Open Access

SCHOLARENA | Journal of Horticulture Science and Forestry

RESEARCH ARTICLE

Screening on DNA Barcodes for Discriminating *Bupleurum* Species

Wen-jing Han, Xuan Liu, He-fang Wan, Bao-sheng Liao, Jian-ping Han, Chun Sui* and Jian-he Wei*

Institute of Medicinal Plant Development (IMPLAD), Chinese Academy of Medical Sciences & Peking Union Medical College (Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education & National Engineering Laboratory for Breeding of Endangered Medicinal Materials), Beijing 100193, China

***Corresponding authors:** Chun Sui, Institute of Medicinal Plant Development (IMPLAD), Chinese Academy of Medical Sciences & Peking Union Medical College (Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education & National Engineering Laboratory for Breeding of Endangered Medicinal Materials), Beijing 100193, China, E-mail: 791537534@qq.com

Jian-he Wei, Institute of Medicinal Plant Development (IMPLAD), Chinese Academy of Medical Sciences & Peking Union Medical College (Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education & National Engineering Laboratory for Breeding of Endangered Medicinal Materials), Beijing 100193, China, E-mail: wjianh@263.net

Citation: Wen-jing Han, Xuan Liu, He-fang Wan, Bao-sheng Liao, Jian-ping Han, et al. (2021) Screening on DNA Barcodes for Discriminating *Bupleurum* Species. J Hortic Sci For 3: 102

Abstract

Objective: Chaihu, roots of some species from *Bupleurum* L., is one of the famous and widely-used traditional Chinese medicines. With the expanding of domestic cultivation, emerged a few of typical cultivated types. The aim of this study was to screen effective DNA barcodes among five universally used, *psbA-trnH*, *matK*, *rbcL*, ITS and ITS2, for identification of *Bupleurum* species and cultivated types.

Methods: A total of 56 samples were used which represent six species and were collected from wild and cultivated fields. Sequences of *psbA-trnH*, *matK*, *rbcL* and ITS of tested samples were cloned and sequenced. CodonCode Aligner was used for sequences analyses. The variable sites in DNA barcodes were compared and phylogenetic trees were constructed.

Results: Compared with other DNA barcodes, ITS and ITS2 showed more discriminable for the tested six species.
Conclusion: Although ITS and ITS2 could be used for discrimination of different *Bupleurum* species, a multifaceted system will be needed to identify different cultivated types of *Bupleurum*.

Keywords: Chaihu; *Bupleurum* L; DNA Barcoding; Identification

Introduction

Bupleurum L. is a large genus represented by about 180-190 species, which are widely distributed in the Northern Hemisphere (She & Watson, 2005). Medicinal Radix Bupleuri (Chaihu; Chinese Thorowax Root), sourced from the dried roots of *Bupleurum* species (Umbelliferae family), has been used in ancient China for about two thousand years with its anti-inflammatory, anti-ulcer, anti-pyretic, antioxidant, anti-tumor and anti-hepatotoxic efficacies (Chinese Pharmacopoeia Committee, 2020; Sui et al., 2020). Nowadays, it is still widely used in China, Japan, Korea and other south Asian countries, as well as in Europe and North Africa. In China, Chaihu has been used as principal or adjuvant agent in more than seventy classical and formulated prescriptions. Most of Chaihu decoction pieces are used as principal agent for production of Chinese patent drugs, such as Chaihu injection, cold heat granules, liver-protecting tablet and Zhengchaihu capsule, and small part used as ingredient of formula, such as Xiao Chaihu Tang (Minor Decoction of *Bupleurum*) (Zhang, 2020) and Chaihu Shugan San (*Bupleurum* Liver-Coursing Powder) (Nie, Deng, & Zheng,2020). Chaihu extracts, including saikosaponins, volatile oils, flavones and polysaccharides, are widely used in pharmaceutical, chemical, food, health care products, cosmetics, pesticide and veterinary drug industry.

In China, a total of 42 species, 17 varieties and 7 variants were recorded in Flora of China. More than 20 species were once surveyed being used with their roots or the whole plant for medicinal purpose through all districts of China (Xiao, 2005). Because uncertain pharmacological differences existing in such multiple species and remarkable negative effects which were ever found, such as toxicity of *B. longiradiatum* Turcz., only two species, *B. chinense* DC. and *B. scorzonerifolium* Willd., are officially specified as medicinal source species (Chinese Pharmacopoeia Committee, 2020). Therefore, effective species discrimination is not only important for taxonomy of *Bupleurum* genus, but also essential for their medicinal usage. Actually, the classical identification of *Bupleurum* species based on plant morphology, such as leaf shape, number and shape of bracts, number of flowers in umbellule, is very difficult (Yao, Chen, Zhang, Li & Yang, 2016). Even though extensive taxonomic research has been carried out from characters in plant morphology, chromosome numbers, fruit dissection, root microscopy and pollen (Ostroumova & Kljuykov, 2015; Wang, He, Zhou, Wu, Yu, & Pang, 2008), it still causes for concern of botanist, pharmacologist and pharmaceutist that it is so difficult to discriminate plants in planting field as well as medicinal dried roots from some of *Bupleurum* species. Several works on molecular identification method were reported, including on nrDNA ITS region (Lin, Chen, & Lin, 2008; Yuan et al., 2017; Yuan, Ma, Yang, Zhou, Lin, & Liu, 2016), restriction site variation of chloroplast DNA (Matsumoto, Ohta, Yuan, Zhu, Okada, & Miyamoto, 2004), DNA fingerprints (Mizukami, Ohbayashi, & Ohashi, 1993) and chloroplast sequence, *rps16* (Wang, Zhou, Liu, Pang, Wu, & He, 2008). It was found that misidentification was not uncommon for herbarium specimens and some *Bupleurum* species in Flora of China are difficult to distinguish (Wang, Ma, & He, 2011).

