metel* caused four cases of herbal poisoning in Hong Kong [4]. In 1996, the herb Weilingxian (Clematidis Radix et Rhizoma) derived from *Clematis* species was substituted by a herb derived from *Podophyllum hexandrum* [9]. Subsequently, several poisoning cases were reported worldwide, drawing global attention to the severe side effects and life-threatening consequences of adulteration of medicinal materials and their products [10, 11]. Adulteration is due to: (1) erroneous adulteration caused by sharing of similar features or absence of distinguishable characters, (2) intentional substitution of high-value materials by inexpensive substances, (3) misuse caused by sharing of similar common names, and (4) historical use of local substitutes. In order to ensure safety, efficacy and quality of traditional medicines and their products, identification of medicinal materials is necessary.

There are a number of effective identification methods which evolve with the improvement of technologies. In the past, identification of medicinal materials is based on the description of morphological features as stated in *Shengnong Bencaojing* (~200 AD). In a later record, *Bencao Gangmu* (1,593 AD), morphological features were graphically illustrated. Nowadays, morphological and microscopic features provide first-line identification of medicinal materials. These methods complemented with chemical profiles obtained from thin-layer chromatography (TLC), high-pressure liquid chromatography (HPLC), or liquid chromatography/mass spectrometry (LC/MS) are applied to increase the accuracy of identification. In 1990s, the introduction of molecular techniques was a major breakthrough in the history of identification of traditional medicines. Recently, identification of living organisms, including medicinal materials, by DNA barcodes has been proposed [12–15]. The DNA barcode initiative provides an international standard reference for organism identification. In the *Pharmacopoeia of the People's Republic of China* (2010 edition), molecular techniques have been added as standard means of

identification for three medicinal materials including Beimu (Fritillariae Cirrhosae Bulbus), Wushaoshe (Zaocys) and Qishe (Agkistrodon). It is foreseeable that molecular protocols will be included for more medicinal materials in future editions. This chapter reviews and comments on the commonly used molecular authentication techniques. An account on the strategies and examples for identifying plant and animal medicinal materials at different taxonomic levels is also included.

3.2 Methodologies of Identifying Medicinal Materials

Assurance of the correct use of medicinal materials is fundamental for the development of traditional medicine industry. The traditional identification methods based on organoleptic and microscopic features, such as shape, color, texture, odor, tissue arrangement and cell components, are simple and inexpensive. However, these methods are subjective and depend heavily on the experience and judgment of the inspector. Also, insufficient informative characters in processed materials may lead to low accuracy and limited resolution. Alternatively, chemical profiling has become a standard practice for species identification and quality control. However, chemical components vary with a number of factors including growing stage, harvest time, locality, storage condition, processing method and manufacturing procedure. The presence of large amounts of proteins, polysaccharides, resins, tannins and thousands of secondary metabolites makes chemical analyses difficult [16].

Molecular authentication based on the variation of DNA sequences in different organisms provides an alternative approach. In principle, the genetic makeup is unique to a species independent to body parts, growing stage, and environment. Therefore, DNA-based identification methods are less sensitive to biological, physiological, physical and environmental factors. In addition, benefited from the development of polymerase chain reaction (PCR), a small amount of sample is sufficient for carrying out the authentication process. These advantages are particularly important in identifying shredded materials or powder, not to mention expensive materials with limited supply [17]. Furthermore, DNA is relatively stable and may be extractable from herbarium specimens, processed food and commercial products. Therefore, DNA technique is applicable to a wide range of forensic issues. Our group has pioneered in using molecular techniques to identify medicinal materials. In the mid-1990s, we applied arbitrarily primed polymerase chain reaction (AP-PCR) DNA fingerprinting to distinguish Oriental ginseng roots (*Panax ginseng*) from American ginseng (*P. quinquefolius*) [18, 19]. Thereafter, the application of molecular techniques has become popular. Various strategies, such as forensically informative nucleotide sequencing (FINS), DNA barcoding and isothermal amplification, have now been introduced to increase accuracy and efficiency [13, 20, 21]. There are three main molecular techniques being used, namely, DNA fingerprinting, DNA sequencing and DNA microarray. A general evaluation of the various identification methods is shown in Table 3.1.

Table 3.1 Comparison of various identification methods

Method	Skill	Sensitivity	Accuracy	Resolution	Reproducibility	Cost
Organoleptic inspection	++	+	+	+	+	+
Microscopic study	++	+	+	+	++	+
Chemical profiling	++++	+++	++	+++	++	+++
DNA fingerprinting	+++	+++	+++	+++	+++	+++
DNA microarray	++++	+++	+++	+++	+++	++++
DNA sequencing	+++	++++	++++	++++	++++	+++

Notes: "+" indicates the lowest degree, and "++++" indicates the highest degree. "++" and "+++" indicate degrees in between.

3.2.1 DNA Fingerprinting

DNA fingerprinting explores the DNA polymorphism in the whole genome or in a specific region of the sample. The polymorphic patterns are usually visualized by agarose or polyacrylamide gel electrophoresis or capillary electrophoresis. Unlike fresh materials, the quantity and quality of DNA in a medicinal material may be poorly preserved due to post-harvest processing and storage. DNA fingerprints without DNA amplification, such as restriction fragment length polymorphism (RFLP), are less applicable because of the low yield of extracted DNA and poor integrity of genomic DNA. Consequently, PCR-based DNA fingerprinting is the preferred practice. These fingerprints include arbitrarily primed PCR (AP-PCR), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), direct amplification of length polymorphism (DALP), inter-simple sequence repeat (ISSR), PCR restriction fragment length polymorphism (PCR-RFLP) and sequence-characterized amplification region (SCAR).

3.2.1.1 Arbitrarily Primed PCR (AP-PCR)

Arbitrarily primed PCR (AP-PCR), or arbitrarily chosen primer PCR (ACP-PCR), is a whole-genome fingerprint approach first reported in 1990 [22]. Multiple loci are amplified using a single primer of approximately 20 nucleotides which anneals to the genomic DNA template at a number of sites and acts as both the forward and reverse primers. When two annealing sites are close enough, such as within two kilobases or less, the DNA in between can be successfully amplified under normal PCR conditions. The number of primer sites and the match of primers with the primer sites contribute to the polymorphic DNA fingerprints among samples. Since the primer anneals to the genomic DNA arbitrarily, AP-PCR does not require prior knowledge of the target genome, and multiple loci can be examined simultaneously. There have been many publications using this approach to identify medicinal materials. For example, our group found that all the AP-PCR fingerprints generated using three primers (M13 forward, M13 reverse, and Gal-K primer) successfully differentiated the dried roots of Oriental ginseng (*Panax ginseng*) from American

ginseng (*P. quinquefolius*) [18]. Similar approach was subsequently applied to the identification of other medicinal species, including Kudidan (Elephantopi Herba), Pugongying (Taraxaci Herba), and Dangshen (Codonopsis Radix) [19, 23–25].

3.2.1.2 Random Amplified Polymorphic DNA (RAPD)

Random amplified polymorphic DNA (RAPD) is another whole-genome fingerprint [26]. The principle of RAPD is quite similar to AP-PCR and differs only from the use of a single primer of 10 nucleotides under reduced stringent conditions. The polymorphic fingerprints are due to the number of primer sites, nucleotide polymorphism in the primer sites, and the distance between adjacent primer sites. RAPD also does not require prior knowledge of the genome and can be used to examine multiple loci simultaneously. In AP-PCR and RAPD, the quality and integrity of genomic DNA remain major concerns. Besides, AP-PCR and RAPD markers are dominant markers and are usually unable to distinguish homozygous loci from heterozygous loci. Our group applied both RAPD and AP-PCR to differentiate medicinal *Panax* species from their adulterants [19]. It was shown that polymorphic fingerprints of *Panax ginseng, P. quinquefolius*, and *P. notoginseng* can be generated with appropriate RAPD and AP-PCR primers, and these fingerprints can differentiate *Panax* species from the adulterants derived from *Mirabilis jalapa, Phytolacca acinosa, Platycodon grandiflorum* and *Talinum paniculatum*. Furthermore, the degree of similarity among the RAPD and AP-PCR fingerprint suggested that *P. ginseng* is more closely related to *P. quinquefolius* than to *P. notoginseng*. Similar approach of RAPD fingerprinting was applied to identify Kudidan, Dangshen, and five medicinal *Dysosma* species [23, 25, 27]. Apart from the identification of medicinal materials, RAPD was also used to assess the genetic diversity of wild populations and cultivars [28–30].

3.2.1.3 Amplified Fragment Length Polymorphism (AFLP)

The principle of amplified fragment length polymorphism (AFLP) is to amplify a subset of DNA restriction fragments from the genomic DNA by restriction enzymes [31]. The genomic DNA is first digested with restriction enzymes (e.g., *Eco*RI and *Mse*I) at various restriction sites in multiple loci to generate restriction fragments with sticky ends. Synthetic adaptors are then ligated to these ends which act as the annealing sites of specific primer for subsequent amplification by PCR under stringent conditions. The amplified fragments are separated by highly resolving polyacrylamide gel and visualized using autoradiography, fluorescence or silver-staining techniques. Similar to AP-PCR and RAPD, AFLP screens multiple loci of the whole genome randomly and simultaneously and does not require prior knowledge of the sequence information. AFLP can detect more loci and generate more polymorphic fragments than RAPD and can be used to differentiate closely related species [32]. However, DNA degradation in medicinal materials may affect the

reproducibility of the polymorphic patterns. Our group used AFLP to differentiate closely related medicinal species such as the Oriental ginseng (*P. ginseng*) and American ginseng (*P. quinquefolius*) from various localities [33, 34]. Other examples of using AFLP include the identification of *P. japonicus*, medicinal *Plectranthus* species, and *Cannabis sativa* [35–37].

3.2.1.4 Direct Amplified Length Polymorphism (DALP)

Direct amplified length polymorphism (DALP) is a modified AP-PCR fingerprinting in which the 5′-end of the forward primer contains the core sequence of a universal primer (e.g., M13 sequencing primer), and thus the resultant fragments can be sequenced directly using the universal primer. DALP is an advanced fingerprinting method which allows simultaneous detection of a large number of polymorphic loci and simplifies the recovery and analysis of polymorphic fragments. Our group adopted this method to distinguish Oriental ginseng (*P. ginseng*) and American ginseng (*P. quinquefolius*) [38]. A 636 bp polymorphic DALP fragment amplified using primers DALP001 and DALPR1 was present in *P. ginseng* but absent in *P. quinquefolius*. This fragment was sequenced and specific primers were designed to allow rapid identification by amplifying this *P. ginseng*-specific fragment.

3.2.1.5 Inter-simple Sequence Repeats (ISSR)

Simple sequence repeats (SSR), also known as microsatellites, are tandem repeats of a few base pairs distributed throughout the genome. ISSR fingerprinting is a whole-genome scanning fingerprint which uses PCR primers designed based on the repeats found in other species [39]. As the PCR primers are based on the sequence repeats, such as $(CA)_n$, or with a degenerate 3′-anchor, such as $(CA)_8RG$ or $(AGC)_6TY$, this method does not require prior knowledge of sequence information to generate a large number of resultant fragments. ISSR fingerprinting is easy to use. It is useful to construct genetic maps and to study generic variation within populations of a species. Our group has recently employed ISSR to differentiate Huajuhong (Citri Grandis Exocarpium) derived from *Citrus grandis* "Tomentosa" from other *Citrus* varieties and cultivars [40]. A total of six ISSR primers ($(CA)_8G$, $(GT)_8A$, $(AC)_8G$, $(CA)_8RG$, $(AC)_8YT$, and $BHB(GA)_7$) were used to reveal the relationship of 23 *Citrus* samples. The six primers generated 57 bands in which 52 (91.2%) of them were polymorphic across the 23 *Citrus* samples. Cladistic analysis based on the band polymorphism of the ISSR fingerprints showed that the cultivar *Citrus grandis* "Tomentosa" was clearly distinguished from *C. grandis* and other *Citrus* species. ISSR fingerprint was also applied to identify *Cannabis sativa* and *Cistanche* species [41, 42]. It was also used to study the genetic diversity of *Salvia miltiorrhiza* and *Vitex rotundifolia* [43, 44].

3.2.1.6 PCR Restriction Fragment Length Polymorphism (PCR-RFLP)

PCR restriction fragment length polymorphism (PCR-RFLP) amplifies a specific region of the genome followed by restriction digestion to produce restriction polymorphic profiles. The specific region should be readily amplified using universal or specific primers. Standard DNA barcodes with high sequence variation, such as the internal transcribed spacer (ITS), are good candidate regions to start with. Restriction digestion of the amplified fragment (e.g., *Hinf*I, *Taq*I and *Sau*3A1) generates restriction fragments of different sizes. Mutations creating or disrupting a restriction site are the key to produce polymorphic fingerprints for sample discrimination. Although data interpretation of PCR-RFLP is simple, the discriminating ability of DNA polymorphism is less than that of ISSR and AFLP. Our laboratory successfully applied PCR-RFLP to discriminate various *Panax* species from the adulterants by amplifying the ITS region followed by restriction digestion using *Hinf*I, *Taq*I, and *Sau*3A1 [45]. We also differentiated Dangshen derived from *Codonopsis pilosula*, *C. tangshen*, *C. modesta,* and *C. nervosa* var. *macrantha* from their adulterants by digesting the ITS region using *Hinf*1 and *Hha*I [46]. Similar approach was applied to identify *Alisma orientale*, *Sinopodophyllum hexandrum,* and *Artemisia* species [47–49].

3.2.1.7 Sequence-Characterized Amplification Region (SCAR)

Sequence-characterized amplification region (SCAR) is a specific region fingerprinting based on the DNA sequences of polymorphic fragments obtained from a whole-genome fingerprint, such as RAPD or ISSR. The polymorphic fragment is cloned and sequenced for designing a pair of specific PCR primers to amplify the concerned polymorphic fragment. The amplification of the polymorphic fragment or the size difference of the fragments in different samples provides a means for differentiating the samples. This technique focuses on a single locus and is usually reproducible under high stringent PCR conditions. To increase the accuracy of differentiation, several SCAR of a sample are analyzed. SCAR requires prior information of the sequence of the polymorphic fragment for specific primer design. Degradation within the DNA fragment and the presence of PCR inhibitors may lead to false-negative results. We found a 25 bp insertion in a RAPD fragment of *P. ginseng* converted to a SCAR marker for differentiating *P. ginseng* and *P. quinquefolius* [50]. We also applied similar approach to identify medicinal snakes and crocodiles [51, 52]. Other similar work included the differentiation of *Artemisia* species, *Phyllanthus emblica,* and *Lycium barbarum* [53–55].

3.2.1.8 Isothermal Amplification

Conventional PCR amplifies DNA fragments through thermocycles for denaturing of double-strand DNA, annealing of primers, and synthesizing of new strand. Isothermal amplification is a technique allowing DNA amplification without

thermocycling, and thus, DNA amplification can be achieved without PCR machines. These techniques are mostly applied for on-site detection of viral and bacterial infections in undeveloped regions where laboratory equipment is limited. There are several ways to perform isothermal amplification. For example, strand displacement amplification (SDA) technique starts with an initial step of denaturing DNA template at 95 °C for 4 min followed by a 2 h incubation at 37 °C for primer annealing and DNA amplification [56, 57]. The ability of exonuclease-deficient Klenow DNA polymerase to extend the 3′-end and displace the downstream DNA strand leads to exponential amplification as the displaced single-strand DNA serves as the template for the synthesis of complementary strands. Double-strand DNA is digested with restriction enzymes *Hinc*II at the recognition site in the SDA primers to create nicks, and Klenow DNA polymerase extends the 3′-end and displaces the downstream strand, and therefore, single-strand DNA templates are continuously produced by strand displacement.

Loop-mediated isothermal amplification (LAMP) is another isothermal amplification technique with impressive specificity, efficiency, and rapidity [58]. Four special primers designed from six alleles (two alleles for the forward and reverse outer primers, respectively, and two alleles for the forward inner primer and two alleles for the reverse inner primer) are used to create "loops" at the end of DNA strands which significantly speed up the process of LAMP, and the whole process can be finished in 1 h. Amplification progress can be accelerated by additional loop primers to achieve amplification in 30 min [21]. Recently, LAMP was applied to identify herbal medicinal materials such as differentiating *Curcuma longa* from *C. aromatica* based on the trnK gene [59]. LAMP was also used to discriminate *Panax ginseng* from *P. japonicus* based on the 18S rRNA gene [21]. LAMP is efficient and sensitive when all the primers match the target DNA. However, primer design is difficult because many combinations of primers are needed. The primer sites should be conserved regions with minimum intraspecific variations. DNA degradation in dried or processed materials may give false-negative results. Integrity control of the amplified region may be necessary to prove that negative amplification is independent to DNA degradation.

Helicase-dependent amplification (HDA) is an isothermal amplification technique that unwinds double-strand DNA by helicase in the presence of single-strand DNA-binding proteins [60]. HDA can be performed with or without an initial denaturation step at 95 °C. Helicase unwinds DNA duplex, and primers anneal to binding sites followed by amplification of complementary strand by DNA polymerase. The double-strand DNA is separated by helicase, and the chain reaction repeats itself. HDA is relatively easy to set up because primer design is not as complicated as LAMP. This technique has been widely applied to detect virus and bacterial strains, but it has not yet been applied to identify medicinal species. The size of amplicon in HDA is restricted to around 100 bp, which is ideal for samples with degraded DNA content such as the processed medicinal materials. However, validation of DNA integrity of the studying region should be carried out to avoid false-negative results.

3.2.2 DNA Microarray

DNA microarray is a hybridization-based technology using labeled nucleotide probes to hybridize single or multiple loci in a target genome. The probes are short nucleotide fragments obtained either from restriction digestion or synthetic oligonucleotides. They are fixed on a supporting matrix where hybridization of probes and tested DNA samples takes place. Our group amplified the internal transcribed spacer (ITS) of 16 *Dendrobium* species and used them as probes to identify medicinal *Dendrobium* species in a prescription with multiple herbs [61]. The ITS2 region of the tested samples were labeled with Cy3 fluorescent dye and allowed to hybridize to the ITS probes. Species-specific fluorescent signal was obtained to clearly identify the five medicinal *Dendrobium* species. We have also applied similar approach using 5S rDNA intergenic spacer as probes to differentiate *D. officinale* from other closely related *Dendrobium* species [62].

3.2.3 DNA Sequencing

DNA sequencing is one of the most definitive means for identification as this technique can directly assess sequence variations on a defined locus. It also provides informative characters to reveal phylogenetic relationship. With the decrease of sequencing cost, identification of medicinal materials using DNA sequencing has become a routine practice. The commonly used DNA regions for medicinal materials identification include nuclear internal transcribed spacer (ITS) and 5S rDNA intergenic spacer (5S), chloroplast trnH-psbA intergenic spacer (trnH-psbA), large subunit of the ribulose-bisphosphate carboxylase (rbcL), maturase K gene (matK), trnL intron (trnL), trnL-trnF intergenic spacer (trnL-F), mitochondrial control region (CR), cytochrome c oxidase subunit 1 (COI), and cytochrome b gene (Cyt b). These regions have different evolutionary rates and therefore possess different variability. For example, the mitochondrial COI region is slowly evolved, and only a few variations were observed in the 1.4 kb COI sequences in flowering plants [63]. However, this region evolve rapidly and is varied enough to discriminate most animal species. To differentiate medicinal materials from adulterants derived from closely related species, it is essential to search for DNA regions with high discriminative power. In 2003, the concept of barcoding global species by selected DNA regions was first proposed [14], and substantial effects have been put on the screening of appropriate DNA barcodes. Until recently, it is generally agreed that the chloroplast rbcL and matK regions are the standard DNA barcodes for higher plants, and the chloroplast trnH-psbA region and nuclear ITS region are the complementary DNA barcodes. For animals and fungi, the mitochondrial COI and nuclear ITS regions are the appropriate DNA barcodes, respectively [12, 14, 15]. These DNA barcodes have been proven to be useful not only

in biodiversity and conservation studies but also in the identification of medicinal materials. For example, our group sequenced the ITS region to differentiate six *Panax* species from their adulterants derived from *Mirabilis jalapa* and *Phytolacca acinosa* [45]. We also used the ITS region to identify medicinal *Dendrobium* species [64], Muxiang (Aucklandiae Radix, Vladimiriae Radix, and Inulae Radix) [65], Baihuasheshecao (Hedyotii Herba) [66], Huajuhong (Citri Grandis Exocarpium) [40], and Leigongteng (Tripterygii Radix et Rhizoma) [67]. Chloroplast trnH-psbA region is another highly varied DNA barcode for identifying Madouling (Aristolochiae Fructus) [20] and Wutou (Aconiti Radix and Aconiti Kusnezoffii Radix) [68]. Apart from the standard DNA barcodes, a few regions are also useful for identifying medicinal materials. For example, the nuclear 5S region was used to identify Dangshen [69], medicinal *Swertia* species [70], Muxiang [65] and Leigongteng [67]. Furthermore, the chloroplast trnL region was used to identify Baibu (Stemonae Radix) [71], and the trnL-F region was used to identify Madouling [20].

3.2.3.1 Forensically Informative Nucleotide Sequencing

DNA sequencing is useful for differentiating groups of materials, but their identities are not known unless their DNA sequences are compared with reference sequences, and this kind of work is called forensically informative nucleotide sequencing (FINS) [72]. FINS was first applied to trace the origin of animals and their products, but its application has now been extended to the identification of unknown medicinal samples [20, 66, 73]. To identify an unknown sample, a selected DNA region was amplified from the DNA extract and sequenced. The FINS approach emphasizes the comparison of the unknown sequence with the sequences of suitable reference species for revealing the identity of a sample. The resolution depends on the discrimination ability of the selected DNA region and the phylogenetic distance between the reference and the unknown sample. With the DNA barcode initiative, many DNA sequences of medicinal materials have now been deposited in public databases. Our group has recently constructed a freely access online database, Medicinal Materials DNA Barcoding Database (http://www.cuhk.edu.hk/icm/mmdbd.htm), which contains approximately 20,000 DNA sequences from 1,300 medicinal materials [74]. Although the accuracy of many of these sequences has not been substantiated, these sequences are nevertheless valuable reference resources for FINS analysis. Our group applied FINS based on ITS region successfully revealed that four samples of Baihuasheshecao retailed in Hong Kong (PR China) and Boston (USA) were adulterants derived from *Hedyotis corymbosa* (Rubiaceae), and the three samples from Guangzhou (PR China) are genuine and derived from *H. diffusa* [66]. Similar approach was used to reveal the identities of the retailed samples of snake meat, Madouling and Leigongteng [20, 67, 75].

3.3 Molecular Identification of Botanical Medicinal Materials

Approximately 90% of medicinal materials recorded in the *Pharmacopoeia of the People's Republic of China* (2010 edition) are derived from botanical sources. The huge international market of herbal medicinal materials suggests the importance of their correct identification. The pharmacological effects of herbal medicinal materials may vary among closely related species, subspecies, varieties, cultivars, and localities, not to mention the adulterants derived from distantly related species. Apart from conventional organoleptic and chemical methods, molecular approach provides an alternative and definite method to identify these samples.

3.3.1 Discrimination at Inter-family and Inter-genus Levels

Adulteration of herbal materials by distantly related species from different families or genera is common. Molecular identification of these adulterants is relatively easy as their genetic makeups are quite different from the genuine species. DNA fingerprinting techniques usually show clear-cut results. For example, AP-PCR, RAPD and RFLP fingerprints of medicinal *Panax* species in family Araliaceae showed different patterns from the adulterants in families Nyctaginaceae, Phytolaccaceae, Campanulaceae, and Talinaceae [19, 45]. DNA sequencing is also useful to discriminate distantly related species. For example, trnL region is able to distinguish medicinal *Stemona* species in family Stemonaceae from adulterants in family Asparagaceae [71]. Similarly, trnL-F and trnH-psbA regions were used to distinguish Madouling derived from *Aristolochia* species (Aristolochiaceae) from the substitute derived from *Cardiocrinum* species (Liliceae) [20]. Identification of materials of different genera can also be achieved by DNA techniques. AP-PCR and RAPD were able to discriminate materials belonging to eight genera in family Asteraceae and identify the herbs Kudidan and Pugongying [23, 24]. Similarly, PCR-RFLP may be applied to differentiate four *Codonopsis* species (Campanulaceae) from two adulterants derived from *Campanumoea* and *Platycodon* species in family Campanulaceae [46]. DNA sequencing of ITS region was applied to distinguish 16 medicinal *Dendrobium* species from *Pholidota* species in the same family Orchidaceae [64]. Although DNA sequencing is useful to differentiate samples derived from distantly related species, such as at the family and genus levels, choosing a suitable DNA region is crucial. Some DNA regions, such as ITS and 5S, evolve rapidly and their sequence similarities at species level in some families are low. For example, the sequence similarity of ITS and 5S regions among Muxiang species (Asteraceae) and the toxic adulterants in Aristolochiaceae were only 56–58% and 20–30%, respectively [65]. Although such low similarity does

not affect the differentiation of samples in different families, it may make sequence alignment and phylogenetic tree construction difficult.

3.3.2 Discrimination at Inter- and Intra-species Levels

One of the major advantages of molecular identification is its high resolution which allows differentiation samples at inter- or intra-species level. DNA fingerprinting, such as AP-PCR, RAPD, SCAR, DALP, and AFLP, readily differentiated closely related species of *P. ginseng* from *P. notoginseng* [18, 34, 38, 50]. DNA microarray with hybridization probes designed based on ITS and 5S sequences successfully detected several medicinal *Dendrobium* species [61, 62]. Choosing an appropriate DNA region with high variability and discrimination power is crucial for differentiation of closely related species by DNA sequencing. For example, trnL is a relatively conserved region which could differentiate medicinal *Stemona* species (Stemonaceae) from adulterants derived from *Asparagus* species (Asparagaceae) but failed to discriminate the medicinal species (*S. japonica, S. sessilifolia* and *S. tuberosa*) and another closely related species *S. parviflora* [71]. On the contrary, the ITS, 5S and trnH-psbA regions are highly varied regions which are commonly used for identification at species level. The ITS region is varied enough to discriminate all 16 medicinal *Dendrobium* species with inter-specific divergences ranging from 2 to 17% [64]. This region was also used to authenticate Baihuasheshecao derived from *Hedyotis diffusa* (Rubiaceae) and resolved all the 14 *Hedyotis* species studied [66]. In fact, the ITS-2 region is highly varied and found useful for discriminating most medicinal species and therefore has recently been proposed to be a DNA barcode for medicinal plants [13]. Although ITS shows high sequence variability among species and is the most frequently used region for species identification of herbal medicinal materials, the presence of multiple copies, which may be non-homogeneous, and the problem of secondary structure resulting in poor-quality sequence data are major drawbacks [76, 77]. Molecular cloning prior to DNA sequencing is necessary to solve these problems. Besides, fungal contamination is common in herbal medicinal materials and would interfere proper amplification of target ITS sequences by universal primers. Specially designed plant-specific primers should be used in such conditions. The 5S region is a highly varied region and frequently used for species and subspecies differentiation. It readily discriminated *Swertia mussotii* from *S. chirayita, S. franchetiana*, and *S. wolfgangiana* with interspecific divergences ranged from 31 to 65% [70]. It also differentiated Dangshen derived from *Codonopsis pilosula* and *C. pilosula* var. *modesta* with intra-specific similarity of 95–98%, respectively, and interspecific similarity ranged from 70 to 73% [69]. In our experience, however, the sequence of 5S region is sometimes too varied, making it difficult for sequence alignment. Moreover, this region has multiple copies and molecular cloning prior to sequencing is essential. TrnH-psbA region

is a complementary DNA barcoding region showing the highest amplification successful rate and discrimination rate among 9 tested loci [15, 78]. It is used to identify 19 *Aconitum* species with an average inter-specific similarity of 85% [68]. The two closely related medicinal species, *A. carmichaeli* and *A. kusnezoffii*, were clearly distinguished by a 56 bp sequence inversion in their trnH-psbA sequences. A disadvantage of the trnH-psbA region is the presence of poly-A structure which reduces the successful rate of DNA sequencing. Besides, sequence alignment may be difficult due to the frequent presence of nucleotide insertion and deletion. In spite of the highly discriminative ability at species level, trnH-psbA could not resolve the relationship between *Cardiocrinum giganteum* and its variety *C. giganteum* var. *yunnanense*, but the trnL-F region could [20]. This example demonstrated that there is no single universal locus suitable for differentiating all taxa at different levels. Searching for a suitable region that suits the purpose is not avoidable.

3.3.3 Discrimination among Cultivars and Geographical Culture Origins

Herbal medicinal materials derived from various cultivars or collected from different geographical origins may be traced using molecular techniques. For example, the herb Huajuhong is derived from *Citrus grandis* or its cultivar *C. grandis* "Tomentosa." ISSR fingerprinting using six primers generated 57 DNA fragments which readily differentiated four samples of *Citrus grandis* from 15 samples of *C. grandis* "Tomentosa." Although there were a few nucleotide substitutions in their ITS sequences, cladistic analysis showed that ITS was unable to differentiate these cultivars as they could not form distinct clusters [40]. AP-PCR fingerprints of Dangshen collected from different geographical origins in China showed that samples from Sichuan and Hubei generated a characteristic fragment of 0.8 kb using the primer OPC-02. Specific fragments of 1.15, 0.63, and 1.15 kb were obtained in samples from Shanxi, Sichuan, and Gansu, respectively, using the primer OPC-04. The AP-PCR primer OPC-05 amplified specific fragments of 1.25 and 1.6 kb for samples from Gansu, while a 0.9 kb fragment is characteristic in Hubei samples [25].

3.4 Molecular Identification of Animal Medicinal Materials

Animals account for 7% of the Chinese medicinal materials in the *Pharmacopoeia of the People's Republic of China* (2010 edition). Some of the materials are expensive. These include Ejiao (Asini Cornii Colla) derived from *Equus asinus* and

Lurong (Cervi Cornu Pantotrichum) derived from *Cervus nippon* and *C. elaphus*. They are frequently adulterated by less expensive products with similar organoleptic features. Correct identification also avoids the misuse of endangered species.

Both DNA fingerprinting and sequencing have been shown to be useful for differentiating animal medicines. For example, in the *Pharmacopoeia of the People's Republic of China* (2010 edition), several snake species are listed. These entries include Jinqianbaihuashe (Bungarus parvus) derived from *Bungarus multicinctus* (Elapidae), Qishe (Agkistrodon) derived from *Agkistrodon acutus* (Viperidae), and Wushaoshe (Zaocys) derived from *Zaocys dhumnades* (Colubridae). The shedded skins of *Elaphe carinata* (Elapidae), *E. taeniura* (Elapidae), and *Z. dhumnades* are used as Shetui (Serpentis Periostracum), and the meat of *A. strauchii* (Viperidae), *B. multicinctus*, and *Opheodrys major* (Colubridae) are used as Sherou. Species-specific RAPD fragments to *A. acutus, B. multicinctus,* and *Z. dhumnades* were used to design SCAR primers to generate specific markers for these species [51]. DNA sequencing of Cyt b region clearly distinguished *Z. dhumnades* and *B. multicinctus* from *N. naja, Ophiophagus hannah, Ptyas mucosus,* and *Python molurus*. It was also found that Cyt b sequence could resolve the phylogenetic relationship of 90 snake species in families Boidae, Colubridae, Elapidae, and Viperidae with sequence similarity ranging from 71 to 93%. FINS approach based on the Cyt b sequence revealed that two of the retailed snake meat samples were derived from *Python reticulates* and four of them were derived from *Python molurus* [75].

3.5 Conclusion and Prospect

With the exponential growth of international market and increasing demand of high-quality traditional medicinal materials, correct identification is a key factor to ensure safety, efficacy, and fair trade. Molecular technology provides a reliable and powerful tool for definite identification of medicinal materials at various taxonomic levels including family, genus, species, varieties, and cultivar (Table 3.2). Molecular identification of traditional medicinal materials may face difficulties in poor quality and quantity of DNA content, presence of secondary metabolites, and fungal contamination. Similar to other authentication methods, substantial efforts are needed to collect authentic reference species for DNA marker development. With the DNA barcode initiative and the reduction of cost in DNA manipulation and sequencing, we expect that more molecular markers will be developed and increased number of laboratories will be able to carry out the tests. Also, benefited from the development of pocket-size PCR machine and speedy isothermal amplification techniques, the trend will be instant on-site molecular identification of medicinal materials in the near future.

Table 3.2 Examples of molecular identification of medicinal materials in our laboratory

Herb	DNA technique	Method	Medicinal species	Distinguished species	Restriction enzyme	DNA region	Reference
Baibu	Sequencing	–	*Stemona japonica, S. sessilifolia,* and *S. tuberosa*	*Croomia* species and *Asparagus* species	–	TrnL	[71]
Baihuasheshecao	Sequencing	FINS	*Hedyotis diffusa*	*Hedyotis corymbosa* and 12 other *Hedyotis* species	–	ITS	[66]
Bajiaolian	Fingerprinting	RAPD	*Dysosma* species	*Dysosma furfuracea, D. majorense, D. pleiantha, D. veitchii,* and *D. versipellis*	–	–	[27]
Dangshen	Fingerprinting	AP-PCR, RAPD	*Codonopsis pilosula*	11 samples of *Codonopsis pilosula* collected from 9 localities	–	–	[25]
	Fingerprinting	PCR-RFLP	*Codonopsis pilosula, C. tangshen, C. modesta,* and *C. nervosa* var. *macrantha*	*Campanumoea javania* and *Platycodon grandiflorus*	*Hinf*1, *Hha*I	ITS	[46]
	Sequencing	FINS	*Codonopsis* species	*Codonopsis dangshen, C. pilosula, C. pilosula* var. *modesta,* and *Platycodon grandiflorum*	–	5S	[69]
Eyu	Fingerprinting	RAPD, SCAR	*Alligator mississippiensis, Caiman crocodiles, Crocodylus niloticus, Crocodylus porosus,* and *Crocodylus siamensis*	Domestic animals, livestock, and snake species	–	–	[52]
Huajuhong	Fingerprinting, sequencing	ISSR	*Citrus grandis* "Tomentosa"	*Citrus grandis, C. chachiensis, C. reticulate, C. medica* var. *sarcodactylis*	–	ITS, trnH-psbA	[40]

(continued)

Table 3.2 (continued)

Herb	DNA technique	Method	Medicinal species	Distinguished species	Restriction enzyme	DNA region	Reference
Kudidan	Fingerprinting	AP-PCR, RAPD	*Elephantopus scaber*	*Elephantopus mollis, Emilia sonchifolia, Paraixeris denticulata, Piloselloides hirsuta, Pseudoelephantopus spicatus, Pterocypsela indica, Sonchus arvensis*, and *Taraxacum mongolicum*	–	–	[23]
Leigongteng	Sequencing	FINS	*Tripterygium wilfordii*	5 other *Tripterygium* species, *Celastrus angulatus, C. hypoleucus* and 6 other *Celastrus* species, *Canotia holacantha, Euonymus alatus, Paxistima myrsinites*, and *P. canbyi*	–	ITS, 5S	[67]
Madouling	Sequencing	FINS	*Aristolochia contorta* and *A. debilis*	*Cardiocrinum giganteum* var. *yunnanense* and 2 other *Cardiocrinum* species; 16 *Lilium* species; 4 *Aristolochia* species	–	TrnL-F, trnH-psbA	[20]
Muxiang	Sequencing	–	*Saussurea lappa, Inula helenium, I. racemosa, Vladimiria berardioidea, V. souliei,* and *V. souliei* var. *mirabilis*	*Aristolochia contorta* and *A. debilis*	–	ITS, 5S	[65]
Pugongying	Fingerprinting	AP-PCR, RAPD	*Taraxacum mongolicum*	*Elephantopus scaber, Emilia sonchifolia, Paraixeris denticulata, Piloselloides hirsuta, Pterocypsela indica*, and *Sonchus arvensis*	–	–	[24]
Renshen	Fingerprinting	AP-PCR	*Panax ginseng*	*Panax quinquefolius*	–	–	[18]
	Fingerprinting	AP-PCR, RAPD	*Panax ginseng, P. quinquefolius*, and *P. notoginseng*	*Mirabilis jalapa, Phytolacca acinosa, Platycodon grandiflorum*, and *Talinum paniculatum*	–	–	[19]
	Fingerprinting, sequencing	PCR-RFLP	*Panax ginseng, P. quinquefolius*, and *P. notoginseng*	*Panax japonicus, P. trifolius, P. major, Mirabilis jalapa*, and *Phytolacca acinosa*	*Hinf*I, *Taq*I, *Sau*3A1	ITS	[45]

	Fingerprinting	SCAR	*Panax ginseng*	*Panax quinquefolius*	–	–	[50]
	Fingerprinting	DALP	*Panax ginseng*	*Panax quinquefolius*	–	–	[38]
	Fingerprinting	AFLP	*Panax ginseng*	*Panax quinquefolius*	*Eco*RI, *Mse*I	–	[34]
Shihu	Sequencing	–	*Dendrobium* species	16 *Dendrobium* species and *Pholidota cantonensis*	–	ITS	[64]
	Microarray	–	*Dendrobium loddigesii, D. fimbriatum, D. chrysanthum, D. officinale, D. nobile*	11 other *Dendrobium* species	–	ITS	[61]
	Microarray	–	*Dendrobium officinale*	7 other *Dendrobium* species	–	5S	[62]
She	Fingerprinting	RAPD, SCAR	*Zaocys dhumnades, Agkistrodon acutus,* and *Bungarus multicinctus*	*Elaphe moellendorffi*, *Naja naja*, *Ophiophagus hannah, Ptyas mucosus, Python molurus,* and *Python regius*	–	–	[51]
	Sequencing	FINS	*Zaocys dhumnades* and *Bungarus multicinctus*	*Naja naja*, *Ophiophagus hannah*, *Ptyas mucosus,* and *Python molurus*	–	Cyt b	[75]
Wutou	Sequencing	–	*Aconitum carmichaelii* and *A. kusnezoffii*	17 other *Aconitum* species	–	TrnH-psbA	[68]
Zangyinchen	Sequencing	–	*Swertia mussotii*	*Swertia chirayita, S. franchetiana*, and *S. wolfgangiana*	–	5S	[70]

References

1. World Health Organization. (2008). *Traditional medicine*. http://www.who.int/mediacentre/factsheets/fs134/en/. Dec 2008.
2. Vanherweghem, J., Tielemans, C., Abramowicz, D., et al. (1993). Rapidly progressive interstitial renal fibrosis in young women: Association with slimming regimen including Chinese herbs. *Lancet, 341*, 387–391.
3. Vanhaelen, M., Vanhaelen-Fastre, R., But, P. P. H., et al. (1994). Identification of aristolochic acid in Chinese herbs. *Lancet, 343*, 174.
4. But, P. P. H. (1994). Herbal poisoning caused by adulterants or erroneous substitutes. *The Journal of Tropical Medicine and Hygiene, 97*, 371–374.
5. Lee, S., Lee, T., Lee, B., et al. (2004). Fanconi's syndrome and subsequent progressive renal failure caused by a Chinese herb containing aristolochic acid. *Nephrology (Carlton, Vic.), 9*, 126–129.
6. Lo, S. H., Wong, K. S., Arlt, V. M., et al. (2005). Detection of Herba Aristolochia Mollissemae in a patient with unexplained nephropathy. *American Journal of Kidney Diseases, 45*, 407–410.
7. Debelle, F. D., Vanherweghem, J. L., & Nortier, J. L. (2008). Aristolochic acid nephropathy: A worldwide problem. *Kidney International, 74*, 158–169.
8. Jha, V. (2010). Herbal medicines and chronic kidney disease. *Nephrology (Carlton, Vic.), 15*(Suppl 2), 10–17.
9. But, P. P. H., Tomlinson, B., Cheung, K. O., et al. (1996). Adulterants of herbal products can cause poisoning. *British Medical Journal, 313*, 117.
10. Zhu, Y. P. (2000). Toxicity of the Chinese herb mu tong (*Aristolochia manshuriensis*): What history tells us. *Adverse Drug Reactions and Toxicological Reviews, 21*, 171–177.
11. Zeng, Z. P., & Jiang, J. G. (2010). Analysis of the adverse reactions induced by natural product-derived drugs. *British Journal of Pharmacology, 159*, 1374–1391.
12. CBOL Plant Working Group. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 12794–12797.
13. Chen, S., Yao, H., Han, J., et al. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PloS One, 5*, e8613.
14. Hebert, P. D., Cywinska, A., Ball, S. L., et al. (2003). Biological identifications through DNA barcodes. *Proceeding of the Royal Society B Biological Sciences, 270*, 313–321.
15. Kress, W. J., Wurdack, K. J., Zimmer, E. A., et al. (2005). Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 8369–8374.
16. Shaw, P. C., Jiang, R. W., & Wong, K. L. (2007). Health food and medicine: Combined chemical and molecular technologies for authentication and quality control. In S. E. Ebeler, G. R. Takeoka, & P. Winterhalter (Eds.), *Authentication of food and wine*. Washington, DC: American Chemical Society.
17. Shaw, P. C., Ngan, F. N., & But, P. P. H. (2002). Molecular markers in Chinese medicinal materials. In P. C. Shaw, J. Wang, & P. P. H. But (Eds.), *Authentication of Chinese medicinal materials by DNA technology*. Singapore: World Scientific.
18. Cheung, K. S., Kwan, H. S., But, P. P. H., et al. (1994). Pharmacognostical identification of American and Oriental ginseng roots by genomic fingerprinting using arbitrarily primed polymerase chain reaction (AP-PCR). *Journal of Ethnopharmacology, 42*, 67–69.
19. Shaw, P. C., & But, P. P. H. (1995). Authentication of *Panax* species and their adulterants by random-primed polymerase chain reaction. *Planta Medica, 61*, 466–469.
20. Li, M., Ling, K. H., Lam, H., et al. (2010). *Cardiocrinum* seeds as a replacement for *Aristolochia* fruits in treating cough. *Journal of Ethnopharmacology, 130*, 429–432.
21. Sasaki, Y., Komatsu, K., & Nagumo, S. (2008). Rapid detection of *Panax ginseng* by loop-mediated isothermal amplification and its application to authentication of ginseng. *Biological and Pharmaceutical Bulletin, 31*, 1806–1808.

22. Welsh, J., & McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research, 18*, 7213–7218.
23. Cao, H., But, P. P. H., & Shaw, P. C. (1996). Authentication of the Chinese drug 'Ku-Di-Dan' (Herba Elephantopi) and its substitutes using random-primed polymerase chain reaction (PCR). *Acta Pharmaceutica Sinica, 31*, 543–553.
24. Cao, H., But, P. P. H., & Shaw, P. C. (1996). A molecular approach to identification of the Chinese drug 'pu gong ying' (herba taraxaci) and six adulterants by DNA fingerprinting using random primed polymerase chain reaction (PCR). *Journal of Chinene Pharmaceutical Sciences, 5*, 186–194.
25. Zhang, Y. B., Ngan, F. N., Wang, Z. T., et al. (1999). Random primed polymerase chain reaction differentiates *Codonopsis pilosula* from different localities. *Planta Medica, 65*, 157–160.
26. Williams, J. G., Kubelik, A. R., Livak, K. J., et al. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research, 18*, 6531–6535.
27. Fu, R. Z., Shaw, P. C., Wang, J., et al. (2000). RAPD differentiation of five medicinal *Dysosma* species. *Journal of Chinese Pharmaceutical Sciences, 9*, 57–60.
28. Chatti, K., Baraket, G., Ben Abdelkrim, A., et al. (2010). Development of molecular tools for characterization and genetic diversity analysis in Tunisian fig (*Ficus carica*) cultivars. *Biochemical Genetics, 48*, 789–806.
29. Li, Y., & Ding, W. L. (2010). Genetic diversity assessment of *Trollius* accessions in China by RAPD markers. *Biochemical Genetics, 48*, 34–43.
30. Lim, W., Mudge, K. W., & Weston, L. A. (2007). Utilization of RAPD markers to assess genetic diversity of wild populations of North American ginseng (*Panax quinquefolius*). *Planta Medica, 73*, 71–76.
31. Vos, P., Hogers, R., Bleeker, M., et al. (1995). AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research, 23*, 4407–4414.
32. Ruselll, J. R., Fuller, J. D., Macaulay, M., et al. (1997). Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theoretical and Applied Genetics, 95*, 714–722.
33. Ha, W. Y., Yau, F. C. F., Shaw, P. C., et al. (2002). Differentiation of *Panax ginseng* from *P. quinquefolius* by amplified fragment length polymorphism. In P. C. Shaw, J. Wang, & P. P. H. But (Eds.), *Authentication of Chinese medicinal materials by DNA technology*. Singapore: World Scientific.
34. Ha, W. Y., Shaw, P. C., Liu, J., et al. (2002). Authentication of *Panax ginseng* and *Panax quinquefolius* using amplified fragment length polymorphism (AFLP) and directed amplification of minisatellite region DNA (DAMD). *Journal of Agricultural and Food Chemistry, 50*, 1871–1875.
35. Choi, Y. E., Ahn, C. H., Kim, B. B., et al. (2008). Development of species specific AFLP-derived SCAR marker for authentication of *Panax japonicus* C. A. Meyer. *Biological and Pharmaceutical Bulletin, 31*, 135–138.
36. Datwyler, S. L., & Weiblen, G. D. (2006). Genetic variation in hemp and marijuana (*Cannabis sativa* L.) according to amplified fragment length polymorphisms. *Journal of Forensic Sciences, 51*, 371–375.
37. Passinho-Soares, H., Felix, D., Kaplan, M. A., et al. (2006). Authentication of medicinal plant botanical identity by amplified fragmented length polymorphism dominant DNA marker: inferences from the *Plectranthus* genus. *Planta Medica, 72*, 929–931.
38. Ha, W. Y., Yau, F. C., But, P. P. H., et al. (2001). Direct amplification of length polymorphism analysis differentiates *Panax ginseng* from *P. quinquefolius*. *Planta Medica, 67*, 587–589.
39. Godwin, I. D., Aitken, E. A., & Smith, L. W. (1997). Application of inter simple sequence repeat (ISSR) markers to plant genetics. *Electrophoresis, 18*, 1524–1528.
40. Su, C., Wong, K. L., But, P. P. H., et al. (2010). Molecular authentication of the Chinese Herb Huajuhong and related medicinal material by DNA sequencing and ISSR marker. *Journal of Food and Drug Analysis, 18*, 161–170.
41. Kojoma, M., Iida, O., Makino, Y., et al. (2002). DNA fingerprinting of *Cannabis sativa* using inter-simple sequence repeat (ISSR) amplification. *Planta Medica, 68*, 60–63.

