



Contents lists available at ScienceDirect

Chinese Herbal Medicines

journal homepage: www.elsevier.com/locate/chmed

Review

Pharmaceutical resource discovery from traditional medicinal plants:
Pharmacophylogeny and pharmacophylogenomicsDa-cheng Hao^{a,*}, Pei-gen Xiao^b^aBiotechnology Institute, School of Environment and Chemical Engineering, Dalian Jiaotong University, Dalian 116028, China^bInstitute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Beijing 100193, China

ARTICLE INFO

Article history:

Received 3 July 2019

Revised 11 September 2019

Accepted 25 September 2019

Available online 12 March 2020

Keywords:

pharmaceutical resource discovery

pharmacophylogenomics

pharmacophylogeny

traditional medicinal plants

ABSTRACT

The worldwide botanical and medicinal culture diversity are astonishing and constitute a Pierian spring for innovative drug R&D. Here, the latest awareness and the perspectives of pharmacophylogeny and pharmacophylogenomics, as well as their expanding utility in botanical drug R&D, are systematically summarized and highlighted. Chemotaxonomy is based on the fact that closely related plants contain the same or similar chemical profiles. Correspondingly, it is better to combine morphological characters, DNA markers and chemical markers in the inference of medicinal plant phylogeny. Medicinal plants within the same phylogenetic groups may have the same or similar therapeutic effects, thus forming the core of pharmacophylogeny. Here we systematically review and comment on the versatile applications of pharmacophylogeny in (1) looking for domestic resources of imported drugs, (2) expanding medicinal plant resources, (3) quality control, identification and expansion of herbal medicines, (4) predicting the chemical constituents or active ingredients of herbal medicine and assisting in the identification and determination of chemical constituents, (5) the search for new drugs sorting out, and (6) summarizing and improving herbal medicine experiences, etc. Such studies should be enhanced within the context of deeper investigations of molecular biology and genomics of traditional medicinal plants, phytometabolites and metabolomics, and ethnomedicine-based pharmacological activity, thus enabling the sustainable conservation and utilization of traditional medicinal resources.

© 2020 Tianjin Press of Chinese Herbal Medicines. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license.

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Contents

1. Introduction	105
2. Chemotaxonomy	105
2.1. Macromolecular systematics and micromolecular systematics	105
2.2. Diversification of biogenic pathways of chemotaxonomic markers: rationale	105
3. Practical application	106
3.1. Looking for domestic resources of imported drugs	106
3.2. Expanding medicinal plant resources	106
3.3. Quality control, identification and expansion of herbal medicines	108
3.4. Help to predict chemical constituents or active ingredients of herbal medicine and assist in identification and determination of chemical constituents	111
3.5. Application of pharmacophylogeny in the search for new drugs	111
3.6. Application of pharmacophylogeny in sorting out, summarizing and improving herbal medicine experiences	113
3.7. Pharmacophylogeny vs. plant systematics	114
4. Application of informatics in the study of pharmacophylogeny	115

* Corresponding author.

E-mail address: hao@djtu.edu.cn (D.-c. Hao).

5. Conclusion and future perspective	115
Declaration of Competing Interest	116
Acknowledgments	116
Supplementary materials.	116
References	116

1. Introduction

Plants provide healthcare for up to 80% of the world's population (Saslis-Lagoudakis et al., 2014). Pharmaceutical plant kinship (plant pharmacophylogeny) is a new subject developed during the practice of medicinal research by the mutual penetration of ethnopharmacology, plant systematics, phytochemistry, pharmacology and bioinformatics (Hao, Xiao, Liu, Peng & He, 2014, 2015). It specializes in studying and discussing the internal relations and laws among plant kinship (phylogenetic/evolutionary relationship), chemical composition and preventive/curative effect, and uses them to guide the research and improvement of medicines. Pharmacological and phytochemical characterization contributes to both the safe use of herbal medicines and the identification of leads for drug development. Plant species with more traditional medicinal uses are studied with modern methods significantly more often (Souza, Williamson & Hawkins, 2018). Ethnobotanical studies, along with pharmacological and phytochemical characterization, are essential for enriching the knowledge base of plant pharmacophylogeny.

Medicinal plants are part of hundreds of thousands of plants in the world, which have evolved over a long period of geological history. In the process of evolution, they form distant or near relationship, which is the concept of plant phylogeny. Similar species are not only similar in morphology, but also, due to genetic links, similar in physiological and biochemical characteristics, so the chemical composition between them is often more similar than that between them and the genetically distant species. Physiologically active components of medicinal plants are mostly secondary metabolites of plants, and their distribution in the plant kingdom is generally regular. Medicinal plants have inherent relationship among genetic relationship, chemical composition and curative effect. The inherent regularity between them may be found through advanced technical means.

The concept of pharmacophylogeny was initially proposed by Prof. Pei-gen Xiao in 1980 (Hao & Xiao, 2017a). At that time, researchers know little about the genetic basis of medicinal plants. It was not well known that “plants within the same phylogenetic groups may share similar genetics, thus possess similar chemical-producing potentials, which in turn may exhibit similar therapeutic effects”; such a logic is not present at the early development stage of plant pharmacophylogeny. Studies of pharmacophylogeny in 1980s and 1990s emphasize the values of phytochemical components in the phylogenetic arrangements of various medicinal plants. The traditional taxonomy relies on the morphological traits, and after 2000, the DNA markers are extensively used in the reconstruction of molecular phylogeny. However, the morphological characters are often misleading, and the use of different DNA markers often results in distinct phylogenetic trees. Researchers of pharmacophylogeny identified quite some phytochemical components of various medicinal plants, and realize that it is better to combine morphological characters, DNA markers and chemical markers in the inference of medicinal plant phylogeny. Some new concepts with regard to the bioprospecting and utilization of medicinal plants have been proposed, e.g., “ethnobotanical convergence” (Garnatje, Peñuelas & Vallès, 2017), which refers to the similar uses for plants included in the same node of a phylogeny. The pharmacophylogenetic approach, together with the “omics” revolution, shows how combining modern technologies with traditional ethnobotanical knowledge could be

used to identify potential new applications of medicinal plants. The following part of this article starts with the introduction and review of chemotaxonomy, followed by the summary of the various practical applications of plant pharmacophylogeny, including the new concept “pharmacophylogenomics”. Then, the application of informatics in the study of pharmacophylogeny is illustrated. This review is ended with the conclusion and future prospects.

2. Chemotaxonomy

2.1. Macromolecular systematics and micromolecular systematics

With the development of chemotaxonomy and chemical systematics, two branches might be distinguished, i.e., macromolecular systematics and small molecular systematics. The former includes some macromolecular compounds, e.g., proteins and enzymes (Johnson et al., 2019; Zhang et al., 2018); the latter includes some compounds with smaller molecular weight, and some secondary metabolites, such as alkaloids (Vanderplanck & Glauser, 2018), glycosides (Pertuit et al., 2018), terpenoids (Frezza et al., 2018) and sterols, which are of great value for classification, belong to this category. For instance, the chemosystematic implications of flavonoids, organic acids and pentacyclic triterpenes of a *Tilia tomentosa* population were discussed (Frezza et al., 2019), together with their pharmacological relevance and the traditional medicinal uses of this plant. The isolated coumarins may be the chemotaxonomic marker of *Psephellus* species (Nawrot, Budzianowski & Nowak, 2019). The germacranolides in *Centaureinae* species with stout apical spine ended bracts of flower (*Stizolophus balsamita* (Lam.) Cass. ex Takht.) and guaianolides in other species with appendages of the bracts without apical spine (*Psephellus sibiricus* (L.) Wagenitz) suggests a possible connection between the chemical structure of sesquiterpene lactones and morphology of flowers in the species of *Centaureinae* subtribe. The VOC (volatile organic compound) profiles could clearly distinguish 12 Brassicaceae vegetables (Liu et al., 2018a), and well reflected the classical morphological classification. Chemotaxonomy, however, has its limitation, e.g. when special habitats led to the evolution of specific secondary metabolites for protection. Cinchona alkaloids are well known to be best produced in a narrow range of altitudes. Chemotypes, ecotypes and genotypes should be considered together in elucidating the phylogeny of Daodi medicinal materials (geoherbs; Hao & Xiao, 2018).

Besides small molecules, the attention has been paid to the study of macromolecule compounds. For example, a transcriptome dataset-driven approach identifies additional diterpene synthase activities in the mint family (Johnson et al., 2019). Rubiaceae-type cyclopeptides of 20 *Rubia* species were systematically characterized and quantified by ultra performance liquid chromatography tandem mass spectrometry (UPLC-QqQ-MS/MS) and chemometrics (Zhang et al., 2018). Few large protein and polysaccharide were investigated from the perspective of chemotaxonomy, leaving a knowledge gap to fill.

2.2. Diversification of biogenic pathways of chemotaxonomic markers: rationale

Under certain external conditions, each plant species is able to produce various substances with its own characteristics. These substances represent its physiological and biochemical characteristics

(Hao & Xiao, 2017a). Each species has the same/similar physiological and biochemical characteristics with its genetically related species. The closer the genetic relationship is, the more common the species has, and vice versa. In the long historical process, when the external conditions of survival change, the physiological and biochemical characteristics of plants accordingly change (Hao & Xiao, 2015, 2018). Whether these changes are slow/progressive or fast/radical, they often result in the addition or alteration of some groups (e.g., oxidation, methylation, ethylation, benzoylation, etc.). These changes in physiological and biochemical characteristics indicate the evolutionary route and process of plants. More complex chemical structures usually have narrower phylogenetic distributions. So far, these laws still have some reference significance. The theoretical explanation of diversification of biogenic pathways of alkaloids is summarized in Table 1. For instance, *N*-methylcytisine is abundant in Berberidaceae, Fabaceae, and Scrophulariaceae, but these three families are obviously not very close, and the parallel evolution of the alkaloid biosynthetic pathway is inferred. The distribution of chemical constituents in the plant kingdom often presents various and complex situations, e.g., similar or even identical constituents are distributed in distinct taxonomic groups; different types of chemical constituents are distributed in the same group, etc. All these can only be explained from the point of view of origin.

3. Practical application

The theory of plant pharmacophylogeny has achieved good results through many specific practices. Here are some examples for discussion.

3.1. Looking for domestic resources of imported drugs

With the assumption of phylogenetic relatives having similar components, domestic resources can be sought from the domestic relatives of imported medicines. Some achievements have been made in this area (Table 2). Attention should be paid to the fact that it must be proven to be available and feasible through practical work and that it must not be done rashly.

Atropine sulfate was successfully made from San Fen San (*Anisodus acutangulus*) of Naxi ethnomedicine, which quickly ended the history of China's dependence on imports (Zeng, 2018). From the "stinky oil" used by both Dai and Han nationalities, a new type of suppository matrix, oleum linderiae, was developed, which was scarce in China and filled in the blank again. It is recorded in the later editions of Chinese Pharmacopoeia. The antihypertensive drug Jiangyaling, which was scarce a few decades ago in China, was successfully developed from *Rauvolfia verticillata* (Lour.) Baill., an ethnomedicinal plant of southern China (Hao & Xiao, 2017a).

3.2. Expanding medicinal plant resources

The active ingredients of some medicinal plants have been elucidated, but their content is too low in the source plants. New resources with high content can be found among the phylogenetically related plants. For example, *Maytenus hookeri* Loes. (Celastraceae) has a good anti-cancer effect, but the yield of the active ingredient, maytansine, was only one to two ten-millionth. Since the analogous alkaloids often have anti-cancer effects, later seven ten-millionth of *Maytenus* alkaloids were obtained from *M. buchananii* (Larson, Schaneberg & Sneden, 1999), and then from *Putterlickia verrucosa* Szyszyl. of Celastraceae 12 mg maytansine per kg were obtained, with the yield about 60 times that of *M. hookeri*. In addition, maytansine, belonging to maytansine alkaloids with the anti-cancer activity, was isolated from *Colubrina texensis* (Torrey & A. Gray) A. Gray (Rhamnaceae), which attracted people's interest in finding such components from Celastraceae and Rhamnaceae.