On the aspect of medical usage, Chaihu has largely sourced from cultivation since 1970's. It was estimated that about one third merchandise Chaihu was derived from cultivation nowadays in China. Cultivated *Bupleurum* was firstly and spontaneously domesticated from wild species by local farmers. The species of *B. kaoi*, *B. falcatum* and *B. chinense* are planted and used in Taiwan. The species of *B. falcatum* is widely used in Japan and Korea (Choi, Kang, Park, & Kim, 1995; Liu, Shyu, Hu, & Chiu, 1989). Still there are much more species cultivated in mainland of China, such as *B. chinense*, *B. scrozonerifolium*, *B. yinchowense*, *B. falcatum*. Up to date, only a few bred cultivars were reported, such as *B. falcatum* cv. Tainung No.1 (Liu, Wang, Lee, & Lee, 1991), *B. chinense* cv. Zhongchai No. 1, No. 2, No. 3 and *B. scorzonerifolium* cv. Zhonghongchai No. 1 in China (Zheng, Sui, Wei, Jin, Chu, & Yang, 2010). It is unclear on the classification of quite most of cultivated *Bupleurum* and systematic discrimination on medicinal materials from different *Bupleurum* species have not been explored. Since some plant traits changed dramatically after longtime domestic cultivation and the hybrid introduction from different populations and different regions also brought some ambiguous morphological characters, it is very necessary to analyze the sources and status of cultivated *Bupleurum* for proper medicinal usage.

The vital goal of plant DNA barcoding research is to identify species by one or several DNA markers. Since 2003, unprecedented great efforts have been made in plant barcoding using for reference techniques initiated in zoology (Kress, Wurdack, Zimmer, Weigt, & Janzen, 2005; Li et al., 2011; Zhang, Chen, Dong, Lin, Fan, & Chen, 2015). Up to 2009, CBOL Plant Working Group proposed to use the combination of *rbcL* + *matK* as a core plant barcode. Both the plastid *psbA-trnH* and ITS (or ITS2) were suggested as complementary markers to the proposed core-barcode of *rbcL* + *matK,* to be further evaluated within 18 months during the Third International Barcoding of Life Conference in Mexico City (Pang, Luo, & Sun, 2012). From then on, much works were conducted to evaluate the candidate barcodes in large-scale plant species, including the whole seed plants, ferns and within some individual families or genera, Rutaceae, Rosaceae, Arecaceae, Fabaceae, Araliaceae*, Fraxinus, Parnassia*, *Panax* (Li et al., 2011; Wang, Lu, Wen, Ebihara, & Li, 2016). It was found that ITS2 cannot solve all the species determination problems in medicinal vines and proposed the use of ITS2 secondary structural information in differentiating species in the family Araliaceae (Liu, Zeng, Yang, Chu, Yuan, & Chen, 2012). In the present study, the usefulness of *rbcL*, *matK*, *trnH-psbA*, ITS and ITS2 in discrimination of *Bupleurum* species were evaluated. Meanwhile cultivated *Bupleurum* was attempted to be classified. Our results will be valuable for further medicinal *Bupleurum* identification.

Materials and methods

Plant materials

A total of 56 samples representing six *Bupleurum* species were collected from nine provinces of China, i.e., Heilongjiang (109°55′E-129°56′E 33°51′N-47°9′N), Liaoning (111°13′E-124°23′E, 35°35′N-40°7′N), Hebei (117°19′E-117°49′E, 40°24′N-40°57′N), Beijing (115°57′E-116°5′E, 39°56′N-40°26′N), Gansu (104°37′E， 34°59′N), Sichuan (105°2′E-105°13′E, 32°12′N-32°34′N), Shaanxi (109°55′E-116°39′E, 33°51′N-39°56′N), Henan (111°36′E-112°29′E, 33°46′N-34°41′N) and Shanxi (110°49′E， 35°24′N) (Table 1S). The samples were collected during the flowering and fruiting period of *Bupleurum*. Clean leaves without disease were collected, dried in color-changing silica gel and then stored in laboratory (4°). Meanwhile, intact plants were pressed into specimens and preserved in the herbarium of Institute of Medicinal Plant Development. All collected samples were identified by Professor Chunsheng Liu (Beijing University of Chinese Medicines) and Jianhe Wei (Institute of Medicinal Plant Development, Chinese Academy of Medicinal Sciences), and the corresponding voucher specimens were deposited in the Herbarium of the Institute of Medicinal Plant Development. The sequences of *psbA-trnH*, *matK*, *rbcL*, ITS and ITS2 were cloned and sequenced for the 56 samples. In addition, sequences from the six species were downloaded from GenBank and were used for further evaluating the discriminability of the *psbA-trnH*, *matK*, *rbcL*, ITS and ITS2.

DNA extraction, PCR amplification and sequencing

Total DNA was extracted from silica gel dried leaves using a Plant DNA Mini Kit (Omega Bio-Tek, Doraville, GA, USA). Primers and reaction conditions were used according to previous reports (Chen et al., 2010) (Table 2S) except that the 2 × PCR Solution PrimeSTAR™ HS Premix (*TaKaRa*, Japan) comprising extreme high-fidelity cloning enzyme was used for PCR. Purified PCR products were sequenced in both directions on a 3730XL sequencer (Applied Biosystems, Carlsbad, CA, USA). Contigs were generated using Codoncode Aligner version 6.0.2 (Codon-Code, Dedham, MA, USA), and quality control of the sequences was performed as previously described (Chen et al., 2010).

Sequences downloading from GenBank and sequences processing

Except for the 56 sample sequences of psbA-trnH, matK, rbcL, ITS or ITS2, which were extracted by ourselves, the remaining sequences were downloaded from GenBank, totaling six species. All sequences were manually checked for retrieval according to GenBank annotations. Organize all downloaded sequences to set up local data sets for data analysis.

Data Analyses

All sequences we cloned and downloaded were aligned using MEGA 7.0, and the variable site and haplotypes of these samples was analyzed by DnaSP version 5.10.1. Genetic distances were calculated as interspecific distance and intraspecific distance and neighbor-joining (NJ) tree analyses with 1000 bootstrap replicates using the K2P (Kimura 2-parameter) model.

