Key Techniques and Application Progress of Molecular Pharmacognosy

XIAO Xue-feng[1], HU Jing[1], XU Hai-yu[2, 5], GAO Wen-yuan[3], ZHANG Tie-jun[4], LIU Chang-xiao[2*]

1. School of Chinese Materia Medica, Tianjin University of Traditional Chinese Medicine, Tianjin 300193, China
2. The National Key Laboratory of Pharmacokinetics and Pharmacodynamics, Tianjin Institute of Pharmaceutical Research, Tianjin 300193, China
3. School of Pharmaceutical Science and Technology, Tianjin University, Tianjin 300072, China
4. Research Center of Modern Chinese Medicines, Tianjin Institute of Pharmaceutical Research, Tianjin 300193, China
5. Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China

Abstract: At the boundary between pharmacognosy and molecular biology, molecular pharmacognosy has developed as a new borderline discipline. This paper reviews the methods, application, and prospect of molecular pharmacognosy. DNA marker is one of genetic markers and some molecular marker methods which have been successfully used for genetic diversity identification and new medicinal resources development. Recombinant DNA technology provides a powerful tool that enables scientists to engineer DNA sequences. Gene chip technique could be used in determination of gene expression profiles, analyses of polymorphisms, construction of genomic library, analysis of mapping, and sequencing by hybridization. Using the methods and theory of molecular biology and pharmacognosy, molecular pharmacognosy represents an extremely prospective branch of pharmacognosy and focuses on the study of systemic growth of medicinal plants, identification and evaluation of germplasm resources, plant metabolomics and production of active compounds. Furthermore, the great breakthrough of molecular pharmacognosy could be anticipated on DNA fingerprint analysis, cultivar improvement, DNA identification, and a global DNA barcoding system in the future.

Key words: DNA marker; genetic diversity; molecular biology; molecular pharmacognosy; pharmacognosy

DOI: 10.3969/j.issn.1674-6384.2011.02.005

Introduction

Molecular biology is an exciting and rapidly expanding field, which could enable numerous other fields to go "molecular". Using the method and technology of molecular cloning, genetic engineering, tissue culture, and molecular markers, pharmacognosy has developed rapidly in recent years and now represents a highly interdisciplinary science (Liu, 2005; Liu *et al*, 2009; Mukherjee *et al*, 2010; Yu *et al*, 2010). Molecular pharmacognosy is an important part of molecular biology. The theory basis of molecular pharmacognosy can be summarized into three points: (1) The development of molecular biology has enabled other biological fields to go "molecular". Using the theory and methods of biology, pharmacognosy focuses on the properties of botanicals and animals; (2) Since the great majority of Chinese medicinal materials originate from either plant or animal sources, DNA which contains the genetic information in every cell of plant and animal materials connects molecular biology and pharmacognosy together. Moreover, DNA analysis becomes an important tool for pharmacognosy; and (3) Researches on medicinal plants and animals in pharmacognosy have developed from the organ, tissue, and cellular level to the genetic level. Consequently, the modernization of pharmacognosy enables its combination with molecular biology and its development (Huang, Chen, and Xiao, 1999; Huang, 2006; Kang, Yang, and Sun, 2002; Xie, 2001).

Pharmacognosy has been built up for 200 years. Looking back to the history, Chinese scholars have made great progress in its development. At the beginning, crude drugs were only identified by characters

* Correspondence author: Liu CX, an academician of Chinese Academy of Engineering Address: 308 Anshan West Road, Tianjin Institute of Pharmaceutical Research, Tianjin 300193, China Tel: +86-22-2300 6863 E-mail: liuchangxiao@vip.163.com
Received: July 21, 2010; Revised: October 16, 2010; Accepted: February 18, 2011

including external morphological characteristics, color, sections, texture, and flavor. Then optical microscopy technology was developed, which was based on the internal cells and tissues of crude drugs. Furthermore, scanning electron microscope (SEM) and ultra-micro identification system emerged as science evolved. Additionally, physics and chemistry analysis were also applied to the quality control of crude drugs and especially the application of spectroscopic techniques in component analysis was presented as a development summit in pharmacognosy (Xie, 2001).

Molecular pharmacognosy is an emerging discipline combining molecular biology and pharmacognosy, and its development depends on systems biology, especially genomics, proteomics and metabolomics and also depends on the development and methods of modern biotechnology (Fig. 1).

Methods based on systems biology

Systems biology is expected on the long term to predict both genomic activations and metabolite flows in complex systems. Their joint application is already

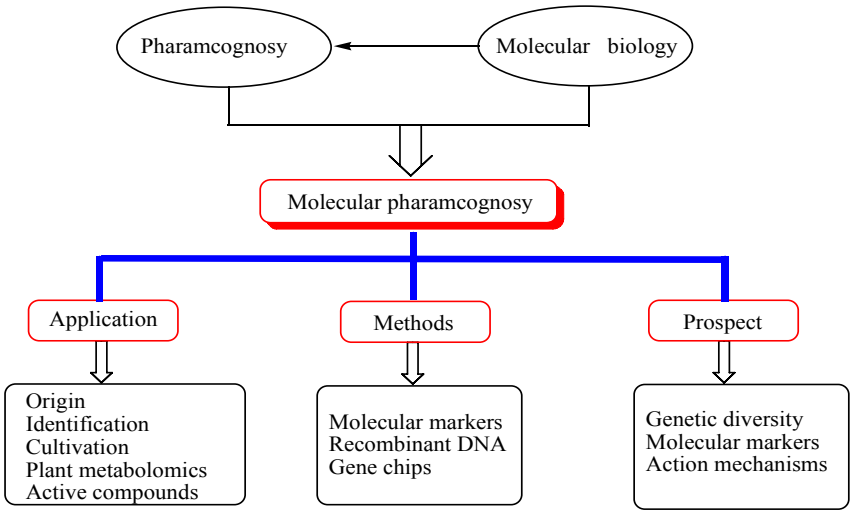

Fig. 1 Development of molecular pharmacognosy

now judged to be the ultimate phenotyping of a cell or plant and considered to have the potential to revolutionize natural product research and to advance the development of scientific-based herbal medicine (Wang *et al*, 2005; Verpoorte, Choi, and Kim, 2005; Patwardhan, 2005; Patwardhan *et al*, 2005). For example, these technologies and methods are likely to change and expedite the study of resources and quality control of medicinal plants. Theory and methods in chemistry, taxology, pharmacodynamics, physics, molecular biology, and genetics are all used in pharmacognosy studies (Ma *et al*, 2003; Huang *et al*, 2009a; Wang *et al*, 2003; Wang and Bi, 2003; Chen *et al*, 2003). The techniques such as molecular markers, recombinant DNA, and gene chips have provided significant support for DNA polymorphism detection in molecular pharmacognosy.