When searching for domestic diosgenin resource plants, it was found that these components were mainly concentrated in the underground part of *Dioscorea* (*Sect. stenophora*) (Table 3), which was rhizome (usually horizontal growth, slightly hard) group; when the underground part was tuber (often vertical growth, slightly soft and juicy) type, the species basically had no these components. In the latter, polysaccharides, tannins and other chemical components are abundant (Huang, Li, Gao & Xiao, 2015), which showed versatile efficacies. According to the known chemical constituents and pharmacodynamics of *Dioscorea* plants, combined with classical morphological taxonomy, *Dioscorea* plants were classified into steroidal saponins group, polysaccharide group, polyphenol group, diosbulbin group, dioscorine group, and miscellaneous group.

Phylogenies reveal predictive power of traditional medicine in bioprospecting (Saslis-Lagoudakis et al., 2012). An explicit phylogenetic framework was developed to explore links between the rich traditions of medicinal use and leaf succulence in aloes (Grace et al., 2015). Large, succulent leaves typical of medicinal aloes arose during the most recent diversification ~10 million years ago and are strongly correlated to the phylogeny and to the likelihood of a species being used for medicine, whereas evolutionary loss of succulence tends to be associated with losses of medicinal use. Species not sharing a typical monosaccharide (glucose-mannose-xylose) profile were grouped outside the core *Aloe* clade in the phylogeny (Grace et al., 2013). Phylogenetic analyses of plant use offer potential to understand patterns in the value of global plant diversity.

The phylogenetic classification was taken into account in evaluating colchicine and related phenethylisoquinoline alkaloids from the family Colchicaceae (Larsson & Rønsted, 2014). Thinking about evolution can influence selection of plant material in drug lead

Table 1
Theoretical explanation of diversification of biogenic pathways of alkaloids.

Relations	Composition	Phylogenetic relationship	Biogenic pathways	Examples
parallelism	same or similar	obviously not very close	unspecified	<i>N</i> -methylcytisine of Berberidaceae, Fabaceae, and Scrophulariaceae
convergence	same or same type	not very close	same	sparteine of <i>Lupinus</i> and <i>Chelidonium</i>
analogy	same type	not closely related	different	anabesine of Chenopodiaceae, Alangiaceae, and Solanaceae
diversification	various	closely related	not yet known	coniine and its congeners are abundant in <i>Cicuta</i> ; other Apiaceae species have monoterpene, sesquiterpene, and polyacetylene, etc.
divergence	different	closely related	different	<i>Calicotome</i> and <i>Genista</i> of Fabaceae are rich in isoquinoline alkaloids; the related genera or species have other Fabaceae alkaloids
homology	distinct	closely related	same	<i>Strychnos vecacona</i> has bakankosine alkaloid. Its relationship with swertiamarin and gentiopicridoside

Table 2

Examples of domestic resources of some imported drugs.

Physicochemical indexes	Imported drugs	Domestic resources	Remarks
	<i>Styrax benzoin</i>	<i>S. macrothyrsus</i> , <i>S. subniveus</i> , <i>S. hypoglauca</i>	The clinical application proves that domestic benzoin has the same effect as imported one, and importing is stopped presently.
total balsamic acid (calculated by alcoholic extract)	26%–35%	25.17%–31.46%	
	<i>Strychnos nux-vomica</i>	<i>S. wallichiana</i>	<i>S. wallichiana</i> has been used for drug production. We have also analyzed many Chinese species of <i>Strychnos</i> .
total alkaloid	2%–5%	2.19%	
strychnine	1%–1.4%	1.34%	
	<i>Acacia senegal</i>	<i>farnesiana</i> , <i>A. decurrens</i> , <i>A. decurrens</i> var. <i>mollis</i>	It is proved that most of the domestic gum can be used clinically.
viscosity of tree gum (centistoke)	1.4855	1.0881–1.7818	
suspension force measurement (hr)	40	24–64	
emulsifying properties	qualified	superior to or close to Arabic gum	
appearance	<i>Picrorrhiza kurroa</i> α (dark brown)- γ -(light gray, yellow and black)	<i>P. scrophulariae</i> α (dark brown)- γ -(light gray, yellow and black)	<i>P. scrophulariae</i> has been used for drug production
alcohol extract content	45.96%	44.21%;49.20%	
water extract content	36.98%	35.05%;37.82%	
bitter ratio	60	60	
	<i>Hydrocarpus anthelmintica</i>	<i>H. hainanensis</i>	According to the physicochemical properties and seed morphology, there is little difference between domestic and imported <i>Hydrocarpus</i> , which needs further clinical confirmation.
oil content	shelled kernel 45.36%	shelled kernel 44.96%	
oil freezing point	yellow oily liquid at room temperature	white solid at room temperature	
iodine value	84.74	76.29	
specific gravity	0.809	0.928	
refractive index (20 °C)	1.4786	1.4758	
coloring reaction of gynocardia oil	① a few drops of oil+ one drop of trichloroacetic acid and 4 drops of HCl; slight heat and show blue color ② a few drops of oil+5 drops of glacial acetic acid and HCl (9:1); heated slightly, dark blue ③ 1 ml oil+H ₂ SO ₄ ; first red-brown, then olive-green	① also blue ② dark blue ③ first red-brown, then olive-green	
	<i>Ferula asafoetida</i>	<i>F. sinkiangensis</i>	<i>F. sinkiangensis</i> has been used for drug production
alkali addition after grinding with water	orange-yellow emulsion, yellowish-green emulsion with alkali	white emulsion, yellow emulsion with alkali	
alcohol extract with a few drops of phloroglucin and concentrated HCl	pink, but a little deep	pink	
umbelliferone reaction	blue fluorescence	blue fluorescence	
volatile oil content	6%–17%	17.5%–18.5%	
sulfur content	17%–38%	20.6%	
specific gravity (20 °C)	0.906–0.973	0.974	
refractive index (20 °C)	1.493–1.518	1.529	

Table 3Distribution and saponin content of rhizome group of Chinese *Dioscorea*.

Species	Distribution	Saponin content/%
<i>Dioscorea althaeoides</i>	Yunnan, Sichuan	0.5–2.3
<i>D. colletti</i>	Southwest, Shaanxi	0.5–1.0
<i>D. colletti</i> var. <i>hypoglauca</i>	Fujian, Zhejiang, Jiangxi, Hunan, Guangdong, Taiwan, Sichuan	0.53–2.02
<i>D. deltoidea</i>	Yunnan, Sichuan, Tibet	1.8–5.4
<i>D. gracillima</i>	Anhui, Zhejiang, Jiangxi, Fujian	1.03–2.39
<i>D. nipponica</i>	Northeast, North China, Northwest, East China, Henan	1.36–2.00
<i>D. panthaica</i>	Southwest, Hunan	1.70–4.20
<i>D. tokora</i>	East China, Central south, Sichuan, Guizhou	1.00–2.10
<i>D. zingiberensis</i>	Shaanxi, Gansu, Sichuan, Yunnan, Hubei, Hunan, Henan	1.05–16.15
<i>D. tenuipes</i>	Zhejiang, Fujian, Jiangxi, Anhui	1.28

discovery, and knowledges of phylogeny may be used to evaluate predicted biosynthetic pathways. A correlation between phylogeny and biosynthetic pathways could offer a predictive approach enabling more efficient selection of plants for the development of traditional medicine and lead discovery. Alkaloid diversity and in vitro inhibition of acetylcholinesterase (AChE) and binding to the serotonin reuptake transporter (SERT) are significantly correlated with phylogeny of subfamily Amaryllidoideae (Amaryllidaceae; Rønsted et al., 2012), validating the use of phylogenies to interpret chemical evolution and biosynthetic pathways, to select candidate taxa for lead discovery, and to make policies regarding traditional use and conservation priorities.

A phylogenetic approach is useful in identifying *Artemisia* species more likely to possess antimalarial properties (Pellicer et al., 2018). At least 117 *Artemisia* species are recorded in ethnobotany. In phylogenetic analysis, lineages with an over-representation of species used to treat malarial symptoms were identified, which could be high priority for further investigation of antimalarial activity. LC-MS/MS analysis was performed to explore artemisinin content in 15 species from both highlighted and not highlighted lineages. Artemisinin was detected in nine species, eight of which are novel antimalarial source plants. Artemisinin may be widespread across the genus, providing an accessible local resource outside the distribution area of *A. annua*.

In ethnomedicine culture, the large, widespread species, rather than small, narrow-ranged species, are preferentially used (Cámara-Leret et al., 2017), and different traits (leaf length, stem volume, fruit volume, geographic range size) are linked to different uses. The reliance on plant size and availability may prevent the optimal realization of wild-plant services, since ecologically scarce yet functionally important (e.g., chemically) clades may be over-

looked. In addition to expand our understanding of how local people use biodiversity (Hao, 2018), the trait- and phylogeny-based approaches help to understand the processes underpinning ecosystem service realization.

In phylogenetic prediction, grouping medicinal plant uses in terms of biological responses to the treatment might be better than according to systems of the human body (Ernst et al., 2016). In the cosmopolitan and pharmaceutically important genus *Euphorbia*, identifying plant uses modulating the inflammatory response highlighted a greater phylogenetic diversity and number of potentially useful species. The medicinally highly valuable subgenus *Chamaesyce* was chemically under-investigated (Ernst et al., 2015), emphasizing the need for further chemodiversity studies. The phylogeny-guided drug discovery at an early screening stage could be more targeted, possibly resulting in higher discovery rates of novel chemistry with functional biological activity.

3.3. Quality control, identification and expansion of herbal medicines

For example, CHMs (Chinese herbal medicines) *Radix Paeoniae Alba* (Bai Shao in Chinese), *Radix Paeoniae Rubra* (Chi Shao in Chinese) and *Cortex Paeoniae* are all derived from *Paeonia* (Wang, He, Peng, Chen & Xiao, 2017). The presence of paeoniflorin and paeonol in domestic species is summarized in Table 4.

It can be seen that the peony bark mainly comes from the woody peony group, characterized by the presence of paeonol. Bai Shao and Chi Shao are mainly from the herbal group, which does not contain paeonol, but paeoniflorin is abundant. The content of paeoniflorin decreased greatly after processing into Bai Shao due to the partial destruction during processing. For another example,

Table 4
Paeoniflorin and paeonol in domestic *Paeonia* species.

Species	Sample sources	Phenological phase	Paeoniflorin content/%	Paeonol
<i>Sect. Moutan</i>				
<i>P. suffruticosa</i>	Hefei, Anhui	vegetative period	1.26	+
	Jinan, Shandong	vegetative period	0.90	+
<i>P. delavayi</i>	Lijiang, Yunnan	fruiting stage	1.93	+
<i>P. delavayi</i> var. <i>lutea</i>	Dali, Yunnan	fruiting stage	1.61	+
(= <i>P. lutea</i>)	Kunming, Yunnan	flourescence	1.45	+
Root bark	Southeast Tibet	fruiting stage	2.52	++
root bark	Xichang, Sichuan	vegetative period	2.17	++
<i>Sect. Paeonia</i>				
<i>P. lactiflora</i>	Fengcheng, Liaoning	vegetative period	7.02	–
	Duolun, Inner Mongolia	vegetative period	6.40	–
	Hangzhou, Zhejiang	vegetative period	5.08	–
	Xuanhua, Hebei	vegetative period	4.96	–
	Jiutai, Jilin	bud appearance	4.34	–
	Linjiang, Jilin	flourescence	4.70	–
	Huairou, Hebei	fruiting stage	3.40	–
	Xinglong, Hebei	bud appearance	10.72	–
<i>P. lactiflora</i> var. <i>trichocarpa</i>	Hangzhou, Zhejiang	vegetative period	5.70	–
	Beijing cultivar	vegetative period	4.96	–
6. <i>P. veitchii</i>	Huzhu, Qinghai	vegetative period	5.76	–
	Minhe, Qinghai	fruiting stage	2.86	–
	Maowen, Sichuan	bud appearance	4.32	–
	Dangchang, Gansu	flourescence	1.86	–
7. <i>P. obovata</i>	Baihua mountain, Hebei	flourescence	2.16	–
	Yangcheng, Shanxi	fruiting stage	1.06	–
	Sichuan	vegetative period	2.46	–
	Ji'an, Jilin	vegetative period	0.27	–
	Fusong, Jilin	fruiting stage	Almost non-existent	±
<i>P. obovata</i> var. <i>willmottiana</i>	Kang County, Gansu	fruiting stage	2.54	–
	Dzoge, Sichuan	fruiting stage	0.09	–
<i>P. sinjingsensis</i> K.Y.Pan	Altay, Xinjiang	flourescence	1.48	–
<i>P. mairei</i>	Hongya, Sichuan	flourescence	2.72	–
<i>P. anomala</i> var. <i>intermedia</i>	Tianshan Mountain, Xinjiang	fruiting stage	1.40	±
	Altay, Xinjiang	vegetative period	2.16	±
	Xinyuan, Xinjiang	flourescence	2.40	±

both Cang Zhu (*Rhizoma Atractylodis*) and Bai Zhu (*Rhizoma Atractylodis Macrocephalae*) belong to *Atractylodes* (Zou et al., 2009), but they are distinct in morphology and composition of volatile oil (Table 5).