Results

PCR amplification efficiency and sequence analysis

DNA barcode regions (*rbcL*, *matK*, *psbA-trnH*, ITS) were amplified and the ITS2 was retrieved from the amplified sequences of ITS. All 56 samples were successfully amplified by universal primers. Assessments of the sequence quality and coverage for the four regions showed that high-quality bidirectional sequences were obtained. Only ITS regions for several samples has lower-quality bases at both ends of PCR amplified sequences, but the bases of ITS1 and ITS2 regions which were retrieved as barcodes were in high quality. The amplicon size was 604-606 bp for ITS, 791-828 bp for *matK*, 379-462 bp for *psbA-trnH*, 553 bp for *rbcL*, and the ITS2 was 226 bp.

K2P distance of Bupleurum species based on five DNA barcodes

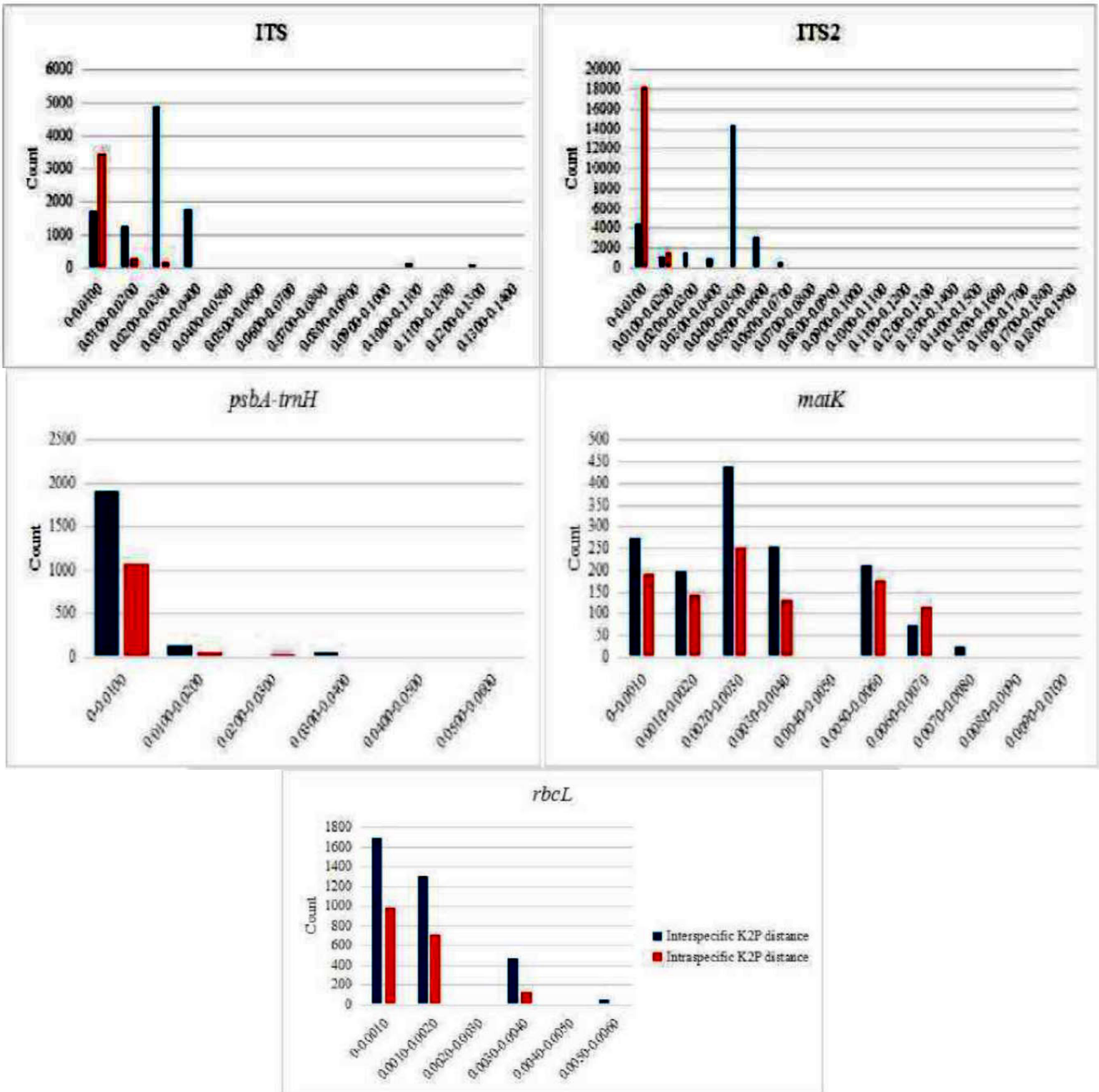

Figure 1: Dendrogram of 26 wild Bupleurum samples and 18 downloaded sequences based on psbA-trnH sequences. Red fonts indicated that these sequences were amplified from Bupleurum samples in this study

Barcode	Interspecific K2P distance		Intraspecific K2P distance	
	Range	Average	Range	Average
ITS	0-0.1400	0.0250	0-0.1259	0.0060
ITS2	0-0.1846	0.0393	0-0.1615	0.0055
psbA-trnH	0-0.0558	0.0060	0-0.0431	0.0057
matK	0-0.0090	0.0028	0-0.0077	0.0030
rbcL	0-0.0055	0.0012	0-0.0055	0.0010

Table 1: Interspecific and intraspecific K2P distance of five barcodes in *Bupleurum*

Interspecific and intraspecific K2P distances of all *Bupleurum* species based on five DNA barcodes were analyzed by MEGA 7.0. The results are shown in Table 1. The average interspecific K2P distance of ITS and ITS2 were much greater than intraspecific K2P mean distance. While the average interspecific K2P distance of *psbA-trnH*, *matK* and *rbcL* were almost equal to the average of their intraspecific K2P distance (Table 1). The distribution of interspecific and intraspecific K2P distance for five DNA barcodes in the six species of *Bupleurum* was shown in Fig. 1. The majority of interspecific distance for ITS and ITS2 in *Bupleurum species* were much greater than intraspecific distance. And there was no obvious difference between interspecific and intraspecific K2P distance of *psbA-trnH*, *matK*, *rbcL*. Consequently, ITS and ITS2 were more suitable for identification of *Bupleurum* species from the aspects of K2P distance.