Molecular markers

DNA marker is one of genetic markers and it is based on polymorphic DNA. Some molecular marker methods have been successfully used for genetic diversity identification and new medicinal resources development. (1) Restriction fragment length polymorphism (RFLP), variable number of tandem repeats (VNTRs), and denaturing gradient gel electrophoresis-RFLP (DGGE-RFLP) are based on molecular hybridization; (2) Randomly amplified polymorphic DNA (RAPD), arbitary primer-PCR (AR-PCR), DNA amplification fingerprinting (DAF), single strand conformation polymorphism-RFLP (SSCP), cleaved amplified polymorphism sequences (CAPS), amplified fragment length polymorphism (AFLP), allele-specific PCR (AS-PCR), single primer amplification reaction (SPAR), simple sequence repeats (SSR), and inter-simple sequence repeat (ISSR) are based on PCR; (3) AFLP is based on the technology PCR and RFLP; (4) Revert transcription PCR (RT-PCR), differential display (DD), representative difference analysis (RDA), and fluorescence quantitative polymerase chain reaction (FQ-PCR) are based on RT-PCR; (5) Single nucleotide

polymorphism (SNP) and DNA barcoding are based on the sequencing; (6) DNA chips, such as cDNA and oligo microarray (Huang *et al*, 2009b; Huang, 2006; Gao *et al*, 2001; Guo *et al*, 2002a; 2002b).

Recombinant DNA

Recombinant DNA technology provides a powerful tool that enables scientists to engineer DNA sequences. It refers to use of molecular techniques to select a specific sequence of DNA from an organism and to transfer it into another organism to code for or alter specific traits. Before production step, target gene is derived from an original organism and integrated to a recombinant DNA. After that, this DNA is transformed into a receptor and replicated. Then target gene is produced in desired amounts and the expression of exogenous gene is identified (Qu *et al*, 1998). As the development of homologous DNA recombination, genetic engineering technology becomes more and more simple, rapid, and accurate. It has the advantages of requiring no use of restriction and combination enzymes and being able to be operated without changing the procedure that is used for recombinant DNA technology.

Gene chips

As a new molecular hybridization and sequencing technique, gene chip technique allows fast determination and assay of a large number of target molecules once. Consequently, it could make up the disadvantages of traditional blotting hybridization techniques, including the complicated operation procedures, lower automatization, fewer sequences, and lower efficiency, *etc*. This technique could be used in determination of gene expression profiles, analyses of polymorphisms, construction of genomic library, analysis of mapping, and sequencing by hybridization (Huang *et al*, 2009b). A novel DNA fingerprinting technique called AFLP is described. AFLP assay was set up by Vos, Hogers, and Bleeker. It is based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA. The technique involves three steps such as restriction of the DNA and ligation of oligonucleotide adapters, selective amplification of sets of restriction fragments, and gel analysis of the amplified fragments (Vos, Hogers, and Bleeker, 1995). The AFLP technique is based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA. Gene chip technology has developed rapidly in the past decade, which places DNA strands on glass computer chips and promises to revolutionize molecular biology. This technology will enable us to analyze the influence of environmental factors in gene expression of geoherbs and determine the gene regulation of active compounds (Ma and Jiang, 2000).

The sequencing and analysis of expressed sequence tags (ESTs) are primary approaches for the discovery of novel genes in plants, especially in non-model plants for which full genome sequences are not currently available EST sequencing represents a rapid and cost-effective method for analyzing the transcribed regions of genomes. EST analysis is also a powerful tool for the discovery of genes involved in plant secondary metabolism. The 454 GS FLX sequencing technology has made EST-based resources more readily accessible for non-model organism transcriptomes (Brautigam *et al*, 2008; Vera *et al*, 2008). Study by Zhang *et al* (2009) focused on *Glycyrrhiza uralensis* Fisch. ex DC., which is one of the most ancient medicinal herbs and has been used as a Chinese herbal medicine to treat infectious diseases for over 3000 years. Using the 454 GS FLX platform and Titanium reagents, their study provides a high-quality EST database for *G. uralensis*. Based on the EST analysis, novel candidate genes related to the secondary metabolite pathway of glycyrrhizin, including novel genes encoding cytochrome P450s and glycosyltransferases, were found. With the assistance of organ-specific expression pattern analysis, three unigenes encoding cytochrome P450s and six unigenes encoding glycosyltransferases were selected as the candidates most likely to be involved in glycyrrhizin biosynthesis. The information from these ESTs represents a significant contribution toward the exploration of the molecular mechanisms of glycyrrhizin biosynthesis. More importantly, a few candidate genes encoding the enzymes responsible for glycyrrhizin skeleton modifications were obtained by screening functional annotations and by organ-specific expression pattern analyses (Li *et al*, 2010).