42. Shi, H. M., Wang, J., & Wang, M. Y. (2009). Identification of *Cistanche* species by chemical and inter-simple sequence repeat fingerprinting. *Biological and Pharmaceutical Bulletin, 32*, 142–146.
43. Hu, Y., Zhang, Q., Xin, H., et al. (2007). Association between chemical and genetic variation of *Vitex rotundifolia* populations from different locations in China: Its implication for quality control of medicinal plants. *Biomedical Chromatography, 21*, 967–975.
44. Song, Z., Li, X., Wang, H., et al. (2010). Genetic diversity and population structure of *Salvia miltiorrhiza* Bge in China revealed by ISSR and SRAP. *Genetica, 138*, 241–249.
45. Ngan, F. G., Shaw, P. C., But, P. P. H., et al. (1999). Molecular authentication of *Panax* species. *Phytochemistry, 50*, 787–791.
46. Fu, R. Z., Wang, J., Zhang, Y. B., et al. (1999). Differentiation of medicinal *Codonopsis* species from adulterants by polymerase chain reaction-restriction fragment length polymorphism. *Planta Medica, 65*, 648–650.
47. Gong, W., Fu, C. X., Luo, Y. P., et al. (2006). Molecular identification of *Sinopodophyllum hexandrum* and *Dysosma* species using cpDNA sequences and PCR-RFLP markers. *Planta Medica, 72*, 650–652.
48. Lee, J. H., Lee, J. W., Sung, J. S., et al. (2009). Molecular authentication of 21 *Korean Artemisia* species (Compositae) by polymerase chain reaction-restriction fragment length polymorphism based on trnL-F region of chloroplast DNA. *Biological and Pharmaceutical Bulletin, 32*, 1912–1916.
49. Li, X., Ding, X., Chu, B., et al. (2007). Molecular authentication of *Alisma orientale* by PCR-RFLP and ARMS. *Planta Medica, 73*, 67–70.
50. Wang, J., Ha, W. Y., Ngan, F. N., et al. (2001). Application of sequence characterized amplified region (SCAR) analysis to authenticate *Panax* species and their adulterants. *Planta Medica, 67*, 781–783.
51. Yau, F. C. F., Wong, K. L., Shaw, P. C., et al. (2002). Authentication of snakes used in Chinese medicine by sequence characterized amplified region (SCAR). *Biodiversity and Conservation, 11*, 1653–1662.
52. Yau, F. C. F., Wong, K. L., Wang, J., et al. (2002). Generation of a sequence characterized amplified region probe for authentication of Crocodilian species. *The Journal of Experimental Zoology, 294*, 382–386.
53. Dnyaneshwar, W., Preeti, C., Kalpana, J., et al. (2006). Development and application of RAPD-SCAR marker for identification of *Phyllanthus emblica* Linn. *Biological and Pharmaceutical Bulletin, 29*, 2313–2316.
54. Lee, M. Y., Doh, E. J., Park, C. H., et al. (2006). Development of SCAR marker for discrimination of *Artemisia princeps* and *A. argyi* from other *Artemisia* herbs. *Biological and Pharmaceutical Bulletin, 29*, 629–633.
55. Sze, S. C., Song, J. X., & Wong, R. N. (2008). Application of SCAR (sequence characterized amplified region) analysis to authenticate *Lycium barbarum* (wolfberry) and its adulterants. *Biotechnology and Applied Biochemistry, 51*, 15–21.
56. Walker, G. T., Little, M. C., Nadeau, J. G., et al. (1992). Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system. *Proceedings of the National Academy of Sciences of the United States of America, 89*, 392–396.
57. Walker, G. T., Fraiser, M. S., Schram, J. L., et al. (1992). Strand displacement amplification–an isothermal, in vitro DNA amplification technique. *Nucleic Acids Research, 20*, 1691–1696.
58. Notomi, T., Okayama, H., Masubuchi, H., et al. (2000). Loop-mediated isothermal amplification of DNA. *Nucleic Acids Research, 28*, E63.
59. Sasaki, Y., & Nagumo, S. (2007). Rapid identification of *Curcuma longa* and *C. aromatica* by LAMP. *Biological and Pharmaceutical Bulletin, 30*, 2229–2230.
60. Vincent, M., Xu, Y., & Kong, H. (2004). Helicase-dependent isothermal DNA amplification. *EMBO Reports, 5*, 795–800.
61. Zhang, Y. B., Wang, J., Wang, Z. T., et al. (2003). DNA microarray for identification of the herb of *Dendrobium* species from Chinese medicinal formulations. *Planta Medica, 69*, 1172–1174.

62. Sze, S. C. W., Zhang, Y. B. K., Shaw, P. C., et al. (2008). A DNA microarray for differentiation of the Chinese medicinal herb *Dendrobium officinale* (Fengdou Shihu) by its 5S ribosomal DNA intergenic spacer region. *Biotechnology and Applied Biochemistry, 49*, 149–154.
63. Cho, Y., Mower, J. P., Qiu, Y. L., et al. (2004). Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 17741–17746.
64. Lau, D. T., Shaw, P. C., Wang, J., et al. (2001). Authentication of medicinal *Dendrobium* species by the internal transcribed spacer of ribosomal DNA. *Planta Medica, 67*, 456–460.
65. Chen, F., Chan, H. Y., Wong, K. L., et al. (2008). Authentication of *Saussurea lappa*, an endangered medicinal material, by ITS DNA and 5S rRNA sequencing. *Planta Medica, 74*, 889–892.
66. Li, M., Jiang, R. W., Hon, P. M., et al. (2010). Authentication of the anti-tumor herb Baihuasheshecao with bioactive marker compounds and molecular sequences. *Food Chemistry, 119*, 1239–1245.
67. Law, S. K. Y., Simmons, M. P., Techen, N., et al. (2011). Molecular analyses of the Chinese herb Leigongteng (*Tripterygium wilfordii* Hook.f.). *Phytochemistry, 72*, 21–26.
68. He, J., Wong, K. L., Shaw, P. C., et al. (2010). Identification of the medicinal plants in *Aconitum L.* by DNA barcoding technique. *Planta Medica, 76*, 1622–1628.
69. Zhang, Y. B., Jiang, R. W., Li, S. L., et al. (2007). Chemical and molecular characterization of Hong Dangshen, a unique medicinal material for diarrhea in Hong Kong. *Journal of Chinese Pharmaceutical Sciences, 16*, 202–207.
70. Yu, M. T., Wong, K. L., Zong, Y. Y., et al. (2008). Identification of *Swertia mussotii* and its adulterant *Swertia* species by 5S rRNA gene spacer. *China Journal of Chinese Materia Medica, 33*, 502–504.
71. Jiang, R. W., Hon, P. M., Xu, Y. T., et al. (2006). Isolation and chemotaxonomic significance of tuberostemospironine-type alkaloids from *Stemona tuberosa*. *Phytochemistry, 67*, 52–57.
72. Bartlett, S. E., & Davidson, W. S. (1992). FINS (forensically informative nucleotide sequencing): A procedure for identifying the animal origin of biological specimens. *BioTechniques, 12*, 408–411.
73. Sahajpal, V., & Goyal, S. P. (2010). Identification of a forensic case using microscopy and forensically informative nucleotide sequencing (FINS): A case study of small Indian civet (*Viverricula indica*). *Science & Justice, 50*, 94–97.
74. Lou, S. K., Wong, K. L., Li, M., et al. (2010). An integrated web medicinal materials DNA database: MMDBD (Medicinal Materials DNA Barcode Database). *BMC Genomics, 11*, 402.
75. Wong, K. L., Wang, J., But, P. P. H., et al. (2004). Application of cytochrome b DNA sequences for the authentication of endangered snake species. *Forensic Science International, 139*, 49–55.
76. Alvarez, I., & Wendel, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution, 29*, 417–434.
77. Baldwin, B. G., Sanderson, M. J., Porter, J. M., et al. (1995). The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden, 82*, 247–277.
78. Kress, W. J., & Erickson, D. L. (2007). A two-locus global DNA barcode for land plants: The coding rbcL gene complements the non-coding trnH-psbA spacer region. *PloS One, 2*, e508.

Suggested Readings

Lou, S. K., Wong, K. L., Li, M., But, P. P. H., Tsui, S. K., & Shaw, P. C. (2010). An integrated web medicinal materials DNA database: MMDBD (Medicinal Materials DNA Barcode Database). *BMC Genomics, 11*, 402.

Shaw, P. C., Wang, J., & But, P. P. H. (2002). *Authentication of Chinese medicinal materials by DNA technology*. Singapore: World Scientific.

Shaw, P. C., Jiang, R. W., Wong, K. L., & But, P. P. H. (2007). Health food and medicine: Combined chemical and molecular technologies for authentication and quality control. In S. E. Ebeler, G. R. Takeoka, & P. Winterhalter (Eds.), *Authentication of food and wine*. Washington, DC: American Chemical Society.

Shaw, P. C., Wong, K. L., Chan, A. W., Wong, W. C., & But, P. P. H. (2009). Patent applications for using DNA technologies to authenticate medicinal herbal material. *Chinese Medicine, 4*, 21.

Chapter 4
The Mechanism of Formation of Dao-di Herbs

Lu-qi Huang and Lan-ping Guo

Abstract "Dao-di herbs" is a term defining those traditional Chinese herbs planted in areas with certain natural conditions and ecological environment, which are produced centrally and cultivated, collected, and processed with particular methods. Therefore, Dao-di herbs have long-standing reputations for their high qualities and excellent treatment effects.

Compared with other herbs, Dao-di herbs have their own distinguishing qualities, including the special geographic variations, having particular quality standards, abundant cultural connotations, and comparatively high economic values.

The biological connotation of Dao-di herbs is the same species but with different origins, and the biological nature is the interaction between nature and genetics. In the view of biology, Dao-di herbs are the products of interaction between genotypes and environment, whose phenotypes are equal to genotypes plus the modification of environment. This is the biological principle of the production of Dao-di herbs.

The chemical constitutions of Dao-di herbs are unique. The more obvious regional characteristics Dao-di herbs have, the more specialized genotypes Dao-di herbs have. Bad environments promote the formation of Dao-di herbs.

The molecular mechanisms of Dao-di herb formations are specific genotypes, that is, the complex regulations in specific environments caused space-time differences in the expressions of key enzyme genes in the secondary metabolism processes. The diversities of patterns and secondary metabolites are caused by the hereditary variations.

Lu-qi Huang (✉) • Lan-ping Guo
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China
e-mail: huangluqi@263.net; glp01@126.com

Lu-qi Huang (ed.), *Molecular Pharmacognosy*,
DOI 10.1007/978-94-007-4945-0_4, © Shanghai Scientific and Technical Publishers and Springer Science+Business Media Dordrecht 2013

4.1 Introduction

"Dao-di herbs" is a term used by ancient Chinese people to describe the inner species variation of Chinese medicinal materials relevant to the geographical variation, and Dao-di herb is the population of a Dao-di herb in specific areas featured with proverbial superior qualities and popularly used in the clinical practice of traditional Chinese medicine (TCM) [1–3]. Dao-di herbalism is the generic term of various merits possessed by Dao-di herbs. As a concept of quality of TCM established by usage in TCM industry in China, Dao-di herbs can be considered as a distinctive standard distinguishing high-quality TCM used by ancient doctor of TCMs. The usage amount of Dao-di herbs is the largest in TCM, and its economic value is the highest. Among the 500 TCMs, there are 200 TCMs with obvious Dao-di herbalism, whose usage amount makes up 80% of the total usage amount of TCMs. *Geoherbs in China* edited by Hu SL records 160 common Dao-di herbs consisting of 132 plant medicines, 20 animal medicines, and 8 mineral medicines [1].

4.2 The Connotation of Dao-di Herbs

4.2.1 History Evolution of Dao-di Herb Concept

The meaning of Dao-di herbs was reflected as early as in the "Shen Nong Ben Cao Jing," in which the importance of origin was emphasized. Among the 365 kinds of herbs which were included in "Shen Nong Ben Cao Jing," many can be distinguished as Dao-di herbs by their names, such as Bajitian, Shujiao, Shuzao, Qinjiao, Qinpi, Wuzhuyu, Ejiao, Daizheshi (Ba, Shu, Qin, Wu, Dong'e, and Daizhou are all names of old places around Western Zhou Dynasty). Although every herb is described without detailed growing area, the records such as growing in the valley, in the marsh, in the pond, in the hill, in the field, and in the flat land are mentioned. It seemed that the details of herbs in "Sheng Nong Ben Cao Jing" were shown with ancient authentic colors. "Huang Di Nei Jing," which is a contemporary medicinal monograph as "Shen Nong Ben Cao Jing," had expounded the meaning of Dao-di herbs in theory, and "Shang Han Lun" was a medicinal monograph which firstly recorded lots of Dao-di herbs such as Ejiao, Badou, Daizheshi, etc.

4.2.2 Concept of Dao-di Herbs

Dao-di herbs are the authentic medicinal materials with excellent quality. This concept came from the clinical practice of Chinese medicine and had been proved by numerous clinical practice of TCM for thousands of years. It has a rich scientific content. As a conventional concept for the standardization of ancient drugs, Dao-di herbs,

which stems from ancient times, is not only a unique high-quality comprehensive standard to identify the Chinese herbal medicines with good qualities, but also a unique comprehensive standard to control the quality of medicinal materials in Chinese material medica. Generally speaking, Dao-di herbs refer to the medicinal materials growing in habitat with particular natural conditions and ecological environment. The production of Dao-di herbs is usually concentrated, and their cultivation techniques and harvesting process usually have certain requirements. Therefore, these drugs usually have better quality and better curative effect than drugs belonging to the same species but from other habitats. The names of Dao-di herbs always show together with their famous region, such as Xining Dahuang, Ningxia Gouji, Chuan Beimu, Chuan Shao, Qin Jiu, Liao Wuwei, Guan Fangfeng, Huai Dihuang, Mi Yinhua, Hao Juhua, Xuan Mugua, Hang Baizhi, Zhe Xuanshen, Jiang Zhiqiao, Su Bohe, Mao Cangzhu, Jian Zexie, Huang Chenpi, Taihe Wuji, A Jiao, Dai Zheshi, etc. There are some special cases. For example, Guang Muxiang is not produced in Guangzhou but imported from Guangdong and Zang Honghua is imported from Xizang, too.

4.2.3 Attributes of Dao-di Herbs

4.2.3.1 Specialized Genotype

Dao-di herbs refers to a special population which is affected by its habitat during long-term evolution. Different groups belonging to the same species, called populations in biology, are formed in different regions. These populations may be divided by ecological trait, geographical trait, local trait, chemical trait, and biological trait or they may be classified into cultivated varieties, forms, lines, farm varieties, etc. Traditionally, Dao-di herbs are considered to have the similar genetic background with non-Dao-di herbs. In fact, they are materials belonging to the same species but growing in different regions, and the qualities of Dao-di Herbs are usually better. These special populations are not divided by human but decided by the environment including soil condition, illumination, wetness, etc. They could be natural populations or cultivated populations growing in the specific space and in the special time, and they were stable. Through DNA analysis technology, in-depth study of specialized genotype of Dao-di herbs could be carried out, and the nature of Dao-di herbs would be clarified. For example, RAPD analysis of angelica (Danggui) growing in famous region showed that the smaller its geographical distribution was, the smaller its genetic diversity would be [4]. Genetic differences between different kinds of Magnolia (Houpu) which originated from famous regions were obvious [5].

4.2.3.2 Obvious Regional Characteristics

The formation and development of concept of Dao-di herbs is closely related to specialized regions. The nature of Dao-di herbs is the materials with best quality among products belonging to the same species but with different origins. So, the

names of Dao-di herbs are always with a region name before them. For example, Yang Chunsha, Xuan Mugua, Mao Cangshu, Min Danggui, Guan Shanbo, Si Da Chuan Yao (four famous drugs from Sichuan), Si Da Huai Yao (four famous drugs from Jiaozuo), Zhe Ba Wei (eight famous drugs from Zhejiang), etc. Environmental factors play an important role in the formation of quality of Dao-di herbs. These environment factors include soil, water, light, temperature, terrain, and so on. The special environment in famous regions is the first question in the studies on Dao-di herbs. There would be no Dao-di herbs without famous regions. Because of the specialized environment in famous regions, the quality of Dao-di herbs is usually very nice and the scale of production is usually easier to be formed. As a result, Dao-di herbs usually have a good reputation in the market.

4.2.3.3 Unique Quality Standards

Regional characteristics result in the differences of Dao-di herbs in characteristics, chemical composition, efficacy, and so on. For example, Chuanxinlian from Guangzhou has a better effect as antibacterial than materials from Fujian and Anhui; Chuan Maidong from Mianyang and Santai is worse in appearance and taste than Hang Maidong from Yuyao and Hushan in Zhejiang province; there is a high content of anthraquinones in Dahuang from northwest of China, while Dahuang from Heilongjiang contains more tannin and has an antidiarrheal effect. Content of active components in Qinpi from Shanxi is higher than those in material from Sichuan. Phenolic content in Houpu from Sichuan is 6 times as those in materials from Jiangxi [6]. Shanxi is the famous region of Huangqi, and characteristics and chemical composition of Huangqi materials from Hubei are much different from Dao-di herbs. Appearance and taste of Gouji from Ningxia are much better than materials from other regions. Dao-di herbs should have a unique chemical composition. Mao Cangzhu is much different from Nan Cangzhu in contents of volatile oil. Total volatile oil content in Mao Cangzhu is significantly lower, but it has an extremely high content of atractylon and atractydin. In addition, the ratio among atractylon, atractydin, hinesol, and β-eucalyptol in Mao Cangzhu shows a unique composition relation [7].

4.2.3.4 Particular Technology or Process of Production

The cultivation (aquaculture), maintenance, harvesting, processing, and other processes of Chinese medicinal herbs affect the formation of qualities of Dao-di herbs and sometimes even become one of the decisive factors. Jiangyou in Sichuan and Hanzhong in Shannxi always have mass production of Fuzi and Fupian, but their qualities are very different, and Jiangyou Fuzi is considered as a Dao-di herb. Famous regions are always together with particular processes and histories. The longer history it has, the more mature the technique will be and the more obvious the authenticity will be.

4.2.3.5 Rich Cultural Connotations

The formation of Dao-di herbs is related to the unique geography and cultural background of our country as well as Chinese medicine theory. There are no concept of Dao-di herbs being mentioned in any other nations and countries in the world. Dao-di herb is a part of ancient cultural traditions in China with a strong cultural reserve. The use of Dao-di herbs reflects the deep understanding of Chinese medicine on "Zheng" concept and the state of the patients' body, as well as understanding of Chinese medicine on knowledge and theory of property, taste, effectiveness, clinical characteristics, and adverse effects of medicinal materials. The formation of Dao-di herbs would promote the development of traditional culture in famous regions to some extent. The appearance of various festival activities named after Dao-di herbs nowadays is one of the important expression forms of cultural connotations of Dao-di herbs, such as the Ningxia Gouqi festival. Therefore, the cultural connotation is an important feature of Dao-di herbs, and it makes the Dao-di herbs become a special class of natural products which are different from ordinary natural products because of their intellectual property attributes. Dao-di herbs are especially valuable because of their cultural feature.

4.3 The Hypothesis of Mechanism of Dao-di Herb Formation

Each Dao-di herb has a typical quantitative trait in its phenotype. The phenotype of quantitative trait resulted from the interaction of heredity and environment, that is, phenotype = gene + environment + interaction of gene and environment. This shows that genes have potency to develop some definite phenotype, but it is not the necessary condition to determine these phenotypes. They could determine a series of possibilities depending on the habitat. According to the meaning of Dao-di herbs, the hypothesis of its formation was proposed.

4.3.1 Chemical Components of Dao-di Herbs and Their Unique Adaptive Feature

Secondary metabolite is the substantial basis of authenticity of Dao-di herbs, and it results in its differential chemical feature and better efficacy comparing to other populations of this species. As an open and complicated system, the chemical feature of Dao-di herbs (one of its phenotypes) is the result of its long-time adaptation to the surrounding environment. Therefore, the unique adaptive feature appears in the chemical composition of Dao-di herbs which grow in unique outside environment and unique habitat. For example, the variation in chemical composition of widespread species is always continuous. Take mint (Bohe), chrysanthemum (Juhua), and atractylodes (Cangzhu), for example, the variation of volatile oil in

them might be changed from continuous to discontinuous when the climate of their distribution domain, especially some dominant factors that can affect the chemical composition in them, changed significantly. This change in chemical composition of Dao-di herbs is the result of its adaptation to habitat, which fully demonstrates the adaptive feature of complicated systems. Composition features of volatile oil in atractylodes of Dao-di herbs and atractylodes from other regions are very different. Volatile oil content in atractylodes of Dao-di herbs is significantly low ($P<0.01$). Components whose contents were higher than 1 % when treated by unitary processing are significantly more ($P<0.01$); the content of atractylon and atractydin is extremely high, and the content of hinesol and β-eucalyptol was extremely low ($P<0.01$) in Dao-di herb atractylodes. Besides, the ratio among atractylon, atractydin, hinesol, and β-eucalyptol shows a unique law, that is, (0.70–2.0):(0.04–0.35):(0.09–0.40):1 [7]. These results indicate that the substance basic of the nice clinical efficacy of Dao-di herbs atractylode was its unique chemical composition.

4.3.2 The More Obvious the Authenticity of Dao-di Herbs Is, the More Obvious Their Gene Specialization Will Be

In long-term process of breeding and domesticating, Dao-di herbs will form a unique genotype. The good quality of Dao-di herbs (e.g., rehmannia populations such as Jin Zhuang Yuan, Beijing No.1 and 85-5; ginseng populations such as Da Ma Ya and Er Ma Ya) is very closely related to its unique genetic background. For wild Dao-di herbs, if it distributes in a large area, the genotypes of populations in different areas are usually different, called local specialized genotype. These genotypes are the result of long-term options under different ecological or geographical conditions, and they are also the essential reasons of Dao-di herbs' formation. That is to say, the more widely the species distributes, the more obvious its authenticity will be and the more specialized its genotype will be. The studies on intraspecific genetic structure and genetic differentiation could provide some proofs for the genotype specialization of Dao-di herbs. For example, authenticity of Magnolia (Houpu) was studied by using RAPD technique. The results show there are obvious genetic differentiation between "Chuanpu" (Dao-di herbs) and "Wenpu," and this phenomenon is also related to active components. This indicates that the authenticity of Magnolia is mainly decided by the genetic factors [8].

4.3.3 Stress Will Promote the Formation of Dao-di Herbs

The active components in Chinese materia medica are usually low molecular weight compounds after secondary metabolism. It is convinced that stress would promote the biosynthesis of secondary metabolites. For instance, dry stress can make the quercetin content in ginkgo higher to some extent [9], polyphenol content in

marigold under dry stress is higher than that under normal conditions [10], polyphenol content in sunflower raise with the increase of nutrition stress [11], and monoterpene content in rosemary become increased with the increase of CO_2 concentration[12]. The most obvious change of secondary metabolism in plants after damaged by pest is the increase of polyphenol content [13].

Many scholars have done researches on the gene expression in plants under stress, and they have revealed the effect of stress on plant secondary metabolism at genetic level. DREB transcription factor, which is induced by the environmental stress, might activate other 12 genes depending on cis-acting elements of DREB to withstand adversity, and then, proline and sucrose content is increased and resistance to a variety of stress (dryness, coldness, and salt) is enhanced [14]. Northern blot is carried out on PvSR6 (Phaseolus vulgaris stress-related) gene which encodes a kidney bean DnaJ-like protein. The results indicate that the expression of PvSR6 gene is weak in kidney bean leaves without any processing. Environmental stress such as heavy metals (Hg^{2+} and Cd^{2+}), mechanical damage, UV, high temperature, and salicylic acid could promote the transcription of this gene. It can be inferred that DanJ-like protein plays an important role in protecting the structure and function of membrane and enzyme as well as increasing the resistance of plants. These researches provide reference and thinking for discovering the effect of stress on expression of gene in Dao-di herbs [15]. Lan-ping Guo et al. [16] had used canonical correlation and stepwise regressive analysis to get the leading factors affecting the content of the essential oil components of *A. lancea* and 6 regressive models between climate factors and 6 main essential oil components. They found that high temperature is the limiting factors for *A. lancea*, and the interaction of temperature and precipitation is the key factor in forming the essential oil components of *A. lancea*. They used GIS software IDIRIEIW to extract the climate characteristic in Mt. Maoshan, the habitat of Dao-di herbs, and found that habitat of *Mao-A. lancea* was characterized by highest temperature, short drought season and more precipitation, and the forming of *Mao-A. lancea* related to the high-temperature stress. They also found that Dao-di herb of *A. lancea* faced nutrient stress and proved that high temperature and K shortage stress could make essential oil compounds more similar to the Dao-di herbs than others by experiments in greenhouse.

Effective constituents are the material basis that could determine the curative effect of materia medicas. Some effective constituents that few or no constituents exist under normal conditions are only produced under external stimulus (e.g., drought, freeze, injury), and these constituents belong to abnormal secondary components called phytoalexin.

In short, on one hand, environment change selects genotype of Dao-di herbs through long-term action, and on the other hand, it affects the formation and accumulation of secondary metabolites in Dao-di herbs by affecting the expression of gene in them. Through long time of adoption, medicinal plants have chosen their favorite natural environment. When environment changes or stress appears, plants' competitiveness for the limited resources by physical means would drop and chemical means would become more important. Secondary metabolites can protect the plants,

and they will be released to the outside to limit the growth of other plants when stress appears. In this situation, it seems that the reaction of Dao-di herbs to environment is more likely an "adverse effect."

4.4 Molecular Mechanism of the Formation of Dao-di Herbs

The morphology of Dao-di herbs and the diversity of secondary metabolites in them are caused by genetic variation of genes. Studies on molecular mechanism of the formation of Dao-di herbs are to reveal the genetic variation between populations of Dao-di herbs at the molecular level, to discover the effect of genotype and environment on gene expression, and then to reveal the contribution of genetic factors to the formation of Dao-di herbs. It is not only an important content of scientific meaning of Dao-di herbs but also the need of cultivation and breeding of Dao-di herbs.

Studying the relationship between genetic variation and secondary metabolites is the most direct way to reveal the genetic essential of authenticity of Dao-di herbs. Analyzing the genetic differentials and genetic structure at population level interspecies can be used to determine genetic variation of Dao-di herbs at molecular level. RAPD is used to analyze the genetic differentials between wild populations and cultivated populations of *Paeonia lactiflora* Pall (Shaoyao) and to explain the formation of Chishao and Baishao. This research has revealed the genetic differentials of Shaoyao populations and provides the basis for the explanation of the authenticity formation of Chishao and Baishao [17].

Some questions should be paid attention to in studies on molecular mechanism of Dao-di herbs:

1. Secondary metabolite is a typical multi-gene trait:
   Steps of plant secondary metabolism are numerous and complex. Each step in the metabolism has at least one gene at work. So, secondary metabolite is a typical character controlled by many genes. As we all know, biosynthesis of the various types of secondary metabolites which has been preliminarily investigated, such as terpenes, alkenes, alkaloids, flavonoids, anthraquinone coumarin, and a series of substances, must go through a considerable number of metabolic steps and involve a large number of key genes whichever secondary metabolic pathways are used. For example, at least 22 genes are involved, and more than 20 steps are gone through in the biosynthesis of aflatoxin B1 [18].
2. Variation of secondary metabolites is a continuous variation:
   According to multi-gene theory, gene passes on in accordance with Mendelian genetics. The effect of each gene is weak, so noncontinuous, quantitative gene mutation may produce a smooth, continuous phenotypic variation. Modification of the environment may cover the noncontinuous variation of genotype; thus, continuous phenotypic variation is shown and the trait variation becomes smooth and difficult to detect. As a typical multi-gene trait, secondary metabolite has shown continuous variation, and it not only brings confusion for the quality evaluation of Dao-di herbs but also increases the difficulty of the detection of its genotype.

3. The interaction of genes and environment should be taken seriously:
   The interaction between genes and environment is an important feature of multi-gene genetics. There are many forms of this interaction, which mainly include the following: ① environment affects the genetic structure of a group through the choices imposed on this group; ② genotype and environment shows interaction in the decision of difference between individuals or populations we observed directly because of the nongenetic effect caused by the interaction of genotype and environment in development of Dao-di herbs. Both of them play an important role in the formation of Dao-di herbs, and the second is an important reason why the production of Dao-di herbs is inseparable from specialized habitat. Studies on the effect of interaction of gene and environment on the formation of authenticity have just begun, but the rapid development of molecular geography, molecular ecology, and other related fields provides good ideas and methods for the researches in this field.
4. Quantitative genetics theory methods should be taken seriously:
   According to the forms of expression, genetic variation can be divided into two types: qualitative variation and quantitative variation. The former, which shows discontinuous variation, is decided by a few major genes and consistent with Mendel's law; the latter, showing as a continuous variation, is decided by many genes and follows the normal distribution. Genetic variation determining the quantitative traits is very susceptible to environmental effects and will interact with the environment. That is to say, quantitative traits is constituted by three parts, that is, genetic variation, environmental variation, and interaction between genetic and environment.

Phenotypic characteristics of famous Dao-di herbs are closely related to the environment firstly and then followed by a trait of continuity, which suggest that the genetic nature of Dao-di herbs may be more performed as the quantitative trait. Therefore, the characteristics of continuous variations should be more considered in looking for the specialized genes of Dao-di herbs. In addition, quantitative genetics should be studied deeply as a main point.

Quantitative genetics is a combination product of genetics and biostatistics. It mainly studies on the continuous variation decided by multi-gene genetic and its relationship with quantitative traits. Interaction between genetic and environment is the most important research field of quantitative genetics. It can be expected that quantitative genetics technology may become an important means for studying the molecular mechanism of the formation of Dao-di herbs.

4.5 Case Study

Impacts of Recent Cultivation on Genetic Diversity Pattern of a Medicinal Plant, Scutellaria baicalensis (Lamiaceae)

Cultivation of medicinal plants is not only a means for meeting current and future demands for large volume production of plant-based drug and herbal remedies

but also a means of relieving harvest pressure on wild populations. *Scutellaria baicalensis* Georgi (Huang-qin or Chinese skullcap) is a very important medicinal plant in China. Over the past several decades, wild resource of this species has suffered rapid declines and large-scale cultivation was initiated to meet the increasing demand for its root. However, the genetic impacts of recent cultivation on *S. baicalensis* have never been evaluated. In this study, the genetic diversity and genetic structure of 28 wild and 22 cultivated populations were estimated using three polymorphic chloroplast fragments. The objectives of this study are to provide baseline data for preserving genetic resource of *S. baicalensis* and to evaluate the genetic impacts of recent cultivation on medicinal plants, which may be instructive to future cultivation projects of traditional Chinese medicinal plants. Cultivation of medicinal plants is not only a means for meeting current and future demands for large volume production of plant-based drug and herbal remedies but also a means of relieving harvest pressure on wild populations. Scutellaria baicalensis Georgi (Huang-qin or Chinese skullcap) is a very important medicinal plant in China. Over the past several decades, wild resource of this species has suffered rapid declines and large-scale cultivation was initiated to meet the increasing demand for its root. However, the genetic impacts of recent cultivation on *S. baicalensis* have never been evaluated. In this study, the genetic diversity and genetic structure of 28 wild and 22 cultivated populations were estimated using three polymorphic chloroplast fragments. The objectives of this study are to provide baseline data for preserving genetic resource of *S. baicalensis* and to evaluate the genetic impacts of recent cultivation on medicinal plants, which may be instructive to future cultivation projects of traditional Chinese medicinal plants.

4.5.1 Methods

A total of 602 and 451 individuals of *S. baicalensis* representing 28 wild and 22 cultivated populations were collected from Northeast to Northwest China, respectively. Seventeen to twenty-four individuals were sampled for each population. Twenty-two Scutellaria rehderiana Diels individuals were sampled from a wild population of Weiyuan county in Gansu province (WYW) as an out-group. Leaves for DNA extraction were dried with silica gel. Voucher specimens were deposited in herbaria of Institute of Chinese Materia Medica (CMMI), China Academy of Chinese Medical Sciences. The precise geographic location of each sampled population was determined using a Garmin GPS unit.

Dried leaves were milled by using RETSCH Mixer Mill (MM301). Genomic DNA was extracted using a modified cetyltrimethyl ammonium bromide (CTAB) protocol [19]. The universal primers described in Hamilton [20] and Sang et al. [21] were used for screening polymorphic cpDNA fragments. After preliminary screening of eight chloroplast fragments, we chose atpB-rbcL, trnL-trnFand psbA-trnH intergenic spacers for the full survey because they contained the most polymorphic sites. PCR amplification was performed in a TC-512 thermocycler (Techne, England)

programmed for an initial 240 s at 94 °C, followed by 30 cycles of 45 s at 94 °C, 30 s at 55 °C (atpB-rbcL), 54 °C (trnL-trnF), or 58 °C (psbA-trnH), 90 sat 72 °C, and a final 240 sat 72 °C. Reactions were carried out in a volume of 20 μL containing 2.0 mm/L MgCl2, 0.5 μm/L dNTP, 10×buffer, 2.5 μm/L primer, 1 U Taq DNA, and 20 ng DNA template. Sequencing reactions were conducted with the forward or reverse primers of the PCR using the DYEnamic ET Terminator Kit (Amersham Pharmacia Biotech), following the manufacturer's protocol. Sequencing was performed on a HITACHI 3130 Genetic Analyzer (Hitachi High-Technologies Corporation, Tokyo, Japan) after the reaction product was purified through precipitation with 95 % ethanol and 3 M sodium acetate (pH 5.2).

Sequences were aligned using Clustal_X version 1.81 [22], and all indels were coded as substitutions following Caicedo and Schaal [23]. All individuals were characterized for cpDNA haplotype. Sequence variation was tested for deviations from neutrality by Tajima's D statistic [24] and by Fu and Li's D^* and F^* statistics [25] using DNASP 4.00 [26]. A haplotype network depicting the mutational relationships among distinct haplotypes was drawn following the principle of parsimony by TCS version 1.13 [27, 28], positing *S. rehderiana* as out-group. The geographical distribution of haplotypes was plotted on a map of China using Arcmap 8.3 (ESRI, Inc.). Significant difference of haplotype frequencies between wild and cultivated populations was quantified using x^2 test. Total diversity (hT), within-population diversity (hS), and level of population differentiation (GST) were calculated using the program HAPLONST, treating wild and cultivated populations as separate groups. The significance of the parameter comparisons (number of haplotype; total diversity, hT; within-population diversity, hS; population differentiation, GST) between wild and cultivated groups were estimated by x^2 test and Wilcoxon two-group tests, being implemented in the statistical package SPSS17.0 (SPSS Inc.).

Hierarchical structure of genetic variation in *S. baicalensis* was determined by an analysis of molecular variance (AMOVA) with ARLEQUIN version 2.000, partitioning genetic variance into three levels: among groups (cultivated and wild groups), among population within groups, and within population. Genetic differentiation between cultivated and wild groups was also evaluated using DNASP 4.00 [25]. To examine the effect of geographical distance on genetic structure, correlations between pairwise genetic distances (Kimura 2-parameter distance generated with Kimura's two-parameter model in MEGA 3) [29–31] and pairwise geographic distances were tested using a mantel test implemented by isolation by distance web service [32].

4.5.2 Results and Discussion

Sequence characteristics of three chloroplast intergenic spacers, the aligned sequences of atpB-rbcL, trnL-trnF, and psbA-trnH spacers in *S. baicalensis* and *S. rehderiana*, were 768, 799, and 502 base pairs in length, respectively. There were 13, 7, and 14 polymorphic sites (including substitutions and indels) for the three

spacers, respectively. Because sequencing poly-N regions could easily lead to homoplasies due to polymerase error [33], the short sequence repeats (poly-C) between 502 and 508 bp of trnL-trnF were not treated as polymorphic sites and were removed from subsequent analyses. Sequence divergence, as measured with Kimura two-parameter algorithm, ranged from 0.000 to 0.262%, 0.000 to 0.257%, and 0.000 to 0.559% for atpB-rbcL, trnL-trnF, and psbA-trnH, respectively. Sequences of the three spacers were consistent with the expectation of neutrality by Tajima's criterion (atpB-rbcL:D = −1.53470, $0.10 > P > 0.05$; trnL-trnF: D = −1.31009, $P > 0.10$ and psbA-trnH:D = −1.38479, $P > 0.10$) and FuandLi's criterion (atpB-rbcL: D* = −1.66525, $P > 0.10$; F* = −1.79736, $P > 0.10$; trnL-trnF: D* = −1.40980, $P > 0.10$; F* = −1.51361, $P > 0.10$; and psbA-trnH: D* = −1.28584, $P > 0.10$; F* = −1.48316, $P > 0.10$). The pooled sequences of the three spacers also complied with the expectation of neutrality ($D = -0.90417$, $P > 0.10$; $D^* = -1.33584$, $P > 0.10$; $F^* = -1.40744$, $P > 0.10$). The sequences of 9 atpB-rbcL, 8 trnL-trnF, and 13 psbA-trnHhaplotypes have been deposited in GenBank databases [GenBank: GQ374124-GQ374155]. Thirty-two haplotypes of *S. baicalensis* (HapA-Y and Hap1-7) and 2 haplotypes of *S. rehderiana* (HapZ1-Z2) were identified when atpB-rbcL, trnL-trnF, and psbA-trnH sequences were combined.

Haplotype network and distribution in the haplotype network, the 32 haplotypes of *S. baicalensis* and 2 haplotypes of *S. rehderiana* differed by at least 10 mutations. The genealogical structure of 32 haplotypes of *S. baicalensis* presented a shallow gene tree with three obvious centers: HapG, HapB, and HapC, which connected to other 10, 6, and 4 haplotypes by just one mutation, respectively. Additionally, both HapB and HapC connected to HapG by one mutation. Exceptionally, several haplotypes detected only in wild populations were relatively remote to the three centers of the shallow gene tree, such as HapA, H, M, T, W, X, and Y. The haplotypes of cultivated *S. baicalensis* (symbolized by ovals and circles) evenly distributed across the haplotype network, not showing any clusters.

The most noticeable pattern was that haplotypes among wild populations were much more structured than those among cultivated populations. For example, HapG in wild populations (WTW, LCW, LPW, CD1W, SXW, and HXW) was mainly restricted to the central range of this species, whereas this haplotype occurred in 17 of 22 cultivated populations across the whole range. Another conspicuous phenomenon was that many wild populations (10 out of 28) were fixed by one unique haplotype, but in cultivated populations, only SXC was fixed by one haplotype (i.e., most cultivated populations characterized by multiple haplotypes). In genetic diversity and genetic structure, of the 32 haplotypes detected in *S. baicalensis*, 25 (78% of the total number of haplotypes) were recovered in wild populations and 22 (69% of the total number of haplotypes) were carried by cultivated individuals. Fifteen haplotypes (47% of the total number of haplotypes) were shared by wild and cultivated populations. Ten haplotypes were found in wild populations but not in cultivated ones and seven were found exclusively in cultivated populations. By comparison, the number of haplotypes and the relative abundance of each haplotype have slightly changed under the anthropogenic influence during the course of cultivation.

However, the change of haplotype frequencies between wild and cultivated groups was not significant by χ^2 test ($P = 0.733$), because cultivated and wild groups only differed in rare haplotypes. As indicated by the haplotype frequencies within population, the within population diversity (hS = 0.649) of cultivated S. baicalensis was significantly higher than that (hS = 0.265) of wild Huang-qin ($P < 0.001$, Wilcoxon two-group test), although the total haplotype diversity was similar between cultivated and wild groups (hT = 0.832 in cultivated vs. hT =0.888 in wild, $P > 0.05$, Wilcoxon two-group test). Consistent with the haplotype distribution, population subdivision of wild populations (GST = 0.701) was significantly higher than that of the cultivated (GST = 0.220; $P < 0.001$, Wilcoxon two-group test). Mantel test analyses showed that genetic distance was significantly correlated with geographical distance ($r = 0.4346$, $P < 0.0010$) in wild populations, however, this pattern was not recovered in cultivated populations ($r = 0.0599$, $P = 0.2710$). AMOVA analysis showed that little genetic variation occurred between cultivated and wild populations (0.09%, $P < 0.001$), and most genetic variance were among populations (56.61%) and within populations (43.30%). The genetic differentiation (FST = 0.022) between the cultivated and wild groups calculated by DNASP 4.00 was well consistent with the result of AMOVA analysis.

4.5.3 Conclusion

Facing the rapidly growing demands for medicinal plants, domestic cultivation is a viable and long-term way of protecting wild medicinal plant resources. However, this study indicates that bringing a species into cultivation may impose profound impacts on genetic diversity patterns and even the evolutionary potentials of medicinal plants. Although the total genetic diversity maintained in cultivated *S. baicalensis* is comparable to wild populations due to the large initial population size and the short cultivation history, substantial rare alleles have lost and extensive seed exchange has caused the homogenization of cultivated populations during the course of cultivation. This study not only provides baseline data for preserving genetic resource of *S. baicalensis* through conservation-by-cultivation approach but also represents a paradigm for evaluating the genetic impacts of recent cultivation on medicinal plants, which may be instructive to future cultivation projects of traditional Chinese medicinal plants.

References

1. Hu, S. L. (1989). *Geoherbs in China*. Haerbin: Heilongjiang Science and Technology Press.
2. Xie, Z. W. (1990). Discussion about geoherbs. *Journal of Traditional Chinese Medicine, 40*(10), 43–46.
3. Huang, L. Q., & Zhang, R. X. (1997). Biological essential about geoherbs. *Chinese Pharmaceutical Journal, 9*(32), 563–566.
4. Gao, W. Y., Qin, E. Q., Xiao, X. H., et al. (2011). Analysis on genuineness of *Angelica sinensis* by RAPD. *Chinese Traditional and Herbal Drugs, 32*(10), 926–929.

5. Guo, B. L., Wu, M., Si, J. P., et al. (2001). Research on DNA molecular marker of *Magnolia officinalis* rehd. Et wils.—RAPD study on certified species. *Acta Pharmaceutica Sinica, 36*(5), 386–389.
6. Wen, F., & Zhuang, J. (2001). The characteristics of geoherbs. *Journal of Beijing University and Traditional Chinese Medicine, 24*(1), 47.
7. Guo, L. P., Liu, J. Y., Ji, L., et al. (2002). The naphtha composing characteristics of geoherbs of *Atractylodes lancea*. *China Journal of Chinese Materia Medica, 27*(11), 814–819.
8. Guo, B. L., Wu, M., Si, J. P., et al. (2000). Genetic study on the geoherbalism of *Magnolia officinalis*. *Journal of Pharmacy Practice, 18*(5), 314–316.
9. He, B. H., & Zhong, Z. C. (2003). Study on variation dynamics of modular population of *Ginkgo biloba* under different conditions of environmental stress. *Journal of Southwest University, 25*(1), 7–10.
10. Tang, C. S., Cai, W. F., Kohl, K., et al. (1995). Plant stress and allelopathy. *ACS Symposium Series, 582*, 142–148.
11. Hall, A. B., Blum, U., & Fites, R. C. (1982). Stress modification of allelopathy in *Helianthus annuus* L. debris on seed germination. *American Journal of Botany, 69*, 776–781.
12. Josep, P., & Joan, L. (1997). Effects of carbon dioxide, water supply, and secondary on terpene content and emission by *Rosmarinus officinails*. *Journal of Chemical Ecology, 23*, 979–993.
13. Brant, J. P., Reichardt, P. B., Clausen, T. P., et al. (1993). Effects of mineral nutrition on delayed inducible resistance in Alaska paper birch. *Ecology, 74*, 2072–2084.
14. Wang, S. X., Wang, Z. Y., & Peng, Y. K. (2004). Dehydration responsive element binding (DREB) transcription activator and its function in plant tolerance to environmental stresses. *Plant Physiology, 40*(2), 7–13.
15. Zhang, Y. X., Chai, T. Y., Li, J. C., et al. (2000). Expression analysis of DnaJ-like protein gene from bean under various environment stresses. *Acta Botanica Boreal – Occident Sinica, 20*(2), 171–174.
16. Guo, L. P., Huang, L. Q., Yan, H., et al. (2005). Habitat characteristics for the growth of *Atractylodes lancea* based on GIS. *China Journal of Chinese Materia Medica, 30*(8), 565–569.
17. Zhou, H. T., Hu, S. L., Guo, B. L., et al. (2002). A study on genetic variation between wild and cultivated populations of *Paeonia lactiflora* Pall. *Acta Pharmaceutica Sinica, 37*(5), 383–388.
18. Huang, Y. F. (2002). *Characterization of aflatoxin biosynthesis genes between aflatoxigenic and non-aflatoxigenic strains of Aspergillus section flavi*. Dissertation, National Chiayi University.
19. Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, 19*, 11–15.
20. Hamilton, M. B. (1999). Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology, 8*, 521–522.
21. Sang, T., Crawford, D. J., & Stuessy, T. F. (1997). Chloroplast DNA phylogeny, reticulate evolution, and biogeography of Paeonia (Paeoniaceae). *American Journal of Botany, 84*, 1120–1136.
22. Thompson, J. D. T. J., Gibson, P. F., Jeanmougin, F., et al. (1997). The windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research, 24*, 4876–4882.
23. Caicedo, A. L., & Schaal, B. A. (2004). Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene. *Molecular Ecology, 13*, 1871–1882.
24. Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics, 123*, 585–595.
25. Fu, Y. X., & Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics, 133*, 693–709.
26. Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X., et al. (2003). DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics, 19*, 2496–2497.
27. Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology, 9*, 1657–1659.

28. Templeton, A. R., Crandall, K. A., & Sing, C. F. (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics, 132*, 619–633.
29. Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution, 16*, 111–120.
30. Kumar, S., Tamura, K., & Nei, M. (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics, 5*, 150–163.
31. Nei, M., & Kumar, S. (2000). *Molecular evolution and phylogenetics*. New York: Oxford University Press.
32. Jensen, J. L., Bohonak, A. J., & Kelley, S. T. (2005). Isolation by distance web service. *BMC Genetics, 6*, 13.
33. Olsen, K. M., & Schaal, B. A. (1999). Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proceeding of the National Academy of Sciences of the United States of America, 96*, 5586–5591.

Suggested Readings

Gao, W., Hillwig, M. L., Huang, L. Q., Cui, G., Wang, X., Kong, J., Yang, B., & Peters, R. J. (2009). A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. *Organic Letters, 11*(22), 5170–5173.

Yuan, Q. J., Zhang, Z. Y., Hu, J., Guo, L. P., Shao, A. J., & Huang, L. Q. (2010). Impacts of recent cultivation on genetic diversity pattern of a medicinal plant, *Scutellaria baicalensis* (Lamiaceae). *BMC Genetics, 11*(1), 29.

Chapter 5
Seeking for New Members of Origin Materials (A New Usage of Plant Species) for CMM

Wen-yuan Gao and Kee-yoeup Paek

Abstract Two major problems of medicinal plants are developing more effective resources and searching for new medicine resources. In recent years, many studies are focused on medicinal plant resources. Molecular systematics of medicinal plants plays an important role in identifying medicinal plants and searching for new medicinal plants. This chapter introduces development of plant systematics, research methods of plant systematics, theoretical basis for molecular systematics of medicinal plants, and some study cases in resources of medicinal plants. Research methods of plant systematics contain morphological taxonomy, anatomy taxonomy, embryo taxonomy, and cell taxonomy. Determination of monophyletic group and character analysis were also based on these theories. Beside the taxonomy, nDNA, cpDNA, and mtDNA, existing in the plant cells and responsible for the differences of the genome structure and function, provide alternative and diversified sources of characters for phylogenetic. These studies all provide the theoretical basis for new members of medicinal plants.

5.1 Introduction

For centuries, efforts have been made to develop safe and effective new medicinal resources. Two problems have been focused on searching for new resources of crude drugs: develop more effective new resources and solve the problem of

Wen-yuan Gao (✉)
Tianjin Key Laboratory for Modern Drug Delivery and High-Efficiency,
School of Pharmaceutical Science and Technology, Tianjing University,
Tianjin, 300072, China
e-mail: pharmgao@tju.edu.cn

Kee-yoeup Paek
Department of Horticultural Science, Research Center for the Development
of Advanced Horticultural Technology (Horttech), Chungbuk National University,
Cheongju, 361-763, Republic of Korea
e-mail: paekky@chungbuk.ac.kr

Lu-qi Huang (ed.), *Molecular Pharmacognosy*,
DOI 10.1007/978-94-007-4945-0_5,

shortage of medicinal resources. Traditional Chinese medicine industry is restricted by lack of medicinal resources. New members of medicinal plants play a significant role in the development of traditional Chinese medicine. The discovery of *Arnebia euchroma* (Royle) Johnst. is a typical classic example of development of new members of medicinal plants.

The roots of some genera of the Boraginaceae family contain Alkanna, Onosma, Arnebia, Lithospermum, and Echium species. *Lithospermum erythrorhizon* belongs to *Lithospermum* species, which was first recorded in a 2,000-year-old Chinese medical material dictionary called "Shen Nong Ben Cao Jing," and is related to the therapeutic actions such as anti-inflammatory, antivirus, antibacterial, and immunomodulatory. Only *Lithospermum* has been broadly applied as a traditional herbal medicine for thousands of years in China [1]. However, *A. euchroma* did not appear in pharmacopoeia until 1990. This herb can be found all over China, including Western Xin Jiang, Tibet, as well as Nepal, Palestine, north of India, and Iran. *A. euchroma* was discovered in Xinjiang in the late 1950s. Most compositions of *A. euchroma* are the same as that of *Lithospermum*. What is more, the content of shikonin and its derivatives are three times as much as that of *Lithospermum* [2]. This advantage might make it have better efficacy and higher quality [3]. *A. euchroma* (Royle) Johnst., *L. eryhrorhizon*, and *A. guttata* Bunge have been recorded in the *Chinese Pharmacopoeia* since 1990.

The development of pharmacognosy is based on the chronic and regional evolution of medicinal resources. Crude drug research not only focuses on authenticity, efficacy, and safety but on the discovery of new Chinese medicine resources and sustainable use of the resources. Therefore, discovery of new members of crude drugs has a great significance.