Efficiency of five DNA barcodes in discriminating wild Bupleurum species

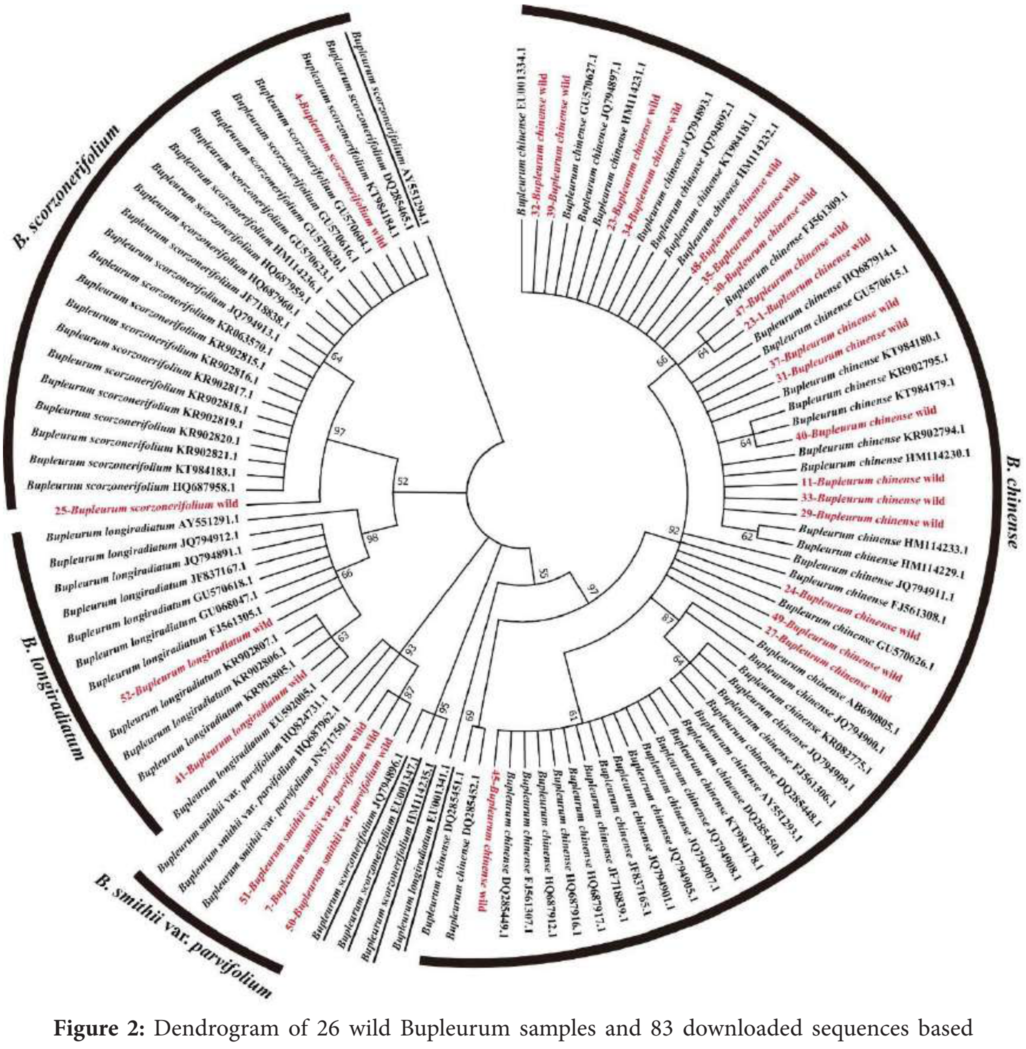

Figure 2: Dendrogram of 26 wild Bupleurum samples and 83 downloaded sequences based on ITS sequences. Red fonts indicated that these sequences were amplified from Bupleurum samples in this study, and the underlined ones were not clustered into their affiliated species