In chemical profiles, Bai Zhu has more atractylon (Fig. 1), while atractylodin is absent or almost absent; atractylodin, β -eudesmol and hinesol are abundant in Cang Zhu. These differences can be

used as chemical indicators for commercial Bai Zhu and Cang Zhu. *A. koreana* contains more atractylodin, but less hinesol, eudesmol and atractylon, so it is not only morphologically, but also chemically distinct from other Cang Zhu and Bai Zhu source plants. *A. japonica* contains more butenolide A and more selina-4 (14), 7 (11)-diene-8-one, but only a small amount of hinesol and β -eudesmol,

Table 5
Rhizome volatile oil of Chinese *Atractylodes*.

Species	Morphology	Origins	Gas chromatography peaks (components)											
			a	bl	cII	d	eIII	fIV	gV	hVI	i	jVII	k	IVIII
<i>Atractylodes macrocephala</i>	leaves 3–5 pinnately cleft or deeply cleft	Pingjiang, Hunan (Bai Zhu)	±	–	+	+	++++	(+)	(+)	(+)	(+)	–	++	++
		Hangzhou, Zhejiang (Bai Zhu)	±	–	(+)	+	++++	–	(+)	(+)	(+)	–	(+)	++
		Changhua, Zhejiang (Yu Zhu)	(+)	–	(+)	+	++	(+)	(+)	(+)	+	+	±	+
<i>A. japonica</i>	leaves undivided or only a few shallow lobes	Zhejiang (Dong Zhu)	(+)	–	+	++	(+)	(+)	(+)	+	+	±	+	+
		Heilongjiang	+	(+)	(+)	(+)	(+)	(+)	(+)	+	(+)	–	–	++
<i>A. lancea</i>	leaves undivided or only a few shallow lobes	Tong Bai, Hunan	±	–	±	±	(+)	+++++	+++++	–	–	±	±	±
		Taiping, Anhui	±	–	±	±	(+)	+++++	+++++	–	–	+	±	(+)
		Hubei	+	(+)	±	±	(+)	+++++	+++	–	–	+	(+)	–
		Jurong, Jiangsu	+	(+)	(+)	+	++	++	+++	(+)	(+)	+	(+)	±
		Huairou, Beijing	(+)	±	(+)	(+)	+	++	++++	(+)	(+)	++	±	–
<i>A. lancea</i> var. <i>chinensis</i>	leaves undivided or only a few shallow lobes	Gu'an, Hebei	(+)	±	(+)	(±)	(+)	++	+++	+	+	++	+	–
		Changping, Beijing	+	±	(+)	(+)	+	++	+++	(+)	(+)	++	–	–
		Mentougou, Beijing	(+)	±	(+)	(+)	+	++	+++	±	(+)	++	–	–
		Shenyang, Liaoning	++	+	(+)	(+)	+	±	(+)	±	–	(+)	++	+
		Qianshan Mountain, Liaoning	++	(+)	(+)	+	+	+	+	(+)	(+)	+++	(+)	(+)

GC percentage content: < 1(trace), ±; 1–5(microscale), (+); 5–10, +; 10–20, ++; 20–30, +++; 30–40, ++++; 40–50, +++++. Examples of chemical constituents of Chinese *Atractylodes*. I. elemol; II. β -selinene; III. atractylon; IV. Hinesol; V. β -eudesmol; VI. selina-4(14),7(11)diene-8-one; VII. Atractylodin; VIII. butenolide A.

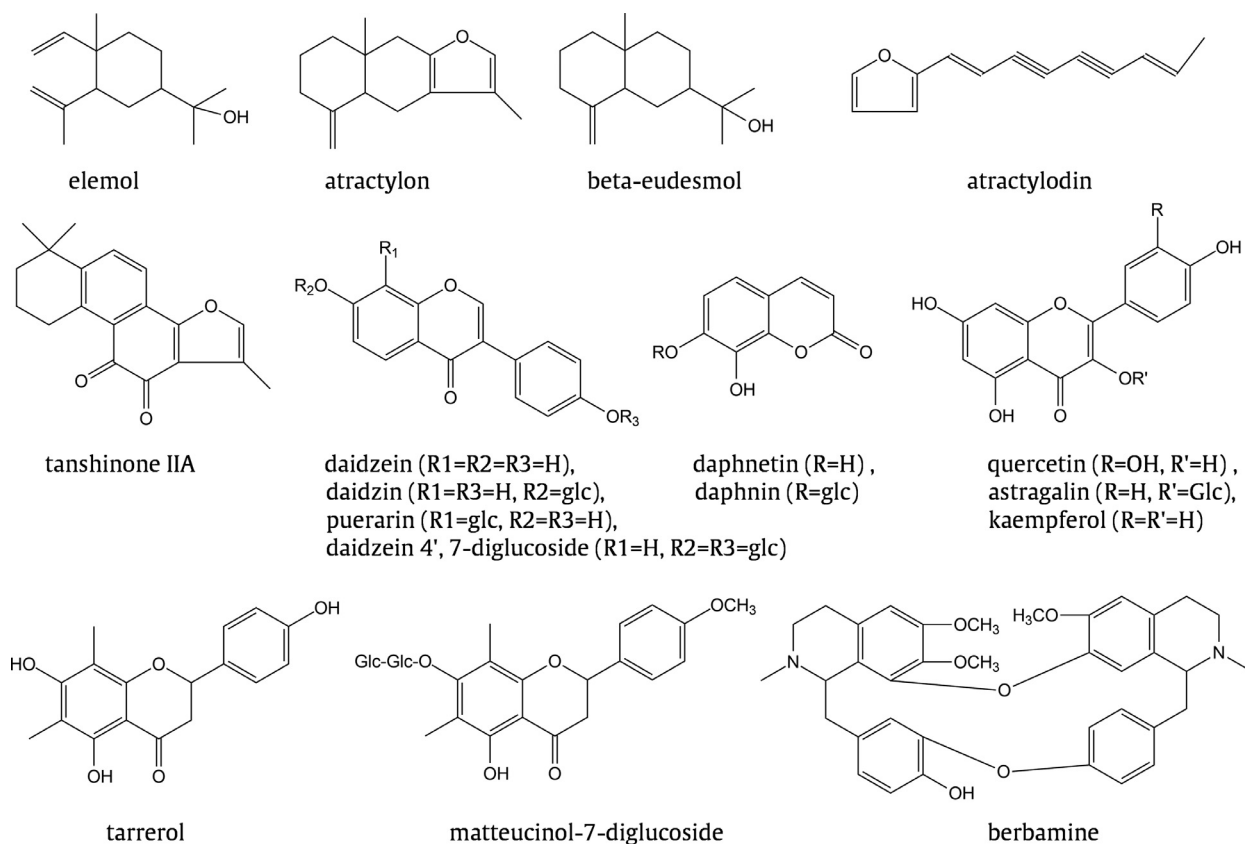


Fig. 1. Examples of herbal compounds mentioned in text.

indicating that it is closer to Bai Zhu in morphology and chemical composition than to Cang Zhu.

According to the similarities and differences of phylogenetic relationships and chemistry, and on the basis of comprehensive research, the effective components of TCM (traditional Chinese medicine) can be used as indicators to expand their resources step by step. Of course, as many TCMs are not fully represented by one or two active ingredients, the work in this regard needs to be carried out very carefully. Table 6 lists some clues for research consideration.

Take *Salvia miltiorrhiza* Bge. as an example. Since tanshinone IIA (Fig. 1) was found to be one of the main effective components for coronary heart disease (Hao & Xiao, 2017b) and cryptotanshinone was one of the main effective components for antimicrobial activity. Scholars have done a lot of work on the distribution and abundance of these components in *Salvia* (Hao, Ge & Xiao, 2018a) (Table 7; Li et al., 2013).

In Shanghai Institute of Materia Medica, tanshinone IIA in the roots of more than 50 *Salvia* species was determined qualitatively by TLC and quantitatively by multi-component spectrophotometry in ultraviolet region (Hao & Xiao, 2017a). More than 20 species were found to contain this ingredient, and eight species were found to contain 0.3%–1.01%. For example, *S. leiaometiensis* had

1.01% tanshinone IIA, *S. miltiorrhiza* f. *alba* 0.73%, *S. trijuga* 0.65%, *S. aerea* 0.51%, and *S. przewalskiae* var. *alba* 0.37%. Tanshinone IIA is abundant in many species of *Subgen. Salvia*, and quite some species of *Subgen. Sclarea* has it; However, it is almost absent in *Subgen. Jungia* and *Subgen. Allagospadonopsis*.

Such research, on the one hand, can provide valuable clues and rules for searching for the resources of such active ingredients; on the other hand, when high-content species are found, in the first step as a drug raw material for the treatment of certain diseases, they can be utilized first.

Sometimes, species belonging to the same family and the same genus differ greatly in the content of active ingredients. For example, the original plant of Ge Gen (*Radix Puerariae*) is *Pueraria lobata* (Willd.) Ohwi, and effective ingredients reflecting part of its curative effect have been found, e.g., daidzein, daidzin, puerarin, and daidzein 4',7-diglucoside (Fig. 1), which belong to isoflavones (Wagle, Seong, Jung & Choi, 2019).

It has been proved by phytochemistry, pharmacology and clinic that total flavonoids of *Pueraria* can reflect its effects on improving cerebral and coronary circulations and treating coronary heart disease (Wei, 2015). Therefore, we compared the content of total flavonoids in Chinese *Pueraria* roots (Table 8). Grasping these data is beneficial in that, on the one hand, high content raw materials

Table 6
Some clues for expanding resources of common TCMs.

TCM names	Identifiable chemical indicators	Plant taxa for resource expansion	Remarks
<i>Radix Glycyrrhizae</i> (Gan Cao)	glycyrrhizic acid	<i>Glycyrrhiza</i> , <i>Sect. Glycyrrhiza</i>	Species without sweet roots cannot be used as licorice.
<i>Radix Gentianae Macrophyllae</i> (Qin Jiao)	gentianine, gentianidine, gentianol	<i>Gentiana</i> , <i>Sect. Aptera</i> (<i>Cruciata</i>)	should have a main root, with fibrous petiole residues in the neck of the root
<i>Radix Arnebiae</i> / <i>Radix Lithospermi</i> (Zi Cao)	naphthoquinone pigments with shikonin as parent	Boraginaceae, Subfam. Boraginoideae, <i>Lithospermum</i> , <i>Arnebia</i> , <i>Macrotomia</i> and <i>Onosma</i>	Species without purple roots cannot be used as Zi Cao.
<i>Radix et Rhizoma Rhei</i> (Da Huang)	conjugated rhein anthrone–sennoside A, B, C; total anthraquinones are mainly rhein-bound anthraquinones, but less free anthraquinones; rhaponticine should not exist	<i>Rheum</i> , <i>Sect. Palmata</i>	Presently, the pharmacological indexes of rhubarb are mainly focused on diarrhea. If other effects are further studied, the medicinal value of other groups of Rhubarb should be further evaluated.
<i>Radix Scutellariae</i> (Huang Qin)	flavonoids such as baicalein and baicalin	<i>Scutellaria</i> , <i>Subg. Scutellaria</i> , <i>Sect. Stachymacris</i> , <i>Subsect. Angustifoliae</i> , <i>Ser. Baicalensis</i> and <i>Ser. Likiangensis</i>	Species without yellow roots cannot be used as Huang Qin.
<i>Radix Salviae Miltiorrhizae</i> (Dan Shen)	phenanthrenequinone derivatives such as tanshinone IIA and cryptotanshinone	<i>Salvia</i> , <i>Subg. Sclarea</i> , <i>Sect. Drymosphaca</i> ; <i>Subgen. Salvia</i> , <i>Sect. Eurysphaca</i>	Species without red roots cannot be used as Dan Shen; the water-soluble effective part needs further study.