DNA sequence and codon alignments were performed using RevTrans. Recently, Barthet and Hilu (2008) evaluated evolutionary constraints on *matK* using protein composition and tried to explain why this

protein coding gene accommodated elevated rates of substitution and yet maintained functionality. Duffy Kelchner, and Wolf (2009) compared *matK* sequences of an intronless fern clade to sequences from seed plants and ferns with the intron and found no significant differences in selection among lineages. The gymnosperm Taxaceae, Cephalotaxaceae, and Pinaceae were used to illustrate the physicochemical evolution, molecular adaptation, and evolutionary dynamics of gene divergence in *matK*s. *matK* sequences were amplified from 27 Taxaceae and 12 Cephalotaxaceae species. *matK* sequences of 19 Pinaceae species were retrieved from Genbank (Hao *et al*, 2010). The phylogenetic tree was generated using conceptual-translated amino acid sequences. Selective influences were investigated using standard d*N*/d*S* ratio methods and sensitive techniques investigating the amino acid property changes resulting from nonsynonymous replacements in a phylogenetic context. Analyses revealed the presence of positive selection in *matK*s (*N*-terminal region, RT domain, and domain X) of Taxaceae and Pinaceae, and found positive destabilizing selection in *N*-terminal region and RT domain of Cephalotaxaceae *matK*. Various amino acid properties were found to be influenced by destabilizing positive selection (Hao *et al*, 2010). Amino acid sites relating to these properties and to different secondary structures were found and had the potential to affect group II intron maturase function. Despite the evolutionary constraint on the rapidly evolving *matK*, this protein evolves under positive selection in gymnosperm. Several regions of *matK* have experienced molecular adaptation which fine-tunes maturase performance.

Metabolomics

Metabolomics of medicinal plants is related to functional genomics and systems biology. A historical account of the introduction and evolution of metabolite profiling into today's modern comprehensive metabolomics approach is provided. Many of the technologies are used in metabolomics, including optical spectroscopy, nuclear magnetic resonance, mass spectrometry, and so on. The role of bioinformatics and various methods of data visualization are used for the study of metabolomics in medicinal plants (Taylor *et al*, 2002). The development of metabolomics, current literature, technological approaches, and bioinformatic tools are applied to molecular pharmacognosy.

Analysis of cells at the metabolic level has a number of advantages over the more conventional transcriptome and proteome analyses (Adams *et al*, 1999; Katona, Sass, and Mo, 1999; Tretheway, Krotzky, and Willmitzer, 1999): (1) Changes in gene and protein expression can cause amplified changes in metabolism, and make it easy in detection; (2) Metabolome technology does not require the complete genome sequence or large EST databases, and do many transcriptome and proteome approaches; (3) There are fewer metabolite types than genes or proteins: in the order of 1000 per organism compared to several thousand genes for the smallest bacterial genomes and tens of thousands of genes for complex multi-cellular organisms; and (4) The technology is more generic, as a given metabolite, unlike a transcript or protein, is the same in every organism.

Application

One of the theoretical principles underlying molecular pharmacognosy is pharmacognosy, which focuses on the identification and authentication of crude drugs (Huang, Yang, and Wang, 2004). Molecular pharmacognosy focuses on the relationship among the genetic diversity, DNA expression, and quality evaluation of crude drugs. Genetic expression is not only one major work of molecular pharmacognosy but also an important link between molecular biology and molecular pharmacognosy (Huang *et al*, 2009a; Zheng, 1990a; 1990b).

Origin

Medicinal herbs have multi-origin involving much of the confusion of species. It is common to find that a given material has been called several different names and that several different herbs share one name (Xiao, 1998; Zhang *et al*, 1995). Morphological characterization is a conventional method for the identification of crude drugs, while the results are controversial (Deng, 1994). With the development of molecular biology and cloning technology, the information on genetic diversity has been used for analyzing biological species. Some molecular identification methods, such as RAPD, SSR, AR-PCR, multiplex amplification refractory mutation system (MARMS), anchored primer amplification polymorphism (APAPD), and PCR-restriction fragment

length polymorphism (PCR-RFLP), have been successfully used for the cultivar identification (Huang *et al*, 2009b).

ISSR assay (Zietkiewicz, 1994) is based on the amplification of regions between inversely oriented closely spaced microsatellites. In recent years, researchers have focused on microsatellites such as SSR, VNTR, simple sequence length polymorphism (SSLP), and short tandem repeats (STR). Microsatellites (SSR, STR, SSLP, and VNTRs) are the class of repetitive DNA sequences present in all living organisms. Particular characteristics of microsatellites, such as their presence in the genomes of all living organisms, high level of allelic variation, co-dominant mode of inheritance, and potential for automated analysis make them an excellent tool for a number of approaches like genotyping, mapping, and positional cloning of genes. The three most popular types of markers containing microsatellite sequences that are presently used are: (1) SSR, generated by amplifying in a PCR reaction with the use of primers complementary to flanking regions; (2) ISSR, based on the amplification of regions between inversely oriented closely spaced microsatellites; and (3) Selective amplification of microsatellite polymorphic loci (SAMPL), which utilizes AFLP methodology, with one exception, for the second amplification, one of the starters is complementary to the microsatellite sequence. The usefulness of the three above-mentioned markers for numerous purposes has been well documented for plants. All of these markers are the class of repetitive DNA sequences present in all living organisms. The key technologies used in proteomics are one- and two-dimensional gel electrophoresis to identify the relative mass of a protein and its isoelectric point (Williams *et al*, 1990; Williams, Coxhead, and Mathers, 2003). Affinity chromatography, fluorescence resonance energy transfer, and surface plasmon resonance are used to identify protein-protein or protein-DNA interactions. X-ray tomography is conducted to determine the location of proteins or protein complexes in labeled cells. Furtherly, fluorescent proteins (FP) like green fluorescent protein (GFP), yellow FP, cyan FP or red FP are frequently used to study cellular events such as localization of proteins to membranes and to cellular organelles. They can mark homogenous populations of specialized cells whose gene expression profiles should be determined by DNA microarray analysis.

It was reported that RFLP had an obvious advantage in identifying the genetic variation, analyzing the species relationship, and selecting new medicinal resources (Dang and Sun, 1986; Cheng, 1994). In addition, the determination of seed purity is a quality control requirement in the production to avoid unacceptable contamination with self-inbred seeds. The trial based on morphological markers is time consuming, space demanding, and usually performed off season (Crockett *et al*, 2000; Zhang *et al*, 2003). Molecular maker techniques provide a new approach for this problem.