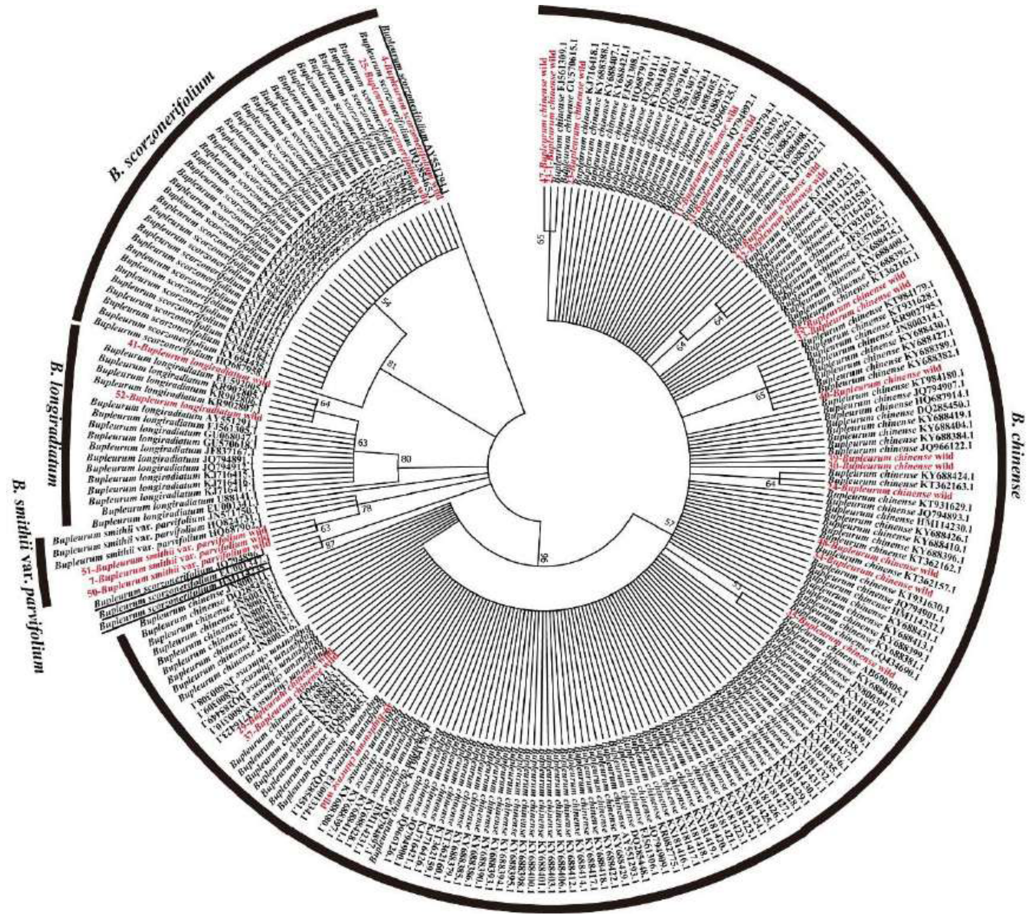

Figure 3: Dendrogram of 26 wild *Bupleurum* samples and 216 downloaded sequences based on ITS2 sequences. Red fonts indicated that these sequences were amplified from *Bupleurum* samples in this study, and the underlined ones were not clustered into their affiliated species. Red font indicated these sequences were amplified from 26 wild *Bupleurum* samples, and the underlined ones were not clustered into their affiliated species

The 26 wild *Bupleurum* samples were morphologically identified as four species, *B. chinense*, *B. scorzonerifolium*, *B. longiradiatum* and *B. smithii var. parvifolium*. To analyze the discriminability of five DNA barcodes, the sequences of 26 samples and the downloaded sequences for the four species were together clustered (Fig. 2, 3 and Fig. 1S, 2S, 3S). For ITS and ITS2 showed high resolution in all five DNA barcodes, all wild *Bupleurum* samples can be identified clearly. Majority of downloaded GenBank sequences were clustered into their affiliated species branches. But some of the downloaded sequences which were underlined in Fig. 2 and 3 were not shown the same rules. While the NJ trees of *psbA-trnH*, *matK*, *rbcL* (Fig. 1S, 2S, 3S) revealed a different pattern, in which four *Bupleurum species* were attributed to same branches promiscuously. Therefore, these three barcodes were not appropriate for discrimination *in Bupleurum* species.

Evaluating the discriminability of ITS and ITS2 for the six species of Bupleurum

In order to see the discriminability of ITS and ITS2 on both wild and cultivated *Bupleurum*, the ITS and ITS2 sequences of all 56 samples and all downloaded sequences were used to construct the phylogenetic tree (Fig. 4 and 5). In both of the barcodes, *B. scorzonerifolium*, *B. longiradiatum*, *B. smithii var. parvifolium* and *B. falcatum* clustered into one branch, respectively, while *B. chinense* and *B. yinchowense* were attributed into one joint branch. And there were no obvious difference between wild and cultivated samples from two NJ trees. In addition, several underlined sequences including sample sequences and downloaded sequences were not clustered into their corresponding species.

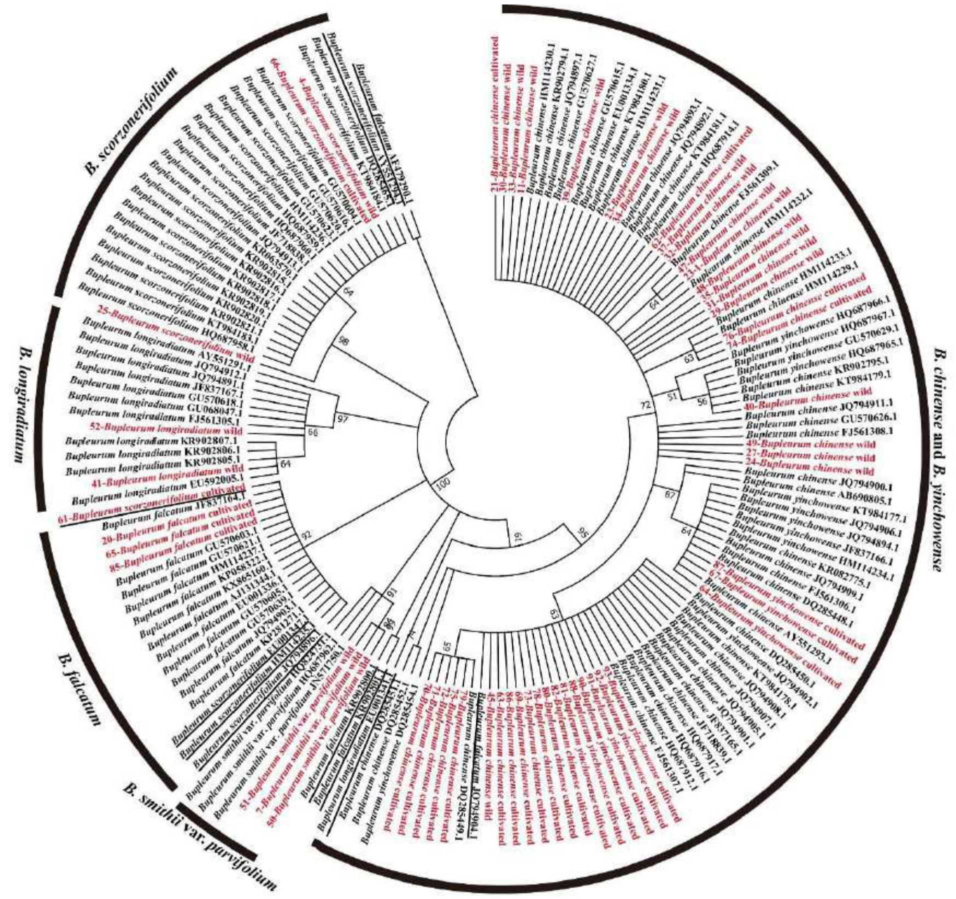

Figure 4: Dendrogram of 56 Bupleurum samples and 110 downloaded sequences in GenBank based on ITS sequences. Red fonts indicated that these sequences were amplified from Bupleurum samples in this study, and the underlined ones were not clustered into their affiliated species