5.2 Theoretical Basis and Research Methods of Searching for New Members of Medicinal Resources

5.2.1 Concept of New Members of Medicinal Resources

The medicinal resources (plants, animals, and minerals) have not been numbered among rude medicinal material and are identified to have biological activity with modern technology. Medicinal plants occupy the most important part in the development of traditional Chinese medicine. New members of medicinal plants can be considered as new medicinal plants in wild, cultivated species, artificial mutation individuals, cultivars of tissue culture, and rapid propagation [4].

5.2.2 *Theoretical Basis and Research Methods on Searching for New Members of Resources*

Theoretical basis of searching for new members of resources is pharmacophylogenetics. Plants of the same genus and family are genetically related, which are associated with certain type of chemical components, and the activity of medicinal plants. According to this discipline, we can search for new medicinal plants.

Resource investigation of medicinal species is an effective measure to search for new medicinal resources. Pharmacophylogenetics is regarded as a guide for wildlife resources through resource investigation of medicinal species. Many medicinal plants were founded, for example, *A. euchroma* [1] and *Ferula sinkiangensis* K. M. Shen from Xinjiang [5, 6], *Lycium* from Qinghai [7, 8], and *Picrorhiza scrophulariiflora* from Tibet [9, 10]. *Salvia yunnanensis* and *S. przewalskii* of the same genus as *S. miltiorrhiza* Bge. were founded across pharmacophylogenetics [11]. However, the contents of effective constituents in the former two plants are higher than the latter one.

5.2.2.1 The Concept of Pharmaphylogeny

During the early evolution history of Earth, plants gradually formed a relative relationship. The species with close relationship show not only a similar kind but also similar physiological and biochemical characteristics. Therefore, the chemical contents as the secondary metabolites are often similar. The species also have similar biological activity and potency. Similarly, species which have the same effective possess close relationships on systematics. Proposition of pharmaphylogeny provides theoretical guidance and direction for medicinal plants [4, 12–17].

Pharmaphylogeny as a discipline is interdisciplinary and marginal, which studies the relationship, chemical constituents, and efficacy of medicinal plants. It ranges over plant taxonomy, plant phylogenetics, phytochemistry, pharmacology, numerical taxonomy, genomics, and informatics.

5.2.2.2 The Background of Pharmaphylogeny

Traditional taxonomy takes the morphology and shape of plants as the basis, which has many limitations and artificiality. There are some complex issues in the field of taxonomy, such as phylogeny. Traditional taxonomy could not analyze such phenomenon. The development of science affords the opportunity for the development of chromosome classification, DNA molecular hybridization, and microscopic classification on the basis of the morphological classification. All of the classifications

supplement the traditional method. The evolvement of pharmaphylogeny went through two stages:

① In the 1970s of twentieth century, plant chemotaxonomy played an important role in exploring the distribution of chemical constituents of medicinal plants, which provided a basis for molecular classification. And plant chemotaxonomy also revealed the law of plant phylogeny at the molecular level. Plant chemotaxonomy which studies the evolvement of medicinal plants is based on the chemical compositions and traditional morphological taxonomy. The distribution of chemical constituents of many species of plants at home and abroad was studied. Previous researches only focused on the distribution of certain chemical compounds. Therefore, there were no systematic study and any theory formed. On the other hand, a combination of pharmacological effects and efficacy was not systematically implemented.

② From the late 1970s, Xiao PG successively conducted a comprehensive study on Belladonna, Berberine [4], rhubarb, Aconitum, Thalictrum, Lithospermum, puhuan, and azaleas by using phytochemistry, pharmacology, plant phylogenetics, numerical taxonomy, and computer technology. From these researches, Xiao PG revealed the correlation of pharmaphylogeny-chemical composition efficacy (biological activity and traditional effects) and applied this theory into practice. This theory makes a great achievement in expanding the resources of medicinal plants, discovering new medicinal resources and finding out substitute for imported drugs, and providing guidance for the basic research of Chinese herbs.

5.2.3 Pharmacophylogenetic Study Method

The determination methods of pharmacophylogenetic study contain resource investigation of medicinal species, chemical constitutes, DNA fingerprint technology, metabonomics, serum pharmacochemistry, and so on.

5.2.3.1 Metabonomics and Serum Pharmacochemistry

Metabonomics involves the composition changes of all kinds of materials in the metabolic system. Through the multiple index analysis, high-throughput detection, and data processing, metabolic constitution changes of organism or tissue cell system are regarded as study objects, especially the characteristics and the influence of endogenous metabolism, genetic variation, environmental change, and all kinds of materials into the metabolic system. Serum pharmacochemistry mainly focuses on the chemical composition of serum after patients take it orally. The methods above provide promising approaches to assess the therapeutic effects of drugs. Urine has been heavily used in metabolomics studies when it is minimally invasive to the animals or human and primarily reveals an overall metabolic state of the given organism.

5.2.3.2 The Application of Modem Separation Technologies

Efficacy is based on chemical compositions. Certain types of chemical components are distributed in certain plant groups of certain families and genus. The aim of pharmacophylogenetic study is to preliminarily reveal the distribution principle in order to guide screening of potential drugs. The study on the chemical composition is not only connected with the types, structure, properties, and content of chemical constituents but also with the metabolic pathway analysis. Chemical composition is critical for pharmacophylogenetic study to determine the chemical compositions.

5.2.3.3 Taxonomy Combined with Technology of Modern Information

Through morphological characterizations and molecular properties, medicinal plants with high medicinal value should be comparably analyzed. Morphology properties from organology, anatomy, embryology, and palynology can be comparably studied to provide systematic evidences for plant groups. Gene or DNA sequence of chloroplast, mitochondria, and nuclear genome is applied to reveal species-pharmacological effect relationship and taxonomy of medicinal plants.

5.2.3.4 Selecting the Right Mathematic Model and Establishing Intelligent Database

Combined with artificial intelligence and KDD (Knowledge Discovery in Database), by selecting the right mathematical model to explore the corresponding software, the data collected was analyzed. In this process, numerical taxonomy and branch taxonomy were applied based on proper mathematical model. Various intelligent database of plant groups-chemical composition, chemical composition-pharmacological effect, plant groups-pharmacological effect, and plant groups-molecular biology characterizations should be set up, which could comprehensively present pharmacophylogenetics of plants of the same genus and family.

5.3 Overview of Phylogeny of Medicinal Plants

The context of evolutionary biology is phylogeny, the connections between all groups of organisms as understood by ancestor/descendant relationships. Not only is phylogeny important for understanding paleontology, but paleontology in turn contributes to phylogeny. Many groups of organisms are now extinct, and without their fossils, we would not have as clear a picture of how modern life is interrelated. Systematics is the study of the diversification of life on the planet earth, both in the past and at present, and the relationships among living organisms through time. Taxonomy is the science of naming and classifying the diversity of

organisms, and it is a major part of systematics that includes four components: description, identification, nomenclature, and classification. Plant systematics is an introduction to the morphology, evolution, and classification of land plants. It contains two parts: study on the classification system and study on the methodology of classification system construction. Molecular system of medicinal plants is an important part of molecular system of plants. The difference between the two is that the objective of molecular system of medicinal plants is the abundant group of molecular system of plants and the classification grade is mainly dominated by species and genera.

5.3.1 Development of Plant Systematics

Study on the plant systematics goes through three stages: artificial systems stage, pre-evolutionary natural systems stage, and phylogenetic systems stage.

5.3.1.1 Artificial Systems Stage

The "Book of Odes" in 700 BC is the first book clearly recording plant species, involving 160 classes of plants, of which 10 corresponds to generals and 112 corresponds to particular species. At the same time, the father of botany *Theophrastus* published "Historian Plantarum," which involves 480 plants. This book refers to arbor, shrub, and herb, and herbs are divided into annual plants, biennial plants, and perennial plants.

"Shen Nong's Herbal," the first documented monograph existing, loads 365 drugs, which are divided into three classes: the superior, the middle, and the inferior. "Compendium of Materia Medica" written by Li Shizhen is the most important in the history of study on the medicinal plants, and it records 1,892 plants, of which 1,195 are plant medicines, which are divided into five dictionaries: straw, corn, vegetable, fruit, and wood.

The appearance of microscope made plant systematics step into a new stage, and then, anatomy, embryology, palynology, etc., form one after another. Plant systematics becomes a special science, resulting in many artificial systems. The "sexual system" proposed by *Carl Linnaeus* is the most representative. In this system, plants are divided into 24 classes mainly according to the presence or absence, amount, and consolidation of stamen. It was believed that numerious species were created by the Almighty God at that time. Therefore, this system has no idea of species evolution, and it is also called artificial systems.

5.3.1.2 Natural Systems Stage

After the second half of the nineteenth century, people mastered more and more knowledge about plants and many scholars made efforts to seek for genetic relationship and development law of plant in nature, thus natural systems stage was established.

These systems, such as Bentham and Hooker system, Jussieu system, and De Candolle, take morphologic characteristics as categorizing base and do not reflect the affinity and evolution relationship. Hence, this stage is called natural systems stage.

5.3.1.3 Phylogenetic Systems Stage

After the theory of biological evolution was proposed by *Darwin* in "Origin of Species," taxonomists reappraised the established system and tried to establish phylogenetic classification systems.

The first system of this kind was published in "Species and Medicine Botany Seminar" by *W. Eichler*, which constituted the basic for Engler classification system. In his system, gymnosperm was placed in front of angiosperm and monocotyledon was arranged before dicotyledon. Engler classification system was first published in "Die Naturichen Ptlanzen-familien" by *A. Engler* and *Prantl* in 1897, and it was the first relatively complete natural classification system in the history of plant classification. Hutchinson classification system is angiosperm classification system raised by *J. Hutchinson* in "The Family of Flowering Plant" in 1926 and was further revised in 1959 and 1973, respectively. Soviet scholar *A. Takhtajian* published Takhtajian system in 1954, which firstly broke through conventional concept, and established "superorder" classification system. American scholar *A. Cronquist* published Cronquist system in 1958 and revised this system in 1981.

With the development of molecular biology, plant molecular systematics was born. Angiosperm phylogeny group (APG) I and II are the most influential molecular systems at the present time.

5.3.2 Research Method of Plant Systematics

The earliest research method of plant systematics took the morphological characteristics of plants as the classification evidences, especially flowers and fruits, and on the basis of these, botanists established many categorization systems. However, the fossil flower is difficult to find, so this system is difficult to establish. With the progress of science and inter-science penetration, plant systematics has rapidly developed during the past decades, resulting in many new research methods. The application of these methods in plant systematics makes the plant classification systems become more rational and closer to the objective reality.

5.3.2.1 Morphological Taxonomy

Morphological taxonomy uses the characteristics of herbarium specimen, such as the leaves, flowers, fruit, branches, roots, and other external morphological characteristics, to classify and name the plant species. In addition to the specimen museum

work, it still needs to investigate, collect, and record the actual variation in nature and the variation degree. This approach is characterized by simple manipulation equipment and practicability in application. But it has the following shortcomings: (1) informative taxonomy with a long history requires taxonomists to memorize the past articles, (2) it is easy to form understanding that breaks away from ecology and heredity and to propose a new species that does not exist at all in nature, and (3) too much emphasis on morphological characteristic.

In using morphology in taxonomy, it is very important to know the range of variability in a particular character under different conditions. Morphometric, allometric, and morphological characters tend to vary under the influence of geographical and ecological conditions which result in ecotypic and host-specific populations. Such populations are sometimes described as new species, although the differences between them are intraspecific variations.

5.3.2.2 Anatomy Taxonomy

Anatomy taxonomy is a method that uses optical microscopy to observe the construction of plants and provide the basis for plant classification. For example, some characteristics of Rutaceae, such as morphology, located position, and density, are helpful for the analysis of the relationships between the subfamily and genus.

5.3.2.3 Embryo Taxonomy

Embryo taxonomy mainly focuses on the ovule shape and male and female gametophyte. The collection of the experimental materials is difficult, and the results have a high similarity in the low taxa. Therefore, this method is limited in the phylogenetic study of medicinal plants.

5.3.2.4 Cell Taxonomy

Chromosome data plays a special role in taxonomy, and chromosome often varies with degrees in the intragenus, interspecies, and intraspecies, which provides an important basis for investigating evolutionary relationship between genus and species and intraspecific variation pattern. The study of cell taxonomy contains chromosome number, karyotype analysis, analysis of reduction division, banding technique in chromosomes, in situ hybridization of genome, etc., and these data can be applied to investigate the relationships of intergenus, interspecies and intraspecific variation, and to reveal the origin of species. The experimental materials of cell taxonomy are live and are more difficult to collect, so its application is limited.

5.3.2.5 Chemical Taxonomy

Morphological characteristics of plants are controlled by genes, which are the result of long-term natural selection. So it has important reference value to use chemical method to study the individual and systematic development of plants. Currently, chemical taxonomy mainly uses plant chemistry and biochemistry to study some compounds with finite distribution and their biosynthetic pathway and provides more objective basis for measuring the similarity and relationship among plant taxa. Chemical taxonomy is divided into serum taxonomy and organic chemical taxonomy. Chemical taxonomy could solve classification problems from the subspecies to the present order. Species that have difference in chemical constituents are not found out and valued by classical taxonomists but concerned by medicinal botanists. The differences of chemical constituents are not only the basis and material for drafting medicine production divisions and medicinal plant breeding but also one of the reasons for the formation of "dao-di herbs."

5.3.2.6 Numerical Taxonomy

Numerical taxonomy is a classification system in biological systematics which deals with the grouping by numerical methods of taxonomic units based on their character states. It aims to create a taxonomy using numeric algorithms like cluster analysis. The concept was first developed by Robert R. Sokal and Peter H.A. Sneath in 1963 and later elaborated by the same authors. Phenetics is a closely related discipline and draws heavily from the methods of numerical taxonomy. Although intended as an objective classification method, in practice, the choice and weighing of morphological characteristics is often guided by available methods and research interests. Furthermore, the general consensus has become that the taxonomic classification should reflect evolutionary (phylogenetic) processes. Some connections between phylogenetic trees and the spectral decomposition of the variance-covariance matrix of quantitative traits subject to Brownian motion over time have been established, providing a theoretical link between phylogenetic methods and numerical taxonomy. The specific phenetic algorithms proposed in numerical taxonomy, however, often failed to properly reconstruct the phylogenetic history of organisms. Numerical taxonomy remains useful in cases where biological species concepts cannot be applied, for example, clonal evolution as in apomictic microspecies like blackberries.

5.3.2.7 Branch Taxonomy

Branch taxonomy is also called cladistics, created by a German entomologist W. Henning. Cladistics is a method of classifying species of organisms into groups

called clades, which consist of all the descendants of an ancestral organism and the ancestor itself. Cladistics can be distinguished from other taxonomic systems, such as phenetics, by its focus on shared and derived characters (synapomorphies). Previous systems usually employed overall morphological similarity to group species into genera, families, and other higher level classifications; cladistic classifications (usually trees called cladograms) are intended to reflect the relative recency of common ancestry or the sharing of homologous features. Cladistics is also distinguished by its emphasis on parsimony and hypothesis testing (particularly falsificationism), rather than subjective decisions that some other taxonomic systems rely upon.

5.3.2.8 Molecular Systematics

Molecular systematics is the use of the structure of molecules to gain information of an organism's evolutionary relationships. The result of a molecular phylogenetic analysis is expressed in a phylogenetic tree.

In short, various evidences could be used to establish plant classification systems, so the research methods is evolving, of which morphology method is the basis, and some methods or techniques, such as plant anatomy, embryology, cytology, chemistry, numerical taxonomy, cladistics, molecular biology, etc., could be also used. These methods have their own characteristics, application, and limitations. In the practical use, several methods are combined, and the most natural-effective relationship and evolution of medicinal plant could be obtained by comparison and analysis.

5.4 Theoretical Basis, Research Method, and Application of Molecular Systematics of Medicinal Plants

Molecular systematics of medicinal plants is an interdisciplinary science developed by the interaction of molecular biology and medicinal plant systematics. It uses various experimental methods of molecular biology to obtain a variety of molecular characters for discussing the classification of medicinal plants, phylogenetic relationships between groups, and processes and mechanisms of evolution. Except for DNA sequence, molecular characters include structure features of genome, characters of proteins, DNA fingerprint characteristics, etc. In practice, according to the different taxon, we can select different molecular characters to obtain the best and the largest resolution about the phylogenetic relationships of specific groups. As the genetic material of plants, DNA is stable, reliable, and free from outside influence, and it can be used in the study of phylogenetic relationships of medicinal plants. Plant molecular systematics is actually cladistics using DNA sequence data.

Methods commonly used in plant molecular systematics include restriction fragment length polymorphism analysis (RFLP) method, simple sequence repeat (SSR)

method, single nucleotide polymorphism (SNP) method, DNA barcoding method, etc. At present, the most commonly used method is sequence analysis.

5.4.1 Characteristics

The most important step in systematics is the selection and processing of biological characteristics. It could be considered that whether the findings are closer to nature is mainly based on whether the selected species and the characteristics used can complement each other. So, we firstly introduce the definition and application of characteristics.

5.4.1.1 The Concept of Characteristics

Characteristics refer to the full form of any organism, structure, physiology, biochemistry, or the performance in response to the environmental factors and are the sum of the kind of all biological features. Characteristics are the differences between one (group) organisms and the other taxa of organisms and share any property with the same unit organism. Different characteristics could be obtained with various approaches.

5.4.1.2 The Application of Characteristics

1. The Variation in Phenotypic Characteristics
   There are no two identical plants in nature. It is polymorphic with difference within the same species. Mastering the variation range of morphological features could avoid regarding the two extremes of the mutations as new species. Abnormalities of morphological should be correctly handled, which are caused by the atavistic, disease, single gene mutation, or genetic recombination. On the other hand, with the increase of enlargement range, the complexity of the discrepancy of the same characteristics will also rise. Therefore, it is important to know the complexity of the same characteristics and grasp the scopes and the correlation within them when using morphological characteristics.
2. The Classification of Phenotypic Characteristics
   Phenotypic characteristics are divided into high and low features. High characteristics vary in a small range and are also not easily altered by the environmental factors, for example, the morphological characteristics of flowers compared to other organs.
3. The Related Factors of Phenotypic Characteristics
   Phenotypes should be the product of the interaction between genes and environment. Genes have the potential to cause a particular phenotype, which is essential for achieving phenotype. However, genes could be affected by environment for the reason that the growth and the traits of organic must rely on the conditions.

5.4.2 Cladistics

Cladistics has raised a series of new concepts and methods for analyzing characters such as plesiomorphy, apomorphy, outgroup, and synapomorphy. As the continually maturation of theory of cladistics, increasingly perfectness of computer programs, and application of new evidence, accurate reconstruction of systematic development becomes real and feasible. In recent years, phylogenetic tree published in various literatures about taxonomy and evolutiology is rapidly increasing.

5.4.2.1 The Basic Theory of Cladistics

1. The Principle of Synapomorphy
   The characteristics of organisms can be divided into three types: ① plesiomorphy is a characteristic inheriting from the remote ancestors of organisms. Synaplesiomorphy is a plesiomorphy which is shared by the two organisms. ② Apomorphy is a characteristic deriving from the direct ancestors of organisms. ③ Convergence is the concordance exhibited by the characteristics with different origins under the similar conditions. Only based on the classification of synapomorphy can a natural parent group be formed, and that is the principle of synapomorphy.
2. The Principle of Strict Monophyly
   A monophyletic group is a group of species derived from the same ancestral species and includes all the offspring it produces. Classification system must clearly show the class structure style of synapomorphy and establish a strict nested series of monophyletic groups.
3. The Principle of Parsimony
   The principle of parsimony is a methodological principle that expresses trait distribution, and it takes the common traits owned by the two sister groups as inherited from their ancestor. This idea is simpler in evolution steps and is more likely evolutionary hypothesis. The principle of parsimony is a procedure, and it does not necessarily mean that the actual process of evolution is simple. In spite of the variety of classification systems, systematics scholars still believe that there is a best system existing, and the principle of parsimony is the mathematical expression of this view.

Of these three principles, the first principle is the most basic one, and other two principles belong to methodology, which have different significance in practical applications.

5.4.2.2 Methods of Cladistics

Determination of Monophyletic Group

Bifurcation analysis requires that the studied taxa should be monophyletic. The meaning of this requirement is that the analyzed traits should have the same performance forms in the group; otherwise, these groups should be further divided until the character does not vary within groups.

Character Analysis

Character analysis has four steps: ① observe the similarity and difference of character, which is the first step to understand the biosphere diversity; ② assume that the similarity is inherited from the same ancestor, that is homology hypothesis, which is the central task of systematics and morphology; ③ draft evolution polar of homologous characters which is the most characteristic and basic step in bifurcation analysis, and the principles are outgroup comparison, commonality principle, principle of character correlation, etc.; and ④ reestablish phylogenetic relationships based on a large number of homologous assumption.

1. Discrimination of Character Homology
   The central task of systematics is to analyze the character similarity and to transfer it into homologous assumption. Homology is the basic view of comparative biology and is one of the core evidences existing during the evolution process. Similarity refers to the corresponding or equivalent relationships between structures. They can be considered as different states of the same character, which forms homologous assumption. Homologous assumption displays a transformation series.

   Three principles to examine homologous assumption should be followed. ① Similarity test – Homologous structures should be similar (exiting corresponding relationships in the location, composition, and structure). Similarity is the basis of the homology, but it is only a requirement of homologous assumption. ② Conjunction test – Different statuses of homologous characters cannot exist in an organism. ③ Congruence test – New information and evidence should not be inconsistent with the original homologous assumption, and homologous assumption of different characters should be coordinated with each other.
2. Discrimination of Character Polar
   It is the most characteristic and basic step in bifurcation analysis. Biologists found that the evidences of phyogenetic speculation are mainly from three aspects: comparative anatomy, paleontology, and ontogeny, which is "threefold parallelism."

   ① *Outgroup comparison*. The method of outgroup comparison means looking at a closely related species which is known to be phylogenetically outside the group of species we are studying. The character state in that outgroup is likely to have been ancestral in the group under consideration. Outgroup comparison is fallible. Sometimes, one possible outgroup will suggest that one character state is ancestral, but another possible outgroup will suggest that a different character state is ancestral. The result will then depend on which outgroup we rely on. The method is most reliable when the closely related species that could be used as outgroups all suggest the same inference, but it is possible to be led astray by the method in particular cases. The inference should be treated with caution and if possible tested against other evidence.

   ② *Recapitulation law*. Recapitulation law, also called the biogenetic law or embryological parallelism and often expressed as "ontogeny recapitulates phylogeny," is a discredited biological hypothesis.

   ③ *Commonality principle*. Assume that the most common state of character is the original state. This principle is closely related to outgroup comparison.

Outgroup comparison confirms character states exiting in outgroups and ingroups as the original states. So outgroup comparison could be considered as a special case of commonality principle.

④ *Principle of character correction*. It confirms and assumes that all characteristics of an original type are original in a group.

Bifurcation Analysis and Computer Programs

After the determination of monophyletic group and character analysis, we can establish mathematical operation in bifurcation analysis. Currently, computer programs in cladistics, including Henning 86, PAUP 4.0, and PHYLIP 3.4. PAUP 4.0, are the most popular programs at present.

Establishment of Cladogram and Biological Taxonomy

Cladists use cladograms and diagrams which show ancestral relations between taxa, to represent the evolutionary tree of life. Although such cladograms traditionally were generated largely on the basis of morphological characters, molecular sequencing data and computational phylogenetics, they are now very commonly used in the generation of cladograms. When cladograms are transferred into classification systems, two principles should be followed: the principle of strict monophyly and principle of categorical ranking.

Reliability Evaluation of Cladogram

As the true phylogeny is unknown, the reliability evaluation of cladogram only bases on the data matrix and the structural characteristics of the cladogram itself and then determines what kind of data and the structure of cladogram are more likely to reflect the true phylogeny based on the principle of evolution and statistics. Currently, the most widely used criteria for reliability evaluation of the entire cladogram are consistency index (CI) and retention index (RI).

5.4.3 Research Method of Molecular Systematics of Medicinal Plants

Through the related knowledge, we know that life is complex system going through lengthy evolution and succession. In the process of biological evolution, evolution of macromolecules is collectively called molecular evolution.

There are nDNA, cpDNA, and mtDNA existing in the plant cells. The evolutionary rates of plants are different in genome structure and function. In general, nDNA evolves

the quickest, about two times of cpDNA; mtDNA evolves the slowest, one-third of cpDNA. Generally, due to less restriction in the sequence, noncoding region shows faster evolutionary rate than coding region; the variations and rates are vastly different among the coding or noncoding sequences within the same genome. These differences in evolutionary rates among sequences provide alternative and diversified sources of characters for phylogenetic study of different taxon. It is now clear that selection of the appropriate molecular fragments contrary to a particular problem in systematics is the most fundamental and most crucial step in the molecular systematic studies.

When selecting a sequence for phylogenetic analysis, we should consider the following points: ① this sequence should be long enough to provide sufficient nucleotide sites with phylogenetic information, and the difference percentage of the selected sequence should be suitable for the system problems to be solved; ② this sequence must be easy to sort, which is important for the correct assessment of character homology; and ③ this sequence must be orthologous.

5.4.3.1 nDNA

The current study of nDNA sequences focuses on the nrDNA encoding rRNA. Ribosomes are made from complexes of RNAs and proteins. Ribosomes are divided into two subunits, one larger than the other. The smaller subunit binds to the mRNA, while the larger subunit binds to the tRNA and the amino acids. Genes encoding rDNA in plant nucleus are multigene family composed of some highly repetitive sequences. 18S encoding small subunit rDNA in ribosome, 5.8S and 26S compose of a transcription unit, and gene intergenics between 18S and 5.8S and 5.8S and 26S are called ITS1 and ITS2, respectively.

① 18S rRNA. The 18S rRNA in most eukaryotes is in the small ribosomal subunit. 18S sequence can provide more information for phylogenetic study of high classification level within angiosperm. Its sequence variation is suitable for discussing the relationship between deep phylogenetic branch within angiosperm and seed plant. Due to the differences of variation of sequences in different groups, 18S can also be used for reestablishment for relationships between subfamily and species.

② Internal transcribed spacer (ITS). ITS between 18S and 26S nrDNA is divided into two parts (ITS1 and ITS2) by 5.8S.

The variations of ITS region in gymnosperms are complex, and only the length variation in ITS1 region may vary several kb. So ITS sequence analysis is one of the most widely used molecular sequences for the molecular systematic study of gymnosperms. In contrast, in angiosperms, ITS region has both a high degree of nucleotide sequence variation and high conservation in length, which indicates that these sequences in septal area are easy to sort in sibling groups. Rich variation in the lower taxon can solve the problem of plant phylogeny.

Each transcription unit contains an external transcribed spacer (ETS), 18S, 5.8S, and 26S. Non-transcribed spacer (NTS) is between genes.

5.4.3.2 cpDNA

cpDNA in plants is closed double-stranded DNA, accounting to about 10–20% of the total DNA. The length of cpDNA is usually 120–220 kb, and its length variation is mainly caused by two inverted repeats (IR). The length of the two inverted repeats is 22–25 kb, and it divides the total cpDNA into large single copy region (LSC) and small single copy region (SSC). cpDNA is exposed and does not form complex with histone, which is similar to *E. coli* DNA.

5.4.3.3 mtDNA

Mitochondrial DNA (mtDNA) is the DNA located in organelles called mitochondria, structures within eukaryotic cells that convert the energy from food into a form that cells can use. In humans (and probably in metazoans in general), 100–10,000 separate copies of mtDNA are usually present per cell (egg and sperm cells are exceptions). In mammals, each double-stranded circular mtDNA molecule consists of 15,000–17,000 base pairs.

The low effective population size and rapid mutation rate (in animals) make mtDNA useful for assessing genetic relationships of individuals or groups within a species and also for identifying and quantifying the phylogeny (evolutionary relationships; see phylogenetics) among different species, provided they are not too distantly related. To do this, biologists determine and then compare the mtDNA sequences in different individuals or species. Data from the comparisons is used to construct a network of relationships among the sequences, which provides an estimate of the relationships among the individuals or species from which the mtDNAs are taken. This approach has limits that are imposed by the rate of mtDNA sequence change.

5.4.4 Application of Molecular Systematics of Medicinal Plants

Under the influence of today's trend of "back to nature," the development and application of medicinal pant resources have aroused widespread interest, so pharmaphylogeny has developed grounded on the medicinal plant systematics. Pharmaphylogeny focuses on the phylogenetic correlation between phylogeny, chemical constituents, and pharmaceutical effects of medicinal plants. Like many other contemporary fields of science, pharmaphylogeny represents a highly interdisciplinary science which is one of the areas of pharmaceutical education. Its scope includes the study of phylogeny, botanical taxonomy, chemical, biochemical, and biological properties of natural medicinal plants as well as the search for new drugs from natural sources. Study of rhubarb is the typical one. Rhubarb is a commonly used Chinese herbal medicine, but the quality of the goods in some areas is found unstable. Study on the pharmaphylogeny of 7 groups 27 species

(44 samples) of domestic rheum showed that authentic rhubarb with obviously genuine cathartic effect distribute in sect. palmate. They contain sennoside and free rhein but no rhaponticin; in morphology, the leaves have palmate division of various degrees. The results of numerical taxonomic and multivariate analysis indicated that the division of leaves has close relation with sennoside, cathartic and purgative effects, which thus provided reliable scientific evidence for rhubarb in quality control, production, and utilization of resources. This proves the importance of medicinal plant systematics.

At present, as an important part of medicinal plant systematics, study of molecular systematics of medicinal plants becomes one of the hot spots in pharmacognosy. Its application mainly focuses on the following two aspects.

5.4.4.1 Identification and Classification of Medicinal Plants

A rigorous scholar, when making study on each discipline of medicinal plants, must have the correct identification and classification of the experimental materials. Without a correct identification, the experimental results would be difficult to be verified and easy to be confused. In practice, taxonomists often apply the concept of morphology based on vision, which has a strong direct viewing effect and can meet the needs of multipurpose classification. But the situation that it can only be understood, not explained, may inevitably appear. Specific groups have more in-depth research, comparative anatomy, palynology, cytology, and other means that can be used to define species more clearly, but for experts in non-species groups, it is still difficult to identify the plant species. The rise of molecular systematics makes it possible to change classification from according to morphological features to DNA sequences. DNA barcoding is a useful attempt.

DNA barcoding as a taxonomic method uses a short genetic marker in an organism's DNA to identify a plant as being a particular species. It differs from molecular phylogeny because the main goal of it is not to determine the classification but to identify an unknown sample in terms of a known classification. Although barcodes are sometimes used in an effort to identify unknown species or assess whether species should be combined or separated, such usage, if possible at all, pushes the limits of what barcodes are capable of. Applications include identifying plant leaves even when flowers or fruit are not available, identifying the diet of an animal based on stomach contents or feces, and identifying products in commerce (e.g., herbal supplements or wood).

5.4.4.2 Search for New Medicinal Plants

In the study of active constituents in medicinal plants, especially by family and genus, certain types of chemical components are found to distribute in certain families and genus. According to this discipline, we can search for new medicinal plants. For example, most of gentian plants contain bitter glycosides; Ephedra

Branch, Ranunculaceae, Papaveraceae, Menispermaceae, Rubiaceae, Loganiaceae, Lilium, and other plants mainly contain alkaloids; Labiatae, Rutaceae, Umbelliferae, Lauraceae, and other plants contain different types of volatile oil; and Polygonaceae mainly contain anthraquinone glycosides. Colchicines was initially found in foreign Liliaceae Colchicum, while in recent years, it was also found in *I. indica Kunth*. in the same family in China. Plants in the same genus may contain the same types of chemical components, for example, Ephedra contains ephedrine and Strychnos contains strychnine and brucine. The results indicate that genetic relationship of medicinal plants provides clues and proof for searching for new medicine. Molecular systematics makes clues for searching for new medicine more accurate. According to the phylogenetic correlation between phylogeny, chemical constituents, and pharmaceutical effects of medicinal plants and on the basis of molecular systematics and genetic relationship of medicinal plants, searching for new medicinal resources from the relationships between species and genus is the most main application of study of molecular systematics of medicinal plants.

5.5 Case Study

A Taxonomic Study on the Original Plant of Radix Astragali

The 2005 Edition of *Chinese Pharmacopoeia* recorded the dry roots of *Astragali membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao or *A. membranaceus* (Fisch.) Bge. as the quality products. *A. pallidipurpureus* stat.nov is the peculiar plant of Astragalus in China. According to Flora of China, *A. pallidipurpureus* stat. nov grows in Gansu, Ningxia, Qinghai, and Sichuan provinces (the northwestern part), which has equal medical values as *A. membranaceus* (Fisch.) Bge.

The original plants of Radix Astragali, that is, *A. membranaceus* (Fisch.) Bge., *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao, and *A. pallidipurpureus* stat.nov, have kinds of variable groups. It is hard to distinguish the three kinds plants for the similar appearances. Scholars or experts hold different ideas on the taxonomy of these three kinds of Radix Astragali. Based on the previous studies, this chapter took the view of population and the modern technologies of molecular biology to investigate the genetic relations among different kinds of the original plants of Radix Astragali at the DNA molecular level. Besides, this chapter aimed to explore the taxonomy and the origin of Radix Astragali.

5.5.1 Materials and Methods

The herbals in this study were collected from Gansu Province, Inner Mongolia, Shanxi Province, Heilongjiang Province, Jilin Province, and Shandong Province, which were fresh leaves and were fast dried by the silica gels in the outdoors, stored

in the specimens' pavilion of Institute of Chinese Materia Medica, Chinese Academy of Chinese Medical Sciences, and identified by Prof. Zhu Xiangyun from Institute of Botany, Chinese Academy of Sciences.

The researchers adapted revised CTAB (cetyltrimethyl ammonium bromide) to extract total DNA, and the amplimers of trnH-psbA fragment were *trnH* (5′-CGCGCATGGTGGATTC ACAATCC-3′) and *psbA* (5′-GTTATGCATGAA-CGTAATGCTC-3′). The amplification experiment had been carried out with the usage of Eppendorf Mastercycler Personal PCR Thermal Cycler. Reaction system was set in the total volume of 20 μL, which included DNA 50 ng, 10×PCR Buffer 2 μL, dNTP 2 μL, 10 μmol/L forward primer 0.5 μL, 10 μmol/L reverse primer 0.5 μL, ExTaqDNA enzyme 1 u, and complementary ddH_2O for the total volume. The procession of the amplification reaction: the predenaturation at 94 °C for 4 min; denaturation at 45 °C for 45 s, renaturation at 58 °C for 30 s, extension at 72 °C for 90 s, and totally 30 repeated cycles; and extension at 72 °C for 7 min and in store at 4 °C.

The products of PCR were sequenced and amplified with the usage of BigDye Terminator (BDT) v3.1 Elisa Kit of ABI. The reaction system was set in the total volume of 10 μL, which included 1 μL above-mentioned PCR products, primer (3.2 μmol/L) 2 μL, BigDye staining 1 μL, and ddH_2O 6 μL. The procession of sequencing: the predenaturation at 96 °C for 1 min, denaturation at 96 °C for 10 s, renaturation at 56 °C for 5 s, extension at 60 °C for 4 min, and in store at 4 °C, totally 25 repeated cycles. The products were purified through alcohol/EDTA/NaAc and sequenced by ABI 3130 Genetic Analyser.

The results of sequencing were analyzed for checking and putting in order by the software of Contig Express and adjusted by the researchers. The base sequences were compared by the software of ClustalX version 1.81, and all the indels and gaps were considered to be the same mutations. After the adjusting and correction of all the sequences, cpDNA haplotypes were analyzed statistically. With the usage of HAPLONST, the genetic variance in the population of cpDNA, the general genetic diversity h_T, and genetic differentiation coefficient among the populations' G_{ST} were calculated. With the usage of TCS version 1.13 software, the mesh evolutionary tree of the haplotypes was drawn.

5.5.2 Results and Discussion

After putting in order, the length of the sequence among *trnH-psbA* genes of 407 individual chloroplasts of 21 populations of Radix Astragali was 331 bp without the length variations. Abundant A/T covered 74.29~74.45% of the bases, as this was consistent with the bases constitutions among DNA genes of most of chloroplasts. Nucleotide mutations appeared in 12 sites, which came into 10 haplotypes (Hap 1~Hap 10). The sequences of these 10 haplotypes had been registered in GenBank Database with the registration number of GQ139474~GQ139483, respectively. After comparison with the sequences of these haplotypes, 8 mutation sites were

found. The Indels and Gaps were 15 bp, 31 bp, 89 bp, 108 bp, 109 bp, 131–132 bp (2 bp), 133–139 bp (7 bp), and 240 bp.

Through the HAPLONST programming software, the genetic variance in the population of h_S, the general genetic variance h_T, and genetic differentiation coefficient G_{ST} of all the populations of the original plants of Radix Astragali were 0.326, 0.766, and 0.575, respectively.

In the 407 individual of 21 populations of Radix Astragali, the frequencies of the ten haplotypes were Hap1 = 0.098, Hap2 = 0.467, Hap3 = 0.130, Hap4 = 0.025, Hap5 = 0.002, Hap6 = 0.032, Hap7 = 0.182, Hap8 = 0.052, Hap9 = 0.002, and Hap10 = 0.010.

With the software of TCS, the mesh evolutionary tree of 10 haplotypes was drawn. The genetic relationships among those haplotypes were inferred in accordance with coalescent theory with the confidence above 95%.

All the *A. membranaceus* (Fisch.) Bge., *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao, and *A. pallidipurpureus* stat.nov belong to membranaceus group, Astragalus subgenus, Astragalus genus, Papilionoideae, and Leguminosae. For the taxonomy of *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao, scholars hold different theories. The main theories are ① as the variant, in terms of the hairiness situations of the ovaries and the sizes and the numbers of the small leaves, the classic taxonomic theory supports that *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao is the variant of A. membranaceus (Fisch.) Bge. ② As the subspecies, according to the morphological characteristics of the pollen, Zhu Xiangyun [5] holds that it is more appropriate to take *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao as the subspecies of *A. membranaceus* (Fisch.) Bge. ③ As the independent species, based on the blooming time of *A. membranaceus* (Fisch.) Bge. and *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao, standing and creeping state of annual stem, consolidation location of bracteole, and chromosome difference, it is proposed that *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao would be taken as an independent species by Wang Ertong, as many scholars and experts support this theory. ④ Some scholars think that *A. membranaceus* (Fisch.) Bge. and *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao would be taken together as one species.

The results of our study demonstrated that 40 individual chloroplasts of 2 populations of *A. pallidipurpureus* stat.nov have specific haplotype Hap1, while all the populations of *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao have Hap2, Hap3, Hap4, Hap5, and Hap10. Most of *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao (190/258, 73.6%) have Hap2, and *A. membranaceus* (Fisch.) Bge. have Hap6, Hap7, Hap8, and Hap9 4 haplotypes as most (74/109, 67.9%) have Hap7. Judged from the distribution of the haplotypes and the mesh evolutionary tree, 407 samples of *A. pallidipurpureus* stat.nov in our study have specific haplotype, Hap1, which distinguishes *A. pallidipurpureus* stat.nov from most of *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao and *A. membranaceus* (Fisch.) Bge. easily. Either *A. membranaceus* (Fisch.) Bge. or *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao has its own exclusive haplotypes without crosses.

Judged by the mesh evolutionary tree drawn based on the haplotypes of chloroplasts, it is obvious that the haplotypes of *A. pallidipurpureus* stat.nov and *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao are diffracted through the center of the evolutionary tree, which is the haplotypes of *A. membranaceus* (Fisch.) Bge. It is inferred that *A. membranaceus* (Fisch.) Bge. is the more original plant and *A. pallidipurpureus* stat.nov and *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao are the variants differentiated from *A. membranaceus* (Fisch.) Bge. The systematic places of sorts of Radix Astragali in the haplotypes evolutionary tree have similarity to the morphologic places, especially the specific haplotypes and purple flowers of *A. pallidipurpureus* stat.nov.

The study is mainly based on the studies of exsiccate stored in specimens' pavilion of Institute of Botany, Chinese Academy of Sciences, with the field studies in the Radix Astragali predominantly distributed areas of Gansu Province. The key indexes form of the differences in the morphologies of Radix Astragali of 3 origins are in below:

① Yellow flowers or yellow white flowers, big-sized, 12~20 mm, loose orders of small leaves, 6~13 pairs, elliptic ovate, big-sized, 15 cm, *A. membranaceus*
② Close orders of small leaves, 12~18 pairs, small-sized, commonly shorter than 10 mm, *A. membranaceus* var. *mongholicus*
③ Purple flowers or purple black flowers, small-sized, 8~15 mm, *A. pallidipurpureus*

Through the conventional morphological identification methods, picking as many properties as possible and mainly observing the variation ranges and the relations of properties help distinguish these three original plants. The methods of modern molecular biological and haplotypes are achieved fully as 100% to distinguish the plants in the labs. *trnH-psb*A fragments have been successfully adapted in the distinguishment of the original plants of Radix Astragali, which is a beneficial experiment for the distinguishment and a rather good supplement and verification to the traditional methods.

Above all, the authors support the theory that *A. membranaceus* (Fisch.) Bge. var. *mongholicus* (Bge.) Hsiao is taken as the variant and propose that *A. pallidipurpureus* stat.nov should be taken as the independent species.

References

1. Wu, D., & Li, C. Y. (2008). Textual research on *Arnebia*. *Lishizhen Medicine of Materials and Medical Research, 19*(8), 2042–2043.
2. Kaori, T., Jin, T., Yuko, O., et al. (2005). Ethylene induced shikonin biosynthesis in shoot culture of *Lithospermum erythrorhizon*. *Plant Physiology and Biochemistry, 43*, 101–105.
3. Han, J., Weng, X. C., & Bi, K. S. (2008). Antioxidants from a Chinese medicinal herb – *Lithospermum erythrorhizon*. *Food Chemistry, 106*, 2–10.
4. Chen, S. B., Peng, Y., Chen, S. L., et al. (2005). Introduction of pharmaphylogeny. *World Science and Technology-Modernization of Traditional Chinese Medicine, 7*(6), 97.
5. Wang, W. J., Shang, Y. H., & Jing, S. (2005). Protection and development of *Ferula sinkiangensis* K.M.Shen. *Capital Medical, 6*, 45–46.

6. Yang, J. R. (2006). *Studies on the chemical constituents of Ferula Teterrima, F. sinkiangensia and F. ferulaeoides*. Dissertation, Beijing: Peking Union Medical College.
7. Li, Y. L., Fan, Y. F., Dai, G. L., et al. (2011). Analysis of genetic diversity for wolfberry germplasms by AFLP technology. *Chininese Traditional and Herbal Drugs, 42*(4), 770–773.
8. Zhang, X. Y. (2003). *Studies on environmental condition of high quality medlar in ningxia autonomous region*. Dissertation, Beijing: China Agricultural University.
9. Wang, H. Z. (2004). *Studies on the chemical constituents and hepatoprotective activities of picrorhiza scrophulariiflora*. Dissertation, Nanjing: China Pharmaceutical University.
10. Zhao, D. M. (2011). *Extraction, separation and pharmacological study on the active components of Tibetan medicine picrorhiza*. Dissertation, Lanzhou: Lanzhou University.
11. Shi, Y. B., Xue, M., Cui, Y., et al. (2002). Pharmacognostical study on *Salvia przewalskii*. *Lishizhen Medicine of Materials and Medical Research, 13*(7), 438–439.
12. Peng, Y., Chen, S. B., Liu, Y., et al. (2006). A pharmacophylogenetic study of the Berberidaceae (s.l.). *Acta Phytotaxonomica Sinica, 44*(3), 241–257.
13. Peng, Y., Chen, S. B., Chen, S. L., et al. (2006). Preliminary pharmaphylogenetic study on Ranunculaceae. *China Journal of Chinese Materia Medica, 31*(13), 1124–1128.
14. Xiao, P. G., Jiang, Y., Li, P., et al. (2007). The botanical origin and pharmacophylogenetic treatment of Chinese materia medica Beimu. *Acta Phytotaxonomica Sinica, 45*(4), 473–487.
15. Xiao, P. G., Wang, F. P., Gao, F., et al. (2006). A pharmacophylogenetic study of *Aconitum* L. (Ranunculaceae) from China. *Acta Phytotaxonomica Sinica, 44*(1), 1–46.
16. Yang, H. J., Tang, S. H., Huang, L. Q., et al. (2008). Research on nature of Chinese materia medica in view of genetic relationship. *China Journal of Chinese Materia Medica, 33*(24), 2983–2985.
17. Ma, J. R., Li, Z. X., Wen, H. Y., et al. (2009). Identification of the original plant of Tibetan material medica Zangyinchen. *Guihaia, 29*(4), 552–555.

Chapter 6
Salvation of Rare and Endangered Medicinal Plants

Lu-qi Huang and Chao-yi Ma

Abstract At present, a large amount of species are facing with serious threats even dying out. Corresponding protective measures set standards for grading precious and endangered animals and plants. Therefore, we need to know the precious and endangered medicinal animals and plants and the currently used standards for grading them. This chapter will firstly introduce the IUCN Red List of Threatened Species Categories and Criteria, CITES appendix standards, U.S. Endangered Species Act, endangered species criteria, grading standards for China Plant Red Data Book, grading standards for the endangered of red list, and the national directory standards of wild animals under special protection.

The genetic diversities are the bases of evolution and species differentiations, while the hereditary varieties are equal to genetic diversities. The genetic composition of species determines its characteristics, which include its adaptability to certain environment and its utilizability for humans. Any specific individuals and species keep plenty of genetic types. To that extent, these individuals and species are considered as independent gene banks. The genetic diversities including the genetic varying degrees in molecular, cellular, and individual levels are the bases of biological evolution and species differentiations.

To keep the genetic diversity of one species, if there is no genetic information available, it will be needed to protect the majority of the species and a vast enough environment for supporting the species and preventing from inbreeding and genetic drifting. Therefore, understanding the genetic diversity will help reduce the number of species for protection, so that it will then reduce the competitive paradox between the reduction of costs and the usage of soil.

For the protection of the endangered plants and animals, in-site and off-site conservation principles will be adopted, that is, in the original sites protect the ecological

Lu-qi Huang (✉) • Chao-yi Ma
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China
e-mail: huangluqi@263.net; machaoyi56@yahoo.com.cn

Lu-qi Huang (ed.), *Molecular Pharmacognosy*,
DOI 10.1007/978-94-007-4945-0_6,

environment for preserving the species and breeding the wild species and off the original sites protect and breed the species. Therefore, understanding the genetic diversity and the extinct mechanism of the rare and endangered plants and animals is of importance for protecting them.

6.1 Introduction

There are a lot of world-famous valuable and rare organisms in our country with vast territory, complicated natural conditions, and rich biological species. It is important to use wildlife resources rationally and protect rare and endangered organisms for the development of economy, the conduction of scientific researches, the improvement of natural environment, the maintenance of ecological balance, and the enrichment of people's life.

For a long time, people have thought of Chinese materia medica resources as being inexhaustible, and they have always used them excessively in the production practice. Meantime, the scope of human activities has been kept expanding. These resulted in the decline of medicinal plant and animal resources, as well as the increase of rare and endangered species. Among more than 800 kinds of commonly used Chinese herbal medicines, there are 10 species in an endangered state, 38 species in a rare state, and 64 species in an easily endangered state. To protect the endangered animals and plants, the state has published books such as "*List of Wild Medicinal Species under Special State Protection," "China Plant Red Data Book: Rare and Endangered Plants*". It is necessary to get some knowledge of rare and endangered medicinal plants and animals in order to strengthen the sustainable development of Chinese herbal medicine resources and further to promote the sustainable utilization of them.

6.2 Meaning and Classification Standard of Rare and Endangered Medicinal Plants

Rare and endangered animals and plants mean that wildlife species are endangered because of their own species conditions or the influence of human activities or natural disasters. In a broad sense, rare and endangered animals and plants refer to those valuable, endangered, or rare wild animals and plants.

At present, there are many criteria for dividing the levels of rare and endangered animals and plants, and this section will focus on the criteria of IUCN (International Union for Conservation of Nature and Natural Resources) Red List of Threatened Species, the CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora) appendix, the U.S. Endangered Species Act, as well as the *China Plant Red Data Book: Rare and Endangered Plants*, and the *List of Wild Animals under Special State Protection*.