Identification

It is difficult to perform DNA analysis in plant and animal materials due to their high DNA degradation. Based on the amplification of nuclear DNA fragments, PCR assay does not require prior knowledge of the target genome, and because of its simplicity, rapidity, high performance, and good quality, the technique is being widely used in all aspects of molecular biology. RAPD which is based on PCR has developed for analyzing DNA polymorphism, separating DNA fragments and forming a specific banding pattern of genuine medicinal materials (Williams *et al*, 1990). RAPD technology has provided a rapid and accurate method for the identification of crude drugs and especially the animals. The identification which is based on the methods of molecular biology is called "molecular marker identification" (Zhang and Yan, 2001).

The term "DNA barcode" for global species identification was first coined by Hebert *et al* (2003) and has gained worldwide attention. The method not only facilitates the work of professional taxonomists but also makes it possible for nonexperts to identify species rapidly, accurately, and efficiently (Pang *et al*, 2010). Recognition of animals, plants, and fungi has been performed using this technique. Mitochondrial gene encoding cytochrome coxidase submit 1 is a favorable region for use as a DNA barcoding in most animal species and even in some fungal species, including those of the groups *Aacomycota*, *Basidiomycota* and *Chytridiomycota*. In plants, several candidate DNA barcodes have attracted the attention of many researchers. For example, Song *et al* (2009)

examined the usefulness of eight candidate barcodes (*trnH-psbA*, *rbcL*, *rpoB*, *rpoC1*, *ndhJ*, *accD*, *YCF5*, and *nr*ITS) on species indentification of Polygnonaceae in *Chinese Pharmacopoeia*. The findings show that the plastid region *trnH-psbA* can not only distinguish ten species of Polygonaceae in *Chinese Pharmacopoeia*, but also recognize eight other species of Polygonaceae including their adulterants. Furthermore, plastid region *rbcL* combined with *rpoC1* or *nr*ITS can also distinguish closely related species (Song *et al*, 2009). Additionally, Kress *et al* (2005) compared 10 loci for authenticating closely related species in seven plant families and 99 species belonging to 88 genera in 53 families, and reported that the *psbA-trnH* spacer and the internal transcribed spacer could be used as a pair of potential barcodes for identifying widely divergent angiosperm taxa. Lahave *et al* (2008) analyzed 1084 plant species and demonstrated that a portion of the plasid *matK* gene could be a universal DNA barcode for flowering plants. In another study, seven potential DNA regions (*psbA-tinH*, *matK*, *rbcL*, *rpoC1*, *ycf*5, *ITS2*, and *ITS*) for their suitability as DNA barcodes across 8557 medicinal plants and closely related samples belonging to 5905 species from 1010 diverse genera of 219 familes in seven phyla (Angiosperms, Gymnosperms, Ferms, Monocotyledons, Dicotyledons, Algae, and Fungi) were compared. The data suggest that the second internal transcribed spacer (ITS2) of nuclear ribosomal DNA can be potentially used as a standard DNA barcode to identify medicinal plants and their closely related species and a novel universal barcode for the identification of a broader range of plant taxa. Hence, DNA barcoding is an efficient and powerful tool for identification of traditional medicines and their adulterants (Chen *et al*, 2010; Yao *et al*, 2009).

Cultivation

Disease and insect control is usually a major concern in cultivation and the conventional pesticide application causes damage to human, livestock, and environment. Molecular technology has been successfully used to obtain transgenic medicinal plants which have disease and insect resistance. In addition, some molecular identification methods have also been used to analyze genetic varieties, and to investigate the formation mechanisms of medicinal herbs with good quality, including RFLP and RAPD (Chen, 1989; Jiangxi College of Traditional Chinese Medicine, 1980).

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 (Canter, Thomas, and Ernst, 2005). Five cultivars of *Aurantii Fructus* (*Citrusaurantium*) were studied to reveal the genetic relationships among the cultivars (Zuo, Zhu, and Liu, 2005). The unweighted pair group method with arithmetic mean (UPGMA) dendrogram yielded by ISSR revealed the relationships within each cultivar. The primers selected could generate enough bands for analysis, and the reproducibility is better than RAPD. So ISSR assay could be a kind of molecular marker which is suitable for the investigation of the variation of inter- or intra-species. And the result could be used to establish the fingerprinting of these herbs.

The genetic diversity of eight species of *Caragana* Fabr. in Ordos Plateau such as *Caragana tibetica* Kom., *C. korshinskii* Kom., *C. Stenophylla* Pojark., *C. roborovskyi* Kom., *C. opulens* Kom., *C. purdomii* Rehd., *C. brachypoda* Pojark., and *C. intermedia* Kuang et H.C. Fu. was analyzed (Yang, Yang, and Yang, 2006) by ISSR assay. And the result showed their genetic diversity (Qiu, Fu, and Wu, 2003) assessed the population of genetic diversity and genetic structure of *Changium smyrnioides* Wolff. and *Chuanminshen violaceum* Sheh et Shan. Fingerprinting was set up for the identification of these kinds of crude drugs by ISSR assay.

Zhou, Jing, and Li (2005) characterized eight cultivars and two virus-free lines micro-propagated by tip tissue culture of *Rehmannia glutinosa* Libosch. in Huai zone and assessed their genetic diversities by ISSR technique. Cao, Li, and Sun (2005) analyzed the sibling relationship and genetic diversity of four certified *Gentianae Radix* species and the result showed the bands yielded by ISSR assay could be used for the analysis of *Gentianae Radix* species in genetic diversity and the setup of its fingerprinting. In addition, the genetic diversities of some species in *Dendrobium* Sw. (Shen, Xu, and Wan, 2005) have also been studied by ISSR assay.