Table 7
Tanshinones of 14 Chinese *Salvia* species.

Species	Root color	Total tanshinone/%	Cryptotanshinone /%	Inhibition zone (mm; paper strip method)
paper strip method)				
<i>Subg. Salvia</i>				
<i>Sect. Eurysphaca</i>				
1. <i>S. evansiana</i>	grayish brown	0.18	—	15
2. <i>S. brachyloma</i>	grayish brown with reddish tint	0.03	—	0
3. <i>S. digitaloides</i>	red	0.67	microscale	23
4. <i>S. przewalskiae</i>	red	1.99	1.60	27
5. <i>S. przewalskiae</i> var. <i>mandarinorum</i>	red	1.30	0.30	23
6. <i>S. aerea</i>	red	1.22	0.17	22
7. <i>S. castanea</i>	red	1.60	0.80	22
8. <i>S. flava</i>	grayish brown	0.069	—	0
9. <i>S. bulleyana</i>	grayish brown with reddish tint	0.028	—	very small
10. <i>S. mekongensis</i>	grayish brown	0.02	—	13
<i>Subg. Sclarea</i>				
<i>Sect. Drymosphace</i>				
11. <i>S. trijuga</i>	red	1.06	0.16	25
12. <i>S. yunnanensis</i>	red	0.55	microscale	20
13. <i>S. miltiorrhiza</i>	red	1.95	0.71	24
14. <i>S. bowleyana</i>	red	0.24	microscale	19

Table 8
Total flavonoids in Chinese *Pueraria* roots.

Species	Sample sources	Total flavonoid content/%
1. <i>P. lobata</i>	12 producing areas of Liaoning, Shandong, Hebei, Gansu, Shaanxi, Anhui, Jiangsu, Jiangxi, Sichuan and Guangxi	5.04–12.30
2. <i>P. thomsonii</i>	8 producing areas of Sichuan, Guangxi and Guangdong	1.42–3.86
3. <i>P. montana</i>	Nandan, Guangxi	0.51
4. <i>P. omeiensis</i>	Mount Emei, Sichuan	0.13
5. <i>P. edulis</i>	Weishan, Yunnan	0.34
6. <i>P. phaseoloides</i>	Cangwu, Guangxi	1.09

can be selected in the production of such drugs; on the other hand, they could be the references for rational drug use.

3.4. Help to predict chemical constituents or active ingredients of herbal medicine and assist in identification and determination of chemical constituents

In recent years, due to the extensive use of new technologies and methods of phytochemistry and metabolomics (Hao, 2018, 2015a), a wealth of basic information has been accumulated for diverse plants. These materials are often accessible in many articles, books or manuals, so that when we study the composition of CHM and other ethnomedicines, there are always some clues to follow.

For example, the herbal medicine *Daphne giraldii* Nitsche (Zu Shi Ma in Chinese) belongs to *Daphne* genus of Thymelaeaceae. At the beginning of studying its constituents, we consulted the chemical data of Thymelaeaceae in Hegnauer's *Chemotaxonomie der Pflanzen* (1973) page 6508, and knew that coumarins, especially daphnetin and daphnin (Fig. 1), are widely contained in this family and this genus. Later it was proved that *Daphne A* is daphnetin (7,8-dihydroxycoumarin) and daphnin is also contained in Zu Shi Ma. The former is an effective component of Zushima for anti-inflammation, analgesia, sedation and vasodilation. An anti-coagulation constituent was isolated from *Daphne mezereum* in the former Soviet Union, and identified as daphnin. The effective component isolated from *D. koreana* in Northeast China was also daphnetin. Therefore, these plants, their components and their pharmacological effects suggest that we can further study them.

Some types of compounds, especially those with special structures, are regularly distributed in the plant kingdom. For example, a special type of flavonoids, carbon-methylated dihydroflavones, is mainly restricted to *Rhododendron* of Ericaceae, *Pinus* of Pinaceae, and *Cyrtomium* of Dryopteridaceae (Zhao, Ding, Zhao, Wang & Gao, 2012). With this in mind, when we study the effective ingredients of Chinese *Rhododendron* (e.g., *R. mariesii*, *R. amesiae*, etc.) (Fig. 1) in the treatment of tracheitis, we will surely get twice the result with half the effort in structure determination.

According to the principle that the chemical structures from the closely related taxa are often very similar, when studying the chemistry of *Aconitum tanguticum* (Maxim.) Stapf in Institute of Medicinal Plant Development, CAMS, given its close relationship with *A. heterophyllum* Wall of India, the rare C-19 diterpene lactone alkaloids benzoylheteratisine and heteratisine in *A. tanguticum* were identified quickly (Xiao, Wang & Gao, 2006). In the structure determination, we should also grasp and understand some components and structural types of plants that are closest to the original plants in which such components exist.

It has been known for a long time that if the structures of plant chemical components are similar, their pharmacological effects are often comparable. If we combine this knowledge with plant genetic relationship, we can often get better results. For instance, in the 1960s and 1970s, in collaboration with the Shaoguan Regional

Health Bureau of Guangdong Province, the Shaoguan Scientific Research Team of Institute of Materia Medica, CAMS unearthed a better antimicrobial herb, *Picrasma quassioides* (D. Don) Benn., from the local mass movement of CHM. Its wood is known to contain two kinds of canthinone alkaloids, nigakinone and methyl-nigakinone (Hao & Xiao, 2017a), a series of keto diterpene lactone bitter substance, etc. In searching for the antimicrobial active ingredients of Simaroubaceae, it has been known that the canthin-6-one (isolated from the bark of *Zanthoxylum elephantiasis* Macf of Rutaceae), which has similar structure to the two alkaloids mentioned above, has strong bacteriostatic effect on *Staphylococcus aureus* and *Mycobacterium spp.* Given that both Simaroubaceae and Rutaceae belong to Rutineae of Rutales, the above two alkaloids may also have bacteriostatic effects. The results showed that they had strong bacteriostatic effects on *Pneumococcus*, *Bacillus subtilis*, *Streptococcus haemolyticus* and *S. aureus*. The injection and tablet based on total alkaloids of *P. quassioides* (including other two alkaloids with lower content) have passed more than 1000 clinical tests, and have good anti-bacterial and anti-inflammatory effects, which have been certified and popularized.

Phylogeny predicted the quantity of antimalarial alkaloids within the iconic yellow Cinchona bark (Rubiaceae: *Cinchona calisaya*; Maldonado et al., 2017). A significant phylogenetic signal was found for the content of two out of four major Cinchona alkaloids (quinine and cinchonidine) and their total content. A clade of high alkaloid producing trees was identified that spanned a narrow range of altitudes, from 1100 to 1350 m. The chemodiversity is primarily driven by phylogeny. Comparisons of the relative effects of environmental and genetic variability in determining plant chemodiversity should be performed at both species and population (e.g., different genotypes and haplotypes between geoherb and non-geoherb) levels, as the extensive genotypic and epigenetic variations in plant biochemistry are directly related with the clinical usefulness of medicinal plants (Hao & Xiao, 2018).

3.5. Application of pharmacophylogeny in the search for new drugs

An important task of drug research is to find new drugs with better efficacy and lower toxicity from nature. The clue that there is a close relationship between plant phylogeny, chemical composition and pharmacokinetics/therapeutic effect can often inspire and instruct the search for new drugs (Hao, 2018; Hao & Yang, 2016). For example, when we study the comprehensive utilization of berberine resources, the first problem to be solved is the utilization of berbamine (Fig. 1) residues in the berberine mother liquor. When we cooperated with Jiangsu Institute of Dermatology Prevention and Treatment to try berbamine on a patient, it was found that the white blood cells of the patient increased rapidly from 1070 to 7600 after six weeks (Hao & Xiao, 2017a). This reminds us of the similar structure of cepharanthine (Xiao et al., 2019), a leukocyte-elevating drug extracted from *Stephania* of Menispermaceae (Hao & Yang, 2016), which is closely related to Berberidaceae in Ranunculales. The pharmacological experiments in the Tumor Laboratory of Institute of Materia Medica, CAMS have proved that berbamine has a good leukocyte-elevating effect. In clinical trial, it was used in 405 cases of various leukopenia, and the total effective rate was 71%. In this way, a new drug was contributed for clinic, opening up a way for the comprehensive utilization of *Berberis* and *Coptis* species. It has been appraised as a new drug. Optimistically, the close relationship between phylogeny and pharmacokinetics has been revealed preliminarily in Ranunculaceae (Hao, Ge, Xiao, Wang & Yang, 2015c), Ranunculales (Hao & Yang, 2016), *Salvia* (Hao & Xiao, 2017b) and *Taxus* (Hao, Ge, Wang & Yang, 2018b). Although simple changes of configuration or methylation could cause dramatic effects in drugs, the predictions by the phylogeny of the

producing plant could give a general pharmacokinetic impression of new drug candidates.

In the mass prevention and treatment of chronic tracheitis, it was found that many *Rhododendron* species have better antitussive and expectorant effects in clinic, and many effective ingredients were isolated (Cai, Hu, Qin, Sun & Li, 2018). For example, farrerol, belonging to flavonoids, has the expectorant effect, hyperin and quercetin (Fig. 1) have antitussive effect, and astragalol has expectorant effect. Two flavonol compounds, kaempferol and myricetin, are abundant in *Rhododendron*; the former has expectorant effect. The coumarin compound scopoletin has the antiasthmatic and expectorant effects. It was also proved that the main toxic component of *Rhododendron* was andromedotoxin 1.

In order to find better and more *Rhododendron* medicines, leaves of 33 Chinese *Rhododendron* species were qualitatively examined with the effective and toxic components as indicators (Table 9). Firstly, species with the effective ingredients and plenty resources, and without andromedotoxin, e.g., *R. racemosum* and *R. rubiginosum*, were selected to cooperate with the local community. Through a large number of clinical tests, it was proved that these species have good curative effect on chronic tracheitis. They have been appraised in clinical application.

Marine Chinese Medicine (MCM) is one of the important parts of traditional Chinese medicine (TCM). The explorations of marine bioresources provide a good base for the development of MCM. However, the evaluation of TCM property (Yao Xing in Chinese) of new MCM resources is challenging. Totally 613 MCM and the related 1091 marine species were screened with regard to the association rules mining and phylogenetic tree (Fu et al., 2015). The MCMs with the same Yao Xing tended to cluster on the phylogenetic tree. The MCMs from the organisms of the same family often had the same Yao Xing. The association rules for the same nature were often distributed on the same branch or near branch of the phylogenetic tree. For instance, the marine plants Chlorophyta, Florideophyceae, and Phaeophyceae were more related with “cold” nature, the marine animals Decapoda, Malacostraca, and Arthropoda had close relationship with “hot” property, while the “neutral” nature was more related with Squamata. There is close relationship between Yao Xing and the phylogenetic relationship. The MCM from the same or close families may have the same or similar Yao Xing. This may be useful in bioprospecting of novel MCMs.

Chinese herbal medicine (CHM) has been used for thousands of years, and has gained valuable human experiences in the treatment of various diseases, including tumors. The abundant resources of

Table 9
Qualitative observation on effective and toxic components in leaves of 33 Chinese *Rhododendron* species.