6.2.1 IUCN Species Endangerment Categories

IUCN (The International Union for Conservation of Nature and Natural Resources) founded in October 1948, is the world's largest union for the conservation of nature. IUCN has been publishing the Red List of Threatened Species since the 1960s. The endangered species have been divided into nine classes according to IUCN Red List of Threatened Species.

6.2.1.1 Extinct (EX)

A taxon is extinct when there is no reasonable doubt that the last individual has died. A taxon is presumed extinct when exhaustive surveys in known and/or expected habitat, at appropriate times (diurnal, seasonal, annual), and throughout its historic range have failed to record an individual. Surveys should be over a time frame appropriate to the taxon's life cycle and life form.

6.2.1.2 Extinct in the Wild (EW)

A taxon is extinct in the wild when it is known only to survive in cultivation, in captivity, or as a naturalized population (or populations) well outside the past range. A taxon is presumed extinct in the wild when exhaustive surveys in known and/or expected habitat, at appropriate times (diurnal, seasonal, annual), and throughout its historic range have failed to record an individual. Surveys should be over a time frame appropriate to the taxon's life cycle and life form.

6.2.1.3 Critically Endangered (CR)

A taxon is critically endangered when the best available evidence indicates that it meets any of the criteria A to E for critically endangered (shown as follows), and it is therefore considered to be facing an extremely high risk of extinction in the wild.

A. Reduction in population size based on any of the following:

① The causes of a population size reduction of ≥90% over the last 10 years or three generations, whichever is the longer, are clearly reversible and understood, and ceased.

② The causes of population size reduction of ≥80% over the last 10 years or three generations, whichever is the longer, may not have ceased or may not be understood or may not be reversible.

③ A population size reduction of ≥80% may be projected or suspected to be met within the next 10 years or three generations, whichever is the longer (up to a maximum of 100 years).

④ A population size reduction over any 10-year or three-generation period, whichever is longer (up to a maximum of 100 years in the future), where the time period must include both the past and the future, is projected or suspected to be ≥80%. The reduction and its causes may not have ceased or may not be understood or may not be reversible.

B. Geographic range in the form of either B① (extent of occurrence) or B② (area of occupancy) or both:

① Extent of occurrence estimated to be less than 100 km^2 and estimates indicating at least two of (a–c):

(a) Severely fragmented or known to exist at only a single location.
(b) Continuing decline, observed, inferred, or projected in any of the following: extent of occurrence; area of occupancy; area, extent, and/or quality of habitat; number of locations or subpopulations; and number of mature individuals.
(c) Extreme fluctuations in any of the followings: extent of occurrence, area of occupancy, number of locations or subpopulations, and number of mature individuals.

② Area of occupancy estimated to be less than 10 km^2 and estimates indicating at least two of a–c under B①.

C. Population size estimated to number fewer than 250 mature individuals and either:

① An estimated continuing decline of at least 25% within 3 years or one generation, whichever is longer (up to a maximum of 100 years in the future)
② A continuing decline observed, projected, or inferred in numbers of mature individuals and at least one of the following (a–b):

(a) Population structure in the form of one of the following:

(i) No subpopulation estimated to contain more than 50 mature individuals
(ii) At least 90% of mature individuals in one subpopulation.

(b) Extreme fluctuations in number of mature individuals.

D. Population size estimated to number fewer than 50 mature individuals.
E. Quantitative analysis showing the probability of extinction in the wild is at least 50% within 10 years or three generations, whichever is the longer (up to a maximum of 100 years).

6.2.1.4 Endangered (EN)

A taxon is endangered when the best available evidence indicates that it meets any of the criteria A to E for endangered (shown as follows), and it is therefore considered to be facing a very high risk of extinction in the wild.

A. Reduction in population size based on any of the following:

① The causes of a population size reduction of ≥70% over the last 10 years or three generations, whichever is the longer, are clearly reversible and understood, and ceased.
② The causes of population size reduction of ≥50% over the last 10 years or three generations, whichever is the longer, may not have ceased or may not be understood or may not be reversible.
③ A population size reduction of ≥50% may be projected or suspected to be met within the next 10 years or three generations, whichever is the longer (up to a maximum of 100 years).
④ A population size reduction over any 10-year or three-generation period, whichever is longer (up to a maximum of 100 years in the future), where the time period must include both the past and the future, is projected or suspected to be ≥50%. The reduction and its causes may not have ceased or may not be understood or may not be reversible.

B. Geographic range in the form of either B① (extent of occurrence) or B② (area of occupancy) or both:

① Extent of occurrence estimated to be less than 5,000 km^2 and estimates indicating at least two of (a–c):

(a) Severely fragmented or known to exist at no more than five locations
(b) Continuing decline observed, inferred, or projected in any of the following: extent of occurrence; area of occupancy; area, extent, and/or quality of habitat; number of locations or subpopulations; and number of mature individuals
(c) Extreme fluctuations in any of the following: extent of occurrence, area of occupancy, number of locations or subpopulations, and number of mature individuals

② Area of occupancy estimated to be less than 500 km^2 and estimates indicating at least two of a–c under B ①.

C. Population size estimated to number fewer than 2,500 mature individuals and either:

① An estimated continuing decline of at least 20% within 5 years or two generations, whichever is longer (up to a maximum of 100 years in the future)
② A continuing decline observed, projected, or inferred in numbers of mature individuals and at least one of the following (a–b):

(a) Population structure in the form of one of the following:

(i) No subpopulation estimated to contain more than 250 mature individuals
(ii) At least 95% of mature individuals in one subpopulation

(b) Extreme fluctuations in number of mature individuals.

D. Population size estimated to number fewer than 250 mature individuals.
E. Quantitative analysis showing the probability of extinction in the wild is at least 20% within 20 years or three generations, whichever is the longer (up to a maximum of 100 years).

6.2.1.5 Vulnerable (VU)

A taxon is vulnerable when the best available evidence indicates that it meets any of the criteria A to E for vulnerable (shown as follows), and it is therefore considered to be facing a high risk of extinction in the wild.

A. Reduction in population size based on any of the following:

① The causes of a population size reduction of ≥50% over the last 10 years or three generations, whichever is the longer, are clearly reversible and understood, and ceased.
② The causes of population size reduction of ≥30% over the last 10 years or three generations, whichever is the longer, may not have ceased or may not be understood or may not be reversible.
③ A population size reduction of ≥30% may be projected or suspected to be met within the next 10 years or three generations, whichever is the longer (up to a maximum of 100 years).
④ A population size reduction over any 10-year or three-generation period, whichever is longer (up to a maximum of 100 years in the future), where the time period must include both the past and the future, is projected or suspected to be ≥30%. The reduction and its causes may not have ceased or may not be understood or may not be reversible.

B. Geographic range in the form of either B① (extent of occurrence) or B② (area of occupancy) or both:

① Extent of occurrence estimated to be less than 20,000 km^2 and estimates indicating at least two of (a–c):

(a) Severely fragmented or known to exist at no more than ten locations
(b) Continuing decline observed, inferred, or projected in any of the following: extent of occurrence; area of occupancy; area, extent, and/or quality of habitat; number of locations or subpopulations; and number of mature individuals
(c) Extreme fluctuations in any of the followings: extent of occurrence, area of occupancy, number of locations or subpopulations, and number of mature individuals

② Area of occupancy estimated to be less than 2,000 km^2 and estimates indicating at least two of a–c under B①.

C. Population size estimated to number fewer than 10,000 mature individuals and either:

① An estimated continuing decline of at least 10% within 10 years or three generations, whichever is longer (up to a maximum of 100 years in the future)

② A continuing decline observed, projected, or inferred, in numbers of mature individuals and at least one of the following (a–b):

(a) Population structure in the form of one of the following:

(i) No subpopulation estimated to contain more than 1,000 mature individuals

(ii) All mature individuals in one subpopulation.

(b) Extreme fluctuations in number of mature individuals.

D. Population very small or restricted in the form of either of the following:

① Population size estimated to number fewer than 1,000 mature individuals.

② Population with a very restricted area of occupancy (typically less than 20 km^2) or number of locations (typically five or fewer) such that it is prone to the effects of human activities or stochastic events within a very short time period in an uncertain future and is thus capable of becoming critically endangered or even extinct in a very short time period.

E. Quantitative analysis showing the probability of extinction in the wild is at least 10% within 100 years.

6.2.1.6 Near Threatened (NT)

A taxon is near threatened when it has been evaluated against the criteria but does not qualify for critically endangered, endangered, or vulnerable now but is close to qualifying for, or is likely to qualify for, a threatened category in the near future.

6.2.1.7 Least Concern (LC)

A taxon is of least concern when it has been evaluated against the criteria and does not qualify for critically endangered, endangered, vulnerable, or near threatened. Widespread and abundant taxa are included in this category.

6.2.1.8 Data Deficient (DD)

A taxon is data deficient when there is inadequate information to make a direct or indirect assessment of its risk of extinction based on its distribution and/or population status. A taxon in this category may be well studied, and its biology well

known, but appropriate data on abundance and/or distribution are lacking. Data deficient is therefore not a category of threat. Listing of taxa in this category indicates that more information is required and acknowledges the possibility that future research will show that threatened classification is appropriate. It is important to make positive use of whatever data are available. In many cases, great care should be exercised in choosing between DD and a threatened status. If the range of a taxon is suspected to be relatively circumscribed, and a considerable period of time has elapsed since the last record of the taxon, threatened status may well be justified.

6.2.1.9 Not Evaluated (NE)

A taxon is not evaluated when it is not yet been evaluated against the criteria.

6.2.2 CITES Appendix Criteria

CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora, also known as the Washington Convention) is an international agreement between governments, drafted as a result of a resolution adopted in 1973 at a meeting of IUCN members. Its aim is to ensure that international trade in specimens of wild animals and plants does not threaten their survival. Each species or population protected by CITES is included in one of three lists called Appendix I, Appendix II, and Appendix III, respectively. The appendix that lists a species or population reflects the extent of the threat to it and the controls that apply to the trade, known as the Berne Criteria. Appendix I lists species that are threatened with extinction and may be affected by trade, which has the same criteria with the endangered grade of IUCN; Appendix II lists species that are not necessarily now threatened with extinction but may become so unless trade is closely controlled, which is similar with the vulnerable grade of IUCN; Appendix III is a list of species included at the request of a party that already regulates trade in the species and that needs the cooperation of other countries to prevent unsustainable or illegal exploitation.

6.2.3 U.S. Endangered Species Act Criteria

The Endangered Species Act (ESA) of 1973 protects plants and animals that are listed by the federal government as “endangered” or “threatened.” An “endangered” species is one that is “in danger of extinction” throughout all or a significant portion

of its range. A "threatened" species is one that is "likely to become endangered" within the foreseeable future.

6.2.4 *China Red Data Book of Endangered Species Criteria and Wild Animals Under Special State Protection Criteria*

The endangered species criteria of *China Red Data Book of Endangered Animals* were established according *to IUCN Red List of Threatened Species* (1996), and China's actual conditions and grades such as extinct in the wild (EX), extinct (ET), vulnerable (V), rare (R), and indefinite (I) were included. Grades such as endangered, rare, and vulnerable were established in *China Red Data Book of Endangered Plants* according to IUCN Red Data Book.

6.2.4.1 Endangered

These are species at risks of extinction at any time within whole or significant part of their distribution domain. Such plants normally grow sparsely with a narrow distribution domain and a small number of individuals and populations. The survivals of plants under endangered grade become threatened usually due to the loss or destruction of their habitat, overexploitation, and so on.

6.2.4.2 Rare

Although there is no immediate danger of extinction in species under this grade, its distribution is very narrow or very scattered, or it belongs to a single or widowed species which is unusual as well.

6.2.4.3 Vulnerable

Survival of the species under vulnerable grade is threatened by human activities and natural causes. It may be considered as endangered grade in the near future because of deforestation, habitat degradation, and overexploitation.

6.2.5 *List of Wild Animals Under Special State Protection*

Two protection grades were used in the *List of Wild Animals under Special State Protection* (1988). Of those growing only in China, the rare or endangered wild

animals are classified into the state's first-grade protection, and those animals with a small number or risking a danger of extinction belong to the state's second-grade protection. When wild animals under the state's first-grade protection need to be hunted for the purposes of scientific research, introduction, and domestication, approval shall be given by the state's wild animal administrative department. When wild animals under the state's second-grade protection need to be hunted for the purposes of scientific research, introduction, and domestication, approval shall be given by the wild animal administrative departments at the provincial level.

6.2.6 *List of Wild Plants Under Special State Protection*

The List of Wild Plants under Special State Protection was advised by state's wild animal administrative department together with state forestry administration and state agriculture department. A total of 419 kinds of plants were included in this list, among which 67 kinds of plants belong to the first-grade protection and 352 kinds belong to the second-grade protection.

There are four criteria for the listed species: first, these species should be endangered species with a very small number and an extremely narrow distribution; second, these species should be endangered or rare species with a major economic, scientific, or cultural values; third, these species should be wild populations of important crop or closely related species of those having genetic value; fourth, these species should have important economic value, but their resources are sharply declined because of overexploitation.

Besides, the list of national rare and endangered conservative plants was issued by the National Committee of Environment in 1984, and the first edition of the list of national precious wood species was issued by Ministry of Forestry in 1992.

6.2.7 *Regulations on the Protection of Wild Medicinal Species Resources*

Medicinal plants are divided into three grades according to the protection and management regulations of wild medicinal resources. Species under the first grade include those rare and specious wild medicinal species under the endangered condition. Species under the second grade include those wild medicinal species with declining distributions and exhausting resources. Species under the third grade include the major commonly used wild medicinal species whose resources are serious reducing.

6.3 Mechanism of Medicinal Plants Being Endangered

To have knowledge of the processes of extinctive species being endangered and mechanisms of being extinct has important inspiration for the protection of the existing endangered species. In general, endangerment and extinction mechanisms of species include the internal mechanisms and external mechanisms.

6.3.1 Internal Mechanisms

Species extinction is a necessary condition for biological evolution. Without species being extinct, biological diversity will not be increased, and the formation of species will be forced to stop. In the history of biological evolution, once a mass extinction of species occurs, there will be a new batch of high-level taxa emerging. Therefore, the role of extinction in evolution is to provide ecological and geographical space for evolution through the eradication of species and the reduction of biodiversity.

Species in young stages or in decay are easier to be endangered or extinct. Every species not only has its process of emerging and expanding but also has its process of decaying, and no single species can always be dominant. A dominant species will gradually decline until extinction when it becomes flourishing to the extreme. Species in the larval stage of its growth period is also easy to become extinct due to lack of adaptation to the environment.

Endemic taxa are easier to be endangered or extinct. The extinction rate of endemic genera is significantly higher than that of the others. Local endemic taxa, especially those at genus level, are easier to be extinct, and therefore the protection of those endemic taxa should be paid more attention.

Taxa with small morphological variation are easier to be endangered or extinct. Taxa with small morphological variation are easier to become extinct, while those diversified in morphology have higher survival rates. Each external morphology of organism is associated with its specific physiological functions. Taxa which are diverse in morphological traits are likely to have diverse physiological functions and better ecological adaptability, so the protection of taxa with small variation should be paid more attention.

Species with a high degree of specialization and a most perfect type of adaptation are easier to be endangered or extinct. Species with a high degree of specialization and a most perfect type of adaptation are easier to become extinct, while the original, conservative, and less differentiated types often become the center of the new spread. Thus, some of the more advanced genera in a larger taxonomic unit are often separated from low-level genera, and the adaptations of these more advanced genera are gradually reduced because of their high-differentiated structure; therefore, extinction is more likely to happen.

Tropical taxa are easier to be endangered or extinct. Tropical rainforest is often considered to have a relatively stable community structure. The richness of species and the complexity of community structure in tropical rainforest have a stronger resistance to extinction. However, when the interference from the environment goes beyond a certain range, the tropical fauna in the community structure appears to be very fragile, and the species in it are more vulnerable to extinction.

6.3.2 External Mechanisms

6.3.2.1 Biological Mechanism

When two species take advantage of one resource, an increase in the number of individuals in one species will lead to a reduction in the number of the others, known as interspecific competition, and it can be divided into two types. If both two species use the same resource while the interaction between them do not occur, it is called resource-use competition; if one species impedes the survival of the other through certain behaviors, such as poisoning, assault, possession of territory, and allelopathy, it is called interference competition. Competition may lead to a reduction in geographic distribution and density of species.

Bacteria are often an important factor for extinction of species. As same as the predators, the survival of bacteria is often based on the existence of host or the prey. This interdependence often brings about a unique balance; therefore, when the bacteria invade into the area which has not been infected before, species extinction often occurs.

6.3.2.2 Physical Mechanism

The destruction or change of environment that organisms depend on for existence will lead to the extinction or threatened extinction of species, such as slow geological change, climate change and catastrophic events, and so on.

Slow geological change mainly refers to the movement of Earth's tectonic plates, the disappearance of waters, and the slow changes of the continent's ecogeographical conditions as a result. For example, the formation of supercontinent leads to the extinction or threatened extinction of large numbers of marine life living on the continental shelf while creating the necessary conditions for the biological evolution of the terrestrial lives.

Climate change has changed the biodistribution range of organisms in the latitude and longitude, and extinction or threatened extinction of a large number of species may occur; for species which have long-term adaptation to a particular climate, their adaptive adjustment range is limited. For example, plants in the tropics have difficulty in adapting to the cold winter and short photoperiod in high-latitude

regions. Therefore, extinction or threatened extinction of species will occur when climate changes exceed the limits of regulation.

Major catastrophic events on Earth often result in mass extinction of biological taxa. Some of these catastrophic events occur locally, while others globally, such as regression phenomena, volcanic eruptions, orogeny, marine role, collisions of asteroids and the Earth, supernova explosions, etc.

6.3.3 Human Activities

Human activities will lead to huge changes in the planet's ecosystems. A number of large animals become extinct because of human slaughter, and more plant and animal species are endangered or extinct mainly due to changes of the environment by human. In addition, to make extinction or threatened extinction of species directly, humans also change plant community structure in a given region through the introduction of species and further break the biological balance in this region; as a result, some species will become endangered or extinct because of losing their original living conditions. For example, creation of plantations will reduce genetic diversity and result in the potential danger of species extinction or threatened extinction. A new parasitic pathogen or predator may completely destroy the species in the plantations.

6.4 Preservation of Genetic Diversity

Genetic diversity is very essential, whether the population can adapt to the environmental changes and have long-term survival. The populations cannot cope with the environmental changes and evolutionary competition without genetic diversity. The loss of population's genetic diversity due to genetic drift results in the increase of homozygous individuals. The increase in the number of homozygous individuals leads to inbreeding, thereby reducing the fitness and the number of individuals; as a result, the populations are more easily to be endangered. Therefore, protection of the vast majority of the rare species is achieved through the protection of genetic diversity. The protection of genetic diversity of endangered species is the basic objective of protecting endangered species.

Under the premise of lacking genetic information, it is necessary to protect the majority of its population and to protect a habitat large enough in order to protect the genetic diversity of a species and to prevent inbreeding and genetic drift. Therefore, the understanding of genetic diversity can reduce the number of protected species, thereby reducing the costs and land for use. In addition, the genetic diversity information could provide guidance for future research. For example, genetic diversity of a species will not be lost so easily if this species has a high genetic diversity. In general, populations with high diversity can be used as

protection objectives, while the declined populations can be used as management objectives, and their diversity can be restored in the meantime.

Genetic information could be used to determine the area of protected areas, as well as which wild relative populations of cultivated medicinal plant species should be protected. For example, 12 populations of *Acacia raddiana* Savi were analyzed by RAPD, and the results showed that a high level of genetic diversity existed in populations and the genetic variation among populations was about 59.4%, which was a fairly high level [1]. Therefore, from a protection point of view, the loss of any population will lead to the loss of genetic variation if the species have a higher inter-population genetic variation, so each population should be protected separately. In addition, the obvious genetic differences between populations may lead to a recession of hybridization, and therefore, the two populations with largest genetic variation should be protected separately. Another example, analysis of *Leucopogon obtectus* was carried out through RAPD and AFLP, and results showed that there was no significant variation between populations, but high level of genetic diversity appeared in interpopulation. Therefore, the protection and management of genetic diversity should focus on the maintenance of a high level of genetic variation by mixing genotypes and promoting hybridization [2].

6.5 Molecular Identification of Species and Effective Population Decline

6.5.1 Molecular Identification of Species

As different plant and animal taxonomists had different points of view, a species sometimes is combined with others and sometimes is isolated. The differences in understanding species scope result in a change in the number of species and then affect the division of species conservation categories and the scope of species protection. Therefore, the molecular identification of species plays an important role in the protection of endangered animals and plants.

In the list of protected priority species, the inclusion of those species whose protective value is uncertain would result in a waste of manpower and material resources; while if an endangered species is not included in the list because of incorrect understanding, it may result in the extinction of this species. For example, Epipactis is a large genus, which has 36 or 56 species in Europe, and the difference between some of them is very small and controversial. The controversy over classification inevitably leads to the controversy over the protection of them. If these local populations were treated as separate species, there would be of strong significance for protection because of their limited distribution; if they were only regarded as variants, not treated as species, their protected status would be greatly reduced. In order to solve this problem, genetic structure analysis was carried out in complex populations of Epipactis using isoenzyme, RFLP based on cpDNA, *trn*L-c

sequence, and *trn*L-d sequence. Results showed that *E. denensis*, *E. leptochila*, and *E. muelleri* are homozygous, and obvious genetic differences exist between these taxa [3]. So the researchers believed that these taxa could be treated as different species and it has a very important significance to protect them.

6.5.2 Molecular Identification of Effective Population Decline

In order to protect organisms effectively, to clarify the viability of various protective objects, and to identify how many protected objects should be protected, we have to estimate or calculate the sizes of effective populations of various protected objects, which is also called the effective population size.

Populations whose effective population size once experienced a larger decline have changed a lot. The loss of their alleles resulted in the coherence of future generations and then resulted in the increase of homozygous forms or the expression of recessive harmful alleles. The ensuing inbreeding depression might increase the possibility of threatened extinction or extinction of populations. Therefore, it is very important to explore whether one population is declining. The traditional method to determine population decline is through direct individual statistics, but the population decline can also be detected according to genetic information. Detection of population decline mainly takes genetic drift into account. Genetic drift may cause the loss of some alleles and the change of allele frequencies. So, to speculate whether the population declines, we should mainly study the changes in allele or the loss of allele. If alleles in one population showed significant reduction compared to other populations, there may be population decay in this population.

6.6 Determination of Superior Conservation Species

In developing the right strategy for the protection of the endangered species, the first to be solved is to determine the protection unit. In this situation, it is of great significance to look for a suitable protection unit. Commonly used protection units include evolutionarily significant unit (ESU), effective population size (EPS), and operational unit (OU).

ESU is a species unit that should be highly protected. An American scholar put forward and discussed the concept of ESU in 1986, and this concept immediately aroused great attention from scientists. Currently, ESU had been widely accepted as the basic unit of biological protection. The earliest proposed concept of evolutionary significant units was a systematic study-based qualitative standard, and they referred to those groups which were mutually monophyletic taxa on mtDNA haplotypes and those which had significant differences in frequency of alleles on nuclear gene locus. Stressing the mutually monophyletic taxa of mtDNA was not only because of its evolutionary importance but also because theory and simulation

experiments showed that isolated groups were able to attain this state after a certain period of time. However, whether the nuclear DNA and mitochondrial DNA should show the same system development structure is still an issue to be discussed. The above concept can be applied to animals, while its application in plants is limited. Emphasizing phenotypic variation when we determine the plant ESU is an effective method to protect meaningful adaptive feature. However, the phenotypic variation can be reflected more objectively through genetic variation.

It is difficult to determine ESU under the situation of a continuous distribution of individuals in space. So, it is a key issue to determine which unit is in need of protection for species within the populations with a continuous distribution in space. A better approach is to assess the genetic continuity of species in geographical distribution, which can be achieved by defining the location of operational units (such as protected areas). It was suggested that an analysis of all genetic variation by spatial autocorrelation could establish operational units of species conservation and management, especially to deal with continuous populations, and this technique was superior. For example, some scholars did some researches on the genetic variations within or between *Eugenia dysenterica* populations with different spatial extent in Cerrado, and these researches have provided a basis for determining the operational unit of *E. dysenterica* [4].

To identify the order of conservation priorities of different populations, we should first consider how to protect the evolutionary potential of species, namely, genetic diversity in the most effective way. This is because the genetic diversity is the basis for species to adapt to the environmental change. Species with a low genetic diversity may be adaptive to one certain type of environment, and their distribution range may be wide. But once the external environment changes, these species may be extinct first. The principles of the establishment of populations' protective priority are shown as follows.

6.6.1 Determination of Superior Conservation Species Based on Gene Mutation

Such an approach is based on the genetic diversity of populations, particularly the allele diversity, to identify order of conservation priorities, that is, the higher level of allele diversity a species has, the more value of prior protection the species has. However, not all populations with a high genetic diversity are in need of protection. In general, there are three reasons for a population to have a relatively high genetic diversity: ① Genetic variations caused by new alleles generate their producing centers because of mutation and recombinant; ② genetic variations from different sources form their cumulative centers through the accumulation of genetic variation by gene flow; ③ genetic variations form their cross-mixing centers because of the cross-mixing of populations in different types. Among these centers, the genetic variation centers have greater value for protection, and the other two categories have relatively weaker protective values. Species populations with marginal distribution

also need adequate concerns. Although the genetic diversities of these populations are relatively low, there are often some characteristic alleles appeared because of the role of harsh environment, and these alleles may be the producing center of genuine medicine's genotype-specific.

6.6.2 Determination of Superior Conservation Species Based on Genetic Distance

Methods to determine the protection unit which are based on genetic distance can also be used to determine the priority of protected populations. Populations given priority of protection are often those with high genetic diversities, that is, the greater genetic differences one population have from the other populations, the higher level of protective priority this population has. Information of mitochondrial DNA and nuclear DNA can be used to judge the priority of protected populations by constructing phylogenetic trees.

6.6.3 Determination of Superior Conservation Species Based on Genetic Contribution Rate of Population

The establishment of a comprehensive index, genetic contribution according to gene richness and uniqueness of various groups, can be used to set priorities in protecting populations. Genetic contribution rate includes the extent of genetic variation and the differential degree (based on measuring differences with other populations). The basic idea is to calculate the contribution rates of various groups to the entire species based on the genetic composition and then to sort the contribution rates, and then it can be determined which populations need priority conservation. In general, populations with high genetic diversity and unique gene show the highest genetic contribution rates to the species; therefore, we can effectively protect the genetic diversity of this species just by protecting these populations.

6.7 Protection Method of Rare and Endangered Medicinal Plants and Animals

The effective methods to protect the endangered animals and plants mainly cover the following types: *in situ* conservation, *ex situ* conservation, and *in vitro* conservation.

6.7.1 In Situ Conservation

It is a method to divide a certain area of land or water including the protective objects into protection and management, such as establishing nature reserves.

6.7.2 Ex Situ Conservation

This method is carried out by moving some species into the zoo, botanical gardens, aquariums, and endangered animal breeding centers for their protection and management. The survival and propagation of these species are often seriously threatened because their living conditions no longer exist, they are rare in number, or it is difficult for them to find a mate.

6.7.3 In Vitro Conservation

This is a method to protect materials which carry all the genetic information of their own species. These materials may be a part of the organs, tissues, cells, or protoplasts of medicinal plants and animals, and so on. We can preserve the germplasm of endangered medicinal plants and animals for a long term by protecting these materials. The establishment of seed bank of endangered medicinal plants is an important way to protect one species *in vitro*. Because of China's vast territory and its richness in herbal medicine resources, genuine herbs with various characteristics are formed. The current application of wild seeds in Chinese herbal medicine GAP cultivation results in a mixture of germplasm of cultivated Chinese herbal medicines. Therefore, the active collection of different germplasm resources and the establishment of germplasm resource database can provide a wealth of genetic resources for future breeding.

6.8 Case Study

Astragalus membranaceus is one of the commonly used Chinese crude drugs with slight warm nature and sweet flavor and has the functions of invigorating qi for strengthening superficies, diuresis, pus draining and toxin expelling, cicatrizing ulcer, and myogenesis. Chinese pharmacopeia (2005 edition) stipulates that *Astragalus* should be the dried root of *A. membranaceus* (Fisch.) Bge .var. mongholicus (Bge.)Hsiao. or *A. membranaceus* (Fisch.)Bge. Because chicken claw root with great variation may be formed during cultivation of *A. membranaceus* (Fisch.) Bge., in recent years, the main plantation and source of commercial *A. membranaceus*

has been from *A. membranaceus* (Fisch.) Bge .var. mongholicus (Bge.) Hsiao. Currently, the quality of the varieties of *A. membranaceus* (Fisch.) Bge .var. mongholicus (Bge.) Hsiao. is uneven with serious states of mixed degeneration, and stable cultivated variety has not been established yet. So, we have to enhance the work of variety selection. It is also necessary to study the genetic relation of the main population of *A. membranaceus* (Fisch.) Bge .var. mongholicus (Bge.) Hsiao. in distinct places of production and thus to provide some arguments on the evaluation and innovation of germplasm resources of *A. membranaceus* (Fisch.) Bge .var. mongholicus (Bge.) Hsiao.

6.8.1 Materials and Methods

A. membranaceus was collected from different places of production of drug material in August 2007 and was rapidly dried with silica gel. All materials were characterized as *A. membranaceus* var. mongholicus by Xue-feng Feng associate investigator, Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences (Table 6.1).

The PCR device is Eppendorf Mastercycler personal; centrifuge is Eppendorf centrifuge 5415D;rTaq DNA polymerase, DNA marker 2000, and dNTPs are purchased from Takara Inc.; primer is synthesized by Shanghai Sangon Biological Engineering Technology & Services Co., Ltd.

Leaf blades collected from 20 single plants for each population were mixed, and the total DNA was extracted using modified CTAB method. The extracted DNA was stored in a freezer at −20 °C.

According to the primer published by Li et al., 8 upstream primers (ME1–ME8) and 10 downstream primers (EM1–EM10) were designed (Table 6.2).

20 μL system: DNA50 ng, 10×PCR buffer 2 μL, 25 mmoL·L-1 $MgCl_2$ 2 μL, dNTP 2 μL, 10 μmoL·L^{-1} upstream primer 2 μL, 10 μmoL·L^{-1} downstream primer 2 μL, *rTaq* DNA polymerase 1 U, and ddH_2O 20 μL were used to reach to the total volume.

Thermal cycling program: initial denaturation at 94 °C for 5 min; 5 cycles with the following conditions: denaturation at 94 °C for 1 min, annealing at 35 °C for 1 min, and primer extension at 72 °C for 1 min; 35 cycles with the following conditions: denaturation at 94 °C for 1 min, annealing at 50 °C for 1 min, primer extension at 72 °C for 1 min and primer extension at 72 °C for 7 min stored at 4 °C. Then electrophoresis was performed with the amplified product on agarose gel (containing EB) in 1×TBE buffer, and the SYNGENE model imaging system was used for observation and taking photos.

According to the electrophoretogram of the amplified product, only the distinct and discriminating bands with polymorphism were subject to statistical process. In region of larger than 800 bp, the amplified bands were less and not easy to be discriminated, so they were not recoded. In region of 100–800 bp, the amplified bands were easy to be identified and recorded, so most polymorphism were found in this region. The amplified bands less than 100 bp were too weak to be identified, so they

Table 6.1 Natural overview of the population

Population	Longitude	Latitude	Altitude/m	Annual mean amount of precipitation/mm	Annual mean temperature/°C	Annual sunshine hours/h	Relative humidity/%
1	35°01′044″	104°37′170″	1,767	515	8.5	1,979	68
2	35°04′359″	104°26′542″	1,894	515	8.5	1,979	68
3	34°52′961″	104°21′004″	2,065	538	8.3	1,891	68
4	41°11′548″	110°36′323″	1,839	290	5.7	2,497	50
5	41°06′733″	110°18′738″	1,538	290	5.7	2,497	50
6	40°54′693″	110°45′156″	1,667	349	5.1	2,441	52
7	40°52′598″	109°45′422″	1,230	249	6.6	2,557	48
8	39°26′834″	113°42′930″	2,018	547	7.2	2,273	58
9	39°27′393″	113°44′917″	1,772	547	7.2	2,273	58
10	39°34′673″	113°43′926″	1,807	547	7.2	2,273	58
11	39°28′704″	113°24′981″	1,759	515	7.6	2,286	58
12	39°25′708″	113°25′073″	1,758	515	7.6	2,286	58

Table 6.2 Primer sequences for SRAP

No.	Upstream primers ME	Downstream primers EM
1	TGAGTCCAAACCGGATA	GACTGCGTACGAATTCAA
2	TGAGTCCAAACCGGAGC	GACTGCGTACGAATTCTG
3	TGAGTCCAAACCGGAAT	GACTGCGTACGAATTCGA
4	TGAGTCCAAACCGGACC	GACTGCGTACGAATTCCA
5	TGAGTCCAAACCGGAAG	GACTGCGTACGAATTAAT
6	TGAGTCCAAACCGGTTG	GACTGCGTACGAATTAAC
7	TGAGTCCAAACCGGTGT	GACTGCGTACGAATTTGC
8	TGAGTCCAAACCGGTGC	GACTGCGTACGAATTTGA
9		GACTGCGTACGAATTGAC
10		GACTGCGTACGAATTGCA

were not recorded. A data table of 0–1 was prepared by recoding the same migration location with band as 1 and the same migration location without band as 0. SPSS software package was used to perform UPGMA cluster analysis of the result.

6.8.2 Results

The 80 pairs of primers were screened with No.1 and No.9 samples, and 16 primer pairs were chosen for the strong signals, good reproducibility, and excellent marks of the bands of amplified products. Amplification was performed with the 16 primer pairs for the *A. membranaceus* var. mongholicus produced in 12 places of production. Altogether, 141 bands were harvested, and, among them, 69 bands of polymorphism were identified with the ratio of polymorphism of 48.9% and 4.31 bands of polymorphism for each primer pair. The range of the ratio of polymorphism was 14.2–80%.

According to the genetic distance among these test articles and regarding 22 as the critical value, these 12 samples could be divided into 2 categories with most of the populations from Gansu and Inner Mongolia classified as category 1 and most of the populations from Hunyuan and Ying Counties of Shanxi classified as category 2. Populations with geographically further locations from each other showed relative longer distance in the cluster map, such as populations in Gansu from populations in Shanxi. Populations of the same location were generally clustered at the same position.

6.8.3 Discussion

In this study, SRAP molecular labeling technique was used to analyze the genetic relationship among the different populations of *A. membranaceus* var. mongholicus, and the results showed the advantages and availability of the method. The technique

offered more amplified bands which were easily identified and had the features of excellent repeatability, simplicity, and primer versatility. Forward and reverse primers could match each other, which might reduce the cost. So, compared with ISSR, RAPD, and other molecular labeling methods, SRAP labeling technique is more suitable for the analysis of the genetic relationship of *A. membranaceus* var. mongholicus.

For amplified product with SRAP, polyacrylamide gel electrophoresis (PAGE) was commonly used, and, sometimes, agarose gel of 1.8 and 2.0% (containing EB) might also be used for electrophoresis in 1×TAE buffer. In this study, electrophoresis with agarose gel of 1.8% (containing EB) was performed in 1×TAE buffer, and the electrophoresis time was 85 min. The buffer capacity of TBE was greater than TAE, so agarose gel in TBE might offer higher resolution and suitable for electrophoresis of longer duration. Compared with PAGE, the separated bands might be reduced, but it was simpler, which might lead to reduction of cost and time.

In consideration of the limitation of sample source, the study just investigated the main populations of *A. membranaceus* var. mongholicus, and it was not sufficient to reflect the population relationship of *A. membranaceus* var. mongholicus; however, the study might provide some initial results. According to the fingerprint of SRAP, the main bands were basically consistent for *A. membranaceus* var. mongholicus from different population, which indicated that they had greatly similar genetic background. However, there were some differences in secondary bands among different populations of *A. membranaceus* var. mongholicus, which demonstrated the abundant genetic diversity. The polymorphism ratio of the selected *A. membranaceus* in this study was 48.89%, which indicated greater genetic diversity and genetic differentiation among the populations. The genetic distances between the cultivated populations and wild populations in the same region are very close, for example, the 2 populations in Longxi (genetic distance = 0.043) and the 2 populations in Hunyuan (genetic distance = 0.083) are both the populations with closer genetic distance in this study. According to the cluster results acquired by tests in this study, the genetic relationship of *A. membranaceus* var. mongholicus is generally consistent with differences in regions, but their kinships could not be fully determined according to their geographic locations. The genetic distance between population in Wulateqian Banner of Inner Mongolia and population in Hunyuan of Shanxi is 0.119, which is not consistent with the geographic distance between them according to the cluster result. After investigating the climatic factors such as annual mean temperature, annual sunshine hours, annual mean rainfall amount, and relative humidity, similar ecotopes are found between the two places. The similarity in ecological microenvironment may be the main cause for far geographic distance and close genetic distance. According to the theory of evolution, higher level of genetic variety means larger range of population distribution. In this study, the result of SRAP analysis on the 12 populations distributed in 3 provinces indicated distinct genetic differentiation among these populations. The main cause for the genetic differentiation is the outcome of long-term adaptation of the plants to the distinctly different environment of distribution regions in sunshine hours, radiation intensity, temperature, rainfall, etc.

The result of cluster analysis shows two main categories, Shanxi and Gansu, and the populations in Inner Mongolia belong to transitional form. The result accurately reflects the history of drug application for *A. membranaceus* var. mongholicus. According to literature of Chinese materia medica, northwest territories of China were the main places of production before Tang Dynasty, especially that produced in Gansu was considered as a famous drug; after Song Dynasty, those produced in Shanxi were considered as good drug, and until Qing Dynasty, *A. membranaceus* produced in Inner Mongolia was also be called "dao-di herbs." In modern time, Hunyuan and Ying counties in Shanxi province and Longxi County in Gansu province were mostly considered as the excellent and primary places of production of *A. membranaceus*. Modern studies show that *A. membranaceus* var. mongholicus produced in Hunyuan County of Shanxi province and Longxi County of Gansu province has higher content of methyl glycoside with stable quality than that produced in other places. Especially, *A. membranaceus* var. mongholicus produced in Shanxi province has the highest content of calycosin glucoside and that produced in Gansu has the highest content of neochanin. The result of the study well supports the herbal textual research and modern chemical constituent studies and indicates that the traditional famous-region drug is consistent with real excellent drug.

References

1. Shrestha, M. K., Golan-Goldhirsh, A., & Ward, D. (2002). Population genetic structure and the conservation of isolated populations of *Acacia raddiana* in the Negev Desert. *Biological Conservation, 108*, 119–127.
2. Zawko, G., Krauss, S. L., Dixon, K. W., et al. (2001). Conservation genetics of the rare and endangered *Leucopogon obtectus* (Ericaceae). *Molecular Ecology, 10*(10), 2389–2396.
3. Squirrell, J., Hollingsworth, P. M., Bateman, R. M., et al. (2002). Taxonomic complexity and breeding system transitions: Conservation genetics of the *Epipactis leptochila* complex (Orchidaceae). *Molecular Ecology, 11*(10), 1957–1964.
4. Zucchi, M. I., Brondani, R. P. V., Pinheiro, J. B., et al. (2003). Genetic structure and gene flow in *Eugenia dysenterica* DC in the Brazilian Cerrado utilizing SSR markers. *Genetics and Molecular Biology, 26*(4), 449–457.

Further Readings

Fu, Y.-M., Jiang, W.-M., & Fu, C.-X. (2011). Identification of species within *Tetrastigma* (Miq.) Planch.(Vitaceae) based on DNA barcoding techniques. *Journal of Systematics and Evolution, 49*(3), 237–245.

Karthikeyan, A., Madhanraj, A., Pandian, S. K., & Ramesh, M. (2011). Genetic variation among highly endangered *Bacopa monnieri* (L.) Pennell from Southern India as detected using RAPD analysis. *Genetic Resources and Crop Evolution, 58*, 769–782.

Yu, H. H., Yang, Z. L., Sun, B., & Liu, R. N. (2011). Genetic diversity and relationship of endangered plant *Magnolia officinalis* (Magnoliaceae) assessed with ISSR polymorphisms. *Biochemical Systematic and Ecology, 39*, 71–78.

Chapter 7
Gene Modification of Pharmic Plant Germplasm Resources

Rong-min Yu

Abstract Germplasm resources refer to the basic materials for selecting and cultivating new species, which include propagating materials of cultivated species and wild species from different plants and genetic materials of various plants artificially created based on the propagating materials above. Pharmic plant germplasm resources are mainly composed of all of the pharmic genetic resources from pharmic plant. According to origin of germplasm resources, pharmic plant can be divided into four categories: local germplasm resources, foreign germplasm resources, wildlife resources, and artificially created germplasm resources. According to genetic relationship, pharmic plant is divided into primary gene pool (Gp-1), secondary gene pool (Gp-2), and tertiary gene pool (Gp-3), while based on the scope of the study area, it can be divided into regional and single type of germplasm resources.

Study on germplasm resources mainly includes the species of wild pharmic plant (including variants), the type of cultivation of medicinal plants, the formation of the species or strains, etc. Besides, photograph and record about their morphological characteristics, biology, medicinal organs, medicinal value, other economic value, etc., should be taken down. Germplasm resources collection refers to the intentional collection of germplasm resources. The purpose of preservation of germplasm resources is for sustainable use of genetic resources. The stored germplasm resources should have the original viability, genetic variation degrees, and the amount of samples maintained.

The evaluation of germplasm includes species identification of germplasm, genetic diversity analysis, and the determination of genetic relationship.

Rong-min Yu (✉)
Biotechnological Institute of Chinese Materia Medica, Jinan University,
Guangzhou 510632, China
e-mail: tyrm@jnu.edu.cn

Lu-qi Huang (ed.), *Molecular Pharmacognosy*,
DOI 10.1007/978-94-007-4945-0_7,

Molecular marker-assisted selection is a modern breeding technology which uses the DNA marker linking closely with target property to select the target property indirectly. The research includes the genetic map of pharmic plant germplasm resources, bulked segregant analysis, and mark for important agronomic traits genes.

7.1 The Concept and Significance of Pharmic Plant Germplasm Resources

7.1.1 The Concept of Pharmic Plant Germplasm Resources

Germplasms are the genetic materials of organisms, which can be inherited to the individual or colony progeny by propagation. In terms of plants, it not only includes seed but also root, stem, embryo, cell, protoplast, etc., and even DNA fragment. The word "germplasm," coined by a German geneticist named Weismann in 1892, originates from "the germplasm theory." It claimed that germplasms were the genetic materials descended from parental generation to offspring, producing habits which were not influenced by the environment.

Germplasm resources, also called "genetic resources," are the basic materials of selecting and cultivating new species. They include propagating materials of cultivated species and wild species from different plants and various plant genetic materials artificially created by the propagating materials above. In a narrow sense, germplasm resources usually refer to a specific species, in which all available genetic vehicles are involved, including cultivated species, wild species, relative species, and special inheritable materials (wildly or artificially mutated materials like polyploid and haploid).

Pharmic plant germplasm resources are composed of the pharmic genetic resources from alga, fungi, lichen, bryophytes, pteridophyte, gymnosperm, angiosperm, etc. Narrowly speaking, pharmic plant germplasm resources refer to a specific species, like "the germplasm resources of *Panax*", "the germplasm resources of *Caladium*," and "the germplasm resources of *Salvia miltiorrhiza*."

7.1.2 The Purpose and Significance of Research on Pharmic Plant Germplasm Resources

China is one of the countries possessing affluent biotic resources in the world, with the longest history of medicinal plant application. The Chinese herbal medicine mainly comes from pharmic plants and covers about 11,146 species (including the units below species). Recently, with the improvement of life quality and the rise of "coming back to the nature," the exploration of pharmic plant resources has become

popular all over the world. Besides being used as herbal materials or supplied as pharmaceutical material, pharmic plants are wildly applied to many novel areas, like health food, beverage, flavoring agents, perfume, cosmetics, plant pesticide, and drugs for birds and domestic animals. Under such a promising situation, pharmacognosy research has to meet a stricter requirement—pharmic plant resources should be protected and used effectively. However, since it is the very first step for making germplasm resources research on pharmic plants, there are many problems during production. Most of the cultivated herbal drugs are introduced blindly and acclimated in a short time without screening evaluation. As a consequence, cultivated plants germplasm resources are mixed up. Even if the species got from farmers go through primary screening, they are usually appraised only according to the appearance and production and in need of further screening.

7.1.2.1 Pharmic Plant Germplasm Resources Are the Basic Stones for Introducing Cultivation and Sustainable Use of Resources

Pharmic plant germplasm resources serve as an essential factor for the quality and production of Chinese herbal medicine and are also the material bases for keeping drug quality stable. Nowadays, the resources shortage of Chinese herbal medicine, we should deal with cultivation of Chinese herbal medicine and protection of core production areas in order to realize sustainable use of pharmic plant. Economic index should be considered before introduction, which means that the most useful genetic resource would be screened out from tremendous genetic resources, so as to obtain more products while consuming the same amount of materials. Take *Taxus* as an example, whose research on germplasm resources has found that this genus contains 11 kinds (4 kinds in China with 1 mutated). Below genus, there are many geographic races and cultivated races—only for *T. baecata* L., it has no less than 40 geographic races, whose active component *Taxinol* is barely 0.01%. In China, the main cultivation type of *T. baecata* L. is the hybridized *Taxus media* from North America, which is selected according to the investigation of germplasm resources, containing 0.014% *Taxinol*. It has a lot of merits, such as strong germination, fast production, easy sticking and propagation, and strong adaptability to the environment.

7.1.2.2 Pharmic Plant Germplasm Resources Are the Material Bases for Cultivation

Modern cultivation is characterized by wide collection, storage, and deep study of germplasm resources. Germplasm resources with different genetic properties are the material bases for cultivation. The more plentiful the germplasm resources are, the more the cultivation could be anticipated. Therefore, it will be more possible for some new and potential races to be obtained. Pharmic plant germplasm resources

play basic roles in cultivation. Whether cultivation has radical achievement depends on the mutation, discovery, and use of these key characters.

In order to avoid genetic base indigence of pharmic species, it is necessary to make use of more germplasm resources. Due to the promotion of good pharmic species and the selection of some germplasm resources with same genetic background during cultivation by some institutions or scientists, it may lead to the following bad results. Firstly, the quality of herbal drugs could decline. Secondly, pharmic plants may have poor resistance and low productivity. Thirdly, the favorable restructuring could be limited by the narrow genetic base, which may slow down the progress of breeding, even leading to stagnation. Therefore, to avoid genetic base indigence of pharmic species, it is necessary to collect more germplasm resources and establish germplasm resources nursery of young plant or gene pool.

Pharmic plant germplasm resource is one kind of important natural resources, which is formed by natural revolution and artificial creation in long history. They have greatly influence on the production of excellent herbal medicine. To some extent, the formation of many effective authentic ingredients is attributed to the effect of excellent regional varieties. Cultivation of pharmic plants pertains to the utilization, processing, and transformation of various germplasm resources by different methods according to people's wills. Thus, pharmic plant germplasm resources are the main origins for people to select novel varieties through plant genetic mutation. In addition, they also serve as the material basis for pharmic plant germplasm cultivation. Whether the cultivation of pharmic plants is successful or not mainly depends on the quantity and quality of the species resources, which we have explored and mastered. And their research extension and depth play key roles as well.

The cultivation of pharmic plants in our country has already established a primary basis. Many Chinese herbal medicines, like *Ginseng*, chrysanthemum, mint, saffron, medlar, Rehmannia, Fritillaria, yams, Polygonatum, Campanulaceae, woad, hemp, *Ginkgo biloba*, Coix, Dendrobium, Motherwort, honeysuckle, Eucommia, Codonopsis, saffron, sage root, Radix, and Angelica, have become good local "varieties." However, as a result that a great number of "varieties" are in lack of the involvement of breeders, their genetic purities are still unknown and they cannot be counted as true species. At present, we are facing an ambitious task for breeding a growing number of medicines, so it is necessary to speed up the introduction and application of new technique and methods for breeding.