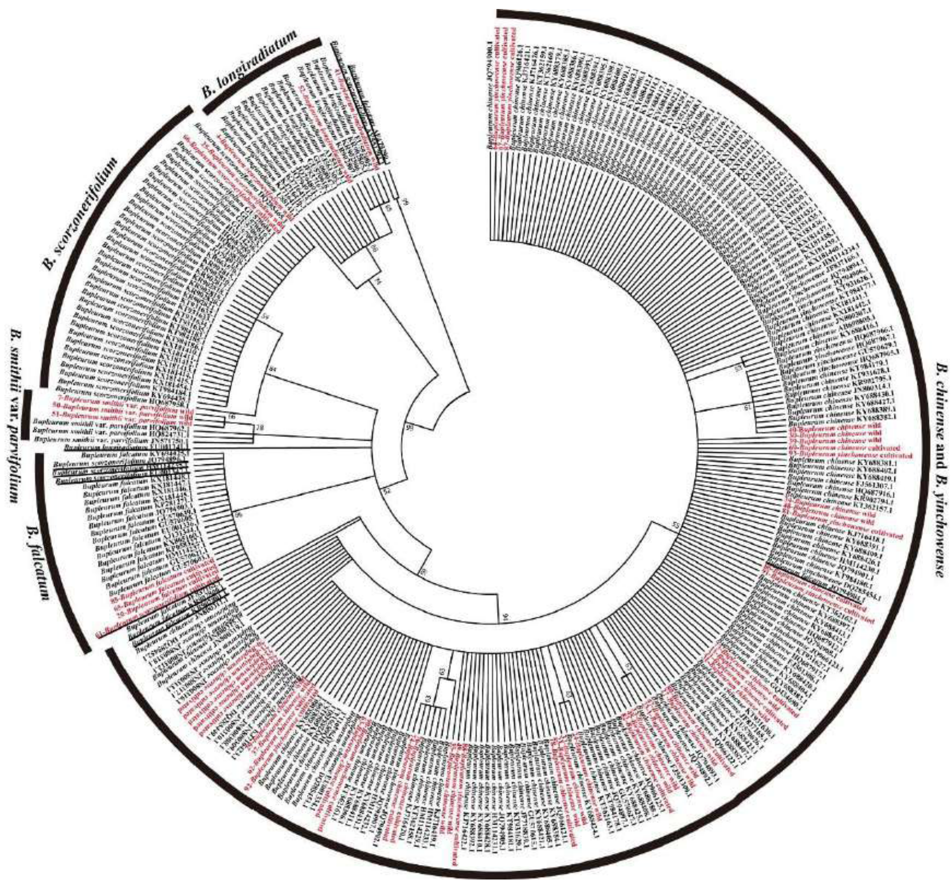

Figure 5: Dendrogram of 56 Bupleurum samples and 250 downloaded sequences in GenBank based on ITS2 sequences. Red fonts indicated that these sequences were amplified from Bupleurum samples in this study, and the underlined ones were not clustered into their affiliated species

Analysis of variable sites and haplotypes of Bupleurum species

Variable sites and haplotypes of *Bupleurum* species were analyzed by MEGA and DnaSP. The results were shown in Table 2 and Table 3S-7S. Specific haplotypes of each species based on five DNA barcodes were analyzed and shown in Table 2. And these specific haplotypes can be used as distinguishable characteristic to six *Bupleurum* species. Additionally, this study obtained some haplotypes which were identified for the first time.

Barcode	Source	Species	No. Variable sites	No. Haplotypes
ITS	Sequenced	*B. chinense*	46	10 (Hap 2, 23, 25, 28, 29, 34, **38**, **41**, **42**, 44)
		B. scorzonerifolium		3 (Hap 9, **37**, **40**)
		B. longiradiatum		2 (Hap 18, 19)
		B. smithii var. *parvifolium*		1 (Hap **39**)
		B. falcatum		1 (Hap 10)
		B. yinchowense		2 (Hap 1, 2)
	Downloaded	*B. chinense*	139	19 (Hap 1, 2, 3, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 43, 44)
		B. scorzonerifolium		6 (Hap 8, 9, 10, 11, 12, 13)
		B. longiradiatum		5 (Hap 18, 19, 20, 21, 22)
		B. smithii var. *parvifolium*		2 (Hap 6, 7)
		B. falcatum		5 (Hap 10, 14, 15, 16, 17)
		B. yinchowense		5 (Hap 1, 2, 3, 4, 5)
	Total	All species	145	44
ITS2	Sequenced	*B. chinense*	26	8 (HT 1, 7, 33, 34, 37, 39, **43**, 45)
		B. scorzonerifolium		2 (HT 2, 5)
		B. longiradiatum		2 (HT 4, 17)
		B. smithii var. *parvifolium*		1 (HT **44**)
		B. falcatum		1 (HT5)
		B. yinchowense		2 (HT 1, 6)
	Downloaded	*B. chinense*	69	29 (HT 1, 6, 7, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 45, 46, 47)
		B. scorzonerifolium		5 (HT 2, 5, 10, 11, 12)
		B. longiradiatum		5 (HT 3, 4, 17, 18, 19)
		B. smithii var. *parvifolium*		1 (HT 9)
		B. falcatum		5 (HT 5, 13, 14, 15, 16)
		B. yinchowense		4 (HT 1, 6, 7, 8)
	Total	All species	70	47