Species	Sample sources	Components								
		X	XI	XII	XIII	XIV	XV	XVI	XVII	
I. Subg. Lepidorrhodium										
1.R. bullatum	Weixi, Yunnan	–	+	+	+	+	–	–	–	
2.R. micranthum	Hebei	ND	+	+	+	+	+	+	+	
3.R. capitatum	Minhe, Qinghai	+	+	+	–	–	–	+	+	
4.R. thymifolium	Minhe, Qinghai	–	+	+	–	+	–	+	–	
5.R. hippohaeroides	Shangri-La, Yunnan	–	+	+	–	+	+	+	–	
6.R. desquamatum ¹ (<i>R. rubiginosum</i>)	Jinyang, Sichuan	–	+	–	+	+	+	+	–	
7.R. siderophyllum	Weixi, Yunnan	–	+	–	–	+	–	+	–	
	Lijiang, Yunnan	–	+	–	+	+	–	+	–	
8.R. chartophyllum	Shangri-La, Yunnan	–	+	+	+	+	+	+	–	
9.R. stereophyllum	Shangri-La, Yunnan	–	+	+	–	+	+	+	–	
10.R. aechmophyllum	Lijiang, Yunnan	–	+	–	+	+	+	+	+	
11.R. chengshienianum	Mount Emei	+	+	–	–	+	+	+	–	
12.R. concinnum	Mount Emei	–	+	+	–	+	+	+	–	
	Taibai Mountain, Shaanxi	–	+	–	–	+	–	+	–	
13.R. polylepis	Jinyang, Sichuan	–	+	+	–	+	+	+	–	
14.R. anthopogonoides	Linxia, Gansu	–	+	+	+	+	+	+	–	
15.R. primulaeflorum	Tibet	–	ND	ND	ND	+	ND	ND	–	
II. Subg. Rhododendron										
16.R. glischrum	Weixi, Yunnan	–	+	–	–	+	–	+	–	
17.R. strigillosum	Mount Emei	–	+	–	–	+	+	+	–	
	Taibai Mountain, Shaanxi	–	–	–	–	–	–	–	±	
18.R. pachytrichum	Mount Emei	–	+	+	–	+	–	–	–	
19.R. argyrophyllum	Mount Emei	–	+	+	–	+	–	+	+	
20.R. calophyllum	Mount Emei	–	+	–	+	+	±	+	+	
21.R. decorum	Lijiang, Yunnan	+	+	–	+	+	+	+	+	
22.R. vernicosum	Shangri-La, Yunnan	–	+	–	+	+	+	–	+	
23.R. orbiculare	Mount Emei	–	+	–	–	+	–	+	–	
24.R. afflectum	Dali, Yunnan	–	+	+	–	+	–	+	+	
25.R. przewalskii	Minhe, Qinghai	–	+	+	+	+	–	–	+	
26.R. purdonii	Taibai Mountain, Shaanxi	–	+	–	–	+	+	+	±	
27.R. faberi	Mount Emei	–	+	+	±	+	–	+	–	
III. Subg. Anthodendron										
28.R. molle	Lushan Mountain	±	+	±	±	+	+	+	±	
29.R. s. imsii	Xuyong, Sichuan	–	+	–	–	+	–	+	–	
30.R. mariesii	Lushan Mountain	–	+	+	–	+	±	+	–	
IV. Subg. Azaleastrum										
31.R. s. tamineum	Mount Emei	–	+	–	+	+	+	+	–	
V. Subg. Pseudorhododendron										
32.R. spinuliferum	Miyi, Sichuan	+	+	–	–	+	+	+	–	
	Leibo, Sichuan	–	+	+	–	+	+	+	–	
33.R. racemosum	Weixi, Yunnan	–	+	+	–	+	–	+	–	
	Shangri-La, Yunnan	–	+	+	–	+	+	+	–	

ND, not determined. X, farrerol; XI, hyperin; XII, quercetin; XIII, astragalol; XIV, kaempferol; XV, myricetin; XVI, scopoletin; XVII, andromedotoxin 1.

¹ The original recorded flowers are purple, but the flower of this specimen is pink.

CHM represent an important material basis for the search of new anticancer drugs. The four ways of further discovering new anticancer drugs based on TCM theory combined with modern scientific discoveries were expounded (Liu et al., 2012a). According to TCM theory, heat-clearing and detoxifying drugs, blood-activating and stasis-removing drugs, immune regulating drugs (invigorating Qi and strengthening the body) and poisonous CHMs (addressing poison with poison) are more likely to possess anticancer activities; anticancer drugs can be searched from anticancer lead compounds found in recent years; the extracts of CHMs can be used to find anticancer drugs via routine screening or high-throughput screening; last but not least, anticancer drugs can be searched according to the principles of plant pharmacophylogeny (Hao et al., 2014, 2015a), which complement other approaches.

Phytochemical taxonomy is based on the fact that closely related plants contain the same or similar chemical constituents. Medicinal plants within the same phylogenetic groups may have the same or similar therapeutic effects, thus forming the core of pharmacophylogeny (Hao & Xiao, 2017a). Many examples of *Glycyrrhiza*, *Mentha*, *Schisandra*, *Fritillaria* (Hao, Gu & Xiao, 2013a), *Cimicifuga* (Hao, Gu & Xiao, 2013b), and *Lonicera*, etc. fully demonstrate these universal principles. According to these principles, the anticancer compounds oridonin and ponocidin were isolated from *Rabdosia* (Bai et al., 2010), daphnane diterpenoids from *Daphne* (Hao & Xiao, 2017a), macrocyclic tannins from Onagraceae and Lythraceae, and curcumin derivatives from *Curcuma*, etc. (Liu et al., 2012a).

3.6. Application of pharmacophylogeny in sorting out, summarizing and improving herbal medicine experiences

Through summarizing and sorting out the chemical composition and curative effect of a large number of herbal medicines, the personality of each kind of CHMs is analyzed by the line of kinship, from which many common things can be summed up. The sorting and analysis of this aspect cannot be achieved without computers (see the next section).

For example, CHMs belonging to Lamiaceae, e.g., mint, perilla, *Elsholtzia*, etc., are aromatic and contain abundant volatile oils, which have the functions of wind-repelling and surface-relieving. Other non-aromatic and almost non-volatile oil species of Lamiaceae often have bitter taste, and often contain diterpene bitter lactones, such as *Rabdosia* plants *R. rubescens* var. *rubescens*, *R. serra* (Maxim.) Kudo, and *R. nervosa* (Hemsl.) C. Y. Wu et H. W. L. Their ingredients often have anti-cancer, anti-bacterial, antipyretic and detoxifying effects. By screening tens of thousands of anticancer plants and applying the viewpoints of pharmacophylogeny, it is concluded that the following components and plant groups deserve further study (Tables S1 and S2).

Through the preliminary comprehensive arrangement of pharmacophylogeny (Bai, Wang, Xiao & Liu, 2012; He, Peng, Xiao & Xiao, 2012; Liu et al., 2012a,b), we can tentatively put forward some Chinese groups of plant taxa worthy of further study (Table S2).

The Chinese Pharmacopoeia (2015) includes 584 plant medicines, 284 of which contain “Daodi” components (Lei, Wu, Leon, Huang & Hawkins, 2018), representing superior clinical properties compared to non-Daodi counterparts. Daodi means good quality, high yield, special processing techniques, and usually produced only in a specific geographic region. In China there are 10 Daodi producing regions. Medicinal preferences for lineages were visualized using phylogenetic mapping. Medicinal parts of species with any Daodi subset were more likely to be “hot” or “warm” (TCM concepts) roots, and less likely to be “toxic”. Roots were over-represented in the Bei (north) region, and whole plants over-represented in Guang (cantonese) region. Both the

Chinese Pharmacopoeia and Daodi indicated preferred families not common in previously studied ethnopharmacopoeias, and fewer endemic species were represented than expected by chance. The phylogenetic and biogeographical results highlighted patterns of plant use and the biological characters of Daodi medicinal plants. Medicinal cultural preferences could gain more scientific explanations via further pharmacophylogenetic investigations (Hao & Xiao, 2017a; Xiao et al., 2006).

Herbal nature (Yao Xing in Chinese) is an essential component of TCM theory, and the matching theory of meridian tropism is a core part of Yao Xing theory of TCM. The accuracy and pertinence of TCM can be improved by the meridian tropism based drug selection. The elusive associations between meridian tropism and family/genus, as well as the distribution pattern of meridian tropism on the phylogenetic tree, were preliminarily investigated in 2435 herbs and related 3044 species to provide a basis for the interpretation and evaluation of traditional meridian tropism (Li, Fu, Li & Wang, 2017). In viridiplantae, 1151 species belong to the liver meridian. In Streptophyta (seed plant), 1109 species belong to the liver meridian. Interestingly, in Liliopsida (monocot), 110 species belong to the lung meridian. In the association rule mining, the association rules for the same meridian tropism were distributed at the same branch or nearby branch of the phylogenetic tree. For example, *Taxus* plants (Hao, Gu & Xiao, 2015b) were related with kidney meridian, Caprifoliaceae and *Rubia* had a close relationship with liver meridian, while large intestine meridian is related with *Punica*. There is a subtle relationship between Yao Xing and the genetic affinity relationship. Some TCMS belonging to species of the same or similar taxonomic groups may have the same meridian tropism. This pharmacophylogeny study contributes a new index and reference for predicting and evaluating Yao Xing, TCM formula compatibility, and precision TCM.

Chemical constituents are the material basis of TCM properties, and their relationship is the basis of revealing the molecular mechanism of drug action. Based on network analysis, the TCM properties of medicines from the same botanical family can be argued with respect to the chemical components therein (Cao & Wang, 2013). Firstly, a drug network corresponding to the plant phylogeny was constructed based on the chemical components. The network structure showed that the relatives of the same family in the network were aggregated, and the network could express the similarity of chemical components between botanical drugs. Then, the Yao Xing distribution of drugs in the network was analyzed. The similar drug properties are generally connected in the network, and there is also the “drug hole” – most adjacent drugs of a drug have a certain drug property, but the center drug itself does not have it. Finally, taking the relationship between Yao Xing of Ranunculaceae and Apiaceae plants and their respective chemical constituents as the example, it is proved that the medicinal properties of closely clustered relatives of the same family are similar; as long as the chemical components of different families are similar and their overlapping ingredients are specific, the medicinal properties of drugs from different families can be similar.

Based on the research progress of TCM theory of pungent (Xin in Chinese) flavor and of *Glehnia littoralis* (Bei Sha Shen in TCM), and using the image (Xiang in Chinese) thinking method of TCM's medicinal property formation, the medicinal property of Apiaceae was compared, analogized and integrated with plant taxonomy and pharmacophylogeny (Tang, 2011). The pungent flavor, the common Apiaceae TCM pharmacological effects and efficacy were analyzed systematically, and the relationship between “Xin flavor – vasoactive activity – chemical composition” of Apiaceae Chinese medicine was probed. Based on this idea and practice, the pungent flavor and the corresponding new efficacy of *Radix Glehniae* were further explored, and the experimental evidence was given, followed by the elaboration of the concept and scientific connota-

tion of TCM Yao Xing. Such studies have important guiding significance for defining the research contents, expanding the research direction and developing the modern research on the medicinal properties of Chinese materia medica.

3.7. Pharmacophylogeny vs. plant systematics

Pharmacophylogeny and plant systematics complement and promote each other. Plant systematics, with developments of centuries, can provide a basis for research and analysis of medicinal plant phylogeny. Through research of the relationship between morphology, chemistry and drug effects in human body, plant pharmacophylogeny can provide new clues and evidence for plant system arrangement. For example, a comprehensive pharmacophylogeny study of Ranunculaceae (Hao et al., 2015a; 2015d) found that the tribe Cimicifugeae, composed of *Beesia*, *Souliea*, *Cimicifuga* and *Actaea*, contains peculiar tetracyclic triterpenoids chemically, but does not contain two characteristic components of Ranunculaceae, ranunculin and magnoflorine, indicating that the tribe is specially present in Ranunculaceae. Special status. For another example, a comprehensive study of Chinese *Aconitum* and related groups has drawn preliminary conclusions on the systematic arrangement of these plants (Hao, Gu & Xiao, 2013c, 2014, 2015b).

The phylogenetic relationship of medicinal plants can be inferred based on the genomic data, therefore we proposed a new concept “pharmacophylogenomics” (Hao & Xiao, 2015; Hao et al., 2014, 2015a), in which phylogenomic approaches are extended to investigate drug discovery and development issues. Conventional Sanger chloroplast (cp) markers provide limited information to resolve species level relationships within plants, in particular within large genera. Cp genome data confidently resolved relationships among major groups of *Ficus* (59 species) and largely support current understanding based on nuclear sequence data (Bruun-Lund, Clement, Kjellberg & Rønsted, 2017). However, conflicts between the new plastome topology and previous nuclear studies are observed for both individual species as well as relationships among some sections at deeper levels. Conflicts could be caused by lack of resolution in the nuclear data or may indicate potential cyto-nuclear discordance. The utility of cp genomes in elucidating the evolutionary relationship is also shown in *Commiphora* (Khan et al., 2019), Betulaceae (Yang et al., 2019), and 689 vascular plant species of Ruili Botanical Garden, Yunnan, China (Liu et al., 2019), etc.