7.1.2.3 Pharmic Plant Germplasm Resources Are the Bases for Ensuring the Quality of Medicine

Germplasm is the source of production of Chinese herbal medicine, whose quality exerts critical effect on the quantity and quality of medicine. With the development of production and scientific technology, human beings will continue exploring more and more novel drug varieties from wild medicinal plant resources, so as to fulfill the growing needs of production and living. The more germplasm resources we have controlled, the more advantages of innovative varieties we can use to occupy the international market. Therefore, many countries spared no expense to collect

large numbers of germplasm resources through multiple channels. Recently, among the 800 kinds of commonly used Chinese herbal medicine, most of them come from wild resources, and only about 300 kinds have been introduced and cultivated, with other 200 kinds being planted in large area. For example, coptis root, cultivated extensively in Sichuan and Chongqing, was found during resources survey and later promoted in the 1960s of the twentieth century. Moreover, according to the research on the origin of long-term-cultivated Chinese herbal medicines like Feng Dan, chrysanthemum, and Chinese yam, it has been uncovered that Feng Dan originated from authentic medicine *Paeonia ostii*, which formed after long-term cultivation in the Tongling area. Thus, its wild species or related species should be analyzed deeply.

At the same time, the main cultivating varieties at present are from the domestication of wild plants in different historical periods. Besides, the major cultivating varieties of Chinese herbal medicine are interwoven seriously without new species cultivation and purification.

7.2 The Types of Pharmic Plant Germplasm Resources

Since there is a wide range of germplasm resources, for convenience of research and use, the methods of classifications by different disciplines according to their use are not exactly the same. Usually, the classification methods are based on kinships, ecological types, sources, and value, etc. In this chapter, on the basis of characters of pharmic plant germplasm resources, there are two following ways to classify them.

7.2.1 Category According to Sources

According to sources, germplasm resources are commonly divided into local germplasm resources, foreign germplasm resources, wild germplasm resources, and artificial germplasm resources.

7.2.1.1 Local Germplasm Resources

Local germplasm resources refer to old regional varieties and improved varieties being currently promoted. They own several characters. Firstly, local germplasm resources are highly adaptive to local natural conditions and ecological characteristics. Secondly, they reflect the needs of production and living of local people. Thirdly, they have special excellent properties. At last, these old local varieties are fertilizer intolerant with low yield and poor disease resistance.

Take ginseng as an example. It can be classified into Malaysia teeth, two horse teeth, round arm, round reed, long neck, and other types or local varieties, according

to the morphology of its roots and rhizomes. In addition, based on commodity-type characteristics, it can be divided into general ginseng, Biantiao ginseng, and Shizhu ginseng. Moreover, to the morphology above ground, it includes yellow fruit ginseng, red fruit ginseng, orange fruit ginseng, purple stem ginseng, green stem ginseng, tight panicle ginseng, loose panicle ginseng, etc.

7.2.1.2 Foreign Germplasm Resources

Foreign germplasm resources stand for the pharmic species and plant types introduced from other countries or areas, whose characters are as follows. Firstly, it can reflect the natural conditions and production properties of original areas. Secondly, most of the foreign germplasm resources are poorly adaptive to the local conditions, such as safflower, gardenia, pomegranate, walnuts, and garlic, which were introduced in ancient time, and modern introduced American ginseng, Japan chuanxiong, and Japanese angelica, etc.

7.2.1.3 Wild Germplasm Resources

Wild germplasm resources refer to the related species of various pharmic types and valuable species of wild plants. Since they formed through long-term natural selection under specific condition, they usually have some essential traits which generally cultivated Chinese medicine are lack of, particularly like stiff resistance, unique quality, and male sterility. At present, most of pharmic plant germplasm resources still belong to wild germplasm type, such as Bidentate, to yellow, red sage root.

7.2.1.4 Artificial Germplasm Resources

Artificial germplasm resources are the hybrid, mutant, and intermediate materials produced by different ways (hybrid, physical, chemical mutagenesis, etc.). For example, when buds of Mongolian astragalus seedlings are applied by the improved agar smear, tetraploid introduction plants will be obtained, which serve as a foundation for further cultivating new varieties of Astragalus polyploidy. American ginseng hybrid generation is produced by a similar way.

7.2.2 Category According to Relationship

According to relationship, gene pool is divided into three types:

(i) Primary gene pool (Gp-1): It includes the materials within the same species, equivalent to the traditional biological species concept. They can mutually hybridize with fertility and transfer genes between relatively simple materials.

(ii) Secondary gene pool (Gp-2): It contains closely related wild species and related genera of plants, including some species and related species which can be hybridized with.
(iii) Tertiary gene pool (Gp-3): It is composed of the materials, which can cause severe cross-infertility and cross-sterility when hybridized with distant species.

7.2.3 Category According to Research Range

Germplasm resources are divided into local germplasm resources and single germplasm resources, based on research range:

(i) Local germplasm resources refer to all pharmic germplasm materials in one district or area.
(ii) Single germplasm resources refer to all pharmic germplasm materials of some specific pharmic species.

7.3 Characteristics of Pharmic Plant Germplasm Resources

China has vast territory, with both land and sea, spanning latitude above 49° from north to south, longitude 60° across east to west. Its topography and climate are complicated, across the frigid, temperate, and tropical zones from south to north, with diverse ecological environment. Besides, China is considered as one of the countries with the most abundant germplasm resources in the world. The core indicator for germplasm resources evaluation is clinical effect, which consequently determines the following characteristics of pharmic plant germplasm resources, compared with farm crops.

7.3.1 Unique Division and Selection Criteria

Crops focus on yield, nutrition, and flavors. However, pharmic plant germplasm resources are in pursuit of drug quality. The most effective components in pharmic plants are secondary metabolic products and abnormal secondary metabolic products, such as alkaloids, saponins, flavonoids, terpenoids, coumarins, and tannins. Their contents in plant are very slight, only according to several percentage or several parts per ten thousand, that can be hardly observed directly, and we have to use physical and chemical analyses to obtain them. Meanwhile, these elements are susceptible to the environment in plants. In the consequence, it is difficult to determine the genetic and environmental effects, respectively. Molecular biology methods to directly analyze DNA of genetic material serve as an important way to study pharmic plant germplasm resources.

7.3.2 Obvious Regional Character of Pharmic Plant Germplasm Resources

In the traditional production of Chinese herbal medicine, genuineness of medicine was stressed on, which shows that pharmic plant germplasm resources put more emphasis on region than any other field crops. In terms of an authentic pharmic plant or farm type, similar to ecological factors, genetic factors also play a significant role as well. Genuine medicines are often produced in a particular area by natural and artificial selection after a long time, like Ludang in Shanxi Changzhi, ShiZhu ginseng in Jiangsu Zhenjiang, and Gongju in Auhui Shexian. These species can best adapt to the conditions of earth and climate, as well as cropping systems. The local conditions can be best suited to the expression of objective and property. More and more evidence have demonstrated that there are some local excellent varieties in traditional authentic ingredients which have been selected for many years but not being purified. They serve as the valuable wealth for us to study pharmic plant germplasm resources.

7.4 Genetic Modification

Genetic modification means the DNA sequence modification by biochemical methods, introduction of target genes into host cells, or deletion of a specific gene from genome to change or enhance original genotype of host cells. Genetic modification has been applied widely to all areas of human life. For instance, in medicine, it can suppress viral replication within some viral host cells so as to achieve the goals of treatment. In agriculture, by method of genetic modification, the characteristics of crops and livestock production have been changed successfully in order to improve varieties and disseminate them. Although genetic modification has been considered as an important issue in scientific area, its safety remains widely questioned. Additionally, its relationships with human health, environment, agricultural development, economy, and politics should be studied further. Only addressing these issues, genetic modification can be really integrated into human life.

7.5 Collection of Pharmic Plant Germplasm Resources

7.5.1 Investigation of Pharmic Plant Germplasm Resources

The pharmic plant germplasm resources are extremely plentiful, whose investigation and collection are the most basic work and also the most important one.

7.5.1.1 Preparations Before Investigation

To investigate the germplasm resources of medicinal plants, especially in large-scale investigation activities, careful plans should be drafted beforehand, and adequate preparations should be made, which are the preconditions to obtain satisfactory outcomes. Before investigation, the following works should be completed:

① Making Investigation Plan

Before making plans, it is necessary to read relevant references to determine location, route, time, expedition, equipment, materials, and so on.

A. Determination of the location and route of investigation
   Investigation of germplasm resources of medicinal plants is mainly on local germplasm resources and single species. Investigation of local germplasm resources should be focused on three regions:

   (a) Primary origin and secondary origin centers of medicinal plants and regions with the most abundant plant diversity.
   (b) Regions where pharmic plant germplasm resources have suffered the most losses and menaces. The investigation location and route of single species of germplasm resources should be designed according to the main producing area, genuine producing area, and area with rich diversity.

   After the investigation location and route have been determined, the staff participating in the study should refer to the data of the area investigated and understand the following information, including landform, vegetation, terrene, hydrology, climate, and also the background data including social structure, ethnic distribution, living habits, economic status, social change situation, arable land, crop species, cultivation methods, and the occurrence of major pests and diseases, to finish the preparation work of investigation.

B. Duration arrangement
   Due to the specificity of Chinese herbal medicine, the investigation duration should be determined according to the growth characteristics and medicinal harvesting period. Enough time is needed to finish the investigation, but it should not be too long for the consideration of cost and working process. The specific length of duration will be determined in accordance to the investigation scope and purpose.
C. Preparation of commodities and equipment
   Commodities and equipment needed in the field investigation for germplasm resources include transportation vehicles, sample collection appliances, record forms of survey, instruments, tools and containers used in investigation, articles for daily use (including medicine), etc.

② Expedition Team Foundation

The number of expedition should be determined according to the investigation tasks, objects, and activity area. The members should receive the skill training on the safety, survey methods, and so on.

7.5.1.2 Investigation Contents of Pharmic Plant Germplasm

This is the most critical step for the investigation of pharmic plant germplasm resources. Investigation contents for regional germplasm resources mainly involve of species of wild medicinal plants (including varieties), species of cultivated medicinal plants, formed species or strains, and so on. We should take photos and records for their morphology, biological characteristics, medicinal organs, medical value, and other economic values.

Character (Video) Records

The characteristics of various types of medical plants have great differences. During the investigation, the record contents and the numbers of descriptor are different, but the following information should be inspected and recorded for all kinds of pharmic plant germplasm resources (Table 7.1).

Accession is the basic unit in the registration of germplasm resources. The so-called one accession of germplasm should have similar genetic unit (such as population/provenance, family, somaclone, etc.) and cannot be repeated with the germplasm heredity that has been registered. For example, the germplasm in the same family collected in different time should not be counted as two accessions, because they have the same genetic composition. Thus, they are two replicated accessions (i.e., genetic homogeneity is one accession, but genetic heterogeneities are multiple accessions). Sometimes, the genetic homogeneity has been recorded as two accessions in actual operations. This is a gradual process, which needs genetic detection at different levels (population relationship, phenotype, physiology, ecology, molecule, etc.), building up the “quasi-core,” “core” germplasm gradually.

A. Information of germplasm habitats

(a) Geographic location
It includes the natural and administrative regions of investigation spot (specify the names of province, county, country, village, and place), latitude and longitude, altitude, direction and distance to the nearest town, topography, slope direction, slope degree, etc.

(b) Conditions of water and soil
These involve the soil type, soil depth, hydrological condition, etc.

Table 7.1 Catalog table of pharmic plant germplasm resources-table of accession

1. Passport information							
1.	Number of germplasm	Plant classifications and codes in China (GB/T14467-93)+XXXXX (5 natural order number)					
2.	Chinese name of germplasm						
3.	Foreign name of germplasm						
4.	Germplasm types	Population		Individual		Cultivated variety	
		Population relationship		Somaclone		Strain	
		Source				Genetic materials	
		Family(Full-sib)		Famous tree		Others	
		Family(Half-sib)		Local variety			
		Others (specified):					
5.	Germplasm source	County (district/city): Province (autonomous region/municipality): Country:					
6.	Storage institute						
7.	Name of family (Chinese name)						
8.	Name of genus (Chinese name)						
9.	Name of species (Chinese name)						
10.	Name of species (Latin name)						

(continued)

Table 7.1 (continued)

2. Extended information													
11.	Internal number in the storage institute												
12.	Germplasm characteristics	Picture 1, 2, 3, 4 (optional)											
13.	Germplasm application												
14.	Germplasm genealogy												
15.	Bred time												
16.	Germplasm habitats	Climate zone:		Longitude: ° ′		″Latitude: ° ′		″Altitude: m		Soil type:			
		Average annual temperature:				Average annual rainfall:							
		Extreme factors:											
17.	Stored material	Plantlet		Root		Bud		Pollen		Cell		Gene	
		Stem		Flagellum		Seed		Tissue		DNA			
		Others(specified):											
18.	Storage method	In situ storage				Ex situ storage				Equipment conservation			
19.	Storage address	County (district/city):											
		Province (autonomous region/municipality):						Country:					
20.	Reproductive mode	Seed		Cutting		Grafting		Layering		Tissue culture			
		Others(Specified):											

21.	Exchangeable material	Seedling		Seed		Cutting		Scion		Pollen		Seminal root	Bud
		Others(specified):											
3. Communication information													
22.	Name of institute												
23.	Address												
24.	Name												
25.	Email												
26.	Tel.												
27.	Fax												
28.	Remark:												

(c) Climate conditions
The average annual temperature, average monthly temperature, highest temperature, lowest temperature, non-frost period, solar radiation, rainfall, etc., belong to climate conditions.
(d) Ecological conditions
They refer to the vegetation, accompanying plants (wild resource), rotation crop (intercropping, relay cropping), etc.

B. Information of germplasm resources

(a) Germplasm types
Germplasm types include wild type (species, varieties, variants), cultivation type (local varieties, breeding materials, bred varieties), etc.
(b) Germplasm source
It contains field, farmland, research institutes, etc.
(c) Germplasm name
Such as Chinese name, Latin name, alias, local name, etc.
(d) Germplasm number
The number involved in the investigation
(e) Utilization values
These values mean the real utilization value, such as yield, quality, usage, storage, and transport characteristics, and the potential utilization value, such as the utilization status of wild resources by local residents, utilizable part of plant, and some specific usage of cultivated plants.
(f) Utilization status
It includes the proportion, cultivation history, and distribution of local economic crops.
(g) Botanical characters
Type and distribution of roots, type and size of stems, plant height, branching situation, type, morphology, and adherent patterns of leaves, morphological characteristics of flowers, type and morphology of fruit, reproduction manners, seeds morphology, etc., are involved.
(h) Biological characters
Biological characters mean the growth habits, fluorescence, fruiting periods, reproductive cycles, and the main stress responses (such as pests, diseases, and various stresses).

C. Other related information

(a) Investigators and their home institutions
(b) Video information

Photos and video should be taken and shot timely and recorded in detail. For the samples that are dehydrated and deformed easily after collection, the video should be recorded as soon as possible.

Real Samples

During the collection of germplasm resource samples, the location, sampling technique, and sampling quantity should be considered.

A. Selection of location
   The locations for collection of germplasm resource samples are mainly the genuine producing areas and main producing areas, including the production farmlands of cultivated species and the natural habitats of wild species.
B. Sampling technique
   As the differentiation incidences in germplasm are imbalanced, the germplasm samples should be collected selectively. If adopting random method, many valuable materials are often omitted. Numbers and distances of sampling sites should be determined according to the following factors: the variety of local plants, distribution manners, the density of target germplasm, differentiation situation of individuals, etc. The purpose is to obtain the largest variability in the smallest sample number. The optimal sampling number should be determined by the diversity degree. Local species, wild species, and wild relatives are all farraginous populations, and the collection numbers of various germplasm materials should be as large as possible. Generally speaking, collection in different sites is better than in the same site for one investigation area.
C. Plant specimen and medicinal parts should be collected during sampling
   As the main basis of plant classification is character, the specimens should be collected with flowers (fruit); the medicinal parts are important proof for evaluating germplasm resources of medicinal plants; thus, the collected medicinal parts should be the crude drug samples in harvest.

7.5.1.3 Sorting of Specimens and Data

Serial Number of Samples

The specimens collected should be labeled and numbered at any time. Various numbers of one kind of specimens should be consistent. One serial number is for one material and should be clear and not repeated in every situation. For example, the number of specimens collected from Mount Emei, Sichuan, is coded as schemsh (the pinyin initials of Mount Emei, Sichuan), and the numbers of the same kind of materials can be set sequentially as schemsh-1, schemsh-2, etc. Fill in the original record card during specimen collection.

Preparation of Specimen

The paper should be changed timely during plant specimen collection in the investigation. After they are dry, they should be prepared into exsiccate specimen, which

benefit for storage; the medicinal specimen should be dried timely to prevent mildewing. It is necessary to prevent moisture, pest, rat, etc., during the storage. If possible, it is better to store them in the specimen museums with necessary facilities.

Original Record Card

The specimen collected could not completely represent all characteristics, producing area, ecological environment, etc. Therefore, the careful observation is needed, and the results should be filled in the original record cards one by one during the field investigation. The original record cards of various medicinal plants should prominently reflect all the featured contents.

Video Information

Photos and videos of the ecological environment of main germplasm materials and their collection spots should be shot, especially for the samples that are dehydrated and deformed easily after collection. These video data can help us to identify and classify the germplasm resources. In addition, for the medicinal plants with high trunks, only certain parts of the plant can be collected during the investigation; thus, the video information must be used to present the whole appearance of plants.

7.5.1.4 Investigation Summary

During the whole process of investigation, we should summarize frequently. If shortages are found, check back to the investigation spot. After investigation, it is necessary to review the investigation work comprehensively and write the summarizing reports. This is a very important process, which can integrate, systematize, and theorize the obtained data and materials during the investigation. As the valuable results of investigation work, the summary should be as detailed as possible, so as to provide references for the current and future use of germplasm resources.

The contents of report should include the following items:

① Basis and purpose of investigation
② Location, objects, and time of investigation
③ The natural and ecological environments of the investigated germplasm resources
④ Local producing situations
⑤ Local planting area, altitude distribution, species, and their changing history
⑥ Plant communities and associated plants of wild species and wild relative plants, their use value, geography and altitude distribution ranges, and changing situation

⑦ Obtained samples, their features and characters, their positions in the plant classification and plant origin evolution, their use values in the scientific researches in genetics and breeding, etc.
⑧ Comprehensive evaluation and recommendation on local agricultural production and developing utilization of germplasm resources of medicinal plants
⑨ The experience and lesson from the investigation

7.5.2 Collection of Germplasm Resources of Medicinal Plants

Collection of germplasm resources is the gathering of germplasm resources with special purpose, including general survey, collection in special class, domestic collection, international exchanges, etc.

7.5.2.1 Collection Objects of Germplasm Resources

The collection objects of germplasm resources include the following:

① Medicinal wild species and cultivated species, especially those species in danger of extinction and local rare excellent species
② Species that were cultivated in the past but has been abandoned
③ Wild relative plants of cultivated medicinal plants
④ Special germplasm resources, such as mutated cultivated species, homozygous, intermediate type of hybridization
⑤ Wild plants with potential value for humans

7.5.2.2 Collection Principles of Germplasm Resources

Comprehensive Property

During the collection of germplasm resources, all varieties should be found, as the germplasm resources of medicinal plants are usually mixed communities. It is necessary to observe carefully and collect the germplasm resources of medicinal plants' existing variants. Only in this way, the integrity of germplasm resources could be represented, especially those valuable germplasm resources, which cannot be omitted in collection. The collection quantities of seeds or propagating materials for each accession of germplasm resources should be based on the actual situations. Generally, an accession of germplasm resources with seeds as carriers ranges from 200 to 2,500.

Integrity

The collected specimen of germplasm resources is required to be integrated, especially the flowers and fruit of medicinal plants, which are important basis for plant classification. When collecting specimens of woody medicinal plants, owing to the big appearance, the whole plants could not be collected. Therefore, it is acceptable to only obtain the branches of flowers and fruits. For the dioecism plants, the female plants and male plants should be collected individually. Specimens of plants whose leaves grow after blooming are collected twice.

Representatives

Each collected specimens, seeds, or asexual propagation materials should come from community plants. Only in such case, they could represent the features and characters of germplasm resources. Each collected materials should be made sure to express genetic variation of this germplasm resources completely. At the same time, the interference of nongenetic variation factors must be avoided. The collected seeds or asexual propagation materials are developed normally and fully matured.

7.5.2.3 Collection Method of Germplasm Resources

Collection of Germplasm Resources

The collection of germplasm resources of medicinal plants is organized by local staff and delivered to the administration units, after receiving the documents or papers sent by the national administration departments, or medicinal or agricultural research departments. However, this method is used rarely.

Investigation Collection of Germplasm Resources

Investigation and collection of germplasm resources by the expeditions is one of the main manners to collect and accumulate the germplasm resources of medicinal plants currently. Three large-scale general surveys were conducted in the twentieth century in China, and basic resource coordination work has been conducted for important medicinal plants in seventh, eighth, and ninth 5-year. At present, basic knowledge in germplasm resources of medicinal resources in China has been obtained, but the collected and stored germplasm resources of medicinal plants are relatively few; there is a lot of work to be carried out in-depth.

7.5.2.4 Establishment of Germplasm Resources Files

Database

Database includes the collection number, species, material names, collectors, collection time, collection channels, source, origin, collection quality, main characteristics, and so on. Plants are classified according to the origin of the name as the first-level classification order and the material names as second-level classification order. Then, the database will be established based on the order of family, genus, species, subspecies, varieties, and cultivated species. Image data and evaluation results over the years should be digitized and input to database. When all collected germplasm resources have been assessed and main characters have been identified, germplasm collection directory should be summarized and printed.

Real Specimens

Real specimens include wax leaf specimens, seeds, fruit and other propagation materials, medicines, and other samples.

7.6 The Molecular Evaluation of Pharmic Plant Germplasm Resources

There are so many methods to identify germplasm, including morphological observation, group determination method, population identification mode, phenotypic determination, pedigree analysis, and molecular measuring method, and so on. Different means and standards could distinguish germplasm in different levels, such as subspecies, varieties, species, population, family, pedigree, clones, and strains, and so on.

The ultimate criterion to evaluate the quality of germplasm resources is clinical effect, whose material foundation is the active components. Therefore, the ratio of the active ingredients is a feasible standard to determine whether pharmic plant germplasm resources are good or not. Since most of the medicinal ingredients of pharmic plants are secondary metabolites with low contents and are controlled by the environment and genes, it is hard to evaluate the quality of pharmic plant germplasm based on active components. In addition, the key point is how to eliminate the effect of environment and evaluate the germplasm. Well, molecular marker technology, which is not affected by the environment, can reflect the germplasm resources differences and genetic differences. This section mainly introduces the molecular evaluation of the pharmic plant germplasm resources.

7.6.1 Species Identification of Germplasm Resources

At most times, when collecting pharmic plant germplasm resources, we usually choose seeds. Thus, it is necessary to identify the seed collected, especially for those with long growing period and a lot of relative species to ensure its authenticity. Given that DNA could essentially reflect the features of pharmic plants, analysis in DNA level could not only be used in species identification of germplasm materials but also clear the disputed genetic relationship and provide evidences for finding new pharmic resource. For example, the genetic diversity and relationship of ten kinds of Dendrobium resources were analyzed by RAPD technology. Its result showed that ten kinds of Dendrobium resources had three categories, which reveals genetic background and genetic relationship of these ten Dendrobium germplasm resources from molecular level. It is obvious that DNA molecular markers technology, as an effective means, can be used to compare the genetic relationship among different pharmic plant species, genus, family, and other taxa or within each individual one. It brings new vitality to the species identification of pharmic plant resources and the discovery of new drug source.

7.6.2 Purity Test of Pharmic Plant Germplasm

In germplasm resources research, analysis of genetic diversity is an important part. At present, the genetic diversity of pharmic plant could be studied in the following aspects.

7.6.2.1 Identification of Species and Purity

As the rapid development of DNA analysis, species and their purity could be identified accurately and reliably, as viewed from DNA. There are many traditional ways to identifying the purity of species, such as using seedling morphology, plant morphology, and seed morphology. Compared with traditional field morphological identification, the application of DNA molecular marker provides an objective, accurate, and rapid channel for the identification of crop varieties. The test object is the DNA fragments (genes) of seeds without organ specificity, which could be detected in various tissues in all developmental stages. They are not influenced by the environment, having nothing to do with expression. Moreover, these DNA fragments are innumerable throughout the entire genome with rich polymorphism. What is more, high accuracy and good repeatability are showed in the process. DNA samples isolated under appropriate conditions can be preserved for a long time, which is very beneficial to retrospective or arbitration identification.

7.6.2.2 Establishment of Germplasm Conservation Strategies

Genetic diversity analysis of pharmic plants provides the theoretical basis for proper development of germplasm collection, conservation, and utilization programs, mainly on the analysis of collected germplasm and narrowing the scope of preservation, to save financial and material resources

7.6.3 Determination of Genetic Relationships

Wild relatives of cultivated pharmic plants are the ancestors of cultivated plants or wild species with close genetic relationship, which serve as an important carrier of disease resistance, insect resistance, and stress resistance genes, due to the long-term survival in natural adversity. There are countless examples of cultivation of high-yield varieties from genes of wild relatives at home and abroad. For example, during the 1970s, China supported three lines of rice to cultivate hybrid rice, substantially increasing rice yield. The key to this success is the discovery and use of male sterile wild rice. In addition, after long-term cultivation and human breeding activities, biological characteristics of some pharmic plants have degraded, and whether they become clones or wild species is unclear, such as Chuanxiong and ginger. During the implementation of the GAP, pharmic plant species and varieties (lines) tend to be single with narrow genetic base, which not only easily causes some disease pandemic but also restricts the production and the quality. Therefore, the protection and use of pharmic plant wild relatives play essential roles in improving the genetic diversity of cultivated pharmic plants.

If we want to protect and use wild relatives of cultured pharmic plants, the primary task is to determine what belong to wild relative species. At present, among numerous identification technologies, DNA molecule appraisal technology with its unique advantages serves as an important means in carrying out this work. For example, the study of the genetic relationship of 26 classifications of Dendranthema and the phylogeny relationship of 7 kinds of wild Chrysanthemum, verified in the molecular level that modern culture Chrysanthemum is the culture hybrid complex. It is the result of the hybridization of natural hybrid based on between Maohua Chrysanthemum and wild Chrysanthemum and purple Chrysanthemum and *Chrysanthemum nankingense* in later hybrid, which is produced after artificial selection. What is more, using RAPD technology to explore the relationship between the genetic differentiation of wild and cultivated *Paeonia lactiflora* Pall and the genuineness formation of red peony root and white peony root, it provides evidences for genuineness formation of red peony root and white peony root in the molecular level. Moreover, the research on the relationship between four types of cultivated *Angelica dahurica* and three wild relative species (*Angelica dahurica*, *A. dahurica var.formasana*, and *A. porphyrocaulis*) by RAPD and ITS sequence shows that the contemporary available wild species of Chinese medicine *Angelica dahurica* come from Taiwan *Angelica dahurica*, which distributes in southeastern areas of China only, mainly in Taiwan. So, the evidence in molecular biology has an important application value in the confirmation of pharmic plant wild relatives.

7.7 Seeking for Essential Functional Genes of Pharmic Plant Germplasm Resources and Molecular-Assisted Selection

7.7.1 Seeking for Essential Functional Genes of Pharmic Plant Germplasm Resources

Functional genes study is one of the most active and core research areas on pharmic plant germplasm genomes in post-genome era. It emphasizes the development and application of experimental methods as a whole (at the genomic level or systemic level) to analyze genome sequence and clarify its function. The basic strategy has changed from a single gene or protein in the past to all genes or proteins at systemic angle. In the study of functional genes in pharmic plant germplasm resources, compared to research of DNA genome and EST, cloning of biosynthesis-related genes is the most active research area, which is the most intimated to activity formation. Due to the increasing practical demand for natural medicines in clinic and continuous developing of gene cloning and expression technology, there is a trend of rapid growth in functional gene cloning of medicinal plants in recent years.

7.7.1.1 Research Methods of Functional Genes in Pharmic Plant Germplasm Resources

Functional gene, as a new area of research, has many study methods, which can be categorized as forward genetics and reverse genetics. Forward genetics starts from a mutant phenotype, focusing on "what is gene"; reverse genetics originates from a mutant gene sequence and studies phenotypic changes. T-DNA insertion, transposon technology, gene expression and analysis, expressed sequence tag, and biochip technology, etc., are involved.

Analysis of Gene Functions by Mutants

Transposon Tagging Technology

Transposon is a certain DNA sequence in chromosomal DNA, which is capable of self-replication and displacement, leading to inactivation of inserted genes by mutation. Undoubtedly, its removal can make them reactivated as well. Therefore, we can use the transposition and insertion of transposon to study the function of inserted target genes, which is called transposon tagging technology. Recently, it has become an effective way on genetic function identification and isolation of pharmic plant germplasm resources. Scientists discovered a male sterile gene from *Arabidopsis* by using this method, while Hirochika used Tos17 homologue recovering transposon detection gene in rice. Takken isolated a large number of tomato genes by Ac/Dc transposon tagging in maize, including genes controlling resistance to different

types of leaf mold locating on Cf-9 and Cf-4, dwarf genes encoding cytochrome P_{450}, DCL genes controlling chloroplast development, and FEEBLY genes involved in metabolism and development.

T-DNA Insertion

With the improvement of gene transfer technology mediated by *Agrobacterium*, T-DNA has been the main method for rapid construction of inserted mutant library. T-DNA insertion can block the expression of normal gene, leading to plant mutation. Specific amplification of both sides of T-DNA genes can isolate wild type genes. This method can isolate genes without products and expressions being known in advance, whose advantages include stability of genomic T-DNA insertion, without site-specific property, being able to construct T-DNA inserted mutant library. The specific strategy is that after cloning marker genes like NPT II, GUS to T-DNA segment, T-DNA is transferred into plant tissue mediated by *Agrobacterium*. Mutants, enhancers, and promoters are discovered from transgenic offspring, resulting in the separation and identification of a number of genes. Wang obtained a male sterile mutant in maize by T-DNA insertion.

Direct Analysis DNA Sequence

System Analysis of Gene Express

System analysis of gene express (SAGE) is a very effective and rapid technology for gene analysis, using oligonucleotide sequence with 9–10 bp on certain region of a transcription as tag to specifically represent a transcription. Several tags (generally 20–60) are linked by ligase and cloned randomly to vector, and SAGE library is established. Through sequence analysis of double labeling, distribution of gene transcription is obtained. Procedures in detail are as follows. Firstly, mRNA is reverse-transcribed to cDNA, which is followed by digestion with the anchoring enzyme. The obtained cDNA is linked with the joint A and joint B, containing anchoring enzyme and tag enzyme restriction sites, respectively. After that, cDNA fragments with joints produced by tag enzyme are linked as double labeling, which are amplified by PCR. PCR products then are digested by anchoring enzyme, and double-labeling are isolated and cloned. Finally, the result is analyzed by SAGE analysis software and compared with the GenBank, EST database, etc., so as to obtain the information of transcript abundance and new expressed gene.

Expressed Sequence Tag Sequencing

Expressed sequence tag sequencing (ESTs) is one of the best ways to study functional genes, which is high-throughput cDNA sequencing technology. There are a large number of ESTs in medicinal plants, providing accurate positioning information of genes. When a single EST is corresponding to two or more exons, ESTs labeling enables to give position information of intron. Combined with software

analysis, ESTs is used to discover and study genetic functions in species with large genomes. Additionally, ESTs is widely applied in gene mapping, finding homologous genes, etc. After establishment of normal cDNA library, through comparison of ESTs from different plants and organizations, gene sequence polymorphisms of different medicinal plants are revealed. Bevan had shown that 56% genes of Arabidopsis genome can be identified with the corresponding EST. Moreover, comparison between EST, intact sequences, and all EST data from other plants can help to recognize and understand the functional genomics of medicinal plant.

Analysis of Gene Function Using Homologous Sequencing

Homologous sequences comparison is the first step of predicting the function of a new gene, which is based on the homology of the tested gene and the corresponding known gene. After homology search of tested gene in DNA and protein sequence database, a series of genes or fragments will be obtained which have high homology with this new gene. Undoubtedly, their functions serve as an access to finding functional information for the new gene. Yi compared the gene sequences of proteins encoded by more than 20 plant disease resistance genes and found that most of these genes contained conserved domains like special nucleotide binding sites, repeated leucine, leucine zippers, and transmembrane structures. According to their conserved domains and PCR primers, a large number of disease resistance gene analogs (RGA) were amplified. With RGA as probes, resistance genes were acquired from the gene pool. U.S. Genome Consortium has combined their research progresses on plant functional genomes of *Arabidopsis*, rice, *Mesembryanthemum*, barley, *Selaginella*, *Dunaliella salina*, etc., and supplied a link with pipe online as data analyzer.

Gene Functional Study by Biological Chips

Biochip is a micro-biochemical analysis system which is built on a solid surface through micromachining technology and microelectronic technology. Combined with software, it can accurately and quickly analyze a great deal of biological information. Nowadays, biochips used in functional genes study of Chinese medicine include gene chip and protein chip.

Gene Chip

A large amount of gene probes or gene fragments are fixed in a particular way on silicon, glass, and plastic sheets, hybridized with the labeled sample. The sequence information of sample will be obtained by detecting the hybridization signal of each probe. According to different preparation methods, gene chips could be divided into DNA chip, microarray, and electronic chip. Gene chip technology is mainly applied in gene expression analysis, the new gene discovery, gene mutation, polymorphism analysis, and monitoring of genome transcription and expression. Notoaki has studied gene expression profiles of chrysanthemum, rice, etc., by cDNA chip, and obtained

a large number of gene expression information. They analyzed the defective mutants and found some determinants whose metabolic and signal mechanisms are closely related with the resistance. Schena constructed cDNA microarray with 45 Arabidopsis cDNA and 3 control cDNA and detected 26 genes with different expressions in roots and leaves. Philippe found a novel gene induced by insect bites, using cDNA chip with 150 ESTs from *Arabidopsis*. Haroni isolated 1701 cDNA fragments to build chips and studied the relationship between the color of fruit and maturity degree. They found that 401 genes in green, white, and red fruits showed significant expression difference, one of which was the gene of strawberry acetyltransferase. This enzyme plays an important role in the synthesis of some special substances in the mature strawberry. Lu combined cDNA with suppression subtractive hybridization to separate and identify a new gene, which was related with affinity and non-affinity reactions of the rice blast. Fang constructed a cDNA with 2,400 genes, making use of human leukocyte cDNA library. The changes of gene expression of human promyelocytic leukocyte cell line HL260 after ajoene treatment indicated that the effect of ajoene was concerned with its anticancer activities in the gastrointestinal tract and leukemia.

Protein Chip

Protein chip is formed by fixing peptides or proteins extracted from a cDNA expression library on the surface of solid medium. When it hybridizes with a variety of antigens, antibodies on the chip will capture corresponding antigens. Following hybridizing with various labeled antibodies, qualitative and quantitative analyses can be carried out according to hybridization signals. In the consequence, it can help to study the relationship between protein structure and function and serve as a rapid and meaningful search database with adequate family classification searching results.

On November 2000, a research team in Yale University reported the first time that functional genes were cloned by protein chips. Five thousand and eight hundred kinds of proteins in yeast were expressed and purified, followed by fixation on slides to produce a microarray chip with yeast and proteome. It was used to screen the protein which can interact with specific proteins and phospholipids. Consequently, a new protein was discovered which can interact with calmodulin and phospholipids. This protein microarray cannot only be used for screening drugs having interaction with proteins but also be used to detect modification of proteins after translation. Their findings confirmed the feasibility of production and use of protein microarray for functional analysis, predicting its wide application prospects in drug development.

Gene Functional Study by Proteomic Techniques

Proteome Research

Proteomics refers to the full set of expressed proteins of a gene group, a kind of species, a cell, or tissue. Proteome research isolates and identifies all the proteins of a

tissue or cell. And protein separation technologies include two-dimensional gel electrophoresis and two-way high-performance column chromatography. In addition, protein identification technologies are Edman degradation method which is used to measure N-terminal sequence, mass spectrometry, amino acid composition analysis, and peptide mass fingerprinting. Among them, mass spectrometry is the most widely used with the fastest growth.

Proteome Research of Medicinal Plant

With the present development of proteomics technology, two-dimensional electrophoresis results between knockout mutants or transgenic recombinants and wild species are compared and analyzed, so as the study of the function of target genes. In 1998, when studying isogenic lines of maize Opaque2 gene, Damerval analyzed the difference of their two-dimensional electrophoresis results. Finally, he isolated and cloned a new transcriptional activator gene. Moreover, through identification of developmental mutants of *Arabidopsis thaliana* by two-dimensional electrophoresis gel, he found cytokinin overexpression in one mutant.

7.7.1.2 Functional Gene Cloning of Pharmic Plant Medicine

Till the end of September in 2003, 255 functional gene sequences had been registered in 42 species, 32 genera of pharmic plants all over the world. In Chinese herbal medicine, *Catharanthus roseus*, licorice root, *Artemisia annua*, and mint have the largest number of registered genes. Moreover, 12 species like yew, Lithospermum, and saffron registered 4–15 genes, respectively. Tianma, ginkgo, lotus, digoxin, and Kudzu in 16 genera and 16 species only registered less than 4 genes, and the other 16 genera and 26 species of plant had 233 registered genes. In the world, the countries making most efforts on plant gene research include Japan, the USA, Germany, and China. China has registered 27 genes of pharmic plant medicine, most of which are artemisinin biosynthesis genes, followed by taxol-related genes and Astragalus polysaccharide synthesis genes. In Germany, scientists paid more attention to cloning of major synthetic enzyme genes of *Catharanthus roseus* and *Rauwolfia*, and indole alkaloids synthesis-related genes. Furthermore, the plants whose genes Japanese scientists cloned the most are Gentian, Eucommia, licorice, shikonin, baicalin, ginseng, and henbane, generating tropane alkaloids. While US scientists cloned yew, mint, and Astragalus-related functional genes, Croteau laboratory of Washington State University carried out the most distinctive work. They have done an excellent research work on terpenoids, especially cloning taxol biosynthesis-related genes.

1. *Cloning of Flavonoid-Related Genes and P_{450} Cytochrome Oxidase Gene*
   Flavonoids possess a variety of pharmacological activities, which are closely related with the flower colors. Thus, the study of flavonoid biosynthesis pathway began in the early times and has made great progress in its biosynthesis steps,

enzymes during each step, and genes. A statistics of flavonoids synthesis genes in 26 plants including Arabidopsis, Antirrhinum, pine, peach, petunia, and grape shows that at least 103 flavonoids biosynthesis genes were cloned, like genes of chalcone synthase, the enzymes catalyzing hydroxylation of flavonoid skeletons and glycosidase.

The other genes studied often were cytochrome P_{450} genes, which belong to a class of oxidases containing heme proteins with a variety of catalytic functions. P_{450} in plants exert an important role—can catalyze a number of primary and secondary metabolic reactions. Due to their essential functions in the secondary metabolism and plant resistance (anti-herbicide) and potential application in biological decontamination, study on P_{450} attracts particular concern. Currently, more than 600 P_{450} genes have been cloned, more than 100 of which were expressed successfully in bacteria, yeast, baculovirus, etc., with function identification as well. Besides the model plant *Arabidopsis thaliana*, the P_{450} enzymes in *Catharanthus* have been studied widely.

2. *Cloning of Taxol Biosynthesis Genes*
   Cloning of taxol biosynthesis genes is the most typical example in cloning of plant medicine functional genes, which represents the research and application trends of natural product chemistry in post-genome era. Croteau laboratory in Washington State University has cloned and expressed 11 genes related with biosynthesis of anticancer drug Taxol from different species of yew by homologous primers PCR, differential display combined with Taxus cell culture, and reverse genetics methods. These genes include skeleton formation genes, three hydroxylase genes catalyzing oxygen replacement of skeleton, and five acetyltransferase genes responsible for the side chain formation. They are all expressed in *E. coli*, yeast, or insect expression system, whose expression products are functionally identified as well. On this basis, it is determined that taxol formation from the common precursor of a diterpenoid at least goes through 20 steps. Besides, the biosynthetic pathway of taxol with eight oxygen substituents has been elucidated. Moreover, the specificity of recombinant enzymes and the order of various substituents in the biosynthesis are determined with different substrates.

7.7.1.3 Conclusion

"Chinese Medicine Genome Project" combines molecular biology, bioinformatics, and other modern biotechnology with modern scientific technology to study and develop traditional Chinese medicine. And the United States put forward the "Human Genome Project" (HGP), "Plant Genome Project" (PGP), and "Microbial Genome Project" (MGP) and has achieved a breakthrough. Proposed and implemented in China, "Chinese Medicine Genome Project" is the significant strategic initiative in modernization of Chinese medicine research and development, among which functional genomics is one of the most important parts. Through study on functional genomics and proteomics, the technical system of important functional

genomics research and the important functional genes discovery from Chinese medicinal plants should be established and improved. The profound meaning of Chinese herbal medicine should be reflected. Furthermore, genes with Chinese independent intellectual property rights, clear function, and potential application should be acquired as well. After understanding the genetic expression and regulation rules in Chinese medicine, we can upregulate the genes responsible for effective components, so as to promote their production. Additionally, through the control of genetic information, the active ingredients of Chinese medicine are analyzed and synthesized, and their mechanism will be clarified, which lays the foundation for the modernization of Chinese medicine and its acceptance in the international market. Last but not least, it can promote the development of related pharmaceutical industry, making contribution to human health.

In functional gene study of plant medicine, compared with genome sequencing and ESTs sequencing, genes related with active ingredients biosynthesis attract more attention. The growing number of their cloned genes suggests the beginning of post-genome era of natural products and possible changes of some natural products production modes. Therefore, our hopes on biotechnological methods may be realized through new technology with combination of genetic engineering and chemistry.

On the basis of *Dendrobium* study for many years, our laboratory carried out some research on its functional genes. Using chemical mutagen EMS (ethylmethane sulfonic acid) to directly induce mutation of protocorm in *Dendrobium* tissue culture, we obtained a stable mutant without *Dendrobium* alkaloids. Moreover, by differential display, antisense RNA technology, and transgenic techniques, genes concerned with the *Dendrobium* alkaloids biosynthesis have been cloned and identified. The work has been completed as follows. mRNA was extracted from Dendrobium mutants and normal materials. The first strand of cDNA was synthesized through reverse transcription with selected 3′-end anchor primer. PCR amplification was carried out with labeled substrates, using 5′-end random primer and 3′-end anchor primer as primer set. At present, a subtractive cDNA library has been established, and the sequencing and bioinformatics analysis are in the process.

7.7.2 The Molecular Marker-Assisted Selection in Germplasm Resources

7.7.2.1 The Concept and Significance of Molecular Marker-Assisted Breeding

Molecular marker-assisted selection is a modern breeding technology which uses the DNA marker linking closely with target property to select the target property indirectly. This method making transfer of target gene cannot only make accurate and stable choices in the early time but also overcome the difficulty on reuse of recessive genes. Therefore, the breeding process will be speeded up with improved

breeding efficiency. Compared to conventional breeding, the technology can improve the breeding efficiency up to about threefold. The key to the technology is the identification of DNA molecule that closely links to agronomic characters. As for the obvious advantages, this technology has attracted great attention in developed countries. The United States, Japan, Western Europe, and other countries have invested a lot to this work in recent years. In China, some universities and research institutes have already mastered this technology and made a number of important achievements after recent years of research in crop breeding. Colleges of Chinese medicine and its research institutes have applied this technology to pharmic plant breeding.

The establishment of large-scale planting base for Chinese herbal medicine indicates the production of herbal medicine in China have changed from relying on wild resources to scale and standardize cultivation. In order to achieve the goal of high-quality, high-yield, standardized production of medicinal plants, the improved varieties of medicinal herbs used play a decisive role in the production. "High content" and "high yield" are the main objectives of medicinal plant breeding. Genetic linkage map is the molecular file of organism, which can provide reference for the breeder. Using RFLP, RAPD, AFLP, SSR, and other molecular genetic markers to construct genetic linkage map for important medicinal plants, essential quantitative trait loci (QTL) research and practice of medicinal plants will be carried out. In addition, good varieties will be selected from the wild types to achieve the sexual hybridization between wild plants and domestic plants. This will become one of the important directions for medicinal plant breeding in the future.

7.7.2.2 Research on the Genetic Map of Medicinal Plant Germplasm Resources

Genetic Map

Linear arrangement of genes on chromosome is known as the genetic map. It is based on the recombination frequency of allele in meiosis to determine its sequence and relative distance in genome. Usually, recombination rate is used to represent the genetic distance between genes, whose unit is centi-Morgan (cM). One cM is accordance with 1% recombination rate. However, cM stands for the relative position of gene in chromosome instead of actual length. There are lots of genetic mapping methods, such as RFLP, RAPD, AFLP, STR, SNP, etc.

Parent Selection

The genetic map is constructed successively by parent selection, composition population production, genetic marker chromosomal location, marker linkage analysis, and so on. As for parent selection, the species or material with distant genetic relationship and genetic variation should be chosen in theory, but it should not be too large, otherwise the seed set rate and the future construction accuracy of map

will be reduced. Molecular marker technique can be applied in detection of selected material polymorphism, and analysis of the result, picking out a pair or several pairs materials as composition parent with certain genetic variation.

7.7.2.3 Bulked Segregant Analysis and Mark for Important Agronomic Trait Genes

Nearly Isogenic Line

Nearly isogenic line (NLL) is the permanently stable lines, which are produced by introducing donor parent chromosome segments with certain target gene or quantitative trait loci into recurrent parent genetic background. It has high efficiency in gene mapping. However, it is nearly impossible to create near-isogenic lines for some woody medicinal plants because of its long cultivation time.

Bulked Segregant Analysis

Bulked segregant analysis (BSA) lays the foundation for quick and efficient screening of important agronomic trait marker genes. This method is often used to research the marker molecule linked to important agronomic trait genes. It can be used in medicinal plant breeding as well. Main procedures of BSA group construction are as follows: e.g., supposed to hybridize disease-resistant varieties with disease-susceptible varieties of a plant, the resistance gene separates in the F2 generation. According to the performance of resistant, the plants are divided into two groups, one of which is susceptible and the other one is disease resistant. Five to ten extreme lines are picked out from each group, respectively. After DNA was extracted, they are equally mixed to form resistant susceptible DNA pools, whose polymorphism are analyzed. The aim is to screen polymorphic markers and then analyze all of the separated plant, the target trait linkage, and the close linkage level.

7.7.2.4 Application of Germplasm Resource Molecular Marker-Assisted Selection

Basic Conditions for Marker-Assisted Selection

Medicinal plants for molecular marker-assisted selection should meet the following conditions:

① Genetic distance between molecular markers and the target gene is isolated or closely linked, generally less than 5 cM chain to be effective for MAS.
② A simple and efficient method of DNA automated extraction and detection should be available, which can be easily used for analysis and operation of large groups. It is better that the molecular marker is PCR-based markers.

③ Detection technology should be reliable, highly repeatable, economical, and practical.
④ A multitask data processing computer software is required. With these conditions, the breeding-assisted selection can be well carried out.

Molecular Marker-Assisted Selection Method

Introduce Favorable Genes

During the cultivation, some characteristics related to degradation of medicinal plants exist, such as decline of component as well as decreased disease resistance, etc. Based on the investigation, it is necessary to select a species having strong resistance and backcrossing several times with cultivated species. In addition, during each backcrossing, favorable genes and satisfying properties should be maintained so as to improve varieties. However, in practice, "linkage drag" often appears, which means the introduction of some negative genes linking with favorable genes brings more difficulty in improving variety. Traditional backcrossing method cannot identify recombination variety directly, due to its blindness, too much time of backcrossing, even more than 20, and low efficiency. On the contrary, molecular marker-assisted selection technology can identify recombination variety rapidly and directly. Using two sides' molecular markers which are linked closely with target favorable genes, it is possible to directly screen the recombination variety on this fragment. Generally speaking, through two or three generations, beneficial genes can be introduced without negative gene to obtain improved variety.

Polymerization of Favorable Genes

There are lots of favorable genes in medicinal plant germplasm resources. In the breeding process, the favorable genes of a parent can be localized and transferred to a variety by hybridizing or backcrossing. By this method, the favorable genes can be combined together so as to cultivate improved varieties. Molecular marker-assisted selection can overcome the shortcomings of conventional breeding and quickly combine the favorable genes together.