psbA-trnH	Sequenced	*B. chinense*	18	9 (TH 2, 3, 7, 8, **13**, **15**, **17**, **19**, 23)
		B. scorzonerifolium		4 (TH 2, 5, **16**, 23)
		B. longiradiatum		2 (TH **14**, **18**)
		B. smithii var. *parvifolium*		1 (TH 23)
		B. falcatum		1 (TH 2)
		B. yinchowense		8 (TH 1, 2, 3, **12**, **13**, **21**, **22**, 23)
	Downloaded	*B. chinense*	35	9 (TH 1, 2, 3, 7, 8, 9, 10, 11, 23)
		B. scorzonerifolium		2 (TH 2, 5)
		B. longiradiatum		2 (TH 2, 6)
		B. smithii var. *parvifolium*		1 (TH 3)
		B. falcatum		1 (TH 2)
		B. yinchowense		4 (TH 1, 2, 4, 20)
	total	All species	41	23
matK	Sequenced	*B. chinense*	12	6 (HA 1, 2, 3, **7**, **10**, **11**)
		B. scorzonerifolium		4 (HA 1, 2, **8**, **10**)
		B. longiradiatum		1 (HA 2)
		B. smithii var. *parvifolium*		1 (HA 1)
		B. falcatum		2 (HA 2, 4)
		B. yinchowense		5 (HA 1, 2, 5, **6**, **9**)
	Downloaded	*B. chinense*	6	4 (HA 1, 2, 3, 5)
		B. scorzonerifolium		0
		B. longiradiatum		1 (HA 2)
		B. smithii var. *parvifolium*		0
		B. falcatum		2 (HA 2, 4)
		B. yinchowense		3 (HA 1, 2, 3)
	Total	All species	12	11
rbcL	Sequenced	*B. chinense*	3	3 (HP 1, **4**, 5)
		B. scorzonerifolium		3 (HP **1**, 3, 4)
		B. longiradiatum		1 (HP **1**)
		B. smithii var. *parvifolium*		1 (HP 1)
		B. falcatum		1 (HP 1)
		B. yinchowense		2 (HP 1, **2**)
	Downloaded	*B. chinense*	5	7 (HP 1, 2, 3, 5, 6, 7, 9)
		B. scorzonerifolium		4 (HP 3, 4, 5, 7)
		B. longiradiatum		2 (HP 2, 7)
		B. smithii var. *parvifolium*		1 (HP 1)
		B. falcatum		2 (HP 1, 7)
		B. yinchowense		4 (HP 1, 7, 8, 9)
	Total	All species	5	9

Table 2: Haplotypes of five barcodes in *Bupleurum*

Specific variable sites which could be used to discriminate two or more species in *Bupleurum* species based on five barcodes were analyzed and shown in Table 3 and 4. No specific variable sites were found based on *matK*, *psbA-trnH* and *rbcL*. For ITS and ITS2, specific variable sites were classified into four kinds (shown in four colors in Table 3 and 4). The identification efficiency of variable sites gradually decreased with different colors from blue, yellow, green to orange. In ITS sequence, four variable sites including that of 13, 105, 430, 530 bp can be used to distinguish *B. smithii var. parvifolium* from the rest five species completely. 57, 79, 419, 584 bp were specific variable sites for *B. scorzonerifolium*, five variable sites (104, 198, 355, 422, 609 bp) were distinctive characters for *B. longiradiatum*, while *B. yinchowense* had three specific variable sites (393, 429, 574 bp). The site of 495 bp can not be used to distinguish *B. scorzonerifolium* and *B. longiradiatum*, but it can be used to distinguish *B. scorzonerifolium* and *B. falcatum*. While no specific variable sites were appropriate for discriminating *B. chinense* and *B. yinchowense*. For the sake of brevity, the analysis of specific variable sites for ITS2 was not written. Obviously, some haplotypes and specific variable sites could be utilized to discriminate some *Bupleurum* species.

Specific variable sites/bp	Species					
	B. chinense	*B. scorzonerifolium*	*B. longiradiatum*	*B. smithii* var. *parvifolium*	*B. falcatum*	*B. yinchowense*
13	T(78)	T(29)	T(14)	**A(6)**	T(19)	T(20)
105	G	G	G	**T**	G/A(17/2)	G
430	T	T	T	**G**	T	T
530	G	G	G	**T**	G	G
57	G/A(77/1)	**T/G(24/5)**	G/A(13/1)	G	G	G
79	G/A(76/2)	**T/G/A(24/4/1)**	G/A(13/1)	G	G/A(18/1)	G
104	C	C	**T/C(13/1)**	C	C	C
198	-/G	-/G	**A/G(13/1)**	G	-/G	-/G
355	C	C	**T/C(13/1)**	C	C	C
419	A/T(77/1)	**G/A(25/4)**	A	A	A	A
584	A	**T/A(24/5)**	A	A	A	A
393	A	A/T(25/4)	A	A	**T/A(15/4)**	A
422	G	G	**T/G(13/1)**	G	G/T(18/1)	G
429	G/T(75/3)	G/A(25/4)	G	G	**A/G(15/4)**	G
574	T	T/C(24/5)	T	T	**C/T(15/4)**	T
609	T	T/C(28/1)	**C/T(13/1)**	T	T/C(18/1)	T
495	T	**C/T(24/5)**	**C/T(13/1)**	T	T	T

Table 3: Specific variable sites in *Bupleurum* species based on ITS

Specific variable sites/bp	Species					
	B. chinense	*B. scorzoneri folium*	*B. longiradia tum*	*B. smithii* var. *parvifolium*	*B. falcatum*	*B. yincho wense*
38	T	T	T	G	T	T
138	G	G	G	T	G	G
27	A/T(192/1)	G/A(39/4)	A	A	A	A
192	A	T/A(38/5)	A	A	A	A
1	A	A/T(39/4)	A	A	T/A(20/4)	A
30	G	G	T/G(17/1)	G	G/T(23/1)	G
37	G/T(190/3)	G/A(39/4)	G	G	A/G(20/4)	G
182	T	T/C(38/5)	T	T	C/T(20/4)	T
217	T	T/C(42/1)	C/T(17/1)	T	T/C(23/1)	T
103	T	C/T(38/5)	C/T(17/1)	T	T	T

Table 4: Specific variable sites in *Bupleurum* species based on ITS2

Discussion

ITS and ITS2 were widely used in identification of medicinal plants. According to the result of this study, ITS and ITS2 sequences were most likely to be used to identify *Bupleurum* species among all of barcodes analyzed. But in all samples tested in the present study and GenBank downloaded of six *Bupleurum* species, there are still a few that were not clustered into the anticipated clades in the phylogenetic trees of ITS and ITS2. For example, the No. 61 *B. scorzonerifolium*, *B. scorzonerifolium* EU001347.1 (Wang, He, Zhou, Wu, Yu, & Pang, 2008), HM114235.1 (Zheng, 2010) and JQ794896.1 (Chao, Zeng, Liao, Liu, Liang, & Li, 2014) was clustered into the clade of *B. falcatum* instead of *B. scorzonerifolium*. *B. longiradiatum* EU001341.1 (Wang, He, Zhou, Wu, Yu, & Pang, 2008) and *B. falcatum* JQ794904.1 (Chao, Zeng, Liao, Liu, Liang, & Li, 2014) were clustered into the branch of *B. chinense* and *B. yinchowense*, respectively. In addition, *B. falcatum* AF479290.1 (Neves & Watson, 2004) and *B. scorzonerifolium* AY551294.1, *B. falcatum* KR902800.1 and KR902801.1 were assigned to one branch, respectively.