The functional phylogenomic analysis in nonmodel species based on homolog groups and inferred ortholog groups was demonstrated in the highly diverse clade Caryophyllales using high throughput transcriptome sequencing (Yang et al., 2015). The species phylogeny was reconstructed using a 1122-gene data set with a gene occupancy of 92.1%. Genes that underwent the greatest gene family expansion were concentrated among those involved in signal transduction and oxidoreduction, including a cytochrome P450 (CYP) gene encoding a key enzyme in the betalain synthesis pathway. Ortholog groups were also extracted from transcriptome datasets of multiple TCM (traditional Chinese medicine) plants to dissect their phylogeny (Hao et al., 2012). Pairwise orthologs were identified by comparative transcriptome analysis in 13 conifer species (Zhao et al., 2018a), from which the rate of diversification was calculated and a phylogenetic tree inferred. Orthologous single-copy genes extracted from transcriptome datasets were also useful in phylogenomic and population genomic analyses of Taxaceae and Cephalotaxaceae anticancer plants (Majeed, Singh, Choudhary & Bhardwaj, 2019; Olsson et al., 2018).

It is noted that the transcriptome sequencing of some representative species of important medicinal families/genera has been performed, e.g., *Clematis* (Hao, Gu & Xiao, 2013d; Liu et al., 2018a, b), *Aconitum* (Hao et al., 2013c; Rai et al., 2017), *Rhododendron* (Choudhary et al., 2018; Hu & Xiao, 1992), and *Epimedium*

(Guo & Xiao, 1999; Ma, Chen & Guo, 2018), etc., inspiring a new wave of large scale sequencing of congeners for transcriptome based phylogeny inference.

Whole cp genomes and transcriptomes can be combined with genotyping-by-sequencing (GBS) to dissect population genetics, phylogenomics and hybrid speciation of food/drug plants. Processes of lineage formation of five native Chinese walnut species were inferred by this approach (Zhao et al., 2018a). The processes of isolation generated diversity during glaciations, and the recent range expansion of *Juglans regia*, probably from multiple refugia, led to hybrid formation both within and between sections of the genus. In southern China, human dispersal of *J. regia* led to the appearance of *J. sigillata*, which could be an ecotype of the former and is now maintained as a landrace. In northern China, walnut hybridized with a distinct lineage of *J. mandshurica* to form *J. hopeiensis*, possibly a horticultural variety. Comparisons between whole cp genomes and nuclear (nr) transcriptome analyses provided conflicting evidence for the divergence time of Chinese *Juglans* taxa. *J. cathayensis* and *J. mandshurica* are poorly differentiated based on the genomic data. Episodes of climatic variation over the past 4.5 to 33.80 million years, as well as glacial advances and retreats and population isolation, have shaped Chinese walnut demography and evolution, and the gene flow and introgression are also present. This approach will be especially useful, if further supplemented with metabolomic/chemotaxonomic and pharmacodynamic data, in the pharmacophylogeny study of Daodi medicinal materials (geoherts; Hao & Xiao, 2018) of China.

So far, the whole nr genome sequencing is still a formidable and costly task, although theoretically it is superior to cp genome and transcriptome sequencing in decoding pharmacophylogeny. The *Potentilla micrantha* genome was sequenced and annotated (Buti et al., 2018), and RNA-Seq data from the different developmental stages of flowering and fruiting were used to predict genes. A 327 Mbp genome sequence of *P. micrantha*, spanning 2674 sequence contigs and covering 80% of the total genome size was obtained. The medicinal genus *Potentilla* has a larger genome size than *Fragaria*, but the recovered sequence scaffolds were obviously collinear at the micro-syntenic level with the genome of *F. vesca*, its closest sequenced relative. A total of 33,602 genes were predicted, and 95.1% of universal single-copy orthologous genes were complete within the presented sequence. The majority of the gene-rich regions have been sequenced. The genome data are a valuable foundation for future studies of berry (a commonly used medicinal part) development in Rosaceae, a medicine-rich family (Hao et al., 2015b). Other recently sequenced genomes include *Tectona grandis* L. F. (Yasodha et al., 2018), *Malaria oleifera* Chun et S. Lee ex S. Lee (Xu et al., 2019), and 689 vascular plant species of Ruili Botanical Garden (Liu et al., 2019), etc.

Hopefully, the assembly and annotation of medicinal plant genome can be driven by multi-omic data. The genome annotation of Sandalwood (*Santalum album* L.) predicted 38,119 protein-coding genes and 27.42% repetitive DNA elements (Mahesh et al., 2018). In-depth proteome analysis revealed the identities of 72,325 unique peptides, which confirmed 10,076 of the predicted genes. The addition of transcriptomic and proteogenomic approaches resulted in the identification of 53 novel proteins and 34 gene-correction events that were missed by genomic approaches. Proteogenomic analysis also helped in reassigning 1348 potential non-coding RNAs as bona fide protein-coding mRNAs. Mass spectrometry (MS)-based proteomic evidence provided an unbiased approach toward the identification of proteins encoded by organellar genomes, which are often missed in transcriptome datasets. The use of integrated omic approaches enhanced the quality of the assembly and annotation of nonmodel plant genome.

4. Application of informatics in the study of pharmacophylogeny

Pharmacophylogeny research is often based on a large number of original images, profiles and data. The storage, retrieval, collation and logical judgment of these data can only be achieved by means of computer technology. In storing and retrieving the original data, it is necessary to establish such databases as NAPRALERT (<https://napralert.org/>) established by College of Pharmacy, University of Illinois at Chicago, and Medicinal Plant Genomics Resource (<http://medicinalplantgenomics.msu.edu/>) of Plant Biology Department, Michigan State University. CMAUP is a database of collective molecular activities of useful plants (Zeng et al., 2019), including 2567 medicinal, 170 food, 1567 edible, 3 agricultural and 119 garden plants collected from or traditionally used in 153 countries and regions. The database HMOD integrates 23 genomic data, 172 transcriptome data, and 55 metabolome data of medicinal plants, and 18 sets of metabolic pathway information (Wang et al., 2018). These comprehensive databases of medicinal plants facilitate the deeper study of pharmacophylogeny and its extension into pharmacophylogenomics.

With the rapidly increasing genome sequencing projects, the professional genome databases have been established for various taxonomic groups. For instance, the GDR (Genome Database for Rosaceae) now houses multiple versions of whole genome assembly and annotation data from 14 species (Jung et al., 2019), and synteny among the newest genome assemblies from different species can be viewed through the new synteny browser, SynView. The phylogenetic distribution of specific traits could be considered with respect to the medicinally relevant ecophysiological strategies of species' tolerance to salinity and alkalinity (Saslis-Lagoudakis, Hua, Bui, Moray & Bromham, 2015), ideally by incorporating more complete, finer-scale geochemical information for modeling and prediction, as well as laboratory experiments.

The computer can store a large amount of raw data, facilitate retrieval, and make further logical judgments. Take "the multivariate analysis of the relationship between the shape, composition and laxative effect of rhubarb plants" as an example. By means of cluster analysis, principal component analysis and regression analysis, the morphology, chemical composition and laxative effect of 33 Chinese species of *Rheum* were transformed into input data for computer (Hao & Xiao, 2017a). Through Q-classification of quantitative taxonomy, a comprehensive and multi-feature comparison of various *Rheum* species was made, and a taxonomic tree was obtained.

The phylogenetic diagram of Q classification shows that the relationships and similarities among species within the wave-leaf group (*Sect. Rheum*), *Sect. Palmata* (*R. officinale*, *R. palmatum*, *R. tanguticum*), spike-sequence group (*Sect. Spiciformia*) and Tahuang group (*R. alexandrae* and *R. nobile*) are close (Hao & Xiao, 2017a), which is consistent with the traditional classification of Rhubarb plants. However, the belongings of species within *Sect. Acuminata* (*R. acuminatum*, *R. kialense*, *R. maculatum*, *R. yunnanense*), *Sect. Deserticola* (*Ser. Nana*, *Ser. Pumilae*, *Ser. Racemiferae*), and *Sect. Globulosa* (*R. globulosum*) are partly overlapping and mixed, suggesting that the classification of these three groups should be further comparatively studied.

Among the commercial medicinal materials of rhubarb, palm-leaf group is considered to be genuine rhubarb, while wave-leaf group is counterfeit rhubarb (Tu Da Huang in Chinese) and inferior rhubarb. Other groups are used sporadically or partially and are not mainstream products. The tree diagram of Q classification shows that there are obvious differences of comprehensive characteristics between palm-leaf group and other groups of Rhubarb. Although *Sect. Rheum* is close to palm-leaf group in systematic ar-

rangement, it can be seen from the tree diagram that there are significant differences of comprehensive features between the two groups. This reveals that the commercial Tu Da Huang (*R. australe*, *R. franzenbachii*, *R. hotaoense*, *R. wittrockii*, etc.), despite similar in appearance to authentic rhubarb, has many essential differences and cannot be mixed clinically.

The minimum spanning tree of Rhubarb characteristic (Fig. S1) shows that there is a close relationship and correlation between characteristics No. 11 (degree of leaf margin division), No. 27 (senescent abundant or not), No. 28 (with/without rhenin) and No. 32 (laxative effect) in the connection of shape-component-laxative effect of rhubarb. In addition, the quantitative relationship between purgative and external morphology can be found by the multiple regression analysis. For instance, for some rhubarb plants not yet been studied, the prediction of purgative role can be obtained by substituting the corresponding morphological values into the regression equation. Since our original data for purgative analysis are only classified into three levels, i.e., obviously: $ED_{50} < 1000$, weak: $ED_{50} 1000-5000$, and no or almost no: $ED_{50} > 5000$, the prediction results are also divided into three levels.

For example, there are other two rhubarb species, *R. coreanum* and *R. tataricum* (*Sect. Orbicularia*). The coding values of two rhubarb morphology characteristics, 3.1 and 0.1, are substituted into the stepwise regression equation, and the purgative values of *R. coreanum* and *R. tataricum* are 3.536~4 and 1.127~1, respectively, suggesting that the former has obvious purgative effect, and the latter has no such effect. The prediction results are in good agreement with the actual situation.

5. Conclusion and future perspective

Pharmacophylogeny is a new and multidisciplinary subject. Its fundamental task is to explore the correlation between plant genetic affinity, chemical components, and pharmacokinetic behavior/efficacy. At present, it can be clearly seen that both in theory and in practice, this new subject has shown a strong vitality. With the continuous cycle of practice-theory-practice in the future, it is bound to become more enlightening and play a greater role in drug development research. When highlighting pharmacophylogeny, we do not intend to reduce everything to phylogeny and ignore other factors. Not adhering to many other factors could lead the drug discovery into a wrong direction with many false judgments. Talking about plant phylogeny here is justified. Comparing genomes (transcriptomes) is better than comparing genes in phylogeny reconstruction (Hao et al., 2014, 2015a; Zhao et al., 2018b), and the chloroplast data are obviously not sufficient for a reliable taxonomy, an aspect closely related to the task of finding new drugs from the perspective of pharmacophylogeny and pharmacophylogenomics.

The phylogenetic comparative methods (PCMs) have been used by anthropologists to gain a sophisticated understanding of the evolution of religious, social and material culture (Teixidor-Toneu, Jordan & Hawkins, 2018), including the medicinal plant use in ethnomedicine. The universal link between plant phylogeny, phytochemistry and traditional/contemporary medicinal utility could be revealed by pharmacophylogeny methods. Optimized methods for routine and cost effective separation and quantification of major medicinal compounds in various taxonomic groups should be developed continuously, so are the easy-to-use pharmacological/pharmacokinetic screening methods, e.g., the fluorescent probes (Ma et al., 2017; Wu et al., 2017).