As for medicinal plant breeding, germplasm resource molecular marker-assisted selection has been rising recently. In the research of germplasm resource, we can learn from the experience of crop breeding, expand research on medicinal plant conventional breeding, and combine the conventional breeding with molecular marker-assisted selection to cultivate good seed of medicinal plant, serving for Chinese herbal medicine GAP production.

7.8 Research Progress of Transgenic Pharmic Plant Resources

Because of the impact of the current environment, medicinal plant resources are in shortage more and more seriously, especially some rare medicinal plants. However, transgenic technology, which is used to change some genetic traits of plants, such as

resistance to insect pests, can improve the quality of medicinal plants and the content of active ingredients. Therefore, it plays a vital role in the increase of Chinese herbal medicine resources. In principle, transgenic plants are a new type of bioreactor, whose pharmaceutical products are known as transgenic plant drugs, regardless of whether the transgenic plants are medicinal plants or not.

Transgenic plants are mainly formed by *Agrobacterium*-mediated gene transfer, particle bombardment, E3F method, and germplasm-system-mediated gene transfer. In addition, *Agrobacterium*-mediated gene transfer is the most popular method for transgenic medicinal plant resources, while others have not been reported in this area yet. Nowadays, hairy roots formed by transforming *Agrobacterium rhizogenes* Ri plasmid and the crown gall formed by transforming *Agrobacterium* Ti plasmid are used to produce effective active ingredients, which are considered as a hot spot in medicinal plant biotechnology research.

7.8.1 Tissue Culture of Transgenic Pharmic Plant Resources

All over the world, hairy roots tissue culture has been established in *Ginkgo biloba*, *Taxus*, periwinkle, tobacco, *Polygonum multiflorum,* shikonin, ginseng, Digitalis, *Artemisia*, etc. Some important secondary metabolites obtained by this system include quinoline alkaloids, indole alkaloids, tropane alkaloids, glycosides (such as ginsenosides, beet saponins), flavonoids, quinones, polysaccharides, proteins, and several important enzymes (like superoxide dismutase). Moreover, domestic scholars have detected some high level active substances in hairy roots of *Lithospermum, Artemisia annua, Gynostemma*, and American ginseng. In foreign countries, some hairy roots have been already applied in industry, such as *Lithospermum*, carrot, and periwinkle.

Some active ingredients of medicinal plants are only synthesized in the leaves and stems; however, the crown galls are able to produce these specific compounds. For example, the root of *Psoralea L.* contains furanocoumarins, which cannot be synthesized in hairy roots but the crown galls. Crown gall tissue culture has been used to produce secondary metabolites in *Asparagus,* Bidens, periwinkle, Cinchona, Digitalis, Lupine, lemon spearmint, spicy mint, *Salvia*, *Taxus brevifolia*, European Taxus, etc. The study by Zhang Yinlin et al., using *A. tumefaciens* to transform Salvia, has shown that the crown gall tissue can produce an active ingredient, tanshinone, exclusive in Salvia roots. In addition, Song Jingyuan et al. used induced crown gall tissue culture of Salvia to produce tanshinone, screening out a high-yielding strain, whose tanshinone content exceeded that in crude drugs. Thus, it may be a feasible way to produce tanshinone in this high-yielding strain with combination of elicitors and liquid static culture. Hank et al. induced a rapid growing crown gall in the hormone-free medium, whose tumor-like tissues contained paclitaxel and its analogues, proven by mass spectrometry and ELISA. Their contents are 0.00008–0.00004% of the dry weight.

However, the synthesis of many secondary metabolites requires the involvement of the leaves, stems, and roots. It is not possible for only the hairy roots or crown galls to produce them. Fortunately, if both of them are combined, the complementation of two metabolic functions can be achieved. Subroto et al. co-cultured genetic transformed belladonna abnormal stems and hairy roots in hormone-free medium *in vitro*. The content of scopolamine produced by co-cultured stems was 3–11 times of that in intact plants. The ratio between scopolamine and hyoscyamine in co-culture was up to 1.9, which was improved significantly compared with that in belladonna hairy roots (0–0.03). Given to low side effects and high value of scopolamine, the study of co-culture conditions of hairy roots and crown galls is possible to be used in mass production of scopolamine. Mahagamasekera found that hairy roots and crown galls from different genera of belladonna can be co-culture, producing scopolamine as well. Nevertheless, when being cultured separately, hairy roots and crown galls of Duboisia hybrid did not produce scopolamine, with only a certain amount of hyoscyamine in the hairy roots. Lei Hetian et al. established a dual transformation system by double transforming Ti plasmid and Ri plasmid into hairy root culture system of *Trichosanthes*. The results indicated that the dual transformation hairy root system grew more rapidly with similar protein content. Therefore, co-culture and dual transformation methods offer a novel approach for producing some special metabolites.

7.8.2 Application of Transgenosis in Feature Modification of Medicinal Plants

Through genetically modifying the characteristics of medicinal plants, the quality and adaptability of plants could be improved. Currently, the study on this area includes the following aspects.

7.8.2.1 Quality Improvement of Medical Plants

Intensively injecting genes to modify the characteristics of drugs, it is even possible to roll drugs and food into one. For example, due to the bitterness of *Gynostemma*, it tastes bad orally. At present, some scientists hope to transfer and express the protein genes from beet into *Gynostemma*, so as to cover the bitterness. Chen et al. introduced and expressed farnesyl pyrophosphate synthase genes from cotton into *Artemisia annua*, obtaining a transgenic plant with artemisinin content 2–3 times higher than that in control group. The flower buds of honeysuckle have the best quality, but their blooming time in the actual production is difficult to control. When using genetic engineering technology to suppress flowering of honeysuckle, it is possible to obtain high-quality medicine at utmost. It is also reported that the fat-soluble components (active ingredients) in the *Salvia* were changed through genetic

engineering technology, in order to increase the active ingredients. Zhao Yahua has integrated metallothionein genes of mice into genome of *Lycium* mediated by *Agrobacterium*, getting a novel zinc-rich transgenic *Lycium* variety. Although research on this area is not much reported, the method used to improve the genetic characteristics of medicinal plants is still promising.

7.8.2.2 Resistance Enhancement of Plants

Virus disease is the most common disease in medicinal plants. The use of gene technology can improve plant disease resistance. Besides, using genetic transformation to acquire antivirus plants has succeeded in beet, *Citrus aurantium* and Ningxia *Lycium*. For example, Luo Qing et al. have transferred lectin genes from snowdrop into Ningqi No.1 stain mediated by *A. tumefaciens* LBA4404, and a transgenic aphid-resistant lyceum is obtained. Similarly, an enhanced aphid-resistant transgenic *Lilium* appeared by introducing *Pinellia ternata* lectin genes into the genome of *Lilium* with *A. tumefaciens*. In addition, in terms of resistance to herbicides, Cun Yueshi et al. introduced the acetyltransferase genes from actinomycetes into belladonna and *Scoparia dulcis* with Ri plasmid as vector. And a herbicide-resistant belladonna was cultivated successfully. Moreover, a resistant strain of woad was obtained by Xu Tiefeng et al., who integrated bar gene into tetraploid woad leaves. This transgenic plant showed significant resistance to herbicide without affecting the production of secondary metabolites.

7.8.3 Genetic Metabolic Engineering of Medicinal Plants

Relatively speaking, molecular biology research on Chinese herbal medicine began a little late, and few genes can be used at present. Thus, it limits the application of genetic engineering in improving medicinal plants. With the development of genetic engineering, determination of the key enzymes during genetic cloning and metabolism in medicinal plants become more and more popular.

Heide et al. study shikonin biosynthesis-related enzymes in shikonin cell culture, and primarily determined p-hydroxybenzoic acid-geranyl transferase is the key enzyme during shikonin biosynthesis. Additionally, Croteau isolated and cloned the cDNA of a diterpene cyclase synthase (taxadiene synthase) from *Taxus brevifolia*, whose DNA and amino acid sequences were analyzed as well. They also successfully transferred other four genes like taxadiene synthase into *E. coli*, resulting in the formation of taxadiene which was only produced in *Taxus.* Furthermore, its yield was equal to that in 7,600-g *Taxus* bark.

To increase the production of artemisinin, mutant squalene synthase (SS) gene successfully replaced the wild-type gene in transgenic *A. annua* by gene targeting, adjusting the flow of metabolites. Ye Hechun et al. have cloned four key genes in

artemisinin biosynthesis pathway and transferred them by *Agrobacterium* Ti and Ri plasmids, so as to obtain a plant material with high artemisinin content.

Nowadays, most of the patents on genes are concerned with flavonoid metabolism, including 7-chalcone synthase, flavonoid skeleton hydroxylases, and glycosides transformation related enzymes.

7.8.4 Safety of Medicinal Plant Resources

The safety of transgenic plants is a controversial issue currently. Since the recombinant proteins in transgenic plants are relatively stable with low expression, difficult to disinfect and determine a structure, whose activity depends on structure, even the folding pattern, it is necessary to establish an appropriate regulation to appraise the safety of transgenic medicinal plants.

7.9 Case Study

Cloning and Characterization of a Novel 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase Gene from Salvia miltiorrhiza Involved in Diterpenoid Tanshinone Accumulation

7.9.1 Materials and Methods

The mature seeds of *Salvia miltiorrhiza* Bunge were surface sterilized by 0.1% mercuric chloride (Sigma–Aldrich, St. Louis, MO, USA) and cultured on solid, hormone-free MS basal medium. The MS medium contained 30 g/L sucrose and 8 g/L agar without ammonium nitrate for germination. Cultures were maintained at 25 °C under a 16-h light/8-h dark photoperiod with light provided by cool white fluorescent lamps at an intensity of 25 μmol/m^2/s.

Roots, stems, and leaves were collected from mature *S. miltiorrhiza*. The 20-day-old hairy roots were treated with MeJA for 0, 12, 24, 48, 72, and 96 h at a final concentration of 100 mM, and roots were harvested for RNA isolation. The hairy root lines were collected 30 days after inoculation. Total RNA was extracted from the tissues by Trizol method (Invitrogen, Carlsbad, CA, USA). Genomic DNA was isolated using the modified cetyltrimethylammonium bromide method.

5′-RACE was performed according to the manual of the SMARTTM RACE cDNA Amplification Kit (Clontech Laboratories Inc., Mountain View, CA, USA). The 5′-RACE PCR was carried out using the 5′-RACE primer and universal primer (UPM, Universal Primer A Mix). The PCR product was purified and cloned into

pMD19-T vectors followed by sequencing. After aligning and assembling the sequences, the full-length cDNA sequence of the SmHMGR2 gene was deduced and subsequently amplified by PCR using a pair of primers. The genome sequence of the SmHMGR2 was confirmed by PCR with the genome DNA as a template.

The nucleotide sequence, deduced amino acid sequence, and open reading frame (ORF) were analyzed using DNASTAR, and the sequence comparison was conducted through a database search using BLAST. SmHMGR2 and other HMGRs retrieved from GenBank were aligned using ClustalW. A phylogenetic tree was constructed using neighbor-joining method. Transmembrane domain was analyzed by TMHMM2.0, and homology-based structural modeling was accomplished by SWISS-MODEL (http://www.expasy.org/).

The entire SmHMGR2 cDNA was amplified by PCR using forward and reverse primers. The PCR product was digested with *EcoRI* and *NotI*, gel-purified, and ligated into the same restriction sites within the pRS406 vector. Positive clones were confirmed by PCR and subsequent sequencing analysis for the presence of the SmHMGR2 gene and used to be transformed into the *S. cerevisiae*.

The *SmHMGR2* gene was PCR-amplified, and the resultant PCR products were purified with the DNA purification kit (Sangon, Shanghai, China).We sequentially subcloned the PCR product into a donor vector (pDONR221) and created an entry vector. The recombination reaction between the entry and destination vectors (pH7WG2D,1) was performed in LR ClonaseTM II enzyme mix (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The *E. coli* DH5α was transformed with the product of LR reaction using the heat shock transformation method. The destination vector pH7WG2D,1 with the ccdB (control of cell death) gene knockout was used as the control vector (pH7WG2D-Control).

The *S. cerevisiae* strain JRY2394 (*MATa, ade2, his3, met-, ura3, hmg1, and hmg2*) was used to examine the function of *SmHMGR2*. The *SmHMGR2* gene was cloned into pRS406 expression vector (pRS406-*SmHMGR2*) and then transformed into JRY2394.

The vectors of pH7WG2D-*SmHMGR2* and pH7WG2D-Control were transformed into *S. miltiorrhiza* through the mediation of the *Agrobacterium rhizogenes* strain ACCC10060 as described previously. Root tissues from three flasks of cultures were collected separately on d 10, 20, 30, 40, and 50 after inoculation.

To detect the presence of Agrobacterium rol (B, C) gene in transgenic hairy root tissues, two pairs of PCR primers were used. Primer sets for the SmHMGR2 gene were designed by the sequence of 35 S promoter with the 5′-end of the target gene and by 3′-end of the target gene with the segments of hygromycin-resistant gene. For GFP checking, the hairy root was examined using a fluorescence microscope. Transgene copy number in transformed *S. miltiorrhiza* root lines was confirmed by DNA gel blot analysis using a nonradioactive digoxigenin-11-dUTP-labeled probe as described previously.

An aliquot (200 ng) of the total RNA was used, and the first strand cDNA was synthesized with total RNA using the reverse-transcription PCR system according to the manufacturer's protocol of PrimeScriptTM 1st Strand cDNA Synthesis Kit (Takara, Tokyo, Japan). To estimate the relative mRNA level, a diluted series of the

reference cDNA sample was used as the standard. The relative amount of the *SmHMGR2* gene was evaluated by the relative expression index of mRNA.

The content of squalene and tanshinones in hair roots was determined by chromatographic methods as described previously.

7.9.2 Results and Discussion

In recent years, there has been a remarkable progress in the understanding of the molecular regulation of diterpenoid tanshinone biosynthesis in *S. miltiorrhiza*. In the present study, the researchers attempted to further dissect the molecular biology of tanshinone biosynthesis pathways in *S. miltiorrhiza* by cloning the full-length cDNA of *SmHMGR2*. This gene was introduced into the leaf explants of *S. miltiorrhiza* under the control of the strong gene promoter with the gateway vectors. In this work, the overexpression of *SmHMGR2* in transgenic hairy roots resulted in increased tanshinone and squalene production. These findings suggest that SmHMGR2 is key enzyme controlling the diterpenoid metabolic flux and can be exploited for higher tanshinone accumulation in *S. miltiorrhiza*.

The HMGR superfamily contains a number of members that responded differently to various external stimuli including pathogen infection and exposure to xenobiotics. Distinct isoforms among HMGRs might be involved in the formation of secondary metabolites of terpenoids. Terpenoid metabolism can be elicited by jasmonic acid or its methyl ester MeJA. The content of tanshinones in hairy roots of *S. miltiorrhiza* was significantly increased after treatment with MeJA. MeJA is not a specific inducer of the *SmHMGR2* gene; it also induces other genes contributing to terpenoid biosynthesis. It is considered that the increased content of tanshinones by MeJA is due to an integrated effect on a cluster of genes related to tanshinone biosynthesis. These results implicate that HMGR plays an important role in the production of tanshinones.

In this study, *SmHMGR2* was successfully cloned from hairy roots of *S. miltiorrhiza*. It is functionally similar to the *SmHMGR1* gene since the transcription of both *SmHMGR1* and *SmHMGR2* was responsive to MeJA treatment. However, the expression of *SmHMGR2* could be detected in all the tissues (leaves, stems, and roots) of *S. miltiorrhiza* but at different levels with the highest expression in leaves, which is different from *SmHMGR1* that has the highest expression in roots. This indicates a tissue-specific regulation of these two genes.

The multiple alignments showed that the deduced *SmHMGR2* sequence had high similarity to other plant HMGRs and contained all conserved substrate-binding motifs of HMGRs. The N-terminal contained two transmembrane domains. The 3D model of *SmHMGR2* represented a typical spatial structure of HMGRs. However, the analysis of phylogenetic tree indicated that *SmHMGR1* and *SmHMGR2* were sorted into different groups, and *SmHMGR2* lacks the "RRRP" motifs. This implicates the functional difference of the two HMGR enzymes.

A detailed analysis of different transgenic lines of *S. miltiorrhiza* showed distinct phenotypic and metabolic variations. This may be due to a position effect or random integration of the transgene at non-specific sites in the plant genome, different number of gene copies integrated into the genome, different and specific posttranslational regulations of the endogenous enzyme, and different methylation patterns and extents. These differences may exist among hairy root lines of *S. miltiorrhiza* due to infection by *A. rhizogenes* ACCC10060, leading to differences in the growth rate, aging rate, size, and branches, especially the color of hairy roots and the content of tanshinones.

The difference in the tanshinone contents in different root lines may be attributed in part to the different growth stages of roots since tanshinone accumulation was high during the stationary phase (40–50 day). Actually, the growth rate of wild-type roots was slower than the transformed hairy roots (generally, 2–3 months per cycle for wild-type roots). As such, the researchers included two additional controls of hairy roots, VCK and HRCK, which could reduce the impact of growth rate. The results support a correlation between the overexpression of the *SmHMGR2* gene and the content of tanshinones.

In transgenic tobacco, overexpressing the HbHMGR can enhance the production of sterol. The cardenolide and phytosterol levels were increased by expression of an *N*-terminally truncated HMGR in transgenic Digitalis minor. Similarly, overexpression of an *N*-terminal-truncated HMGR increased the production of essential oils and sterols in transgenic *Lavandula latifolia*. The high expression of CrHMGR in these transgenic plants was associated with higher artemisinin content. Amorphadiene production was improved by 50% through transferring the HMGR gene into engineered yeast. However, the MVA pathway can also contribute to mono- and sesquiterpene production in *L. latifolia*. The study showed that transgenic hairy roots were morphologically distinguishable from wild-type plants, and there was a significant correlation between the increased expression level of *SmHMGR2* mRNA and the overproduction of diterpenes and triterpenes in the hairy roots of *S. miltiorrhiza*. These data suggest an important role of *SmHMGR2* in isoprenoid biosynthesis of *S. miltiorrhiza* via the MVA pathway.

Determining transgene copy number is an important step in transformant characterization and can differentiate between complex and simple transformation events. The number of inserted *SmHMGR2* gene copy in H14, H19, and H24 genomes was 2–3. This is in agreement with increased levels of *SmHMGR2* transcripts and contents of tanshinones or squalene. It provides further evidence that overexpression of SmHMGR2 through stable gene transfer enhanced the biosynthesis of tanshinones.

Interestingly, the deducted *SmHMGR2* sequence did not contain the motif of "RRRP." The motifs rich in arginines (RRR) are considered to be important for endoplasmic reticulum retention. On the other hand, proline residues are known to strongly modify the secondary structure of proteins. Plant HMGR has been shown to insert *in vitro* into the membrane of microsomal vesicles, but the final *in vivo* subcellular localization remains controversial. Endogenous *Arabidopsis* HMGR

was found to be localized at steady state within endoplasmic reticulum but also predominantly within spherical, vesicular structures located in the cytoplasm and within the central vacuole in differentiated cotyledon cells. The *N*-terminal region, including the transmembrane domain of *Arabidopsis* HMGR, was necessary and sufficient for directing HMGR to endoplasmic reticulum and the spherical structures. Further studies are needed to explore the subcellular distribution of HMGR in *S. miltiorrhiza*.

Further Readings

Arntzen, C. J. (2008). Using tobacco to treat cancer. *Science, 321*(5892), 1052–1053.

Celis, C., Scurrah, M., Cowgill, S., Chumbiauca, S., Green, J., Franco, J., Main, G., Kiezebrink, D., Visser, R. G., & Atkinson, H. J. (2004). Environmental biosafety and transgenic potato in a centre of diversity for this crop. *Nature, 432*, 222–225.

Ro, D.-K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J., Michelle, C. Y., Chang, S. T., Withers, Y. S., Sarpong, R., & Keasling, J. D. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature, 440*, 940–943.

Chapter 8
Regulation of the Active Constituents' Production of Medicinal Plants

Lu-qi Huang, Tsai-yun Lin, and Chao-yi Ma

Abstract The biosystem includes microbe, suspended cells of plants, transgenic organs, and other transforming systems. The biological transformation is to use the enzymes produced by the biosystem for catalyzed reactions of exogenous compounds, which has been widely used in the structural modifications of secondary metabolic products of plants, the search for the active leading compounds, and the explorations of drug structure-activity relations. This is called green chemistry.

The plant genetic engineering includes the development of crown gall nodules, hairy roots, and other transgenic organs. After being infected by *Agrobacterium tumefaciens*, the T-DNA fragments in the Ti plasmid of plant are integrated into the genome of the nucleus to induce and form into the crown gall nodules. The hairy roots are the special type induced by the infection of *A. rhizogenes* to the dicotyledonous plants. The plants with crown gall nodules and hairy roots will have a large amount of secondary metabolic products.

8.1 Introduction

Secondary metabolites of medicinal plants are usually the main active components of traditional Chinese medicine, but their contents are quite low in normal conditions. Nowadays, majority of these active components are used in clinic and medicine

Lu-qi Huang (✉) • Chao-yi Ma
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China
e-mail: huangluqi@263.net; machaoyi56@yahoo.com.cn

Tsai-yun Lin
Department of Life Science, Institute of Bioinformatics and Structural Biology, National Tsing Hua University, No 101, Sec 2, Kuang Fu Road, Hsinchu 30013, Taiwan
e-mail: tylin@life.nthu.edu.tw

Lu-qi Huang (ed.), *Molecular Pharmacognosy*,
DOI 10.1007/978-94-007-4945-0_8,

production with traditional Chinese medication decoction pieces as carriers. Wide collection and artificial cultivation of medicinal plants are the main ways to obtain active components. However, wild sources of medicinal plants are limited, and overcollection may place a certain pressure on biodiversity as well as destroy the ecological environment in a certain degree. Besides, field cultivation tides up a lot of farmland, and medical material from field cultivation, which has a low growth rate and a long growing period, is susceptible to pesticide residue and heavy metals. As a result, the production of Chinese herbal medicine active components using biotechnology has become the hotspot of molecular pharmacognosy. The biological production of Chinese herbal medicine active components includes biotransformation and gene engineering production.

8.2 Production of Active Component of Traditional Chinese Medicine by Biotransformation

8.2.1 The Meaning of Biotransformation

Biotransformation is a series of physiological and biochemical reactions which modify the structures of foreign compounds by biological systems and the enzymes produced in them to procure valuable products. The essence of biotransformation is enzymatic reactions which use foreign compounds as substrates and enzymes produced in biological systems as catalysts. Foreign compounds may react with the biological systems and the enzymes produced in them after these natural or synthetic organic compounds were added into and cocultured for a while with growing biological systems, and this process may result in some structure changes, such as hydroxylation, glycosidation, epoxidation, and so on. Biotransformation includes microbial transformation, plant suspension cell transformation, and transgenic organ transformation. By extension, drug metabolism belongs to biotransformation, too.

With the development of drug-screening technique and the application of high-throughput screening (HTS) technology, the discovery of lead compounds with novel structures and high medicinal values from traditional Chinese medicine has become more and more difficult; although some compounds with high activity exist in some traditional Chinese medicine, it is still hard to explore and utilize them very well because of their low contents or limited resources. So, finding the method for promoting the value of available traditional Chinese medicine resources has been the urgent affair. Treating traditional Chinese medicine with biotransformation, combining the method of quick and effective drug screening, and exploring effective natural active compounds with self-owned intellectual property rights and Chinese characteristics are considered as an important way to improve the value of traditional Chinese medicine. On the other hand, structure modification of some active components in traditional Chinese medicine by biotechnology is also one of the ways to solve the problem of losing value for exploring and utilizing because of low bioavailability, poor stability, or strong poisonous effect.

Research on biotransformation of traditional Chinese medicine *in vivo* is helpful to disclosing the absorption mechanism of traditional Chinese medicine compounds and explaining their active precursors, and it also can deepen the cognition of action mechanism of traditional Chinese medicine. Traditional Chinese medicine, which is usually taken orally, interacts with intestinal bacteria in digestive tract inevitably. So, the role of intestinal bacteria metabolism on traditional Chinese medicine could not be neglected. Some traditional Chinese medicines may begin to take effect only after being metabolized by digestive enzyme or intestinal bacteria in human body. For example, saikosaponin is absorbed and takes effect after being converted into metabolites by intestinal bacteria in rats; metabolite of icariin in intestine has a stronger promotion effect on IL-6; 20(*S*)-protopanaxadiol-20-O-β-D-glucopyranoside, the metabolite of ginsenoside in intestine, can suppress angiogenesis in tumor tissue, inhibit the tumor growth and metastasis of melanoma B16-F10, and induce apoptosis of human lung cancer cells. Now, there are many researches on the metabolism of amygdalins, coumarins, flavonoids, anthraquinones, and terpenoids by intestinal bacteria, and the identification of their metabolites suggests that it is conceivable to explore preparations of traditional Chinese medicine which can be used by human body directly and to improve the value of traditional Chinese medicine through doing research on intestinal bacteria.

8.2.2 Biotransformation System

Biotransformation system is usually to be microbe culture, plant suspension cell culture, plant transgenic organ culture, refined enzyme preparation or raw enzyme preparation, etc. Besides, seaweed and some insect larvas are also used for biotransformation. Among these systems, microbe culture and plant culture system are the most commonly used systems. There are also some reports in which plant transgenic organ is used for biotransformation.

8.2.2.1 Microbial Transformation System

Different microorganism strains can be bought from culture collection center, for example, China General Microbiological Culture Collection Center in Institute of Microbiology Chinese Academy of Science and Agricultural Culture Collection of China in Soil and Fertilizer Institute of Chinese Academy of Agricultural Sciences. You can isolate the strains from soil samples by yourself, too.

Microorganism culture method is usually done through two-stage fermentation. Firstly, microorganism is inoculated from solid slant medium into sterile liquid medium, and the seed bacteria, also called stage I culture, is obtained after 3 days of fermentation on a shaker. Secondly, proper amount of seed bacteria is inoculated into new liquid medium with a normal inoculum size of 5–15% V/V. One to two days later, stage II culture is obtained, and substrate could be added now as the microorganism grows swiftly.

8.2.2.2 Plant Suspension Cell Transformation System

1. Induction and Cultivation of Calli
   Inoculate the explants into proper solid medium after normal disinfection. Calli appear from the incisions of explant and then expand to whole explant surfaces after a period of culture. Pick up the calli from explants and transfer them into the same but fresh medium for further cultivation. Steady callus culture will be obtained after several subcultures.
2. Suspension Cell Culture
   Suspension cell culture is a process to culture single cells or small cell clusters in liquid medium. Transfer loose calli into liquid medium and cultivate on a shaker for some time in order to obtain homogeneous and suspended single cells or small cell clusters for expanding cultivation.

Suspension culture usually uses the method of shaking horizontally with its shaker rotate speed 30–150 rpm and its stroke 2–3 cm. The temperature for cultivation is usually 24–30 °C. The liquid medium usually occupies 1/3 of the flask volume. Calli for suspension culture should be fragile so as to obtain dispersed single cells or small cell clusters under shaking condition. We can adjust the components of medium, kinds, and concentration of hormone in order to establish the callus culture appropriate for suspension culture. For example, raising the concentration of 2,4-dichlorophenoxyacetic acid (2,4-D) or reducing the concentration of cytokinin, which includes 6-benzolaminopurine (6-BA), kinetin (KT), and so on, can make the calli loose. Medium optimized for callus culture may not be suitable for cell suspension culture. So, it is necessary to find the medium which would promote both the growth of suspension culture and the dispersion of cells. In this process, the type and concentration of hormone usually play an important role. As a general rule, one-cell suspension culture can be steady in its growth situation and its physiological characteristics only after several times of subculture, and then it can be suitable for biotransformation.

8.2.2.3 Transgenic Organ Transformation System

Induction of hairy roots: After *Agrobacterium rhizogenes* infects plant, T-DNA in Ri plasmid of *A. rhizogenes* will be transformed into the nuclear genome of plant and then hairy roots are induced.

Induction of crown gall nodule: After *A. tumefaciens* infects plant, T-DNA in Ti plasmid of *A. tumefaciens* will be transformed into the nuclear genome of plant, and then crown gall nodule is induced.

Transfer the transgenic hairy roots and crown gall nodule which are cultured in liquid MS medium into solid MS medium and cultivate them in darkness under 25 °C for 9 days and 28 days, respectively. And then transfer the culture into liquid MS medium for another 9 or 28 days in darkness under 25 °C, 110 rpm. After solid culture and liquid culture are alternately used for 3 times, the hairy roots and crown gall nodules could be used for biotransformation.

8.2.3 Addition of Substrate

Inoculate the plant cells or microorganism into flasks for liquid preculture and add the substrate when the growth of cells or microorganism is at the end of its logarithmic growth phase. The substrate could be added in the powdered form or be added after being dissolved by proper volume of solvent. We can use ethanol, acetone, dimethyl sulfoxide, etc., as little as we can to dissolve the substrate if its water-soluble capacity is really poor. However, it should be noted that the terminal concentration of these organic solvents should not be bigger than 1% to ensure the activities of enzymes. The added quantity of substrate may influence the transformation efficiency, and it is usually 30–100 mg/L. It is usually a little big in industry production, and sometimes it may be as large as 1.0 g/L, even larger. Some poorly water-soluble substrates, such as sterides, are usually added in the form of powder, and they can also be added into culture systems after being treated with surfactant or formed into inclusion complex of β-cyclodextrin.

8.2.4 Screening of Biotransformation Systems

Before biotransformation, we should screen the biotransformation systems at first. The aim of the screening process is to find the optimal plant cell culture, microorganism strain, and transgenic organ which have the strongest capacity of biotransformation. It is very important to set control group in the process of screening biotransformation system, and it may be helpful for you to exclude false positive and false negative cases and at last ensure the correctness of transformation results.

8.2.4.1 Control Group of Lacking Substrate

Inoculate and culture the plant cells, microorganism, or transgenic organs as same as the experimental groups, but replace the substrate with the same volume of solvent. Setting this kind of control group may be helpful to exclude the interference of organism's own secondary metabolites.

8.2.4.2 Control Group of Sterile Blank Medium Lacking Plant Cells, Microorganism, or Transgenic Organs

Add the substrate into sterile blank medium without plant cells, microorganism, or transgenic organs in, and then culture and treat this control group as same as the experimental groups. Setting this kind of control group may be helpful to ensure if the reaction is the result of biotransformation or only the result of the instability of substrate in medium.

8.2.4.3 Biotransformation Group

After the substrate is added into the culture system with a certain concentration, it will react for some time, usually 2–10 days, and then isolate the culture from the medium by centrifugation and filter. It is better to clear the culture for 2–3 times with solvent. On the one hand, extract the filtrate 3–5 times using proper solvent such as chloroform, ethyl acetate, and n-butanol after being saturated by salt. Mix the extracts, dry the mixture by sodium sulfate anhydrous, and then treat it with reduced pressure distillation. These samples and the samples in above control groups are comparatively detected by TLC or HPLC. On the other hand, extract the dry culture using appropriate solvent (e.g., cold-leached and ultrasonic extraction) for 2–3 times, treat the extract with reduced pressure distillation, and then, repeat the processes as above. Based on the work mentioned above, the culture which has the strongest capacity of biotransformation or has a specific capacity of transforming the substrate to our target product will be screened.

8.2.5 Biotransformation and the Extraction, Separation, and Identification of Products

After the selection and scale-up of the optimal culture system, biotransformation could be carried out in the same culture condition. Extract the filtrate and the culture as same as above after the biotransformation finish, separate the extracts of both the filtrate and the culture to obtain transformation products and substrate, and then calculate the yield of product and the recovery of the substrate. The following work is structure identification of biotransformation product using NMR, MS, IR, UV, CD, etc.

8.2.6 Investigation of Different Influence Factors and Reaction Dynamics

Biotransformation product will be influenced by many factors, such as concentration and addition time of substrate, culture temperature, etc. It is necessary to investigate the effects of these factors on biotransformation reaction to find the best combination.

To ascertain the culture time, we can take samples at various time points, detect the products in them, investigate the reaction dynamics in different periods of time, and find the optimal culture time at last.

8.3 Production of Active Component of Traditional Chinese Medicine by Gene Engineering

Plant gene engineering is to transfer the target genes into plants' genome by various methods in order to obtain steady inheritances and new characters. Plant gene engineering has entered into a rapid developing phase since the first batch of transgenic plants came out and carried out the field tests. According to incomplete statistics, there are almost 90 kinds of transgenic plant products approved for commercial application abroad, among which more than 50 kinds are in America and Canada. Great progress has been made in our country in the research and exploitation of plant gene engineering, and some researches have come up to an international advanced level. Since the 1990s, twentieth century, scholars both in home and abroad have shifted the focus of research to the area of character analysis, gene regulation, and the clone and separation of genes related to active component biosynthesis in medicinal plants. Gene regulation studies of active components in dozens of medicinal plants were carried out, including in *Catharanthus roseus* (changchunhua), *Papaver somniferum* (ying su), *Lithospermum erythrorhizon* (zicao), *Artemisia annua* (qinghao), *Taxus chinensis* (hongdoushan), *Berberis thunbergii* (Xiaobo), *Hyoscyamus niger* (tianxianzi or langdang), *Datura stramonium* (mantuoluo), *Atropa belladonna* (qian qie), *Camptotheca acuminata* (xi shu), *Nicotiana tabacum* (yan cao), *Rauwolfia serpentina* (shegen mu), etc. In these years, there were about 20 kinds of transgenic plants entering into the field experiment phase or environment release phase. Gene engineering of medicinal plant mainly includes crown gall nodule and hairy roots.

8.3.1 Production of Active Component of Traditional Chinese Medicine by Crown Gall Nodule Cultivation

8.3.1.1 Concept of Crown Gall Nodule

Crown gall nodule is a special phenotype which is induced after *A. tumefaciens* infects plant and T-DNA in Ti plasmid of *A. tumefaciens* being transformed into the nuclear genome of plant. Crown gall nodule can grow rapidly in hormone-free medium and can keep growing after the *A. tumefaciens is* being removed.

8.3.1.2 *A. tumefaciens*

A. tumefaciens is gram-negative, and it is present in soil in extremely high amounts. It has the capacity to infect wound of dicotyledonous plants, even monocotyledonous plants according to recent researches. T-DNA, which is a special segment of Ti plasmid in *A. tumefaciens*, can be randomly integrated into the genome of plants and then induce the tumor. According to the different types of opine, we can divide the *A. tumefaciens* into three groups, namely, nopaline type, octopine type, and agropine type.

8.3.1.3 Mechanism of the Infection by *A. tumefaciens*

The infection of *A. tumefaciens* and then resulting in gene transformation are the result of the interaction of plant cells and *A. tumefaciens*. The T-DNA carries genes for the biosynthesis of the plant hormones, auxin and cytokinins. By altering the hormone balance in the plant cell, the division of those cells cannot be controlled by the plant, and tumors form. There are two small (25 base pair) incomplete repeats in both ends of T-DNA, namely, border sequences, and at least one of them is needed for transformation and integration. The transformation of T-DNA also has to do with the gene product of vir region. Following an initial weak and reversible attachment, the bacteria synthesize cellulose fibrils that anchor them to the wounded plant cell. These fibrils also anchor the bacteria to each other, helping to form a microcolony.

8.3.1.4 Transformation Methods of Crown Gall Nodule in Medicinal Plants

The induction of crown gall nodule is a process in which the target gene (T-DNA) is integrated into and then expressed in plant cells. This process basically includes the purification and fermentation of *A. tumefaciens*, the coculture transformation of *A. tumefaciens* and plant explants, and the sterilization of plant materials.

1. Conversion
   Main steps: first, use a punch on sterile leaves to obtain 2–5 mm blades, also called leaf discs. Second, soak the leaf discs in *A. tumefaciens* suspensions which grow swiftly after overnight cultivation, and then coculture them on medium for 2–3 days. Third, when colonies around leaf discs have grown large enough to be visible to the naked eye, transfer the leaf discs to a medium containing bacteriostat and antibiotics to remove out *A. tumefaciens* and select transformed ones, and the transformed regeneration plants will be obtained after 3–4 weeks of culture.
2. Coinfection Through Inoculation to Whole Plants
   The infection process of this method mimics the natural process of *A. tumefaciens* infection. Create wounds artificially on whole plants, inoculate the *A. tumefaciens* to the wounds to infect plants in vivo, and then, the transformed cells of plants will be obtained. This method is also called in vivo transformation.
3. Transformation Through Coculture of Protoplasts and *A. tumefaciens*
   Coculture the *A. tumefaciens* and protoplasts on which the new cell walls are just regenerated in order to bring about the transformation of genetic materials between them. Therefore, this method can also be seemed as an in vitro transformation method which is carried out by the infection of *A. tumefaciens* to single cells under artificial conditions.

8.3.1.5 Application of Crown Gall Nodule in Researches of Plant Secondary Metabolites

Since the 1980s, twentieth century, with the development of plant biotechnology, studies on crown gall nodule have developed very quickly, and great progress has been achieved in production of secondary metabolites of medicinal plants by transgenic crown gall nodule. This technique has achieved more and more attention as a new way in the mass production of plant active components. For example, *A. tumefaciens* is used to infect *Asparagus officinalis* (shidiaobai) to produce crown galls with high content of auindine alkaloids, *Bidens sp.* (gui zhen cao) to produce lots of polyacetylenes compounds, *Catharanthus roseus* (changchunhua) to produce abundant alkaloids, and *Digitalis purpurea* (maodihuang) to produce cardiotonic steroids.

8.3.2 Production of Active Component of Traditional Chinese Medicine by Hairy Root Cultivation

8.3.2.1 Concept of Hairy Roots

Hairy root is a special phenotype which is induced after *A. rhizogenes* infects plants and T-DNA in Ri plasmid of *A. rhizogenes* being transformed into the nuclear genome of plants. Hairy root culture has become a new culture system these 10 years. Compared to plant cell suspension cultures, the advantage of hairy root cultures is that they retain differentiation while growing rapidly in hormone-free medium. The synthesis of many active components of traditional Chinese medicines has been related to the differentiation of organs. As a result, some active components may be contained little or even not existed in plant cell suspension cultures. In this situation, hairy root culture seems to be more important to traditional Chinese medicines as 1/3 of traditional Chinese medicines are the roots of plant. To date, there have been more than 40 kinds of hair root culture systems of medicinal plants, including *Catharanthus roseus* (changchunhua), *Nicotiana tabacum* (yan cao), *Lithospermum erythrorhizon* (zicao), *Panax ginseng* (renshen), *Datura stramonium* (mantuoluo), *Atropa belladonna* (dian qie), *Salvia miltiorrhiza* (danshen), *Astragalus membranaceus* (huang qi), *Glycyrrhiza uralensis* (gancao), *Artemisia annua* (qinghao), etc.

8.3.2.2 *A. rhizogenes*

The key point of hairy root culture is *A. rhizogenes*. It is a kind of gram-negative bacteria, and it has the capacity of infecting most of the dicotyledonous plants, some monocotyledonous plants, and very few gymnospermous plants to induce hairy roots in their wounds.

8.3.2.3 Mechanism of the Infection by *A. rhizogenes*

The Ri plasmid of *A. rhizogenes* is the cause of hairy roots induced in wound of infected plants. Ri plasmid is a large (about 250 kb) extrachromosomal plasmid in *A. rhizogenes*. It carries genes for the biosynthesis of opine. There are two regions relating to transformation in Ri plasmid, namely, T-DNA region (region of transferred DNA) and vir region (virulence region). Actually, genes in vir region are not transferred, but they are very important for the transfer of T-DNA. When *A. rhizogenes* infects plants, bacterial genes are turned on leading to the transfer of its T-DNA from its Ri plasmid into the plant through the wound. After integration and expression, in vitro or under natural conditions, the hairy root phenotype is observed.

Ri plasmid of *A. rhizogenes* can be divided into three types, namely, agropine type, mannopine type, and cucumopine type. Strains containing agropine-type Ri plasmid have more hosts and a stronger root-induction capacity. The hairy roots of plant formed after infected by *A. rhizogenes* can be cultured in vitro to regenerate plants. Hairy roots of many plants have the capacity to biosynthesize secondary metabolites when cultured in vitro, and their products' yields are usually higher than those of normal plants and suspension cells.

8.3.2.4 The Commonly Used Strains of *A. rhizogenes* for Hairy Root Induction

Commonly used strains of *A. rhizogenes* for hairy root induction include ATCC15834, ATCC39207, G58P GV3296, A4, NCPPB2659, R1500, R1601, LBA9402, TR105, etc., and all of them contain Ri plasmid. *A. rhizogenes* strains could be stored for several months on plating medium at 4 °C and for long period in low-temperature refrigerator at −70 °C.

8.3.2.5 Transformation Methods of Hairy Roots

Ri plasmid of *A. rhizogenes* is a natural vector for gene transfer, and it is often used for transformation to plants. The commonly used transformation methods are shown as follow.

1. Direct Inoculation to Whole Plants
   After disinfecting the plant seeds, put them on suitable medium to make them germinate and grow into sterile plantlets, and then take shoot tips for further culture. When sterile plants have grown for a certain period, cut out the shoot tips and leaves, and retain the stems and roots for further usage. Draw some wounds on the stems, inoculate *A. rhizogenes* with Ri plasmid into the wounds and incision on the top of stems, and then continue to culture the infected plants. After a period of cultivation, hairy roots will appear in the inoculation sites. This method is the most convenient.

2. Inoculation to Explants
   Coculture the plant leaves, stems, petioles, and other sterile explants with *A. rhizogenes* for 2–3 days, and then move the plant explants to selection medium containing antibiotics for further cultivation. After continuous subculture, *A. rhizogenes* will be killed, and calli will be induced on transformed cells, and then, hairy roots may be induced.
3. Coculture of Protoplasts and *A. rhizogenes*
   The protoplasts can be prepared from calli by conventional methods. Mix and then coculture the regeneration protoplasts and *A. rhizogenes*, and the transformation will be carried out. After the screening of transformed cells on selection medium containing antibiotics, transformed cells cloned will be obtained, and whole plants will finally be gotten on the differentiation medium.

8.3.2.6 Identification of Hairy Roots

Whether the induced hairy roots are genetically modified products still needs to be identified. Hairy roots can be identified from the morphological level. The roots with many root hairs and branches can grow in clusters vigorously toward all the direction on hormone-free medium. These phenotypic characteristics are different with those of the untransformed roots. We also can prove whether T-DNA of Ri plasmid has been transferred and integrated into nuclear genome of plant cells at tissue level by detecting the activity of GUS or NPT II assay in hairy roots. The detection of opine also can be used as a method for hairy root identification, which is because genes for the biosynthesis of opine will not be expressed in *A. rhizogenes* but in plant cells. The existence of opine can be regarded as an indicator of plant cell transformation. High-pressure paper electrophoresis method with alkaline solution of silver nitrate staining and sodium thiosulfate fixing can be used for the detection of opine.

8.3.2.7 Application of Hairy Roots in Researches of Plant Secondary Metabolites

The production of secondary metabolites of medicinal plant by plant tissue culture or cell engineering method has been limited by many factors in the process of industrialization, for example, the long production cycle, the large front load, and no or little target product in calli. The character of hairy roots is prominently shown in aspects such as rapid growth, short cycle, as well as the high and stable levels of target products; in addition, it is still possible to find new compounds from hairy root culture. Therefore, hairy roots have the potential value for industrial production.

Alkaloids, anthraquinone, naphthoquinone, polysaccharides, saponins, proteins, and other metabolites can be produced by hairy root cultures. For example, the growth rate of hairy roots which were induced from calli of *Panax ginseng* (renshen) and grew without hormones was twice of the calli, and the maximal content of

ginsenosides in these hairy roots reached 0.95% DW; vinblastine and vincristine, which have anticancer effect and present in *Catharanthus roseus* as bisindole alkaloids, were unable to be detected right now in cell cultures of *Catharanthus roseus* (changchunjian) but have been detected in its hairy root cultures. That is to say, hairy roots have the capacity of synthesizing some secondary metabolites which cannot be synthesized in suspension cells. The maximal contents of hyoscyamine and scopolamine in hairy roots culture of *Scopolia carniolicoides* (sailangdang) have reached as much as 0.35% DW and 0.10% DW, respectively; total saponin content in transformed hairy roots of *Gynostemma pentaphyllum* (jiaogulan) was about twice that of the natural roots; the content of trichosanthin produced by hairy roots of *Trichosanthes kirilowii* (Tianlou) was up to 8.16 mg/g WW. Taxol is a diterpenoid compound isolated from Taxus species. To produce taxol using hairy root cultures has been regarded as a most promising method for providing sources of taxol drugs, because the content of taxol in hairy roots was about 0.218 mg/g DW, and it was 1.3–8.0 times that of calli; in addition, taxol was also secreted from hairy roots to medium of 0.01–003 mg/L. Studies on hairy roots of *Astragalus membranaceus* (huang qi) indicated that contents of saponins, flavonoids, polysaccharides, amino acids, and other components were similar to those of natural materials, and the crude saponin content and soluble polysaccharide content even were slightly higher than those of natural materials. Studies on hairy roots of *Salvia miltiorrhiza* (danshen) proved that 7 kinds of tanshinone compounds did not only exist in hairy root organizations, and about 40% of them were secreted into the medium. Besides, hairy roots of *S. miltiorrhiza* also had the ability to produce water-soluble phenolic compounds, in which salvianolic acid A had a content of 2.7 times that of original plants.

Using *A. rhizogenes* to transform hairy roots provides a new way for the production of those plant secondary metabolites which are complex, difficult to be synthesized, low in content level, and poor in natural resources.

8.4 Case Study

Selection and Optimization of a High-Producing Tissue Culture of Panax ginseng C. A. Meyer

8.4.1 Material and Methods

Panax ginseng C. A. Meyer (5 years old) was supplied by Institute of Special Economic Animal and Plant Science, and the roots were selected as explants.

Selected roots were cleaned by water and then put into the drying oven for 30 min at 50 °C. After being surface sterilized by 75% aqueous EtOH for 5 min, the roots were rinsed repeatedly with sterile distilled water. The middle tissue of distilled roots were cut in sections of 5 mm and inoculated into Murashige and Skoog (MS) solid media containing 2 mg/L 2,4-dichlorophenoxyacetic acid (2,4-D), 0.5 mg/L

kinetin (KT), and 3% sucrose. Each petri dish (9 cm in diameter and 1.5 cm in height) was sealed with a wrap and cultured at (23 ± 2) °C in dark. After 4 weeks of culture, calli were induced, which proliferated in the same media and hormones as well as under the same conditions.

Adventitious roots of *P. ginseng* were isolated from calli and rooted in solid MS medium supplemented with 3% sucrose and 5 mg/L indole-3-butyric acid (IBA). Separated roots were transferred to liquid 3/4 strength MS medium supplemented with 4% sucrose and 5 mg/L IBA. The cultivation was held in 500-mL Erlenmeyer flasks containing 200 mL of the same liquid medium on a rotary shaker (120 rpm) for 4 weeks. Roots were subcultured every 4 weeks.

The solid-cultured hairy root was maintained on hormone-free MS medium solidified with agar at (25 ± 2) °C in darkness and was subcultured every 4 weeks. After cutting the solid-cultured hairy roots into little minor segments and then transferring them to liquid hormone-free MS medium, we got the liquid-cultured hairy root. It was maintained in 500-mL Erlenmeyer flasks containing 200-mL hormone-free liquid medium on a rotary shaker (120 rpm) for 30 days. Roots were subcultured every 4 weeks.

Ginseng calli, adventitious roots, solid-cultured hairy roots, and liquid-cultured hairy roots were collected after 4 weeks of culture. The growth index and ginsenoside production were analyzed, and the ginsenoside content in each culture was compared with that in field-cultivated ginseng roots.

Media of 0.5, 0.75, 1.0, 1.5, and 2.0 MS salt strength containing 4% sucrose were tested. The root growth index, ginsenoside production, and polysaccharide production were analyzed after 4 weeks of culture.

Media of 0.75 MS salt strength containing 2, 3, 4, 5, and 6% sucrose were tested. The root growth index, ginsenoside production, and polysaccharide production were analyzed after 4 weeks of culture.

Ammonium chloride (NH_4Cl) replaced ammonium nitrate (NH_4NO_3) when we prepared the media, and the concentration of the total nitrogen sources was maintained at 45 mM, which was basically the level of that in media of 0.75 MS salt strength. The ratios of ammonia/nitrate (NH^{4+}/NO^{3-} mM) varied at 0/45, 9/36, 15/30, 22.5/22.5, 36/9, and 45/0. The root growth index, ginsenoside production, and polysaccharide production were analyzed after 4 weeks of culture.