Plant metabolomics

Metabolomics has been defined as the technology

designed to give us the broadest, essentially nontargeted insight into the richly diversed population of small molecules present in living things. Recent progress of plant metabolism and the generic value of the information were gathered in metabolomics and its application. Metabolomics is predicted to play a significant role in bridging the phenotype-genotype gap and thus in assisting us in our desire for full genome sequence annotation as part of the quest to link gene to function (Hall, 2006).

The genus *Strychnos* Linn. is very well known as the plants providing one of the most famous poisons. Because of their toxicity, many of these species have been used as arrow poisons or in ordeals. The genus *Strychnos* Linn. comprises about 200 species in the world and can be subdivided into three geographically separated groups of species: one in Central and South America (at least 73 species), one in Africa (75 species), and one in Asia including Australia and Polynesia (about 44 species) (Xu, Si, and Liu, 2009).

In a study on metabolomic analysis of *S. nux-vomica* L., *S. icaja* Baill., and *S. ignatii* Bereius. extracts by ^{1}H-NMR spectrometry and multivariate analysis techniques, ^{1}H-NMR spectrometry and multivariate analysis techniques were applied for the metabolic profiling of three *Strychnos* Linn.: *S. nux-vomica* (seeds, stem bark, and root bark), *S. ignatii* (seeds), and *S. icaja* (leaves, stem bark, root bark, and collar bark). The principal component analysis (PCA) of the ^{1}H-NMR spectra showed a clear discrimination between all samples, using the three first components. The key compounds responsible for the discrimination were brucine, loganin, fatty acids, and *S. icaja* alkaloids such as icajine and sungucine. The method was then applied to the classification of several "false angostura" samples. These samples were, as expected, identified as *S. nuxvomica* by PCA, but could not be clearly discriminated as root bark or stem bark samples after further statistical analysis. Their study then showed that the major compounds responsible for the discrimination were brucine, fatty acids, loganin, and several Strychnos icaja alkaloids (mainly icajine and sungucine). Strychnine, though present in various amounts in all extracts analyzed, was not the key compound for the discrimination of samples (Frédérich *et al*, 2004).

Production of active compounds

The purpose of molecular pharmacognosy is to obtain high production of active compounds in crude drugs. With the development of molecular biology, techniques such as transgenic and antisense technology offer new perspectives for improving the production of active compounds. Plants secondary metabolism, such as the bioactive natural products in medicinal plants, and their biosynthesis are generally regulated by transcriptional control of the relevant enzymatic genes, which is one of the key research areas of molecular pharmacognosy (Wang, Morris-Natschke, and Lee, 2007; Cui *et al*, 2007; Wei *et al*, 2009).

Decursin and its structural isomer decursinol angelate are major secondary metabolites in the root of *Angelica gigas* Nakai which possess several chemo-therapeutic properties. The results suggested that strain PYR1001 could be successfully used to transform decursin for the production of decursinol, a compound known to have cancer chemopreventive activity (Kim, Lee, and Cha, 2010). The hairy roots of *Polygonum multiflorum* Thunb. were able to convert the aromatic exogenous substrate, thymol, into its glycoside. The glycosylated product, namely DMP, was isolated and purified, the structure of which was determined as 5-methyl-2-(1-methylethyl) henyl-β-*D*-glucopyranoside. And the distribution of DMP in the medium or culture was varied in different co-cultivated periods, and co-cultivated for 5 d, it mainly existed in the medium (Dong *et al*, 2009). Two aromatic substrates, paeonol, and emodin, were biotransformed by using transgenic crown galls of *Panax quinquefolium* L. Four biotransformed products were isolated and identified by physicochemical and spectral methods. A β-glucoside and a 1-(2,4-dimethoxyphenyl)-ethanone were isolated from the suspension cultures after 7 d incubation of substrate. Another β-glucoside and a hydroxylated derivative, citreorosein, were also obtained. The results demonstrated that transgenic crown galls of *P. quinquefolium* had the capacities to catalyze glycosylation, hydroxylation, and methylation reactions in the plant cells on those aromatic compounds (Ma *et al*, 2010).

Prospect

Molecular pharmacognosy has developed rapidly

in recent years and is playing an important role in the identification, production, and evaluation of medicinal plants and animals. The development of molecular pharmacognosy could be anticipated in the following areas.

Genetic diversity and molecular systematics

The information on genetic diversity could guide the identification and classification of plant and animal species. However, the genetic background of most medicinal materials is unknown. Molecular marker technology has an obvious advantage in evaluating genetic diversity of medicinal plants at population, individual, and gene levels, and constructing phylogenetic trees which are based on the chloroplast and nuclear genome sequences by dividing the evolutionary significant units, analyzing the species relationship, selecting endangered medical plants and animals, and developing new medicinal resources. At present, the main genes in molecular systematics include *rbcL*, *matK*, *rps4*, *18S rDNA*, and *ITS*, *etc*. Moreover, the development of biosystematics and molecular systematics provides a powerful weapon for the classification and identification of crude drugs, compared with the conventional morphological assay, which is affected by man-made factors (Fan *et al*, 1996; Deng *et al*, 1994; Fan, Qi, and Shi, 1983; Qiu *et al*, 1998).

Molecular markers and important traits

Genetic diversity is the foundation, while molecular markers and important traits including production, quality (active compounds and secondary metabolites) and resistance (insect, disease, drought, and salinity) are the purpose of natural resources exploitation, and transformation during the practice of molecular pharmacognosy (Deng, 1994). The technology of cell and genetic engineering accelerates the application of molecular pharmacognosy in economic construction.

Molecular marker can be utilized to select the fine germplasm, assist breeding or construct genetic maps. The quality of medicinal materials is highly affected by their genetic basis and ecological environment. Genes of medicinal materials play an important role in the biosynthesis of secondary metabolites, which are mainly used to obtain active compounds. Genetic material is the foundation and ecological factors that affect the accumulation of secondary metabolites. Therefore, genetic engineering of secondary metabolites and high production of active compounds play an important role in the protection of medicinal resources.