Studies of medicinal plant associated microbial communities shed further light on the diversity and distribution of the world's "hidden biodiversity" (Hao, Song, Mu, Hu & Xiao, 2016, 2018c; Hao & Xiao, 2017c), and help understand the intricate link between plant taxa, their chemical profiles and relevant

drug efficacy. Boundaries between plant and microbes should be broken at the interface of rhizosphere and in the name of holobiont. The plant-microbe/insect codiversification may shed new light on the correlation between biodiversity, chemodiversity and pharmacotherapy. Transcriptomes and ultraconserved elements (UCEs) were combined to illuminate the phylogeny of Apidae (Bossert et al., 2019). Novel phylogenetic evidence supports the placement of two enigmatic, oil-collecting genera (*Ctenoplectra* and *Tetrapedia*). This approach can also be used to generate large, multigene datasets for the analysis of phylogenetic relationships in medicinal families and genera.

Declaration of Competing Interest

The authors declare no conflict of interests.

Acknowledgments

This study is supported by the National Natural Science Foundation of China (No. 41977048), Scientific Research Funds Project of Liaoning Education Department (JDL2019012) and Natural Science Fund of Liaoning Province (20180550190). We thank Dr. Xiaojie Gu for her help in drawing the molecular structures of some compounds.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chmed.2020.03.002.

References

- Bai, N., He, K., Zhou, Z., Tsai, M. L., Zhang, L., Quan, Z., et al. (2010). Ent-kaurane diterpenoids from *rabdosia rubescens* and their cytotoxic effects on human cancer cell lines. *Planta Medica*, 76(2), 140–145.
- Bai, Z. F., Wang, X. Q., Xiao, P. G., & Liu, Y. (2012). Investigation on ethnopharmacology of gesneriaceae in Guangxi. *Journal of Chinese Medicinal Materials*, 2012(1), 20–23.
- Bossert, S., Murray, E. A., Almeida, E. A. B., Brady, S. G., Blaimer, B. B., & Danforth, B. N. (2019). Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of apidae. *Molecular Phylogenetics and Evolution*, 130, 121–131.
- Bruun-Lund, S., Clement, W. L., Kjellberg, F., & Rønsted, N. (2017). First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *ficus* L. (Moraceae). *Molecular Phylogenetics and Evolution*, 109, 93–104.
- Buti, M., Moretto, M., Barghini, E., Mascagni, F., Natali, L., Brilli, M., et al. (2018). The genome sequence and transcriptome of *potentilla micrantha* and their comparison to *fragaria vesca* (the woodland strawberry). *GigaScience*, 7(4), 1–14.
- Cai, Y. Q., Hu, J. H., Qjn, J., Sun, T., & Li, X. L. (2018). *Rhododendron molle* (Ericaceae): Phytochemistry, pharmacology, and toxicology. *Chinese Journal of Natural Medicines*, 16(6), 401–410.
- Cámara-Leret, R., Faurby, S., Macía, M. J., Balslev, H., Gödel, B., Svenning, J. C., et al. (2017). Fundamental species traits explain provisioning services of tropical American palms. *Nature Plants*, 3, 16220.
- Cao, J., & Wang, Y. (2013). Relationship between chemical constituents and herbs' properties of relative plant herbs. *China Journal of Chinese Materia Medica*, 38(3), 453–458.
- Choudhary, S., Thakur, S., Najjar, R. A., Majeed, A., Singh, A., & Bhardwaj, P. (2018). Transcriptome characterization and screening of molecular markers in ecologically important Himalayan species (*Rhododendron arboreum*). *Genome National Research Council Canada Genome Conseil National de Recherches Canada*, 61(6), 417–428.
- Ernst, M., Grace, O. M., Saslis-Lagoudakis, C. H., Nilsson, N., Simonsen, H. T., & Rønsted, N. (2015). Global medicinal uses of *euphorbia* L. (Euphorbiaceae). *Journal of Ethnopharmacology*, 176, 90–101.
- Ernst, M., Saslis-Lagoudakis, C. H., Grace, O. M., Nilsson, N., Simonsen, H. T., Horn, J. W., et al. (2016). Evolutionary prediction of medicinal properties in the genus *euphorbia* L. *Scientific Reports*, 6, 30531.
- Frezza, C., De Vita, D., Spinaci, G., Sarandrea, M., Venditti, A., & Bianco, A. (2019). Secondary metabolites of *tilia tomentosa* moench inflorescences collected in central Italy: Chemotaxonomy relevance and phytochemical rationale of traditional use. *Natural Product Research*, 12, 1–8.
- Frezza, C., Venditti, A., Sciabba, F., Tomai, P., Antonetti, M., Franceschin, M., et al. (2018). Phytochemical profile of *euphorbia peplus* L. collected in central Italy and NMR semi-quantitative analysis of the diterpenoid fraction. *Journal of Pharmaceutical and Biomedical Analysis*, 160, 152–159.
- Fu, X. J., Wang, Z. G., Wang, C., Li, X., Wang, H., & Zhao, J. (2015). The distribution and association relationships of marine Chinese medicine with different nature in the phylogenetic tree of marine organisms. *Modernization of Traditional Chinese Medicine and Materia Medica-World Science and Technology*, 17(11), 2189–2196.
- Garnatje, T., Peñuelas, J., & Vallès, J. (2017). Ethnobotany, phylogeny, and 'omics' for human health and food security. *Trends in Plant Science*, 22(3), 187–191.
- Grace, O. M., Buerki, S., Symonds, M. R., Forest, F., van Wyk, A. E., Smith, G. F., et al. (2015). Evolutionary history and leaf succulence as explanations for medicinal use in aloes and the global popularity of *Aloe vera*. *BMC Evolutionary Biology*, 15, 29.
- Grace, O. M., Dzajic, A., Jäger, A. K., Nyberg, N. T., Önder, A., & Rønsted, N. (2013). Monosaccharide analysis of succulent leaf tissue in *Aloe*. *Phytochemistry*, 93, 79–87.
- Guo, B., & Xiao, P. G. (1999). The flavonoids in *epimedium* L. and their taxonomic significance. *Acta Phytotaxonomica Sinica*, 37(3), 228–243.
- Hao, D. C. (2018). *Ranunculales medicinal plants: Biodiversity, chemodiversity and pharmacotherapy*. London: Academic Press.
- Hao, D. C., Ge, G. B., Wang, P., & Yang, L. (2018b). Impact of drug metabolism/pharmacokinetics and their relevance upon taxus-based drug development. *Current Drug Metabolism*, 19(11), 930–959.
- Hao, D. C., Ge, G. B., & Xiao, P. G. (2018a). Anticancer drug targets of salvia phytochemicals: Chemistry, biology and omics. *Current Drug Targets*, 19(1), 1–20.
- Hao, D. C., Ge, G. B., Xiao, P. G., Wang, P., & Yang, L. (2015c). Drug metabolism and pharmacokinetic diversity of ranunculaceae medicinal compounds. *Current Drug Metabolism*, 16(4), 294–321.
- Hao, D. C., Gu, X. J., & Xiao, P. G. (2013a). Phytochemical and biological research of *fritillaria* medicinal resources. *Chinese Journal of Natural Medicines*, 11(4), 330–344.
- Hao, D. C., Gu, X. J., & Xiao, P. G. (2013b). Recent advance in chemical and biological studies on cimicifugeae pharmaceutical resources. *Chinese Herbal Medicines*, 5(2), 81–95.
- Hao, D. C., Gu, X. J., & Xiao, P. G. (2013c). Recent advances in the chemical and biological studies of *aconitum* pharmaceutical resources. *Journal of Chinese Pharmaceutical Sciences*, 22(3), 209–221.
- Hao, D. C., Gu, X. J., & Xiao, P. G. (2013d). Chemical and biological research of clematis medicinal resources. *Chinese Science Bulletin*, 58(10), 1120–1129.
- Hao, D. C., Gu, X. J., & Xiao, P. G. (2015b). *Medicinal plants: Chemistry, biology and omics*. Oxford: Woodhead Publishing.
- Hao, D. C., Ma, P., Mu, J., Chen, S., Xiao, P., Peng, Y., et al. (2012). De novo characterization of the root transcriptome of a traditional Chinese medicinal plant *polygala cuspidatum*. *Science China Life Sciences*, 55(5), 452–466.
- Hao, D. C., Song, S. M., Mu, J., Hu, W. L., & Xiao, P. G. (2016). Unearthing microbial diversity of *taxus rhizosphere* via MiSeq high-throughput amplicon sequencing and isolate characterization. *Scientific Reports*, 6, 22006.
- Hao, D. C., Xiao, P. G., Liu, M., Peng, Y., & He, C. N. (2014). Pharmaphylogeny vs. pharmacophylogenomics: Molecular phylogeny, evolution and drug discovery. *Yao xue xue bao Acta pharmaceutica Sinica*, 49(10), 1387–1394.
- Hao, D. C., & Xiao, P. G. (2015). Genomics and evolution in traditional medicinal plants: Road to a healthier life. *Evolutionary Bioinformatics*, 11, 197–212.
- Hao, D. C., & Xiao, P. G. (2017a). *An introduction of plant pharmacophylogeny*. Beijing: Chemistry Industry Press.
- Hao, D. C., & Xiao, P. G. (2017b). Impact of drug metabolism/pharmacokinetics and their relevance upon salvia based drug discovery. *Current Drug Metabolism*, 18(12), 1071–1084.
- Hao, D. C., & Xiao, P. G. (2017c). Rhizosphere microbiota and microbiome of medicinal plants: From molecular biology to omics approaches. *Chinese Herbal Medicines*, 9, 199–217.
- Hao, D. C., & Xiao, P. G. (2018). Deep in shadows: Epigenetic and epigenomic regulations of medicinal plants. *Chinese Herbal Medicines*, 10, 239–248.
- Hao, D. C., Xiao, P. G., Liu, L. W., Peng, Y., & He, C. N. (2015a). Essentials of pharmacophylogeny: Knowledge pedigree, epistemology and paradigm shift. *China Journal of Chinese Materia Medica*, 40(17), 3335–3342.
- Hao, D. C., Xiao, P. G., Ma, H. Y., Peng, Y., & He, C. N. (2015d). Mining chemodiversity from biodiversity: Pharmacophylogeny of medicinal plants of ranunculaceae. *Chinese Journal of Natural Medicines*, 13(7), 507–520.
- Hao, D. C., & Yang, L. (2016). Drug metabolism and disposition diversity of ranunculales phytometabolites: A systems perspective. *Expert Opinion on Drug Metabolism and Toxicology*, 12(9), 1047–1065.
- Hao, D. C., Zhang, C. R., & Xiao, P. G. (2018c). The first *taxus* rhizosphere microbiome revealed by shotgun metagenomic sequencing. *Journal of Basic Microbiology*, 58(6), 501–512.
- He, C. N., Peng, Y., Xiao, W., & Xiao, P. G. (2012). A preliminary study on ethnopharmacology of *scutellaria* in China. *Modern Chinese Medicine*, 2012(1), 16–20.
- Hegnauer, R. (1973). *Chemotaxonomie der pflanzen: Eine Übersicht über die verbreitung und die systematische bedeutung der pflanzenstoffe*. Germany: Birkhäuser.
- Hu, M., & Xiao, P. G. (1992). Chemotaxonomic study of *rhododendron*. *Acta Phytotaxonomica Sinica*, 30(3), 226–238.
- Huang, H. H., Li, X., Gao, W. Y., & Xiao, P. G. (2015). Studies on genetic relationship of *dioscorea*. *Zhongguo Zhongyao Za Zhi*, 40, 3470–3479.
- Johnson, S. R., Bhat, W. W., Bibik, J., Turmo, A., Hamberger, B., & Consortium, Evolutionary Mint Genomics (2019). A database-driven approach identifies additional diterpene synthase activities in the mint family (Lamiaceae). *Journal of Biological Chemistry*, 294(4), 1349–1362.
- Jung, S., Lee, T., Cheng, C. H., Buble, K., Zheng, P., Yu, J., et al. (2019). 15 years of