Media of 0.75 MS salt strength containing 0, 0.625, 1.25, 2.5, and 3.75 mM phosphate source were tested. The root growth index, ginsenoside production, and polysaccharide production were analyzed after 4 weeks of culture.

8.4.2 Results and Discussion

The production of particular ginsenosides as well as their sum in different tissue cultures differed considerably from that of native roots. The sum content of ginsenosides was found to be 0.82% in dry fibrous roots and only 0.27% in dry taproots of native ginseng. Generally speaking, ginsenoside contents in ginseng tissue cultures

were not so much as that in native ginseng. The sum contents of ginsenosides in callus culture and hairy root culture were much lower than that of native roots, while in adventitious root culture, the sum content of ginsenosides reached 0.67%, which basically reached the level of that in native fibrous roots but was much higher than that of native taproots. The results showed that ginseng adventitious root culture was a nice tissue culture system for the production of ginsenosides.

In terms of polysaccharides, all the four tissue cultures appeared with similar production capacity. It is obvious that the content of polysaccharides in native ginseng root was much higher than that in tissue cultures, and it may be ascribed to the reason that there were lots of starches in crude extract of polysaccharides from native ginseng roots while little in that from ginseng tissue cultures.

The data of growth ratio showed that root cultures in liquid media, including adventitious root culture and liquid-cultured hairy root culture, grew much faster than those cultures under solid-cultivation conditions. Hairy root culture growing in liquid medium proliferated very rapidly, and the average growth ratio reached 11.84 after 4 weeks of culture. However, in consideration of the low content of ginsenosides in liquid-cultured hairy roots, it regarded the adventitious root culture, which also proliferated in a high rate and contained much more ginsenosides, as the optimum ginseng tissue culture for the production of active components.

The effects of salt strength on adventitious root culture seem that higher salt strength in media was better for the production of polysaccharides than lower salt strength but conversely led to low growth ratio and ginsenoside contents. For example, the contents of polysaccharides in adventitious roots growing in 1.5–2.0 salt strength MS medium were higher than 2.5%, but the growth ratios and ginsenoside contents of them were much lower than those of adventitious roots growing in other levels of salt strength. 0.75 salt strength MS medium was the most favorable for ginsenoside production. However, the polysaccharide yields in different salt strengths did not show obvious differences.

The appropriate concentration of medium strength is one of the critical determinants in controlling root growth and active component production. The optimum concentration of minerals has been implicated in regulating the rate of organ elongation, morphology, and secondary metabolite accumulation. This was due to the interactions among the nutritional salts that enhance the availability of ions to the roots. MS medium, which is widely used in ginseng tissue culture, is often reported as a saline-rich medium. For this reason, we tested different salt strengths of MS medium from 0.5 to 2.0 salt strength, and the results showed that the maximum dry biomass and growth rate, as well as the sum of ginsenoside contents and ginsenoside yield, were obtained in 0.75 salt strength MS medium.

The increase in sucrose level from 2 to 6% remarkably enhanced the root growth and polysaccharide production. It seemed that 3% sucrose was the most benefit, and the sum content of ginsenosides reached 0.94% in dry mass. Further increase in sucrose concentration appeared to repress the accumulation of ginsenosides.

The most favorable sucrose concentration of medium was 4% according to the yield of ginsenosides and 6% according to the yield of polysaccharides, which reached 91.35 and 407.63 mg/L in medium, respectively.

The researchers have tried to determine the optimum sucrose level of the medium for the optimal growth and ginsenoside as well as polysaccharide production of ginseng adventitious roots. The increase in sucrose level from 2 to 6% remarkably enhanced the root growth and polysaccharide production. In terms of ginsenoside production, high levels of sucrose concentration appeared to repress the accumulation of ginsenosides, which can be attributed to the high osmotic pressure.

The largest amount of root growth ratio was achieved at a ratio of 9 mM ammonium to 36 mM nitrate, and root biomass decreased very rapidly when ammonium was used as the sole nitrogen source. The effects of ammonia/nitrate ratio on ginsenoside content and polysaccharide content appeared with similar trend while without significant differences between each other. The maximum ginsenoside and polysaccharide yield achieved at the ratio of 9 mM ammonium to 36 mM nitrate, which was 69.98 and 261.42 mg/L, respectively.

It is a general trend that a low NH^{4+} to NO^{3-} ratio is more favorable for the plant tissue and cell growth, and our study also supported this viewpoint. Nitrate rather than ammonium was necessary for root growth and ginsenoside production, as well as polysaccharide production.

The growth of roots and biosynthesis of ginsenosides and polysaccharides were severely inhibited in the case of phosphate starvation. The root growth ratio reached its peak at the concentration of 0.625 mM phosphate, and it decreased under higher phosphorus concentrations. Generally speaking, contents of ginsenosides in high levels of phosphate source were higher than those in low levels, for example, in a range of 1.25–3.75 mM compared to concentrations of 0 and 0.625 mM, and the maximum ginsenoside content achieved at the concentration of 1.25 mM phosphorus. In terms of polysaccharide accumulation, the content of polysaccharides seemed to be raised with the increase of phosphate concentration. According to the yields of ginsenosides and polysaccharides in adventitious roots under different levels of phosphate concentration, 1.25 mM phosphate was favorable for ginsenoside production, and the yield of polysaccharides was raised with the increase of phosphate concentration from 0 to 3.75 mM. Consequently, excessively lower phosphate concentration inhibited the roots' growth and active components' synthesis; high concentration of phosphate was favorable for the accumulation of polysaccharides while disadvantageous to growth.

8.4.3 Conclusion

In conclusion, results of present study demonstrate adventitious root culture as an optimum culture for the accumulation of biomass and active components compared to callus and hairy root culture. Medium containing 0.75 MS salt strength, 4% sucrose, 9 mM ammonia to 36 mM nitrate, and 1.25 mM phosphate was found suitable for ginsenoside production, while the optimum culture condition for polysaccharide accumulation seemed to be 0.75 salt strength, 6% sucrose, 9 mM ammonia to 36 mM nitrate, and 3.75 mM phosphate source in media. Appropriate conditions allowed for a maximum ginsenoside yield of up to 132.90 mg/L and polysaccharide yield of 407.63 mg/L to be obtained after 4 weeks of culture.

Chapter 9
Molecular Mechanism and Regulation on Biosynthesis of Active Ingredients of Medicinal Plants

Lu-qi Huang, Xue-yong Wang, and Chao-yi Ma

Abstract This chapter introduces the basic principles and methods for the studies on molecular mechanism and gene regulation of biosynthesis of active ingredients in medicinal plants, including biosynthetic pathway of active ingredients in medicinal plants; clone, isolation, and expression of genes related to biosynthesis of active ingredients in medicinal plants; and the manual regulation of the biosynthesis of the active ingredients of medicinal plants. In addition, the authors explain the problems in gene cloning, stability, and transportation and proposed the development prospects and the researching direction.

9.1 Overview on the Studies of Molecular Mechanism and Regulation on Biosynthesis of Active Ingredients of Medicinal Plants

The use of medicinal plants in prevention and therapy of diseases by human has been thousands of years. Today, about 25 % of the prescription drugs come from the medicinal plants. With the thought of returning to the natural becoming more and more popular, medicinal plants attract much attention. Domestic and foreign production and research units have conducted a lot of fruitful work and achieved encouraging outcomes in the classification, cultivation, ecology, analysis of active ingredients, action mechanism, and processing methods for medicinal plants. Many of

Lu-qi Huang (✉) • Chao-yi Ma
Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China
e-mail: huangluqi@263.net; machaoyi56@yahoo.com.cn

Xue-yong Wang
College of Chinese Materia Medica, Beijing university of Chinese Medicine, Beijing, China
e-mail: wangxydoctor@yahoo.com.cn

Lu-qi Huang (ed.), *Molecular Pharmacognosy*,
DOI 10.1007/978-94-007-4945-0_9,

the active ingredients in medicinal plants can be directly used as drugs, and even for those ingredients that cannot be used directly as drugs frequently become the raw materials of chemical semisynthesis and the targets for structural transformation, so the active ingredients of medicinal plants have always been the focus to be studied.

The chemical structures of the active ingredients of medicinal plants usually contain multiple chiral centers with relative complex structure; it frequently requires a very long period from the discovery to elucidate its chemical structure for many active ingredients. Take morphine as the example, which contains five asymmetric centers, its chemical structure was determined 146 years after the isolation of the compound, that is, 1952. The structural complexity of the active ingredient in medicinal plants reflects the magical mechanism of biosynthesis and prompts the extent of difficulty in relevant studies which might be beyond our imagination. In the late 1950s, the researchers fed the medicinal plants with radiolabeled precursor, which started the experimental biology research stage for the biosynthetic pathway of the active ingredients in medicinal plants [1]. Later, with the emergence and use of mass spectrometry, spectroscopy, NMR spectrometer, and other advanced equipments, in combination with the radiolabeled precursor feeding experiments, the studies on the biosynthetic pathway of the active ingredients in medicinal plants have become much easier. Since the 1970s, with the use of cultured cells to the studies on the biosynthetic pathway of pharmaceutical active ingredients, especially the enzyme involved in the biosynthetic pathway of the active ingredient, the studies on active ingredients in medicinal plants have deeply reached into the protein level. It is estimated that, from then to now, more than 80 enzymes involved in alkaloid biosynthesis have been successfully isolated and identified. In the late 1980s, scholars had shifted the focus to the research field of the clone, isolation, characterization, and gene regulation of the genes related to the biosynthesis of the active ingredients in medicinal plants. Gene regulation studies on the active ingredient have been conducted in tens of plants including periwinkle, poppy, lithospermum, sweet wormwood herb, yew, berberis, henbane, datura, belladonna, black poison, tobacco, Camptotheca acuminata Decne, Java devipepper rootspecies, etc. In the 1990s, encouraged by the advances in molecular biology research techniques and the research results, the studies on the gene regulations of the active ingredients in medicinal plants achieved further development. The biosynthetic pathways and gene regulation details of some active ingredients with extremely complex structures in medicinal plants have been gradually revealed, and primary success in manual control of some active ingredients has been acquired, indicating a promising development prospect in this research field.

9.2 Significance of Studies on Molecular Mechanism and Regulation of Biosynthesis of Active Ingredients in Medicinal Plants

The contents of alkaloids, saponins, flavonoids, glycosides, terpenes, and other active ingredients, which can be used in therapy, in medicinal plants are usually minimal. It is very difficult to significantly elevate the contents of the targeted active

ingredients through planting and cultivation; however, if we regulate the key genes for biosynthesis at molecular level to promote the expression, the contents of the target products could be elevated. Furthermore, through regulation at molecular level, the proportions of various ingredients could also be altered, and some toxic ingredients could also be reduced or even be removed completely, which is beyond the reach of ordinary cultivation practices. The metabolic engineering based on such regulation could produce new leading compounds for the development of new drugs and be used to enhance the ability of plants in resisting diseases, increase the content of beneficial chemical components in herbs, and change the taste and color of herbs.

Since the 1980s, the plant cell engineering, which produces active ingredients in medicinal plants through plant cell culture technology, has achieved great development. In the middle of 1990s, more than 400 kinds of plant have been studied in this field, and more than 600 kinds of secondary metabolites have been isolated from the cultured cells; however, among them, only a few of cells culture, such as cells from ginseng and lithospermum, reached the industrial production scale. Considering the reason, it is mainly because very low contents of active ingredients are contained in the cultured plant cells; even if occasionally high contents are found, unstable yields might also become the problem. Therefore, clarification of the mechanism of the biosynthesis of the active ingredients at molecular level and conducting regulation on the biosynthesis at gene level will increase the gene expression of key enzymes in the process of biosynthesis of active ingredients, improve, and stabilize the production of the active ingredients in suspension to harvest cell lines of high and stable production, thus completely solving the problems of low output and instability in active ingredients of suspension culture plant cells and truly realizing the goal of industrial production of active ingredient by the use of medicinal plant cell culture technology. The studies on the gene regulation of the active ingredients in medicinal plants may provide unprecedented possibilities in purposely regulating the content of active ingredients in medicinal plants and realizing the goal of mass production of the active ingredients of medicinal plants by cell cultivation, so it has a very important significance. In addition to technically ensuring the purpose of regulating the content of active ingredients in medicinal plants and realizing the goal of mass production of the active ingredients in medicinal plants by cell vitro cultivation, the genetic engineering researches on the active ingredients in medicinal plants will provide unprecedented possibilities in improving the quality of herbs and cultivation of new varieties. It is particularly worth mentioning that in addition to the regulation on key genes in biosynthesis of target active ingredients and the improvement of the content of the target products, the regulation at molecular lever can also be used to alter the proportions of the relevant ingredients in medicinal plant to reduce or remove some undesirable components, especially some toxic ingredients; to resolve the toxic effects of traditional Chinese medicine; and to ensure the quality of traditional Chinese medicine technically. Therefore, the studies on the molecular mechanism and gene regulation of biosynthesis of the active ingredients in medicinal plants have great significance both in theory and in production practices.

Because the synthesis of the active ingredients in medicinal plants is realized through a number of enzymatic reaction steps and completed in the specific differentiated cells, even with modern molecular biological measures, there are still many difficulties

in fully clarifying the biosynthetic mechanism. An important prerequisite for target regulation of the active ingredients in medicinal plants is the necessary understanding on the biosynthetic pathway of this active ingredient, in particular, the understanding of the key steps. In consideration of the not yet clear understanding on the molecular mechanism of biosynthesis for most of the active ingredients in medicinal plants, it would be much more difficult to purposely regulate the active ingredients in medicinal plants. Even so, in the past few years, considerable advances have been achieved in the studies on the molecular mechanism and the regulation for the active ingredient biosynthesis in medicinal plants. In subsequent sections, we will firstly introduce the basic principles and methods for the studies on the gene regulation of the active ingredients in medicinal plants and then, with terpenes as the example, describe the progress in gene regulation for this class of compounds.

9.3 Basic Principles and Methods for the Studies on Molecular Mechanism and Gene Regulation of Biosynthesis of Active Ingredients in Medicinal Plants

Active ingredients contained in traditional drug are secondary metabolites, and their biosynthetic pathways are very complicated, often involving several enzymes or dozens of enzymes during the reaction. Thus, to identify a key enzyme in the formation of a particular product has become one of the key steps in utilization of genetic engineering technology to produce active ingredients of medicinal plants, and in order to find out gene for the key enzyme, the selection of appropriate carrier and the promotion of the expression of exogenous genes have become the most important problems.

The study on the molecular mechanism and regulation of the biosynthesis of medicinal plants generally starts with the research on the available biosynthetic pathways of the active ingredients in medicinal plants; then with the research on the extraction, analysis, purification, and characterization of each enzyme involved in the enzymatic reactions in the biosynthesis of the active ingredients in medicinal plants; followed by cloning and separation of the corresponding genes and conducting research on their expression characteristics; and finally performing manual regulation on the relevant genes, especially the key enzymes, with transgenic technology. In the following sections, we will individually discuss the involved basic theory and basic method in the above-mentioned four aspects.

9.3.1 Biosynthetic Pathway of Active Ingredients in Medicinal Plants

The study on the biosynthetic pathway of active ingredients in medicinal plants is the prerequisite for carrying out the genetic regulation on the active ingredients. In 1970s, the magical effect of isotope was discovered. People begin to use isotope

to label the possible precursors or the identified precursors and then use the labeled precursors to feed the whole plant and study the biosynthetic pathway of the active ingredients in medicinal plants. Later, it was discovered that utilization of isotope-labeled possible or identified precursors to the cultured plant cells might be a better method. Although cultured cells biosynthetic capacity and overall plant biosynthetic capacity may not be identical, the culture of cells can be cultured in sterile conditions under manual control and proliferation; therefore, it has many advantages which the overall plant does not have. In one hand, the research findings with cultured cells might be used as the reference for the studies on the biosynthetic pathway with the whole plant. On the other hand, the research findings on the biosynthetic pathway of active ingredients in cultured plant cells could be directly used in the industrial production of targeted active ingredients by utilization of the cell culture.

In the following paragraph, we will introduce the basic steps of the biosynthetic pathway of active ingredients in medicinal plants.

In carrying out studies on the biosynthetic pathway of active ingredients in medicinal plants, the first step is to infer and label the intermediates and precursors. Firstly, the possible intermediates and initial precursors in the biosynthetic pathway of the product should be inferred according to the structure of the end product first, and then, the intermediates and initial precursors were chemically synthesized and labeled with ^{3}H, ^{13}C, ^{14}C, or other isotopes. Isotopically labeled compounds (such as the possible precursors) might have the following sources: the commercial product (if market available), enough relevant compounds isolated from plants or cultured cells and isotopically labeled, or manually synthesized (chemical synthesis) product and isotopically labeled.

The second step is to directly feed plants or the cultured cells with the isotope labeling (such as ^{3}H, ^{13}C, ^{14}C) compounds. Then, radioactive high-performance liquid chromatography (Radio-HPLC), spectroscopy, mass spectrometry, and nuclear magnetic resonance (NMR) technology are to be used to analyze the extracts from plants or plant cells. Through radioactive high-performance liquid chromatography (Radio-HPLC) analysis on the extracts from plants of cultured cells fed with isotope for some time, the retention time and quantities of all the peaks with radioisotope incorporation will be recorded; according to the known standards or through mass spectrometry and NMR identification of the collected peaks, the structure and properties of the compounds represented by radioisotope incorporation peak will be determined. Note that the peak must be a pure peak containing only one compound. In case of a pure peak incorporated with radioactive isotope corresponding to a compound with the same structure as the final product, the fed compound will unquestionably be the precursor of the final product because the radioactivity of the fed compounds can be incorporated into the end product.

After fed with isotopically labeled compounds, with the exception of the peaks of the final product, other peaks with radioisotope incorporation may be the intermediate products produced in the pathway of biosynthesis of the final product, or compounds outside the biosynthetic pathway of the final product. How can we determine the peaks with radioisotope incorporation corresponding to the intermediate products of the final product? Namely, how can we determine the peaks with radioisotope incorporation representing the compounds involved in the biosynthesis

of the targeted final product? The methods for these determination processes are same with previously described methods, that is, isolation of adequate amount of compounds represented by these peaks with radioisotope, labeling the compound with radioisotope or labeling the manually synthesized compound with isotope, and then feeding relevant plants or cultured cells with the labeled compound. If the radioactivity of the fed compound can be effectively incorporated into the final product, then the fed compound is the intermediate product of the targeted final product.

Similarly, the isotope-labeled exact precursor of target product can also be used to feed cultured plant cells, in combination with modern analytic techniques, such as spectroscopy, mass spectrometry, and nuclear magnetic resonance, especially with radioactive high-performance liquid chromatography (Radio-HPLC), to carry out research in this area. After analysis with radioactive high-performance liquid chromatography (Radio-HPLC), we can see the peak retention time and the number of radioisotope incorporation and also determine the structure and properties of the compound at the peak with radioisotope incorporation according to the known standards or through identification with mass spectrometry and NMR on the collection of the peaks. One of the peaks with radioisotope incorporation has its corresponding compounds with the same structure as the end product. Similarly, after feeding with the isotope-labeled specific precursor, with the exception of the final product peaks, other peaks with radioisotope incorporation may be the intermediate products produced in the pathway of biosynthesis of the final product, or compounds outside the biosynthetic pathway of the final product. Next, similarly, it should be determined whether the compounds corresponding to the peaks with radioisotope incorporation between the precursor and final products are the intermediate product, that is, whether it participates in the biosynthesis of the end products. With the same method, that is, isolation of adequate amount of compounds represented by these peaks with radioisotope, labeling the compound with radioisotope, or labeling the manually synthesized compound with isotope, the relevant plants or cultured cells will be fed with the labeled compound. If the radioactivity of the fed compound can be effectively incorporated into the final product, then the compound used for feeding is certainly the intermediate product of the targeted final product.

After determining the chemical structure of the final product and some intermediates, there are many difficulties in acquiring the complete profile of the biosynthetic pathway of an active ingredient in medicinal plant (of course, the final product acquired after fed with specific precursor must be active ingredient). These difficulties are mainly reflected in that the sites of various intermediates in the biosynthetic pathway are relatively difficult to determine. Generally, it is believed that the compound with stronger radioactivity occurs after the plant or cultured cells fed with the labeled intermediate product is the next intermediate product after the intermediate product in the biosynthetic pathway. Though the principle for the studies on the biosynthetic pathway of the active ingredients in medicinal plants is simple, specific training is required on the isotope labeling technique and the identification technique of compound structure. Further detailed prescription on relevant technique can be found in the relevant books and articles.

Many of the active ingredients in medicinal plants are secondary metabolites. The biosynthetic pathways of the active ingredients in higher medicinal plants are still unclear, but the primary profile of the biosynthetic pathways of various types of secondary metabolites has been described in many published articles. According to the biogenesis pathway, the active ingredients of medicinal plants are generally divided into three classes, that is, terpenes, aromatic compounds, and alkaloids. The active ingredients of medicinal plants can be subdivided into seven categories, including phenylpropanoids, quinones, flavonoids, tannin, terpenes, sterols and their glycosides, alkaloids, etc. The main biosynthetic pathways include (1) acetate–malonate pathway, fatty acids, phenolic compounds, and anthraquinone produced in this way; (2) mevalonate pathway, terpenoids produced in this way; (3) cinnamic acid pathway, phenylpropanoids, coumarins, lignin, and flavonoids produced in this way; (4) amino acid pathway, alkaloid composition produced in this way; (5) the combined pathway of above pathways, and natural product with more complex structure produced in this way. Among the above-mentioned compounds, there are two biosynthesis pathways for isopentenyl pyrophosphate (IPP), the important intermediate product in the biosynthetic pathway of terpenoid. The first is the mevalonate pathway (MVA pathway) in cytoplasm, which exists in almost all organisms in nature. The second is a newly discovered pathway: 1-deoxyxylulose-5-phosphate pathway, DXP, which exists in plastids. Currently, it is considered that the enzymes in MVA pathway, which are associated with the synthesis of triterpenes, carotenoids, and steroidal in higher plants, exist in cytoplasm; however, the synthesis of monoterpenes and diterpenes occurs in plasmids through DXP pathway; for the synthesis of sesquiterpene, there is no clear conclusion yet. Phosphoenolpyruvate, produced in 4-phosphate erythritol glycolytic pathway, will form 7-phosphate ketoheptose by condensation; then through a series of transformation, it will enter into shikimic acid pathway; subsequently, aromatic amino acid will be formed through many branches; and finally, various aromatic active ingredients will be produced. Alkaloids are produced by the transformation of amino acids produced in Krebs cycle pathway, and indole alkaloids can be synthesized by the transformation of tryptophan produced in shikimic acid pathway.

9.3.2 Clone, Isolation, and Expression of Genes Related to Biosynthesis of Active Ingredients in Medicinal Plants

Studies on the characteristics of clone, isolation, and expression of the genes related to biosynthesis of the active ingredients in medicinal plants are the prerequisite of the purposely artificial regulation to the relevant targeted ingredient, so it has very important significance. We will introduce the research methods for three aspects below.

9.3.2.1 Clone Method for the Relevant Genes in Biosynthesis of Active Ingredients in Medicinal Plants

The so-called gene cloning is to artificial link in vitro the gene from different organisms with the vector DNA, which can be automatically replicated, to form a new recombinant DNA, and then deliver it into the recipient cells to perform unlimited asexual reproduction process. As same as cloning method for other genes in plants, the cloning methods of the relevant gene in biosynthesis of active ingredients in medicinal plants include the following two sections:

cDNA Gene Cloning

Eukaryotic genomic DNA is very large; its complexity is about 100 times the protein and mRNA and contains a lot of repetitive sequences. Therefore, it is difficult to isolate the target gene fragment using either electrophoretic separation or hybridization method. This is one of the major difficulties in the direct cloning of the target gene with the chromosomal DNA as the starting material. This problem can be partly resolved through the cloning of cDNA generated from mRNA. Although higher organisms generally have about 10^5 kinds of different genes, in a specific phase and in single cell or individual, expression occurs by only about 15 % genes, producing about 15,000 kinds of different mRNA. The basic processes of cDNA clone are, firstly, through a series of enzyme catalysis to form double-stranded complementary DNA (cDNA) colony from the total poly (A) mRNA under the effect of reverse transcriptase and then transfect the host *E. coli* cells after connected into an appropriate carrier. By this way, a cDNA gene library containing all the coding sequence of the genes will be established [2]. In practical work, guanidine thiocyanate–phenol–chloroform extraction method can be used to extract RNA; then the mRNA will be isolated through the poly (dT) cellulose and chromatography column; with poly (dT) as primer, double-stranded cDNA will be synthesized by reverse transcription process; after end-filling of the double-stranded cDNA and connecting it with an EcoRI connector, the resulted product will be cloned into the λ-ZAP II of EcoRI site of the vector. A cDNA library containing recon is constructed by in vitro packaging of the constructed vector and then using the packaging product infecting the *E. coli*×L 1-blue MRF host strain. cDNA library has many advantages and uses: first, the screening of the cDNA library is a simple thing. Proper choice of the source of mRNA will have the clones of a specific sequence raised to a relative high proportion in the constructed cDNA gene library, which would decrease the working load for screening of a specific gene clone. Second, each cDNA clone will contain a kind of mRNA sequence, so the probability is relatively low for positive choice during choosing the choice, and the positive clone chosen in positive hybridization signal generally contains target gene sequence. In addition, cDNA cloning can also be used in the structure determination of gene sequences [2, 3].

Genomic DNA Cloning

cDNA cloning method has many advantages, but there are also some limitations in practical use. cDNA clones can only reflect the molecular structure of the mRNA, but not including the interval of genomic DNA sequence. For exogenous eukaryotic gene expression in eukaryotic cells, genomic clone is more effective than cDNA clone. In cDNA library, the distribution of different clones reflects the mRNA substate: cDNA clones corresponding to the high-abundance mRNA have higher proportion and are easier for separation, while the cDNA clones corresponding to low-abundance mRNA account for relative lower ratio and are harder for separation. In addition, cDNA clone cannot be used for the clone of non-transcribed sequences in genomic DNA. Therefore, cDNA cloning is unable to meet the requirements for study on the structure and function of the regulatory sequences outside the coding region [2].

For genomic DNA cloning, the first step is to cut the intact plant genomic DNA into a group of DNA fragments of various lengths with restriction enzymes and ultrasound or other physical methods, and then these fragments of different lengths will be linked into appropriate DNA vector. After the recombinant DNAs are transferred into competent cells or transfect host cells; they can be amplified to great quantities, and thus, we will harvest the clone group consisting of all of the DNA fragments of different sizes in this medicinal plant genome DNA [4].

An ideal genomic DNA clone library should fulfill the following several conditions: (a) storage of the DNA sequences composed of genome in entire chromosomes in the organism; (b) total quantity of clones is not too large, thus reducing the screening workload; (c) the size of the cloned fragment must be large enough to contain a complete gene; (d) in order to facilitate restriction endonuclease analysis and adapt to the capacity of carrier, the cloned fragment cannot be too large; (e) the library should be easy to build from a small amount of starting material (exogenous DNA of high molecular weight) and easy to filter the target fragment; (f) partially overlapping sequence should be existed between different cloned fragments, thus taking advantage of the chromosome-walking technology to understand the linking sequence of different cloned fragments; (g) the cloned fragment should be easy to be cut from the vector and not carry any vector sequence; and (h) the library should be able to be amplified without any loss of the clones and also stored for many years without significant deterioration. At present, only phage λ and cosmid can meet or partially meet these requirements. λ phage can accommodate an exogenous fragment with the size up to 20–25 kb and is a simple and efficient tool for library construction. Cosmid has stronger cloning capacity (nearly 45 kb), but it is more difficult to be manipulated than λ phage. The basic steps for gene library construction with λ phage as the vector are the following: (a) using the restriction enzymes to cut the vector DNA, thus forming the carrier to form sticky ends for linking exogenous DNA; (b) annealing the cos region at both arms of the vector and removing the central part unnecessary for the proliferation of λ phage; (c) both annealed arms linking to the exogenous DNA fragments, forming a concatenate; (d) *in vitro* packaging of concatenated DNA molecules; and (e) recombinant phage infecting *E. coli*, forming a large number of plaques [5].

The average distance between the two RFLP markers (restriction length polymorphism markers) is approximately 400 kb. If you use conventional λ clone (containing the exogenous DNA fragment of about 20 kb) or cosmid clone (containing about 40 kb of exogenous DNA fragment) to build the overlapping regions (contigs) containing two RFLP markers, the working load is hard to imagine. Yeast artificial chromosome clone (YAC), developed by Burke et al., can satisfy this demand. During YAC clone, exogenous plant DNA is attached to the site between the arms of the YAC vector. The right and left arms of YAC vector contain the necessary genes for the stabilization and normal replication and separation of the whole clones in yeast cells, and they also contain some genes which can be used as selective marker (e.g., nutritional needs); thus, the YAC clone in yeast cells can survive, replicate, and spread to the offspring as the chromosomes of yeast itself. YAC clone can accommodate exogenous DNA fragments with the size up to 1 Mb, so it is suitable to be used as chromosome-walking bridge to build the overlapping region of chromosome. YAC clone itself can be subcloned into λ vectors and cosmid vectors, which provides material for cloning genes and DNA sequencing [6].

Compared to the construction of other types of gene bank (such as λ gene bank or cosmid gene bank), the construction of a YAC gene pool needs preparing DNA fragments of high molecular weight at first and then transferring the prepared YAC clone to the yeast. The basic processes of construction of YAC bank of medicinal plant genome with yeast artificial chromosomes include dual enzyme digestion of YAC vector $pYAC_4$ with the selected endonuclease EcoRI and BamHI, removing the fragments between the two telomeric sequences and resulting in two arm molecules with EcoRI sticky ends. In this way, each arm molecule will carry a fragment of telomere sequence and a selective marker (such as amp^r). At the same time, a restriction endonuclease with few cleavage sites (such as EcoRI) should be selected to locally digest the genomic DNA of high molecular weight in medicinal plants. DNA fragments with molecular weight less than 200 kb in digestion reactants can be removed by pulse alternative field gel electrophoresis or density gradient centrifugation, and the remaining high molecular weight DNA fragments will be harvested and purified and then linked to the YAC arm. Or, the gel block containing DNA fragments within the desired molecular weight range is cut out directly from the gel electrophoresis gel, and then DNA is extracted from it and linked to the YAC arm. After another pulse-field gel electrophoresis, the YAC vector containing the insert sequence will be separated and transferred into the yeast cells with the cell walls removed (i.e., yeast spheroplasts). The selective genes existed in both arms can be applied for screening yeast clone containing both YAC arms [2].

9.3.2.2 Separation of Genes Related to Biosynthesis of Active Ingredients in Medicinal Plants

After completion of the gene cloning related to the biosynthesis of active ingredients in medicinal plants, we will separate the cloned genes from the cDNA library and genomic library. According to the results of current studies, the most commonly

used separation methods for the relevant genes to the biosynthesis of active ingredients in medicinal plants are summarized below:

1. Initial separation of protein sequence
   This is the most "original" method for gene isolation. The method uses the separated and purified protein. First, the amino acid sequence of a polypeptide fragment in the protein is obtained, and then the corresponding nucleotide sequence is designed and synthesized. Then, using above nucleotide sequence as the molecular probe to hybridize the chromosomal gene or cDNA library, the corresponding gene will be isolated. Amino acid sequence can also be used to design degenerate nucleotide primers, and then PCR technique is used to amplify a nucleotide fragment, using it as screening gene or the probe of cDNA library. So far, there have been some gene sequences isolated through initial separation of protein sequence [6]. However, due to the low protein yield of the relevant genes to biosynthesis of the active ingredients in medicinal plants, the separation and purification of these proteins become very difficult, so this separation method is very poor in availability for the separation of genes in medicinal plants.
2. Homologous gene sequence method
   According to the homologous conservation of gene sequences, the molecular probe prepared by the gene isolated form a kind of organism can be used to separate the homologous genes in medicinal plants. For example, P450 commonly exists in a variety of eukaryotic organisms; using the P450 genes, cDNAs, or their conserved region to prepare molecular probe and then hybridizing the cDNA library or gene library, some P450 genes in plants have been isolated successfully. Homologous sequence hybridization method can be used to separate the homologue genes with the known genes in another medicinal plant.

   Additionally, the conservative region of gene can be used to design a PCR primer of dozens of nucleotides, and then the PCR technique is used to amplify the corresponding fragment in chromosomal DNA or cDNA in plant. Then with the obtained PCR product as the molecular probe to hybridize and screen the gene bank, the clones of the gene to be isolated will be harvested. Compared with homologous hybridization, this technology not only reduces the length required for homologous sequences but has also accelerated the process of gene isolation. This sequence homologous gene PCR separation technology is widely used. Many P450 genes involved in the hydroxylation of active ingredients in medicinal plants are isolated by this method [6].

   For genes that are expressed under the induction by fungi or jasmine acids and involved in the biosynthesis of the active ingredient in medicinal plants, we can design the PCR primer through the conserved region of gene sequences and then amplify the cDNA fragment related to the medicinal plant with RT-PCR (reverse transcriptase polymerase chain reaction). Then with the obtained PCR product as the molecular probe to hybridize and screen the gene bank, the clones of the gene to be isolated will be harvested.
3. Protein functional complementation cloning method
   Some genes involving the most basic biological functions are very conservative in the process of biological evolution. Such genes are commonly called "house-

keeping genes." Some of the proteins produced by housekeeping genes have the same functions in different species, which can be exchanged among them. According to the principle of proteins with the same function, gene isolation method of protein functional complementation was developed in *E. coli* or yeast. For example, E J Corey et al. had the cDNA of *Arabidopsis thaliana* expressed in yeast mutant lacking of lanosterol synthetase which confirmed the full-length cDNA clone encoding *Arabidopsis thaliana* cycloartenol synthetase (with the similar function to lanosterol synthetase) correct [7].

4. Hybridization separation by differential expression
   Many genes have unique expression patterns. Different gene expression level would be found under different growth conditions, in different tissue or at developmental stage, and such difference in expression has been utilized to develop the separation method by differential hybridization. With this separation method, assume two different sources, A and B; firstly, construct cDNA gene bank by using mRNA of source A, and then prepare molecular probes by using cDNA of source A and source B mRNA, respectively; in the next step, above two probes are used to hybridize the cDNA gene bank of source A, respectively. To screen the differential clones from the two source probes, and finally, with the cloned cDNA as probe, compare the expression pattern of the hybridized mRNA with the expression pattern of the targeted gene. If the conformity in expression pattern can be confirmed, the cDNA is the cDNA of the gene to be cloned [6]. P450 gene CYP72A1 involved in the synthesis of vinca alkaloids is separated by this way [8]. The practice shows that many limitations exist in the separation of target gene with hybridization separation by differential expression: First, the sensitivity of differential hybridization is relatively low, especially for mRNA of low abundance. Because the hybridization probes used in differential hybridization are the cDNA group reverse-transcribed from mRNA and the proportion of the probes exactly complementary to the nucleotide sequence of target gene is very low, it is very difficult to detect the cDNA clones from mRNA of low abundance. Secondly, in differential hybridization, lots of hybridization filters have to be screened, and lots of plaques or cloned fragments have to be identified, so this is a time-consuming and labored work. Moreover, the membrane between the two sets of parallel transferring the retained quantity of DNA is frequently different, so the acquired hybridization signal intensities will not be consistent; you need to repeat dot blot hybridization for further positive clone identification. So, poor reproducibility is another defect of differential hybridization screening [2].
5. mRNA differential display separation
   mRNA differential display was a new molecular biology technique established in 1992 by Liang Pardee, which intended to search for genes of specific expression by comparing mRNA in different cells. The basic strategy of mRNA differential display is as follows: A specific small fragment of mRNA is to be amplified through reverse transcription and PCR amplification, and then the DNA sequence analysis gel (6–8 % polyacrylamide gel) is used to conduct synchronization separation, showing the amplified product for comparison. Firstly, using the 5′-T-(n) MN-3′ (3′ fixed primer, in which $n = 10–20$, M- = dA/dC/dG, N = dA/dC/dT/dG)

as the primer, the mRNA to be compared is reversely transcribed to get the single-stranded cDNA. After reverse transcription, the single-stranded cDNA will be used as a template, and then PCR process will be immediately performed with above-mentioned fixed primer as 3′ primer and random primer of 10–13 bases as the 5′ primer. Using the different combinations of 12 3′ primers and 25 5′ primers, 15,000 different genes can be analyzed by 95 % confidence interval, which basically include all the genes expressed in a single cell [3].

The basic procedures of mRNA differential display include the following: the first step, separation of total mRNA from the material to be compared and synthesizing the first strand of cDNA with the selected 3′ fixed primer by reverse transcription; the second step, performing PCR amplification with the pairs of random primer consisting of 5′ random primer and 3′ end fixed primer pairs of primers under the conditions of adding radiolabeled dNTP; and the third step, electrophoresis separation through adding the amplified sample into the denatured DNA sequencing gel. After X-ray film exposure, the DNA amplification bands from the differential expressed mRNA in the material to be compared will be displayed. The fourth step is to cut the DNA band related to differential expression out of the sequencing gel and recover the amplified DNA bands. The fifth step, due to the small amount of DNA in the gel band, it cannot be directly used for cloning, so it is necessary to perform the second PCR amplification with the same primers. Only after achieving a certain amount of DNA, recombinant clone has been completed. The sixth step is to perform the Southern and Northern hybridization process of the specific DNA clones, respectively, with the genomic DNA and total mRNA, to verify the specificity of its expression, and then with this fragment as a probe, screen full-length cDNA or genomic clones from cDNA library or gene library [2].

mRNA differential display technology has the features of high sensitivity, simplicity, intuition, and being able to compare multiple samples simultaneously, so it becomes very popular among the researchers of molecular biology. However, after several years of practice, the mRNA differential display technology also exposes some obvious shortcomings, such as high false positive ratio and small fragments. Recently, through the efforts of the scholars, these two shortcomings have basically been overcome.

6. Separation by expression library
cDNA is cloned in the expression vector and then introduced into *E. coli* host cells, thus building a cDNA expression library. Prepare antibodies from the purified protein, then screen the gene clone from gene expression cDNA library with the prepared antibodies.

As nucleic acid hybridization screening, in the expression library screening, the first step is also to perform in situ reproduction of the plaques or bacterial colonies to a nitrocellulose membrane. When cultured phage develops into a plaque, each plaque contains a lot of the fusion protein produced by the phage, and in the reproduction process, they will be transferred to the membrane. After properly treating the filter membrane, the transferred proteins from each plaque will be exposed, and then perform the incubation with the antibodies containing the target protein (the first antibody). After an appropriate time, rinse the membrane

to remove the unbound antibodies, then perform incubation again after adding second antibody. The second antibody can be radiolabeled, be coupled with biotin, or be combined with a certain enzyme (such as alkaline phosphatase). Through the combination of the first antibody and the second antibody, based on observable markers (such as a radioactively labeled or color), we can determine the site of the positive clone which can be specifically identified by the first antibody. Finally, compare the film (or the X-ray negatives) with the original plate control, pick positive clone, and perform sequencing and identification of the isolated recombinant phage DNA [2]. Lois et.al used the method to separate the casbene synthase gene successfully from the expression library [9].

7. Database cloning

   Database cloning is a very useful method of gene isolation. It is the use of expression sequence tags (ESTs) to isolate the gene, also known as expression sequence tags method. Expressed sequence tag is actually a part of the exon.

   The basic steps of database cloning include sequencing of cDNA clones by using DNA sequent or manual sequenator, acquiring a full range of cDNA (i.e., full length of cDNA). Then, the sequence of related cDNA cloning will be compared with that of related gene sequences in gene bank for the homology, providing an initial determination on the nature of the gene. Then, the relevant full-length cDNA will be constructed into a suitable expression vector, and then protein expression will be achieved in a heterologous expression system, such as *E. coli*, yeast, or insect. Finally, in combination with the relevant substrate, we can identify the functions of the protein expressed according to the relevant gene.

8. Gene position cloning

   Gene position cloning is also called map-based cloning, which is an efficient way to separate the genes unknown of the coding products.

   Gene position cloning requires yeast artificial chromosome (YAS) as the carriers to build YAC bank containing DNA of large fragments and DNA probes closely linked with the same target genes, which should be in the genetic map distance of hundreds of kb (kilobase). In addition, gene position cloning needs to build the high-density RFLP (DNA restricted fragment length polymorphisms) or RAPD (random amplified polymorphic DNA) molecular markers map. The so-called RFLP molecular mark means the length variety of DNA fragments produced by cutting DNA molecules in the utilization of specific restriction endonucleases enzymes. DNA restriction fragments are separated through agarose gel electrophoresis (AGE) and stained by ethidium bromide (EB). After these steps, in the ultraviolet light, the visual specific electrophoresis bands translated from RFLP can be seen.

   The elementary steps of gene position cloning are the following: firstly, positioning the target genes in the chromosomes and determining a pair of closely linked RFLP or RAPD molecular makers in both sides of the target genes; secondly, using the most closely linked pair of two-side molecular markers as the probes and through the genome walking technique, the specific genomic fragments located between these two molecular markers and containing the target gene are cloned and separated; and finally, by the techniques of sequencing, genetically modified complementary experiments, site-directed mutagenesis, and gene silencing (such as RNA interference, RNAi), the functions of the gene are determined.

The starting point of chromosome-walking technique is the identified molecular marker (RFLP or the known gene cloning), which is adjacent to the separating target gene as closely as possible. Take RFLP marker study as example: firstly, draw up the restriction map marked and cloned by the starting RFLP, and then the restriction fragment which is closest to the target gene is subcloned. This restriction fragment is marked as the molecular hybridization probe using the radioisotope, and the new cloning (one-step cloning) is chosen from the genome banks for overlapping the starting point cloning. The above-illustrated steps are repeated to get a second new cloning (two-step cloning). And the like, the new clonings more closely linked to the target gene will be obtained, and the chromosome walking is continued until the cloning of the target gene is gotten. Because this method takes slow steps to get close to the target gene through the overlapped sequences just as walking along the chromosome, it is called chromosome walking. Although the mechanism of this method is easy, the practice is complicated [2].

9. Gene chip

Gene chip is an advanced technology developed from 1980s, which is also called DNA chip or DNA microarray. Through in situ synthesis, point contact, or ink-jet method, a large amount (the reticular density per square centimeter usually reaching to thousands of, even ten thousands of molecules) of DNA molecules (target) is fixed on buttress such as membrane, glass, or silicon slice. Because the matrix is similar to the computer's chip, it is called gene chip or DNA chip.

After the hybridization of gene chip and marked molecule sample, all the tested genes could be analyzed and examined qualitatively and quantitatively by testing the hybridization signal intensity of every probe molecule. Therefore, transcriptome will be studied in a larger scale, even in the level of genomes, that is, the models and laws of genetic expressions, to reveal the nature of the life. The chip with cDNA cloning fixed in the support holders (membrane, glass, silicon slice, etc.) is called cDNA chip, and the related technique is called cDNA chip technique. The elementary steps of cDNA chip analysis are illustrated in the chapters of methods.

10. Proteomics technology

Although tens of thousands of gene expressions can be tested by functional genome technologies—DNA microarray analysis and serial analysis of gene expression (SAGE)—the real levels of gene expressions and proteins are often not correlated due to the transportation, degeneration, translation, regulation, and procession after translation of mRNA. Thus, we cannot get the complete information. One of the ways is to study the proteins which are the products of the related gene expressions. Some species, individuals, organs, tissues, and even all the proteins of the cells are studied together. Corresponding to the genome, it is called proteome.

Proteome was firstly proposed by Wilkins and Williams from Australia in the meeting of dimensional electrophoresis held in Siena, Italy, in the year 1994. The subject aimed to the study of proteome is proteomics, which is a new one in the study field. The main method for proteomics studies is using two-dimensional polyacrylamide gel electrophoresis (2D PAGE) to separate the complicated

components of proteins, scanning tens of thousands of protein spots tested through silver staining, fluorescent staining, and autoradiography and then obtaining and analyzing the data from the images using the special computer software. In addition, the protein spots recollected from the gel need to be identified through the analysis of amino acid composition, mild protein sequencing, and mass-spectrometric technique to get the nature of the proteins, the changes of expressions, and the procession after the translation.

The proteomics technique includes the elementary steps listed below:

1. Extraction of the proteins: The preparation of protein samples is carried out through the methods of acetone precipitation, trichloroacetic acid (TCA)–acetone precipitation, etc. Sometimes, organelles (such as microsomes) are needed to be separated and purified to enrich the organelle proteins.
2. Dimensional electrophoresis of the proteins: Using the system of IPGphor for isoelectric focusing (IEF) electrophoresis, after balancing, conducting vertical SDS-PAGE electrophoresis, and then the gel was stained by silver staining, fluorescent staining, and Coomassie brilliant blue (CBB) staining.
3. Scanning of the protein spots and image analysis: stained gel is scanned by light-density spectrometer to digitalize the images, and the data is analyzed by the softwares pf PDQUEST, LIPS, HEMES, and GEMINI. The differential proteins are screened out initially.
4. Identifying the candidate proteins by mass spectrum: Using the peptide mass fingerprint (PMF) to identify the proteins. PMF technique uses enzymes (commonly used pancreatin) to process the proteins separated from dimensional electrophoresis in gel or membrane for enzymolysis in C-ends of arginine or lysine. Using MALDI-TOF mass spectrum for testing, look up the proteins database to identify the proteins through comparing with the masses of peptides in the database.
5. Mild sequencing of the proteins: The proteins separated through the gel are blotted in PVDF membrane, and after being stained and cut, the proteins are sequenced by the protein sequencer. N-end sequencing method is used to identify the proteins.

Besides the methods listed above, other methods to separate the genes are developed on the basis of the interactions of large biological molecules, such as yeast one-hybrid system, yeast two-hybrid system, mammalian two-hybrid system, bacteria two-hybrid system, yeast three-hybrid system, reversal two-hybrid system, and reversal one-hybrid system.

9.3.2.3 The Study of the Gene Expressions of Medicinal Plants Active Ingredients Biosynthesis

After finishing the cloning and separating the genes of medicinal plants active biosynthesis, the expression characteristics of the genes need to be studied, which include the expression characteristics of the plants or plants cells in their own systems and in the heterologous systems. These two kinds of expression characteristics involve in the studies of the transcription and translation of the genes. The commonly

used method for studying the transcription of the genes is Northern blot, and the method for studying the translation of the genes is Western blot.

Northern blot means the hybridization of RNA and DNA molecules. Its elementary steps include blotting RNA in the degenerated gels on the nitrocellulose membrane or on the filter paper with special processions and using DNA probes for blotting. For the intuitive studies, in situ Northern blot is used sometimes to reveal the transcriptions in the scales of time and space.

Western blot is the hybridization of the proteins and the antibodies, the mechanism of which is as follows: the proteins in the gel electrophoresis are blotted on the nitrocellulose membrane, on the nylon membrane, or on the filter paper with special processions, and the antibodies marked with isotopes and the blotted proteins interact. If the proteins have antigens, the antibodies marked with isotopes will bond to the proteins, and then the autoradiography will show the result [4].

In a suitable expression system, the related genes to the biosynthesis of the active ingredients of the medicinal plants are expressed to study the expression characteristics and the biological functions, which are the main objects for studies of its molecule mechanism. At present, the expression systems include bacteria, fungus, and insects.

The expressions of the related genes to the biosynthesis of the active ingredients of the medicinal plants in bacteria are usually developed in *Escherichia coli* DN5α as hosts. The choice of the host bacteria is essential to the expression of eukaryotic genes because the exogenous proteins expressed by the eukaryotic genes in the bacteria are not stable so that they are easily degenerated by the protease in the bacteria. In addition, it also requires the suitable expression carriers. The carriers have the following features: (1) having a strong pronucleus promoter and the regulation sequence in both sides, (2) SD sequence located in the upstream of reading frame has suitable distance from starting codon AUC, (3) the right reading frame exists between eukaryotic genes and the promoter, and (4) a transcription terminal area not dependent on ρ factor should be added in the downstream of the eukaryotic genes. pBV220 is such a primary expression carrier. After having the suitable host bacteria and expression carrier, to get more efficient gene expressions, the space between SD sequence and AUG should be adjusted, or some bases are changed in the method of point mutations for eliminating the secondary structure of mRNA and improving the completeness of translation. The expression levels are raised by inserting repeatability palindromes (RFP) sequences and DNA fragments with reversal repeated sequences. Sometimes, through the induction expression, the growth of the bacteria and the transcriptional activities are controlled by humans [3].