A transgenic organism is one which has been modified with genetic materials from other species. Transgenic plants have been engineered to possess several desirable traits, including expression of complex natural protein, continual acquisition from the blood and milk of animals, increased production of active compounds, and decreased expression of toxic genes. In addition, they can also pass the modification onto future generations by breeding with other members of the same species (Huang, 2006). The development of an alternative production method using tissue culture is a topic of great interest. However, there is still no such transgenic drug on the market. Recently, the techniques such as hairy roots and crown gall cultures provide new methods for fast growth and stable high-level production of active compounds.

With pesticide pollution increasing annually, there has been an increasing demand for breeding of green medicinal plants which contain no toxic or harmful elements for human health and implementing of good agriculture practice (GAP) in production (Wei, Chen, and Guo, 2004). Transgenic technology has been successfully used to obtain transgenic crops protected from virus, insect, and disease. The coat protein of tobacco mosaic virus (TMV) has been introduced into tobacco and provided resistance to virus. The bacterium *Bacillus thuringiensis* (Bt) produces proteins which are specifically toxic to a variety of insect species. Modified genes have been derived from Bt2, a toxin gene cloned from one *B. thuringiensis* strain. Transgenic tobacco plants expressing these genes synthesize insecticidal proteins which protect them from feeding damage by larvae of the tobacco hornworm (Vaeck *et al*, 1987). It is believed that there will be much progress in commercial sale and widespread planting of transgenic medicinal herbs in the near future.

Conclusion

The molecular pharmacognosy is growing up with the development of molecular biology. The methods of molecular marker technology have developed rapidly in recent years, and there has been much progress on

identification, resource protection, formation mechanism, quality evaluation, and production of active compounds in crude drugs. However, each method has certain limitations, which includes validity, accuracy, normalization, and cost, so there are still difficult problems needing to be solved in molecular marker assays, such as standard operation procedure, accurate determination, and complete automation. Furthermore, the great breakthrough of molecular pharmacognosy could be anticipated in DNA fingerprint analysis, cultivar improvement, DNA identification, and a global DNA barcoding system in the future.

References

Adams MA, Chen ZL, Landman P, Colmer TD, 1999. Simultaneous determination by capillary gas chromatography of organic acids, sugars and sugar alcohols in plant tissue extracts as their trimethylsilyl derivatives. *Annu Biochem* 266: 77-84.

Barthet MM, Hilu KW, 2008. Evaluating evolutionary constraint on the rapidly evolving gene *matK* using protein composition. *J Mol Evol* 66: 85-97.

Brautigam A, Shrestha RP, Whitten D, Wilkerson CG, Carr KM, Froehlich JE, Weber AP, 2008. Low-coverage massively parallel pyrosequencing of cDNAs enables proteomics in non-model species: Comparison of a speciesspecific database generated by pyrosequencing with databases from related species for proteome analysis of pea chloroplast envelopes**.** *J Biotechnol* 136(1-2): 44-53.

Canter PH, Thomas H, Ernst E, 2005. Bringing medicinal plants into cultivation: Opportunities and challenges for biotechnology. *Trends Biotechnol* 23: 180-185.

Cao YN, Li Q Z, Sun Y, 2005. Analysis of genetic diversity in certified *Radix Gentianae* by RAPD and ISSR. *Chin Tradit Herb Drugs* 36(1): 100-103.

Chen MJ, Wu YJ, Fan XH, Chen YY, 2003. Identification of the habitat of *Linguisticum chuangxiong* with chromatographic fingerprinting. *Chin J Chin Mater Med* 28: 606-610.

Chen SL, Yao H, Han HP, Liu C, Song JY, Shi LC, Zhu YJ, Ma XY, Gao T, Pang XH, Luo K, Li Y, Li XW, Jia XC, Lin YL, Leon C, 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLos Biol* 5: e8163.

Chen Y, 1989. Achievement and experience in the introduction of medicinal plants. *Chin J Chin Mater Med* 14: 3-6.

Cheng LG, 1994. RFLPs and its principle. *J Biol* 3: 7-8.

Crockett PA, Bhalla PL, Lee CK, Singh MB, 2000. RAPD analysis of seed purity in a commercial hybrid cabbage (*Brassica oleracea* var. *capitata*) cultivar. *Genome* 43: 317-321.

Cui GH, Huang LQ, Tang XJ, Qiu DY, Wang XY, Fu GF, 2007. Functional genomics studies of *Salvia miltiorrhiza* I-Establish cDNA microarray of *S. miltiorrhiza*. *Chin J Chin Mater Med* 32: 1137-1141.

Dang JJ, Sun ZX, 1986. DNA restriction fragment length polymorphism. *Prog Biochem Biophys* 4: 12-16.

Deng WG, 1994. Development of the methods of genetic polymorphisms. *Bull Biol* 29: 7-9.

Deng YC, Li J, Hua HM, Yang BZ, Fu BY, 1994. Approach to some problems in the present research on the exploitation and utilization of plant resources in China. *J Plant Res Environ* 3: 56-69.

Dong QF, Jia JZ, Zhu JH, Yu RM, 2009. Biotransformation of thymol by hairy roots of transgenic *Polygonum multiflorum*. *J Chin Med Mater* 32: 1495-1499.

Duffy AM, Kelchner SA, Wolf PG, 2009. Conservation of selection on *matK* following an ancient loss of its flanking intron. *Gene* 438: 17-25.

Fan JA, Yi SP, Zhang AJ, Zhao PP, Luo ZZ, Xiao XH, Li ZJ, 1996. Restricting effects of geologic background system on genuine crude drugs in Sichuan. *Chin J Chin Mater Med* 21: 12-14.

Fan JF, Qi DL, Shi ZL, 1983. Distribution and utilization of *Oleum Viticis Negundo* resources. *Chin Tradit Herb Drugs* 14(5): 35-36.