- GDR: New data and functionality in the genome database for rosaceae. *Nucleic Acids Research*, 47(D1), 1137–1145.
- Khan, A., Asaf, S., Khan, A. L., Al-Harrasi, A., Al-Sudairy, O., AbdulKareem, N. M., et al. (2019). First complete chloroplast genomics and comparative phylogenetic analysis of *commiphora gileadensis* and *C. foliacea*: Myrrh producing trees. *PLoS one*, 14(1), e0208511.
- Larson, G. M., Schaneberg, B. T., & Sneden, A. T. (1999). Two new maytansinoids from maytenus buchananii. *Journal of Natural Products*, 62(2), 361–363.
- Larsson, S., & Rønsted, N. (2014). Reviewing colchicaceae alkaloids - perspectives of evolution on medicinal chemistry. *Current Topics in Medicinal Chemistry*, 14(2), 274–289.
- Lei, D., Wu, J., Leon, C., Huang, L. F., & Hawkins, J. A. (2018). Medicinal plants of Chinese pharmacopoeia and Daodi: Insights from phylogeny and biogeography. *Chinese Herbal Medicines*, 10(3), 269–278.
- Li, J. Y., Fu, X. J., Li, X., & Wang, Z. (2017). Association relationships of traditional Chinese medicine with different meridian tropism in phylogenetic tree. *Chinese Journal of Experimental Traditional Medical Formulae*, 23(12), 194–200.
- Li, M. H., Li, Q. Q., Liu, Y. Z., Cui, Z. H., Zhang, N., Huang, L. Q., et al. (2013). Pharmacophylogenetic study on plants of genus *salvia* L. from China. *Chinese Herbal Medicines*, 5(3), 164–181.
- Liu, H., Wei, J., Yang, T., Mu, W., Song, B., Yang, T., et al. (2019). Molecular digitization of a botanical garden: High-depth whole genome sequencing of 689 vascular plant species from the Ruli botanical garden. *GigaScience*, 8(4) pii: Giz007.
- Liu, H. T., Qi, Y. D., Xu, L., Peng, Y., Zhang, B. G., & Xiao, P. G. (2012a). Investigation on ethnopharmacology of schisandraceae plants in China. *China Journal of Chinese Materia Medica*, 2012(10), 1353–1359.
- Liu, Y., Zhang, H., Umashankar, S., Liang, X., Lee, H. W., Swarup, S., et al. (2018a). Characterization of plant volatiles reveals distinct metabolic profiles and pathways among 12 brassicaceae vegetables. *Metabolites*, 8(4) pii: E94.
- Liu, Y. Z., Chen, S. L., Ma, P., & Xiao, P. G. (2012b). Ways to find new anticancer drugs from Chinese herbal medicines. *Drugs & Clinic*, 27(4), 323–337.
- Liu, Z., Shao, W., Shen, Y., Ji, M., Chen, W., Ye, Y., et al. (2018b). Characterization of new microsatellite markers based on the transcriptome sequencing of *clematis finetiana*. *Hereditas*, 155, 23.
- Ma, H. Y., Yang, J. D., Hou, J., Zou, L. W., Jin, Q., & Hao, D. C. (2017). Comparative metabolism of DDAO benzoate in liver microsomes from various species. *Toxicology In Vitro*, 44, 280–286.
- Ma, Y., Chen, X., & Guo, B. (2018). Identification of genes involved in metabolism and signalling of abscisic acid and gibberellins during *epimedium pseudowushanense* B. L. Guo seed morphophysiological dormancy. *Plant Cell Reports*, 37(7), 1061–1075.
- Mahesh, H. B., Subba, P., Advani, J., Shirke, M. D., Loganathan, R. M., Chandana, S. L., et al. (2018). Multi-omics driven assembly and annotation of the sandalwood (*Santalum album*) genome. *Plant Physiology*, 176(4), 2772–2788.
- Majeed, A., Singh, A., Choudhary, S., & Bhardwaj, P. (2019). RNAseq-based phylogenetic reconstruction of taxaceae and cephalotaxaceae. *Cladistics: The International Journal of the Willi Hennig Society*. doi:10.1111/cla.12362.
- Maldonado, C., Barnes, C. J., Cornett, C., Holmfred, E., Hansen, S. H., Persson, C., et al. (2017). Phylogeny predicts the quantity of antimalarial alkaloids within the iconic yellow cinchona bark (Rubiaceae: *Cinchona calisaya*). *Frontiers in Plant Science*, 8, 391.
- Nawrot, J., Budzianowski, J., & Nowak, G. (2019). Phytochemical profiles of the leaves of *stizolophus balsamita* and *psephellus sibiricus* and their chemotaxonomic implications. *Phytochemistry*, 159, 172–178.
- Olsson, S., Pinosio, S., González-Martínez, S. C., Abascal, F., Mayol, M., Grivet, D., et al. (2018). De novo assembly of English yew (*Taxus baccata*) transcriptome and its applications for intra- and inter-specific analyses. *Plant Molecular Biology*, 97(4–5), 337–345.
- Pellicer, J., Salsis-Lagoudakis, C. H., Carrió, E., Ernst, M., Garnatje, T., Grace, O. M., et al. (2018). A phylogenetic road map to antimalarial artemisia species. *Journal of Ethnopharmacology*, 225, 1–9.
- Pertuit, D., Mitaine-Offer, A. C., Miyamoto, T., Tanaka, C., Delaude, C., & Lacaillè-Dubois, M. A. (2018). Terpenoid glycosides from the root's barks of *ericoelium microspermum* radlk. ex engl. *Phytochemistry*, 152, 182–190.
- Rai, M., Rai, A., Kawano, N., Yoshimatsu, K., Takahashi, H., Suzuki, H., et al. (2017). De novo RNA sequencing and expression analysis of *aconitum carmichaelii* to analyze key genes involved in the biosynthesis of diterpene alkaloids. *Molecules*, 22(12) pii: E2155.
- Rønsted, N., Symonds, M. R., Birkholm, T., Christensen, S. B., Meerow, A. W., Molander, M., et al. (2012). Can phylogeny predict chemical diversity and potential medicinal activity of plants? A case study of amaryllidaceae. *BMC Evolutionary Biology*, 12, 182.
- Salsis-Lagoudakis, C. H., Hawkins, J. A., Greenhill, S. J., Pendry, C. A., Watson, M. F., Tuladhar-Douglas, W., et al. (2014). The evolution of traditional knowledge: Environment shapes medicinal plant use in Nepal. *Proceedings of the Royal Society B: Biological Sciences*, 281(1780), 20132768.
- Salsis-Lagoudakis, C. H., Hua, X., Bui, E., Moray, C., & Bromham, L. (2015). Predicting species' tolerance to salinity and alkalinity using distribution data and geochemical modelling: A case study using Australian grasses. *Annals of Botany*, 115(3), 343–351.
- Salsis-Lagoudakis, C. H., Savolainen, V., Williamson, E. M., Forest, F., Wagstaff, S. J., Baral, S. R., et al. (2012). Phylogenies reveal predictive power of traditional medicine in bioprospecting. *Proceedings of the National Academy of Sciences of the United States of America*, 109(39), 15835–15840.
- Souza, E. N. F., Williamson, E. M., & Hawkins, J. A. (2018). Which plants used in ethnomedicine are characterized? Phylogenetic patterns in traditional use related to research effort. *Frontiers in Plant Science*, 9, 834.
- Tang, S.H. (2011). Exploration and study on "Spicy" of radix glehniae based on plant pharmacophylogenetics. PhD dissertation, China Academy of Chinese Medical Sciences.
- Teixidor-Toneu, I., Jordan, F. M., & Hawkins, J. A. (2018). Comparative phylogenetic methods and the cultural evolution of medicinal plant use. *Nature Plants*, 4(10), 754–761.
- Vanderplanck, M., & Glauser, G. (2018). Integration of non-targeted metabolomics and automated determination of elemental compositions for comprehensive alkaloid profiling in plants. *Phytochemistry*, 154, 1–9.
- Wagle, A., Seong, S. H., Jung, H. A., & Choi, J. S. (2019). Identifying an isoflavone from the root of *pueraria lobata* as a potent tyrosinase inhibitor. *Food Chemistry*, 276, 383–389.
- Wang, X., Zhang, J., He, S., Gao, Y., Ma, X., Gao, Y., et al. (2018). HMOD: An omics database for herbal medicine plants. *Molecular Plant*, 11(5), 757–759.
- Wang, Z., He, C., Peng, Y., Chen, F., & Xiao, P. G. (2017). Origins, phytochemistry, pharmacology, analytical methods and safety of *cortex moutan* (Paeonia suffruticosa andrew): A systematic review. *Molecules*, 22(6) pii: E946.
- Wei, S. Y. (2015). Progress on cardiovascular protections and mechanism research of pueraria. *China Journal of Chinese Materia Medica*, 40(12), 2278–2284.
- Wu, J. J., Cao, Y. F., Feng, L., He, Y. Q., Hong, J. Y., Dou, T. Y., et al. (2017). A naturally occurring isoform-specific probe for highly selective and sensitive detection of human cytochrome p450 3A5. *Journal of Medicinal Chemistry*, 60(9), 3804–3813.
- Xiao, J., Pan, Y., Zhang, L., Wang, X., Han, Y., Sun, L., et al. (2019). High performance liquid chromatography determination and optimization of the extraction process for the total alkaloids from traditional herb *stephania cepharantha* hayata. *Molecules*, 24(3) pii: E388.
- Xiao, P. G., Wang, F. P., & Gao, F. (2006). A pharmacophylogenetic study of *aconitum* L. (Ranunculaceae) from China. *Journal of University of Chinese Academy of Sciences*, 23(1), 1–3.
- Xu, C. Q., Liu, H., Zhou, S. S., Zhang, D. X., Zhao, W., Wang, S., et al. (2019). Genome sequence of *malania oleifera*, a tree with great value for nervonic acid production. *GigaScience*, 8(2) pii: Gij164.
- Yang, Y., Moore, M. J., Brockington, S. F., Soltis, D. E., Wong, G. K., Carpenter, E. J., et al. (2015). Dissecting molecular evolution in the highly diverse plant clade Caryophyllales using transcriptome sequencing. *Molecular Biology and Evolution*, 32(8), 2001–2014.
- Yang, Z., Wang, G., Ma, Q., Ma, W., Liang, L., & Zhao, T. (2019). The complete chloroplast genomes of three betulaceae species: Implications for molecular phylogeny and historical biogeography. *Peer Journal*, 7, e6320.
- Yasodha, R., Vasudeva, R., Balakrishnan, S., Sakthi, A. R., Abel, N., Binai, N., et al. (2018). Draft genome of a high value tropical timber tree, teak (*Tectona grandis* L. f): Insights into SSR diversity, phylogeny and conservation. *DNA Research*, 25(4), 409–419.
- Zeng, Q. Y. (2018). From looking for new medicines from ethnomedicines to modernization of traditional Chinese medicine—Construction and practice of Zeng Yulin's ethnopharmaceutical theory. *Chinese Journal of Ethnomedicine and Ethnopharmacology*, 27(12), 134–138.
- Zeng, X., Zhang, P., Wang, Y., Qin, C., Chen, S., He, W., et al. (2019). CMAUP: A database of collective molecular activities of useful plants. *Nucleic Acids Research*, 47(D1), 1118–1127.
- Zhang, X., Bi, Q., Wu, X., Wang, Z., Miao, Y., & Tan, N. (2018). Systematic characterization and quantification of rubiaceae-type cyclopeptides in 20 *rubia* species by ultra performance liquid chromatography tandem mass spectrometry combined with chemometrics. *Journal of Chromatography A*, 1581–1582, 43–54.
- Zhao, J., Ding, H. X., Zhao, D. G., Wang, C. M., & Gao, K. (2012). Isolation, modification and cytotoxic evaluation of flavonoids from *rhododendron hainanense*. *Journal of Pharmacy and Pharmacology*, 64(12), 1785–1792.
- Zhao, P., Zhou, H. J., Potter, D., Hu, Y. H., Feng, X. J., Dang, M., et al. (2018a). Population genetics, phylogenomics and hybrid speciation of *Juglans* in China determined from whole chloroplast genomes, transcriptomes, and genotyping-by-sequencing (GBS). *Molecular Phylogenetics and Evolution*, 126, 250–265.
- Zhao, Y. J., Cao, Y., Wang, J., & Xiong, Z. (2018b). Transcriptome sequencing of *pinus kesiya* var. *langbianensis* and comparative analysis in the *pinus* phylogeny. *BMC Genomics*, 19(1), 725.
- Zou, X. X., Huang, L. Q., Cui, G. H., Yuan, Q. J., Peng, Y., Liu, Y., et al. (2009). Genetic relationships of *atractylodes* plants. *Acta Pharmaceutica Sinica*, 44(6), 680–686.