The way that the eukaryotic genes of the medicinal plants are expressed in pronucleus host cells is cheap and easy, which is an efficient way at present. However, the eukaryotic proteins expressed by the prokaryotes have the folding mistakes, and the low efficiency of folding or after the translation and the formation of disulfide bonds, glycosylation, and phosphorylation which cannot be performed in the prokaryotes cells can result in the low biological activities of the expression products; especially when the membrane proteins or secretory proteins are expressed, the problems go worse [3]. Therefore, it would be better to mimic the expression of the related eukaryotic genes to the biologic synthesis of the active ingredients of the medicinal plants in the eukaryotic cells.

Saccharomyces cerevisiae of the fungus is the eukaryotes of the single cell, which is also one of the ideal eukaryotes for studying the gene expressions. It has several advantages: firstly, its genome is small, which is as 4 times as large as the genome of *Escherichia coli* (4,000 kb); secondly, its reproduction is rapid as the production of one generation in several hours; thirdly, it is present in both haploid (n) and diploid (2n); and fourthly, just as the prokaryotes, the experiments are easily performed. Many yeast strains can be considered as the recipients of the gene cloning. When the exogenous genes are expressed in the yeast cells, it would be better to place the structural genes of the exogenous genes in the control of the promoter of the yeast genes. Although the exogenous genes are in control of the yeast promoter (such as ADHI) and expressed successfully in the yeast cells, the expression levels are commonly lower than the expression levels of endogenous genes. The factors impacting the expressions of the exogenous genes are many. For example, the space between the promoter and the starting codon ATG is smaller than normal, and the base sequence of the adjacent ATG codon is also different from normal. These situations would lead to the low translation levels. Pauli and Kutchan used the yeast expression system for the heterogeneous expression of CYP80B1 in P450 gene of poppy *E. californica* and successfully got P450 protein with biologic activities. This reveals the biologic functions of this gene—participation in the hydroxylation reaction of *N*-methylcoclaurine in *E. californica* [10].

Besides the yeast, insects' cells are the ideal eukaryotic expression system. The virus is commonly used as the carrier, and the genes of eukaryotes (medicinal plants) are expressed in the insects' cells. Insects' nuclear polyhedrosis virus (NPV) is the carrier for highly efficient expressions of eukaryotic genes, which are the Baculoviridae and the causative agent of many insects. The genome of this virus is double-strand circle DNA, about 130 kb. The expression carrier system of BacPAK Baculovirus is an efficient expression system especially for recombinant proteins. The elementary steps of expressing the eukaryotic genes with the carrier of baculovirus are the following: firstly building the intermediate carrier, using it for transforming DH5α competent cells, extracting the intermediate carrier plasmids, purifying the plasmids by spin column, recombining the intermediate carrier and the virus, making transfection of the cultured cells of *Spodoptera frugiperda* (Sfg) of the insects, and finally testing the expression products and their biologic activities [3].

9.3.3 Study of the Manual Regulation of the Biosynthesis of the Active Ingredients of Medicinal Plants

9.3.3.1 The Fundamental Principles of the Manual Regulation of the Biosynthesis of the Active Ingredients of Medicinal Plants

The main methods for improving the metabolic engineering of the plants secondary metabolites are: (1) improving the flow of synthesis approaches of the target compounds, (2) increasing the amount of the production cells, and (3) reducing the

catabolism of the target products. The fundamental principles of these three ways are illustrated below:

Improving the Flow of Synthesis Approaches of the Target Compounds

In the synthesis approach of the target compound, many factors can control the flow, such as rate-limiting enzyme, feedback inhibition, and competition ways. As for the control of rate-limiting enzymes, the protein-coding genes controlled by the strong promoter are introduced to improve the levels and increase the flows. In addition, the heterologous genes coding enzymes of similar functions are introduced from other plants or microbes to increase the flow. To overcome the feedback inhibition, a gene coding the similar enzymes is used, and this enzyme is not sensitive to the feedback inhibition. This aim can also be achieved by protein engineering. In these efforts, the expression of a functional protein in the cells is needed.

Through the analyses of the natures of the biologic synthetase of the active ingredients of the medicinal plants and the enzymatic activities and the reaction dynamics of the related genes' heterologous expressions, we can determine which enzymes and genes are essential for the biologic synthesis of the active ingredients and which step is the committed step of the biologic synthesis. With this knowledge, we can control the related enzymes and genes to improve the biologic synthesis levels of target active ingredients. The expression amount or activities of the essential enzymes are increased through some elicitors. Meanwhile, some inhibitors are used to inhibit the key enzymes in the branch approach of the biologic synthesis of the active ingredients or the expression amount or activities of the enzymes participating in the degeneration of the target active ingredients to further improve the levels of the biologic synthesis of the active ingredients.

The choice of the regulation step depends on the total flows of primary and secondary metabolism. If the total flow is high but the output of the target compounds is low, the regulation target is the secondary metabolism, taking the production of vinblastine in the culture of *Catharanthus roseus*. If the amount of the secondary metabolites is low, the prime regulation target will focus on the enzymes from the primary to secondary metabolism.

The competition approach can be inhibited through the antisense genes. Genes can overcome the influence of the competition approach through improving the enzymatic levels of catalyzing the target products or introducing the enzyme with similar functions which has high affinity with the target products. A totally different concept is to introduce a gene to code an antibody, which can inhibit the activities of an enzyme which competes with target products for the same substrate in the metabolism of the target products.

Obviously, these efforts need a comprehensive understanding of the related secondary metabolism for all the midbodies, enzymes, and protein-coding genes and the metabolic regulation and principles at different levels. However, a few approaches in the secondary metabolism are understood. Falconoid-anthocyanidin metabolic approach is best known, which has been used in changing the colors of

the flowers. For most of the metabolic approaches, the study levels still leave in midbodies. In recent years, the further studies have been studied in the levels of the enzymes, such as for the study of the metabolic approaches of terpenoids, indole alkaloid, isoquinoline, tropane, and some terpenoids. The first step is cloning the coding genes. The following text will illustrate on this.

Increasing the Amount of the Production Cells

Different tissues of the plants produce different secondary metabolites. Only a small part of cells participates in the synthesis of the target compounds. Some of the plants cells culture reveal that not all cells synthesize the target compounds. The study of *C. roseus* cell culture shows that the amount of anthocyanidin depends on the cells producing the compounds and that the levels of anthocyanidin in the cells synthesizing anthocyanidin are similar. Through the method of increasing the production cells by genetic regulations, the production rate of anthocyanidin can increase five to six times in the process of *C. roseus* cells culture. However, for the process of a cell producing a secondary metabolite, we know little about it. This needs more basic studies to complete in the process of cell differentiation. Actually, it is a meaningful research topic of studying the production of secondary metabolites in cells differentiation.

Inhibiting the Catabolism

From some study reports, the cells culturing the final products degenerate. For the study of *C. roseus* cells cultures, in the last loop of the growth, the rate of degenerated alkaloid is similar to the rate of synthesis [11]. If the degeneration metabolism is wanted to be prevented, it is essential to find the related enzymes. However, the reversion of the chemical degeneration is inevitable.

9.3.3.2 The Basic Methods of Manual Regulation of the Target Active Ingredients of Medicinal Plants on the Basis of Gene Levels

Transformation Technologies of Medicinal Plants

Using transgenic technologies for the manual control of biologic synthesis of the active ingredients of the medicinal plants on the basis of the gene levels, another basis is to build ideal transformation technologies besides building target genes. Plants transgenic technologies include agrobacterium-mediated approach, PEG-induction method, electrization method, virus vector method, liposome method, microinjection method, gene gun method (or micro-bombardment method), DNA solution directly cultured method, pollen transferring gene method, pollen tube track, electrophoresis method, laser holing method, and ultrasound transferring

method [12]. For the medicinal method, the most commonly used method is agrobacterium-mediated approach.

There are two kinds of agrobacterium in the medicinal plants transgene, which are *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes*. The former induces tumor in plants, and the latter induces hairy roots. The transgenic mechanisms of both are similar, which are based on a kind of plasmid (Ti or Ri plasmid). This plasmid has one fragment which can be integrated in the plant genome and the expressed T-DNA. T-DNA area of which is 20 kb and contains the genes for controlling the differentiation of stems and roots and opine synthetic enzyme genes. In addition, in both ends of T-DNA area, there is a terminal repeat sequence around 25 bp, which is called either RB or LB. RB is essential for the integration of T-DNA into plants chromosomes. A series of genes in Vir area of Ti plasmid are helpful for the integration of T-DNA into plants chromosomes.

In the upstream and downstream of the exogenous genes, the suitable promoter (such as CaMV-35S) and terminator are connected, and then the connected sequences are inserted in T-DNA area of Ti plasmid in which the oncogene is knocked out. Because Ti plasmid is about 200 kb, it is difficult to insert the exogenous genes directly into T-DNA area of Ti plasmid. Therefore, in the practice, the exogenous genes are cloned to the intermediate carriers in *Escherichia coli*, and then the intermediate carriers with exogenous genes are transferred into *Agrobacterium tumefaciens*. There are two kinds of carrier systems which use this method, which are cointegrate vector system and binary vector system. The oncogene sequence in T-DNA area of cointegrate vector Ti plasmid is replaced by DNA of pBR322 plasmid. When pBR322-derived intermediate carriers with the exogenous genes enter *Agrobacterium tumefaciens* from *Escherichia coli*, the exogenous genes can be integrated into T-DNA of noncarcinogenic Ti plasmid because of the homologous recombination of intermediate carriers and pBR322 sequence in noncarcinogenic Ti plasmid. The common cointegrate vector systems include pGV3850, pGV2260, and pMON220 systems. Binary vector system is constituted with two compatible Ti plasmids. One is shuttle plasmid, which belongs to the multifunctional cloning vector plasmid and is replicated in both *Escherichia coli* and *Agrobacterium tumefaciens*. This plasmid has the features: (1) in its T-DNA, there are cloning points and plants selected marker genes; (2) the areas outside T-DNA area contain bacteria selected marker genes; (3) it can be transported from *Escherichia coli* to *Agrobacterium tumefaciens*. The other is Ti-derived plasmid, such as pGV2260, which can provide functional areas of trans-toxicity to help T-DNA integrate into the plant chromosomes. Binary vector system can transport larger than 40 kb DNA into the plants. Either cointegrate vector system or binary vector system needs to transport the intermediate vectors from *Escherichia coli* to *Agrobacterium tumefaciens*. To finish this process, there are two methods. One is triparental mating, taking pMON220 system as an example. The triparental mating method needs an *Escherichia coli* with pMON220 intermediate plasmids, an *Escherichia coli* with transportation plasmid pRK2013, and an *Agrobacterium tumefaciens* with noncarcinogenic Ti plasmid pTiB6S3SE. The steps are the following: firstly, transporting pRK2013 plasmid into the *Escherichia coli* with

pMON220 plasmid; secondly, code transportation proteins RK2 and movement proteins combining pMON220 plasmid to help pMON220 plasmid transferring into *Agrobacterium tumefaciens*; and thirdly, by LIH sequence, it is homologous recombined in pTiB6S3SE plasmid. The other method is to use $CaCl_2$ to prepare competent cells of *Agrobacterium tumefaciens* to change the permeability of the cell membranes and to transfer the intermediate vectors into *Agrobacterium tumefaciens* by freeze-thawing method [3].

Besides the usage of *Agrobacterium tumefaciens*, *Agrobacterium rhizogenes* can be used for the transformation of exogenous gene. Ri plasmid of *Agrobacterium rhizogenes* can induce the growth of the hairy roots of the plants, and the hairy roots can differentiate into the complete plants in the effect of plants hormones. In *Agrobacterium rhizogenes*, cointegrate vector system containing Ri plasmid $pRiA_4$ is used. pCGN529 plasmid (containing exogenous genes) can form into cointegrant with T_LDNA of $pRiA_4$ plasmid due to its T_LDNA segment. In the transformational plants tissues, the integrant $pRiA_4$: pCGN529 induces the formation of the hairy roots. These uncertain roots can be cut and placed in the medium containing kanamycin to select the cells with recombinant T-DNA, and the suitable plants hormone can be added into the medium for inducing the development of fertile plants [2]. In practice, it is less common for the usage of *Agrobacterium rhizogenes* than *Agrobacterium tumefaciens*. People use the wild *Agrobacterium rhizogenes* to induce the plants to develop the hairy roots for useful secondary metabolites [13, 14]. In recent years, this topic is a hot issue in the research field and gets a lot of valuable results.

There are many ways for Agrobacterium to transform the plants, and the most used technology is leaf disc method, which takes the leaves as the recipient cells with hereditary consistence and develops the complete plants by tissue cultures. The steps are: firstly, the bacteria-free leaves slices and the selected soil Agrobacterium are cultured together for 24–36 h, and then these things are transferred to the selective medium containing antibiotics for culturing. In this medium, the untransformed cells are killed by the antibiotics, and the transformed cells will develop into complete plants [3].

Several Approaches of Manual Regulation of Biologic Synthesis of the Active Ingredients of the Medicinal Plants

To raise the biologic synthesis levels of the active ingredients of the medicinal plants and based on the known knowledge, we can start from several approaches listed below:

Firstly, some key genes in biologic synthesis of the active ingredients are imported into this medical plant to increase the expression amount of the key genes and to increase the biologic synthesis of the target active ingredients.

Secondly, the usage of antisense technology. Based on the principle of base complementarity, DNA or RNA fragments (or their chemical modification products) of certain complementation by manual synthesis or biologic synthesis inhibit or close some genes expressions. Through this technology, antisense RNA

fragments can be introduced into the plant cells for inhibiting the expressions of the key enzymes catalyzing a branch of metabolism. Thus, the amount of target compound can be increased, and the synthetic approaches of other compounds are inhibited. At present, this method has produced successful results, for example, Mol JNM et al. successfully regulated the activities of cinnamyl alcohol dehydrogenase in the hairy roots of *Linum flavum* and inhibited the synthesis of the lignin molecule in the branch metabolism for increasing the amount of anticancer compound 5-methoxypodophyllotoxin [15].

Thirdly, using RNAi technology for inhibiting the related genes expressions to the biologic synthesis of the active ingredients of the plants.

As we know, there are three kinds of RNA: tRNA (transfer RNA), rRNA (ribosome RNA), and mRNA (messenger RNA). RNA is commonly single-strand linear molecule as well as double strand and loop single strand. In 1983, RNA molecules with branched chains were found. According to the early genetic points, RNA appeared as affiliated to DNA in the hereditary process of biologic characters, which took simple functions that RNA transfers DNA's genetic information from cellular nucleus to ribosomes synthesizing proteins as the messenger and template of the biologic synthesis of proteins. With deeper studies of RNA, it is found that RNA is not only the messenger of hereditary information but also the controller of the gene expressions especially some small molecules with the nucleotides length of 21–28. RNA can close the genes or change the expression levels of the genes and even have the catalytic activities of enzymes, which are unknown functions previously.

Actually, the similar studies had been reported. Dated back to 1920s, it was known that once the plants were infected by the wild virus, the plants could produce the resistance of virulent virus with relations. This is the phenomenon of virus-induced gene silencing (VIGS). In the beginning of 1980s, Napoli C et al. imported the key gene of synthesizing anthocyanin-chalcone synthase to petunia to get deeper purple petunia. On contrary to their expectations, some flowers became mottled, even white. The test results showed the expressions of both the imported exogenous chalcone synthase, and the endogenous chalcone synthase of the petunia had been suppressed. This gene-silencing phenomenon of transgene suppressing the related endogenous gene and its expression was called co-suppression of the plant genes [16].

Several years after this, Satan et al. found the similar phenomenon. They tended to import the replicate gene of the replication of tomato virus X to the tobacco, and the purpose was to block the life cycle of the virus and control the growth of the virus. The results showed some tobacco presented the antivirus characters, and the other presented the opposite. Waterhouse et al. from Commonwealth Scientific and Industrial Research Organisation (CSIRO) also found the combination of sense and antisense RNA could suppress the expression of potato virus Y (PVY) and create the anti-PVY in potato plants [17]. Besides the plants, Cogoni et al. from Italy imported *al-I* or *al-3* into *Neurospora crassa*, and endogenous *al-I* or *al-3* in cells transformed by *Neurospora crassa* had been suppressed. They called this gene inactivation as gene quelling [18].

In 1995, Dr. Su Guo et al. from Cornell University used antisense RNA to block Par-1 gene of *Caenorhabditis elegans*. Out of their expectations, sense RNA as the control in the experiment just like antisense RNA could specifically block the expressions of the related genes, and the suppression effects were better than antisense RNA. How did the plants endogenous virus resist the infection of exogenous virus? How did the exogenous genes and sense RNA suppress the expressions of the endogenous genes? These questions had been mysteries in the past for a long time.

In February, 1998, Andrew Fire from Washington Carnegie Institute, Craig Mello from University of Massachusetts, and other researchers firstly revealed the truth. They found that in various life beings endogenous or exogenous double-stranded RNA were imported into the cells, mRNA with the same origins as the double-stranded RNA would be degenerated, and the corresponding genes expressions would be suppressed. What is more, the suppression effects could last to the next generation [19]. Because this is the suppression of gene expressions in the level of RNA, they called this double-stranded RNA-mediated posttranscriptional gene silencing (PTGS) as RNA interference (RNAi). In various life beings, such as fruit fly, Arabidopsis, and mouse, double-stranded RNA-mediated RNAi was found. Until now, RNAi has been found widely in fungus, plants, invertebrate, and mammalian animals.

At this point, plant virus-mediated gene silencing (VIGS), co-suppression of plant cells, gene quelling in fungi, and RNA interference widely existing in the biosphere have been gradually harmonized because the same basic molecular mechanism, namely, 21–23 nt small RNA-mediated posttranscriptional gene silencing (PTGS), was discovered. Actually, RNA interference is a naturally occurring mechanism in eukaryotic organisms, monitoring and defending exogenous and movable genetic materials. Therefore, RNAi technology forms using this mechanism purposefully and creating specific gene silencing artificially. A series of major breakthroughs have occurred in studies on small RNA molecules represented by RNAi technology in the past 3 years, and in the annual selection of "the world's top ten scientific and technological breakthroughs" of "SCIENCE," all RNA studies have occupied an important position and won the championship in 2002.

Unlike antisense RNA technology and co-suppression, RNA interference (RNAi) has the following four important advantages: (1) RNAi target sequence is selective. After double-stranded RNA corresponding to exon sequences of a gene is imported into biological cells, mRNA of the corresponding target gene can be specifically degraded, thus inhibiting the expression of the gene. But double-stranded RNA importing intron or promoter sequences of the gene show no interference effect. (2) RNA interference is highly specific. Small RNA molecules can specifically degrade mRNA of a single endogenous gene corresponding to its sequence, thereby specifically inhibiting the expression of target genes. (3) RNA interference shows a high efficiency in inhibiting the expression of genes: Only a very small amount of double-stranded RNA molecules (the number is far less than that of endogenous mRNA) can completely inhibit the expression of the corresponding genes, and even the substoichiometric level of only a few dsRNA in each cell is still sufficient to inhibit target gene function. Phenotype generated by RNAi can achieve the extent of the lack of mutant phenotype. (4) RNAi has a magnifying effect. RNAi effects

can enlarge; the RNAi effect can still be maintained after 50–100 times of cell proliferation. The effect of RNA interference inhibiting gene expression can also go through cell boundaries, delivering and maintaining through long distances among different cells. Interference effects in *C. elegans* can even be passed to offspring. In plants, the effect of silencing gene expression of RNAi can be passed from local to the whole plant.

RNAi technology has been widely used in the identification of biological gene function, screening, and breeding of new genes. In plants, Professor Waterhouse from Commonwealth Scientific and Industrial Research Organisation (CSIRO) made 100 % inhibition of the expression of Arabidopsis δ12-desaturase gene (FAD2) using the technology; the activity of FAD2 proteins was significantly decreased, and one of the transgenic strains has passed through five generations, the inhibitory effect of the phenotype, that is, FAD2 gene, still has been very stable. This is the first example in the world to use RNAi technology to successfully make plant seed traits stabilized genetic changes. In addition to the work of Waterhouse P, RNAi technology has been widely used in the worldwide research of plant gene function and the work of genetic breeding: Chuang CF et al. have successfully, effectively, and specifically inhibited functions of multiple genes of *A.thaliana* using iHP RNAi technology invented by Waterhouse, and the effect of RNA inhibiting gene expression can be passed down to future offspring [20]; Schweizer P et al. produced effective interference in the function of dihydroxyflavonol-4-reductase gene encoding maize and barley at the single-cell level using RNAi method, and the accumulation of red anthocyanin related to cells was decreased [21]; Wang E et al. successfully identified the function of two genes which play an important role in the metabolism of terpenes using RNAi: A gene is responsible for cyclizing geranyl pyrophosphate (GPP) into cembratriene-ol (CBT-ol), and the other gene of the P450 family is responsible for transforming CBT-ols into CBT-diols [22]; Ogita et al. successfully inhibited CaMXMT1 gene involved in caffeine biosynthesis using RNAi technology and reduced 70 % of the caffeine content in transgenic plants, which makes the future production of decaffeinated coffee beans possible [23]. Recently, Davuluri et al. inhibited the regulatory gene DET1 in tomato photomorphogenesis by RNAi so that carotenoids and flavonoids in tomato fruits were greatly improved [24]. No wonder, the Nobel laureate Director and Professor Phillip A. Sharp from the McGovern Institute for Brain Research, the Massachusetts Institute of Technology (MIT) praised RNAi technology as the most important and most exciting scientific breakthroughs over the past 15 years and even the past few decades. The discovery of the key role of the double-stranded RNA in a variety of gene-silencing phenomena indicates the arrival of a new era of RNA.

Four, the content of the target product can be increased by controlling the expression in regulatory genes of biosynthetic genes of the effective ingredients of medicinal plants (such as transcription activator). Because of the plasticity of secondary metabolites in plants, it is often very difficult to improve target products through artificial regulation on genes of key enzyme biosynthetic genes of the effective ingredients. In this case, we can regulate regulatory genes of related gene to achieve our objective.

9.4 Problems

9.4.1 Problems in Gene Cloning

First of all, it is very complex to clone a gene from the secondary metabolic pathway. Though a variety of molecular biology means are available for gene cloning, such as transposon tagging and differential screening, these methods are often unsuccessful in the cloning of secondary metabolic pathway genes. This is mainly due to the lack of understanding to secondary metabolic pathways and their intermediates.

Presently, many studies have been conducted along the path: identification of metabolic intermediates, separation, purification, and identification of metabolic enzymes, followed by cloning of genes. The low level of most secondary metabolic synthetase is a very complex factor in the study. G10H, a cytochrome P450 enzyme, is a good example. The enzyme shows only a very low yield after purification. Obtained antibodies are not sufficient to be selected for the cloning of expressed proteins. Using a probe designed combining PCR with the conserved region of the cytochrome P450 enzyme, 16 very closely related genes were obtained from *C. roseus* [25]. It is unrealistic to express all these cDMA to identify genes encoding G10H. Schroeder and his colleagues used the differential screening of *C. roseus* CDNA library to isolate a putative G10HcDNA cloning, but the expression of the cDNA in different systems has shown no G10H activity presently [26]. Cloning genes encoding NADPH have been successful.

9.4.2 Problems in Stability

According to reports, with the reproduction of plants from generation to generation, transferred genes gradually become silent. In plant cell culture process, the fastest-growing cells should be screened continuously. Under such a continuous selection pressure, the degree of the stability of the transgene remains unclear. The continuous growth of cells in a selective medium is very expensive for industrial production. Initial results show that the transgenic cell line selected in the culture of transgenic *C. roseus* cells is relatively stable [27].

9.4.3 Problems in Transportation

The synthesis of secondary metabolites involves different organelles or even different tissues and organs, so the transportation of intermediates plays an important role. For an only physical and mechanical transportation, the factor to affect biosynthesis rate is mainly diffusion rate, while metabolic engineering can only be

beneficial to increasing concentration gradient. If the transportation is selectively active, transport proteins are involved in the transport process, or the transportation is driven by pH gradient; genes encoding transport proteins or genes controlling pH will require overexpression [28]. The storage of alkaloids in vacuoles of *C. roseus* cells is a good example. The low pH value in vacuoles makes vacuoles as an ion trap for alkaline alkaloid. In this case, it is necessary to identify pH-regulatory genes, which ultimately affect the transfer and promote the accumulation of alkaloids [29].

9.5 Prospects for Development and Future Direction of Efforts

Metabolic engineering has wide application prospects in the modification of secondary metabolic pathways, and that there will be many successful examples. In particular, a study in which only one step of response changed can lead to high yield of target compounds is likely to become the first commercial application. For example, scopolamine is produced by *Atropa belladonna* L. cell culture, but it is to produce hyoscyamine under normal circumstances. It can also be envisaged to allow plants to produce entirely new compounds, for example, by the method of introducing oxidase which has broad substrate specificity.

In short, we have entered a new area of plant culture and can assume factory production of plants which have the diversity of comprehensive useful chemical compositions. Moreover, we can also change secondary metabolic pathways in plants so that they have special quality in aspects of color, taste, toxicity, health, and resistance to diseases and insects. Studies on the genetic engineering of secondary metabolites in medicinal plants to obtain secondary products of medicinal plants using transgenic technology have shown very attractive development prospects and become the focus of attention of many scholars. Nevertheless, biosynthesis pathways of effective ingredients of many important medicinal plants have not been thoroughly studied clearly. There have been few enzymes and genes, especially key enzymes and genes involved in the biosynthesis of effective ingredients of medicinal plants for separation and identification, and it can be said that it is blank for the understanding about the regulatory mechanism of biosynthetic gene of effective ingredients of most of medicinal plants. In order to make better use of molecular biology techniques to achieve the purpose of the artificial regulation of effective ingredients of medicinal plants, we should focus on the following aspects in the future while further clarifying biosynthetic pathways of effective ingredients:

Secondary metabolic pathways of most medicinal plants are not yet very clear, so the study of metabolic pathways should be strengthened.

More key enzymes and their genes of biosynthesis of specific effective ingredients should be isolated and identified. The total flow of increasing a metabolic pathway may require a series of genetic engineering. Each gene encodes another rate-limiting step. Therefore, it is necessary to study the overall regulation of alternative ways and strive to find key steps to control the overall metabolic pathway. For example, studies on the regulation of expressions of PDC genes and STR genes

in *C. roseus* cells show that both were governed by the downregulation of auxin and upregulation of inducer. It can be seen that a gene is likely to control the expression of the two genes at least and may also control a greater part of metabolic pathways and even the entire metabolic pathways of alkaloids.

Tissues and development regulatory mechanism of biosynthetic gene expression of effective ingredients should be further studied.

We should further make clear various cis-and trans-acting factors which play the spatial and temporal regulation role in biosynthetic genes of effective ingredients. Conclusions can be drawn from all of the above experiments: the gene cloning of secondary metabolites is feasible and can show overexpression, but a limitation should be overcome. Therefore, the identification and cloning of regulatory genes may eventually lead to greater success to improve the content of secondary metabolites. This is because these genes can control a large part of secondary metabolism if they cannot control the whole. Of course, we should start the study from genes encoding special steps.

Promoters with the function of tissue-specific expression are separated to achieve the directional expression of effective ingredients of medicinal plants in particular tissues or cells.

Development of new type of bioreactor and establishment and perfection of efficient cell culture techniques;

Studies on functional genomics and systems biology of plant secondary metabolites are carried out. As we all know, most of the previous studies on the plant quality focus on the clarification of biosynthetic pathways of secondary metabolites in related plants, the cloning, transformation, and regulation of related genes, such as paclitaxel in *Taxus* and arteannuin in *Artemisia apiacea*. However, these studies remain at the stage of studying individual genes one by one. Studies conducted simultaneously on multiple genes and even all genes and proteins as well as their interaction network of biosynthetic pathways of secondary metabolisms in plants such as paclitaxel in *Taxus* and arteannuin in *Artemisia apiacea* are expected to reveal the mechanism and regulation rhythm of biosynthesis of secondary metabolites of plants fundamentally and clarify plant quality and reasons for the formation of the resistance to achieve the purpose of artificially controlling plant (herbs) quality and breeding new varieties. This will be the new trend of the development in research areas of biosynthetic pathways of secondary metabolites. It can be expected that functional genomics and systems biology techniques including gene chip technology, proteomics, and metabolomics technologies will be widely used in the study of the biosynthesis and regulation of plant secondary metabolites.

We have every reason to believe that in the near future, we will get more new varieties of transgenic medicinal plants of which the secondary metabolites content is increased or the ratio between secondary metabolite compositions is changed, and use them in productive practice. In addition, we may also get good cell lines more by genetic engineering technology and conduct large-scale production of effective ingredients so that more large-scale cultivations of plant cells tend toward industrialization, thus providing new protection for the health of majority of people.

9.6 Case Study

The Regulation of Tanshinone Accumulation by Overexpression of 3-Hydroxy-3-methylglutaryl-CoA Reductase in Savia miltiorrhiza Hairy Roots

Tanshinones are abietane-type norditerpenoid quinones found in a commonly used Chinese medicinal herb *Salvia miltiorrhiza* Bunge. This group of diterpenoids mainly includes dihydrotanshinone I, tanshinone I, tanshinone IIA, and cryptotanshinone [30]. These compounds have exhibited diverse pharmacological activities, including antibacterial, antioxidant, anti-inflammatory, cytotoxic, neuroprotective, cardioprotective, antiplatelet, and antitumor effects [30, 31]. However, the content of tanshinones in *S. miltiorrhiza* is low. Biotechnology strategies are a good way to efficiently increase the yield of tanshinones from the cultured hairy roots. Here, we introduce our research work on overexpression of 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) gene to increase tanshinone accumulation in hairy roots.

9.6.1 Materials and Methods

9.6.1.1 Plant Materials Culture

The mature seeds of *S. miltiorrhiza* Bunge were surface-sterilized by 0.1 % mercuric chloride (Sigma-Aldrich, St. Louis, MO, USA) and cultured on solid, hormone-free MS basal medium. The MS medium contained 30 g/L sucrose and 8 g/L agar without ammonium nitrate for germination. Cultures were maintained at 25 °C under a 16-h light/8-h dark photoperiod with light provided by cool white fluorescent lamps at an intensity of 25 μmol/m^2/s and then induced hairy roots for the following experiments.

9.6.1.2 DNA and RNA Isolation

Total RNA was extracted from the hairy roots of *Salvia miltiorrhiza* by Trizol method (Invitrogen, Carlsbad, CA, USA). Genomic DNA was isolated using the modified cetyltrimethylammonium bromide method [32].

9.6.1.3 Cloning of SmHMGR2 Full-Length cDNA by Rapid Amplification of cDNA Ends (RACE)

5′-RACE was performed according to the manual of the SMARTTM RACE cDNA Amplification Kit (Clontech Laboratories Inc., Mountain View, CA, USA). The 5′-RACE PCR was carried out using the 5′-RACE primer and universal primer (UPM, Universal Primer A Mix). The PCR product was purified and cloned into pMD19-T

vectors followed by sequencing. After aligning and assembling the sequences, the full-length cDNA sequence of the SmHMGR2 gene was deduced and subsequently amplified by PCR using a pair of primers. The genome sequence of the SmHMGR2 was confirmed by PCR with the genome DNA as a template.

9.6.1.4 Construction of the SmHMGR2 Binary Expression Vector pH7WG2D-SmHMGR2

The SmHMGR2 gene was PCR-amplified, and the resultant PCR products were purified with the DNA purification kit (Sangon, Shanghai, China).We sequentially subcloned the PCR product into a donor vector (pDONR221) and created an entry vector. The recombination reaction between the entry and destination vectors (pH7WG2D,1) was performed in LR ClonaseTM II enzyme mix (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The *E. coli* DH5α was transformed with the product of LR reaction using the heat shock transformation method. The destination vector pH7WG2D,1 with the ccdB (control of cell death) gene knockout was used as the control vector (pH7WG2D-Control) [33].

9.6.1.5 Plant Transformation in Cultured Hairy Roots of *S. miltiorrhiza*

The vectors of pH7WG2D-SmHMGR2 and pH7WG2D-Control were transformed into *S. miltiorrhiza* through the mediation of the *Agrobacterium rhizogenes* strain ACCC10060 as described previously [34]. Root tissues from three flasks of cultures were collected separately on d 10, 20, 30, 40, and 50 after inoculation.

9.6.2 Results

9.6.2.1 Growth, mRNA Level of SmHMGR2, Enzyme Activity of HMGR, and Contents of Tanshinones in the Hairy Roots of *S. miltiorrhiza*

All hairy root lines examined the mRNA level, enzyme activity, and tanshinone accumulation. Results showed the H lines significantly increase in the mRNA level of SmHMGR2 than the other lines, and there were no differences between VCK, HRCK, and wild-type roots (Fig. 9.1a). Consistently, the enzyme activity of HMGR was significantly increased in H14, H19, and H24 compared to the wild type or HRCK roots (Fig. 9.1b). The VCK lines had a slightly higher enzyme activity than HRCK and wild-type roots. During first 20 days of culture, the accumulation of all tanshinones determined was low and then slowly increased and reached the maximum on d 50 (Table 9.1). Generally, H lines had higher levels of individual tanshinones

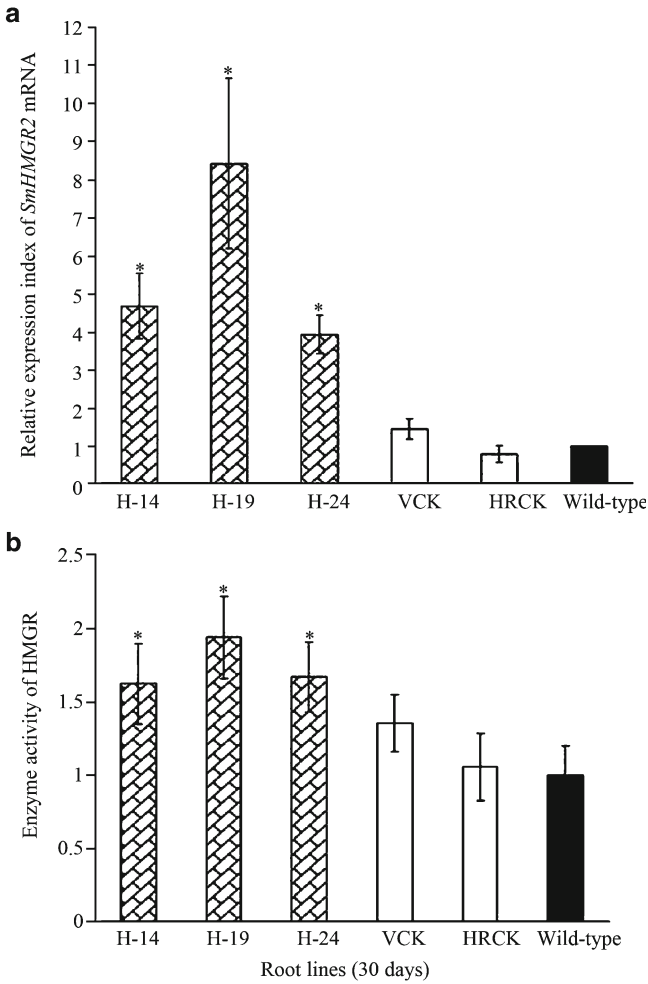

Fig. 9.1 The mRNA level of SmHMGR2 and enzyme activity of HMGR in hairy roots of *S.miltiorrhiza*. (**a**) Expression pattern of SmHMGR2 in different hairy roots. (**b**) HMGR activities of hairy roots. *Vertical bars* indicate the mean ± SD from three independent experiments. *$P < 0.05$ compared to the wild-type, VCK, or HRCK lines

in H lines compared to the wild-type, VCK, or HRCK roots. For dihydrotanshinone and cryptotanshinone, the increase folds in H lines after 50 days of culture were 2.40–2.97 and 3.41–5.46, respectively (Table 9.1 A, B). The level of tanshinone I varied from 0.69- to 1.35-fold, while the level of tanshinone IIA in H lines was

Table 9.1 Contents of four tanshinones in nontransformed and three transgenic lines of hairy roots of *S. miltiorrhiza*

	20d	30d	40d	50d
A				
Root line	Content of dihydrotanshione I (mg/L)			
H14	0.42 ± 0.25[a]	3.89 ± 0.65 [a,b]	6.17 ± 1.99[a,b]	23.72 ± 1.63[a,b]
H19	0.90 ± 0.36	3.80 ± 0.74[a,b]	7.90 ± 1.63[a,b]	25.66 ± 0.55[a,b]
H24	0.21 ± 0.07[a]	3.77 ± 0.83[a,b]	8.31 ± 1.47[a,b]	20.73 ± 0.69[a,b]
VCK	0.13 ± 0.12[a]	2.96 ± 0.28[a,b]	6.05 ± 0.70[a,b]	13.84 ± 1.18[a,b]
HRCK	0.09 ± 0.10[a]	1.91 ± 0.36[b]	5.36 ± 0.44[a,b]	13.65 ± 2.86[a,b]
Wild-type	1.12 ± 0.05	1.66 ± 0.75[b]	3.21 ± 0.38[b]	8.64 ± 0.78b
B				
Root line	Content of cryptotanshinone (mg/L)			
H14	0.26 ± 0.19[a]	1.92 ± 0.46[a,b]	3.72 ± 1.54[a,b]	18.67 ± 1.43[a,b]
H19	1.02 ± 0.51	2.38 ± 1.10[a,b]	7.44 ± 1.35[a,b]	18.05 ± 5.54[a,b]
H24	0.13 ± 0.05[a]	1.75 ± 0.68[a,b]	4.81 ± 1.48[a,b]	11.66 ± 1.09[a,b]
VCK	0.04 ± 0.04[a]	0.49 ± 0.11[a,b]	2.16 ± 0.53[a,b]	6.85 ± 1.40[a,b]
HRCK	0.06 ± 0.05[a]	0.56 ± 0.13[a,b]	2.63 ± 0.62[a,b]	7.90 ± 1.52[a,b]
Wild-type	0.91 ± 0.03	0.93 ± 0.31	1.28 ± 0.33[b]	3.42 ± 0.30[b]
C				
Root line	Content of tanshinone I (mg/L)			
H14	0.73 ± 0.36[a]	2.72 ± 0.59[b]	3.58 ± 0.84[a,b]	6.24 ± 1.12[b]
H19	2.00 ± 0.81[a]	4.82 ± 0.53[a,b]	6.93 ± 0.30[a,b]	8.36 ± 1.32[a,b]
H24	0.55 ± 0.17[a]	1.77 ± 0.75[a,b]	2.34 ± 0.31[a,b]	4.29 ± 0.22[a,b]
VCK	0.26 ± 0.22[a]	1.83 ± 0.22[a,b]	3.29 ± 0.72[a,b]	4.04 ± 0.88[a,b]
HRCK	0.21 ± 0.14[a]	1.25 ± 0.20[a,b]	2.82 ± 0.53[a,b]	3.82 ± 0.75[a,b]
Wild-type	1.52 ± 0.22	2.36 ± 0.61[b]	5.71 ± 0.79[b]	6.19 ± 0.56[b]
D				
Root line	Content of tanshinone IIA (mg/L)			
H14	0.24 ± 0.16[a]	0.77 ± 0.09[a,b]	1.35 ± 0.17[a,b]	2.35 ± 0.26[a,b]
H19	1.19 ± 0.56	2.50 ± 0.74[a,b]	3.92 ± 0.23[a,b]	3.22 ± 1.13[b]
H24	0.15 ± 0.15[a]	0.38 ± 0.20[a,b]	0.60 ± 0.14[a,b]	0.97 ± 0.14[a,b]
VCK	0.04 ± 0.05[a]	0.23 ± 0.05[a,b]	0.71 ± 0.33[a,b]	1.66 ± 0.93[a,b]
HRCK	0.05 ± 0.04[a]	0.21 ± 0.07[a,b]	0.65 ± 0.17[a,b]	0.97 ± 0.17[a,b]
Wild-type	0.99 ± 0.12	1.75 ± 0.57[b]	3.18 ± 0.97[b]	3.19 ± 0.29[b]

Note: H and VCK refer to the transgenic hairy root lines generated from the SmHMGR2 gene and vector control, respectively. Hairy root lines generated from transformations with ACCC10060 containing no pH7WG2D-SmHMGR2 and pH7WG2D-Control are denoted as HRCK. Values are reported as the mean ± SD from at least three regenerants with three determinations each

[a] $P < 0.05$ vs. the wild-type VCK or HRCK lines at different culture time, by two-way ANOVA with Tukey's test

[b] $P < 0.05$ vs. 20d in different root lines, by two-way ANOVA with Tukey's test

slightly decreased or increased (0.30- to 1.01-fold) when compared to the wild-type roots (Table 9.1). Overall, the total levels of all four tanshinones were increased 1.76- to 2.58-fold in H lines compared to the wild-type line. The maximum increase in total tanshinones (2.58-fold) was recorded in the transgenic lineH19. These results indicate that tanshinone levels were increased in hairy roots of *S. miltiorrhiza* overexpressing the SmHMGR2 gene.

References

1. Kutchan, T. M. (1995). Alkaloid biosynthesis—The basic for metabolic engineering of medicinal plants. *The Plant Cell, 7*, 1059–1070.
2. Wu, N. H. (1998). *Principle of gene engineering* (2nd ed.). Beijing: Science Press.
3. Hao, F. Y., Zhu, Y. X., Zhu, S. G., et al. (1998). *Experimental techniques of molecular biology*. Beijing: Peking University Press.
4. Chen, Z. L. (1994). *Plant gene engineering*. Changchun: Jilin Scientific and Technical Press.
5. Xie, Y. J. (1992). *An introduction to genetic engineering*. Beijing: Beijing Agricultural University Press.
6. Xu, Z. H., & Liu, C. M. (1998). *Molecular mechanism of plant development*. Beijing: Science Press.
7. Corey, E. J., Matsuda, S. P. T., & Bartel, B. (1993). Isolation of an *Arabidopsis thaliana* gene encoding cycloartenol synthase by functional expression in a yeast mutant lacking lanosterol synthase by the use of a chromatographic screen. *Proceedings of the National Academy of Sciecnes of the United States of America, 90*, 11628–11632.
8. Vetter, H. P., Mangold, U., Schroder, G., et al. (1992). Molecular analysis and heterologous expression of an inducible cytochrome P-450 protein from periwinkle (*Catharanthus roseus L.*). *Plant Physiology, 100*, 998.
9. Lois, A. F., & West, C. A. (1990). Regulation of expression of the casbene synthetase gene during elicitation of castor bean seedlings with pectic fragments. *Archives of Biochemistry and Biophysics, 276*, 270–277.
10. Pauli, H. H., & Kutchan, T. M. (1998). Molecular cloning and functional heterologous expression of two alleles encoding (s)-N-methylco aurine 3′-hydroxylase (CTP 80B1), a new methyl jasmonate-inducible cytochrome P-450-dependent mono-oxygenase of benzylisoquinoline alkaloid biosynthesis. *The Plant Journal, 6*, 793–801.
11. Dos Santos, R., Schripdema, J., & Verpoorte, R. (1994). Ajmalicine metabolism in *Catharanthus roseus* cell cultures. *Phytochemistry, 35*, 677–681.
12. Yb, Bi. (1993). *Transfer method for plant gene*. Haerbin: Northeast Forestry Univeristy Press.
13. Qiu, D. Y., Zhu, C., & Zhu, Z. Q. (1996). Production of Trichosanthin from the hairy roots of *Trichosanthes kirilowii* Maxim. *Acta Botanica Sinica, 6*, 439–443.
14. Zhou, L. G., Wang, J. J., & Yang, C. R. (1998). Progress on plant hairy root culture and its chemistry. *Natural Product Research and Development, 3*, 87–95.
15. Oostdam, A., Mol, J. N. M., & van der Plas, L. H. W. (1993). Establishment of hairy root cultures of *Linum flavum* producing the lignan 5- methoxypodophyllotoxin. *Plant Cell Reports, 12*, 474–477.
16. Van der Krol, A. R., Lenting, P. E., Veenstra, J., et al. (1988). An antisense chalcone synthase gene in transgenic plants inhibits flower pigmentation. *Nature, 333*, 866–869.
17. Waterhouse, P. M., Graham, M. W., & Wang, M. B. (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proceeding of the National Academy of Sciences of the United States America, 95*, 13959–13964.

18. Cogoni, C., & Macino, G. (1997). Conservation of transgene induced post-transcriptional gene silencing in plants and fungi. *Trends in Plant Science, 2*, 438–443.
19. Fire, A., Xu, S., Montagomery, M. K., et al. (1998). Potent and specific genetic interference by double-stranded in Caenorhabditis elegans. *Nature, 391*, 806–811.
20. Chuang, C. F., & Meyerowitz, E. M. (2000). Specific and heritable genetic interference by double- stranded RNA in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States America, 97*, 4985–4990.
21. Schweizer, P., Pokorny, J., Schulze, P., et al. (2000). Double-stranded RNA interference with gene function at the single-cell level in cereals. *The Plant Journal, 6*, 895–903.
22. Wang, E., & Wagner, G. J. (2003). Elucidation of the functions of genes central to diterpene metabolism in tobacco trichomes using posttranscriptional gene silencing. *Planta, 216*, 686–691.
23. Ogita, S., Uefuji, H., Yamaguchi, Y., et al. (2003). Producing decaffeinated coffee plants. *Nature, 423*, 823.
24. Davuluri, G. R., van Tuinen, A., Fraser, P. D., et al. (2005). Fruit-specific RNAi-mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes. *Nature Biotechnology, 7*, 890–895.
25. Meijei, J. J., ten Hoopen, H. J. G., & KchAM, Luyben. (1993). Effects of hydrodynamic stress on cultured plant cells: A literature survey. *Enzyme and Microbial Technology, 15*, 234–238.
26. Mangold, U., Eichel, J., Batschauer, A., et al. (1994). Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from *Catharanthus roseus*, and transgenic expression of the gene and a cDNA in tobacco and *Arabidopsis thaliana*. *Plant Science, 96*, 129–136.
27. Verpoorte, R., van der Heijden, R., Ten Hoopen, H. J. G., et al. (1999). Metabolic engineering of plant secondary metabolite pathways for the production of fine chemicals. *Biotechnology Letters, 6*, 467–479.
28. Zhang, W., Curtin, C., & Franco, C. (2002). Towards manipulation of post-biosynthetic events in secondary metabolism of plant cell cultures. *Enzyme and Microbial Technology, 30*, 688–696.
29. Blom, T. J. M., Sierra, M., van Vliet, T. B., et al. (1991). Uptake and accumulation of ajmalicine into isolated vacuoles of cultured cells of *Catharanthus roseus (L.)* G. Don. and its conversion into serpentine. *Planta, 183*, 170–177.
30. Zhou, L., Zuo, Z., & Chow, M. S. (2005). Danshen: An overview of its chemistry, pharmacology, pharmacokinetics, and clinical use. *Journal of Clinical Pharmacology, 45*, 1345–1359.
31. Yuan, Y., Liu, Y., Lu, D., et al. (2009). Genetic stability, active constituent, and pharmacoactivity of *Salvia miltiorrhiza* hairy roots and wild plant. *Zeitschrift für Naturforschung Section C, 64*, 557–563.
32. Del Sal, G., Manfioletti, G., & Schneider, C. (1989). The CTAB-DNA precipitation method: A common mini-scale preparation of template DNA from phagemids, phages or plasmids suitable for sequencing. *BioTechniques, 7*, 514–520.
33. Coleman, D., Knights, J., Russell, R., et al. (1991). Insertional inactivation of the Staphylococcus aureus beta-toxin by bacteriophage phi 13 occurs by site- and orientation-specific integration of the phi 13 genome. *Molecular Microbiology, 5*, 933–939.
34. Komarovska, H., Giovannini, A., Kosutha, J., et al. (2009). *Agrobacterium rhizogenes*-mediated transformation of *Hypericum tomentosum* L. and *Hypericum tetrapterum* Fries. *Zeitschrift für Naturforschung, Section C, 64*, 864–868.

Index

Lu-qi Huang (ed.), *Molecular Pharmacognosy*,
DOI 10.1007/978-94-007-4945-0,