Frédérich M, Choi YH, Angenot L, Harnischfeger G, Lefeber AW, Verpoorte R, 2004. Metabolomic analysis of *Strychnos nux-vomica*, *Strychnos icaja* and *Strychnos ignatii* extracts by ^{1}H nuclear magnetic resonance spectrometry and multivariate analysis techniques. *Phytochemistry* 65: 1993-2001.

Gao WY, Qin EQ, Xiao XH, Yu H, Gao GY, Chen SB, Zhao YJ, Yang SL, 2001. Analysis on genuineness of *Angelica sinensis* by RAPD. *Chin Tradit Herb Drugs* 32(10): 926-929.

Guo BL, Lin S, Feng YX, Zhao YJ, 2002a. Primary research on genetic relationship among main populations of *Salvia miltiorrhiza* and genuineness of herb. *Chin Tradit Herb Drugs* 33(12): 1113-1116.

Guo LP, Liu JY, Ji L, Huang LQ, 2002b. The naphtha composing characteristics of geoherbs of *Atractylodes lancea*. *Chin J Chin Mater Med* 27: 814-819.

Hall RD, 2006. Plant metabolomics: From holistic hope, to hype, to hot topic. *New Phytologist* 169: 453-468.

Hao D, Mu J, Chen SL, Xiao PG, 2010. Physicochemical evolution and positive selection of the gymnosperm matK proteins. *J Genet* 89: 81-90.

Hebert PDN, Ratnasingham S, deWaard JR, 2003. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proc R Soc Biol Sci SerB* 270: S96-S99.

Huang LQ, 2006. *Molecular Pharmacognosy*. Beijing Medical University Press: Beijing.

Huang LQ, Chen ML, Xiao PG, 1999. An approach to some problems on utilization of medicinal plant resource in China. *Chin J Chin Mater Med* 24: 70-73.

Huang LQ, Xiao PG, Guo LP, Gao WY, 2009b. Molecular pharmacognosy: A new borderline discipline. *Sci China Ser C* 39: 1101-1110.

Huang LQ, Yang B, Wang M, 2004. Modern biological basis and model hypothesis of geoherbs. *Chin J Chin Mater Med* 29: 494-496.

Huang LQ, Yuan Y, Cui GH, Dai ZB, Xiao PG, 2009a. Molecular pharmacognosy: A new borderline discipline. *Nat Prod Commun* 4: 1611-1613.

Jiangxi College of Traditional Chinese Medicine, 1980. *Cultivation of Medicinal Plants*. Shanghai Science & Technology Press: Shanghai.

Kang AL, Yang YS, Sun JW, 2002. The development of molecular

pharmacognosy. *Acta Chin Med Pharm* 30: 56-58.

Katona ZF, Sass P, Moln´ar-Perl I, 1999. Simultaneous determination of sugars, sugar alcohols, acids and amino acids in apricots by gas chromatography-mass spectrometry. *J Chromatogr A* 847: 91-102.

Kim KY, Lee S, Cha CJ, 2010. Biotransformation of plant secondary metabolite decursin by *Mycobacterium* sp. PYR1001. *J Agric Food Chem* 58: 2931-2934.

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH, 2005. Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci USA* 102: 8369-8374

Lahave R, Van der Bank M, Bogarin D, Warner J, Pupulin F, 2008. DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci USA* 105: 2923-2928.

Li Y, Luo HM, Sun C, Song JY, Sun YZ, Wu Q, Wang N, Yao H, Steinmetz A, Chen SL, 2010. EST analysis reveals putative genes involved in glycyrrhizin biosynthesis. *BMC Genomics* 11: 268-278.

Liu CX, 2005. The understanding drug metabolism and pharmaco-kinetics of traditional Chinese medicines. *Asian J Drug Metabol Pharmacokinet* 5: 25-38.

Liu CX, Xiao PG, Peng Y, Song NN, 2009. Challenges in research and development of traditional Chinese Medicines. *Chin Herb Med* 1: 1-28.

Lou ZQ, 1955. *Pharmacognosy*. People's Medical Publishing House: Beijing.

Ma LR, Jiang ZH, 2000. *Biological Chips*. Chemical Industry Press: Beijing.

Ma LY, Zhao HQ, Tian ZK, Wang XY, Qu Y, Cui HB, 2003. Studies on chemical pattern recognition of quality assessment of *Radix Astragali*. *Chin Tradit Herb Drugs* 34: 460-462.

Ma WL, Yan CY, Zhu JH, Duan GY, Yu RM, 2010. Biotransformation of paeonol and emodin by transgenic crown galls of *Panax quinquefolium*. *Appl Biochem Biotechnol* 160: 1301-1308.

Mukherjee PK, Pitchairajan V, Murugan V, Sivankaran P, Khan Y, 2010. Strategies for revitalization of traditional medicine. *Chin Herb Med* 2: 1-15.

Pang XH, Song JY, Zhu YJ, Xie CX, Chen SL, 2010. Using DNA barcoding to identify species within Euphorbiaceae. *Planta Med* 76: 1784-1786.

Patwardhan B, 2005. Ethnopharmacology and drug discovery. *J Ethnophamacol* 100: 50-52.

Patwardhan B, Warude D, Pushpangadan P, Bhatt N, 2005. Ayurveda and traditional Chinese medicine: A comparative overview. *Evid Based Complement Alternat Med* 2: 465-473.

Qiu F, Fu JM, Jin DM, Wang B, 1998. The molecular detection of genetic diversity. *Chin Biodiver* 6: 143-150.

Qiu YX, Fu CX, Wu FJ, 2003. Analysis of population genetic structure and molecular identification of *Changium smyrnioides* and *Chuanminshen violaceum* with ISSR marker. *Chin J Chin Mater Med* 28: 598-603.

Qu LJ, Gu HN, Hu P, Chen ZL, 1998. *An Introduction to Modern Biological Technology*. China Higher Education Press & Springer-Verlag Press: Beijing.

Shen Y, Xu C, Wan XF, 2005. Application of ISSR-PCR to identification of *Dendrobium* Sw. species. *Chin Tradit Herb Drugs* 36: 423-427.

Song JY, Yao H, Li Y, Li XW, Lin YL, Liu C, Han JP, Xie CX, Chen SL, 2009. Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique. *J Ethnopharmacol* 124: 434-439.

Taylor J, King RD, Altmann T, Oliver F, 2002. Application of metabolomics to plant genotype discrimination using statistics and machine learning. *Bioinformatics* 18(S2): S241-S248.

Tretheway R, Krotzky AJ, Willmitzer L, 1999. Metabolic profiling: A rosetta stone for genomics? *Curr Option Plant Biol* 2: 83-85.

Vaeck M, Reynaerts A, Hofte H, Jansens S, De Beuckeleer M, Dean C, Zabeau M, Van Montagu M, Leemans J, 1987. Transgenic plants protected from insect attack. *Nature* 328: 33-37.

Vera JC, Wheat CW, Fescemyer HW, Frilander MJ, Crawford DL, Hanski I, Marden JH, 2008. Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Mol Ecol* 17(7): 1636-1647.

Verpoorte RYH, Choi YH, Kim KK, 2005. Ethnopharmacology and systems biology: A perfect holistic match. *J Ethnopharmacol* 100: 53-56.

Vos P, Hogers R, Bleeker M, 1995. AFLP: A new technique for DNA fingerprint. *Nucleic Acids Res* 23: 4407-4414.

Wang M, Lamers RJ, Korthout HA, van Nesselrooij JH, Witkamp RF, van der Heijden R, Voshol PJ, Havekes LM, Verpoorte R, van der Greef J, 2005. Metabolomics in the context of systems biology: Bridging traditional Chinese medicine and molecular pharmacology. *Phytotherapy Res* 19: 173-182.

Wang R, Zhang Q, Jia ZP, Fan JJ, Xu LT, Xie H, 2003. Application of chromatograph-digitized fingerprint spectrum in identification of indigenous medicinal materials *Radix* et *Rhizoma Rhei* from Gansu province. *J Chin Med Mater* 26: 484-486.

Wang X, Morris-Natschke SL, Lee KH, 2007. New developments in the chemistry and biology of the bioactive constituents of Tanshen. *Med Res Rev* 27: 133-148.

Wang Y, Bi KS, 2003. Fingerprint construction of *Panax notoginseng*. *Chin J Chin Mater Med* 28: 316-320.

Wei G, Matthew LH, Huang LQ, Cui GH, Wang XY, Kong JQ, Yang B, Reuben J, Peters RJ, 2009. A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. *Org Lett* 11: 5170-5173.

Wei JH, Chen SL, Guo QS, 2004. Current situation and development of GAP in China. *Res Inform Tradit Chin Med* 6: 4-8.

Williams EA, Coxhead JM, Mathers JC, 2003. Anticancer effects of butyrate: Use of micro-array technology to investigate mechanisms. *Proc Nutr Soc* 62(1): 107-115.

Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Ting SV, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18: 6531-6535.

Xiao PG, 1985. Review and development of Chinese medicinal plant material resources. *Chin Tradit Herb Drugs* 16: 33-34.

Xiao PG, 1988. Utilization of Chinese medicinal plant material resources. *Acta Bot Yunnanica* 10: 1-3.

Xie ZW, 2001. A brief introduction to molecular pharmacognosy. *Chin J Chin Mater Med* 26: 216-224.

Xu YY, Si DY, Liu CX, 2009. Research on bioresponse of active compounds of *Strychnos nux-vomica* L. *Asian J Pharmacodyn Pharmacokinet* 9: 179-201.

Yang JY, Yang J, Yang MB, 2006. ISSR analysis on medicinal plants of *Caragana Fabr*. in Ordos Plateau. *Chin Tradit Herb Drugs* 37: 1562-1566.

Yao H, Song JY, Ma XY, Liu C, Li Y, Xu HX, Han JP, Duan LS, Chen SL, 2009. Identification of *Dendrobium* species by a candidate DNA barcode sequence: The chloroplast psbA-trnH intergenic region. *Planta Med* 75: 667-669.

Yu HH, Yang ZL, Yang X, Tan ZF, Shu X, Liu RN, 2010. Advances in studies on ITS sequence of medicinal plants germplasm resources. *Chin Tradit Herb Drugs* 41(3): 491-496.

Zhang GJ, Yan NY, 2001. *Identification of Chinese Materia Medica*. Science Press: Beijing.

Zhang HC, Liu JM, Lu HY, Gao SL, 2009. Enhanced flavonoid production in hairy root cultures of *Glycyrrhiza uralensis* Fisch by combining the over-expression of chalcone isomerase gene with the elicitation treatment. *Plant Cell Rep* 28: 1205-1213.

Zhang HY, Zhao RH, Yuan CQ, Sun CQ, Zhang ZY, 1995. Species of Chinese medicinal material resources. *Chin J Chin Mater Med* 20: 387-390.

Zhang ZY, Li P, Qi H, Li J, 2003. Analysis on the geologic background and physicochemical properties of cultivated soil of *Flos Lonicerae* in the geo-authentic and non-authentic pruducing areas. *Chin J Chin Mater Med* 28: 114-117.

Zheng JS, 1990a. Formation and development of authentic medicinal materials (I). *J Chin Med Mater* 13: 39-40.

Zheng JS, 1990b. Formation and development of authentic medicinal materials (II). *J Chin Med Mater* 13: 43-45.

Zhou YQ, Jing JZ, Li ZY, 2005. ISSR identification of genetic diversity of *Rehmannia glutinosa* in Huai zone. *Chin Tradit Herb Drugs* 36: 257-261.

Zietkiewicz E, Rafalski A, Labuda D, 1994. Genomic fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20: 176-183.

Zuo YJ, Zhu PL, Liu Q, 2005. Genetic relationships among *Fructus Aurantii* cultivars revealed by ISSR. *Chin J Chin Mater Med* 30: 1416-1419.