Compared with ITS, ITS2 was shorter and easier to amply, thus, it was most conveniently used for discrimination of traditional Chinese medicine, especially specimen, crude drugs, Chinese patent medicine. However, ITS with longer sequence may supply more information to discriminate *Bupleurum* samples. It was proposed that ITS could be used to discriminate *B. chinense* and *B. yinchowense* precisely, and demonstrated several species specific haplotypes between *B. chinense* and *B. yinchowense* (Yuan et al., 2017). While it was very complicated when much more sequences of samples were considered and compared. KT 984177.1, KT 984178.1 and KT 984179.1 in the report were not specific haplotypes to *B. chinense* and *B. yinchowense* anymore. It is hard to know whether because of obscure morphologic identification or the mixed species samples or something else. But one point is clear that most of samples from both *B. chinense* and *B. yinchowense* could be discriminated using their species specific haplotypes of ITS or ITS2.

Conclusion

Our study suggested that ITS and ITS2 could be used to discriminate *Bupleurum* species. The distinct sites and some new found haplotypes were presented which were valuable for identification of *Bupleurum* species. As to cultivated types, it was hard to discriminate using DNA barcodes like we used to test in our study. New methods are proposed to develop to identify cultivated *Bupleurum* types and their corresponding products within the same species.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We are grateful to Professor Chunsheng Liu (Beijing University of Chinese Medicines) for his assistance on morphological identification of *Bupleurum* species. The present work was supported by the National Key R&D Program of China (Grant No. 2019YFC1710800), CAMS Innovation Fund for Medical Sciences (CIFMS) (2016-I2M-2-003; 2019-I2M-3-001) and the Sichuan Province "13th Five-Year" Crop and Livestock Breeding Key Varieties Breeding Projects (2016NYZ0020).

Appendix

References

1. Chao Z, Zeng WP, Liao J, Liu L, Liang ZB, & Li XL (2014) DNA barcoding Chinese medicinal *Bupleurm*. Phytomed 21: 1767–73.

2. Chen SL, Yao H, Han JP, Liu C, Song JY, Shi LC, et al (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 5: e8613.

3. Chinese Pharmacopoeia Committee (2020) Pharmacopoeia of the People's Republic of China. Vol I. Beijing: Chin Med Sci Press.

4. Choi BR, Kang SW, Park KY, Kim DH (1995) Effects of mulching on emergence and yield of *Bupleurum falcatum* L., RDA. J Agric Sci (Korea Republic), 37: 106–10.

5. Sui C, Han W, Zhu C, Wei J (2020) Recent progress in saikosaponin biosynthesis in *Bupleurum*. Curr Pharm Biotechnol 22: 329-40.

6. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. P Natl Acad Sci USA 102:8369–74.

7. Li DZ, Liu JQ, Chen ZD, Wang H, Ge XJ, et al. (2011) Plant DNA barcoding in China. J Syst Evol *49*: 165–68.

8. Li M, Cao H, But PPH, Shaw PC (2011) Identification of herbal medicinal materials using DNA barcodes. J Syst Evol 49: 271-83.

9. Li DZ, Gao LM, Li HT, Wang H, Ge XJ, et al. (2011) Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. P Natl Acad Sci USA 108: 19641-46.

10. Lin WY, Chen LR, Lin TY (2008) Rapid authentication of *Bupleurum* species using an array of immobilized sequence-specific oligonucleotide probes. Planta Med 74: 464–9.

11. Liu SY, Wang JY, Lee JC, Lee TY (1991) Studies on the adaptability of *Bupleurum falcatum* L. cv. Tainung No. 1 and its officinal value of various plant parts. J Agric Res China 40: 28-36.

12. Liu ZH, Zeng X, Yang D, Chu GY, Yuan ZR, et al. (2012) Applying DNA barcodes for identification of plant species in the family Araliaceae. Gene 499: 76-80.

13. Matsumoto H, Ohta S, Yuan CQ, Zhu YC, Okada M, et al. (2004) Phylogenetic relationships among subgroups in *Bupleurum falcatum* L. sensu lato (Umbelliferae) based on restriction site variation of chloroplast DNA. J Jpn Bot 79: 79-90.

14. Mizukami H, Ohbayashi K, Ohashi H (1993) *Bupleurum falcatum* L. in northern Kyushu and Yamaguchi prefecture are genetically distinguished from other populations, based on DNA fingerprints. Biol Pharm Bull 16: 729–31.

15. Neves SS, Watson MF (2004) Phylogenetic relationship in *Bupleurum* (Apiaceae) based on nuclear ribosomal DNA ITS sequence data. Ann Bot 93: 379–98.

16. Nie H, Deng YJ, Zheng CY (2020) A network pharmacology-based approach to explore the effects of Chaihu Shugan powder on a non-alcoholic fatty liver rat model through nuclear receptors. J Cell Mol Med 24: 5168-84.

17. Ostroumova TA, Kljuykov EV (2015) Fruit structure and microsculpture in the annual species of the genus *Bupleurum*, section *Perfoliata* (Umbelliferae). *Phytologia Balcanica*, 21: 117–27.

18. Pang XH, Luo HM, Sun C (2012) Assessing the potential of candidate DNA barcodes for identifying non-flowering seed plants. Plant Biol 14: 839-44.

19. She ML, Watson MF (2005) Flora of China. Beijing: Science Press.

20. Wang CB, Ma XG, He XJ (2011) A taxonomic re-assessment in the Chinese *Bupleurum* (Apiaceae): Insights from morphology, nuclear ribosomal internal transcribed spacer, and chloroplast (*trnH-psbA*, *matK*) sequences. J Syst Evol 49: 558-89.

21. Wang FH, Lu JM, Wen J, Ebihara A, Li DZ (2016) Applying DNA barcodes to identify closely related species of ferns: A case study of the Chinese *Adiantum* (Pteridaceae). PLoS One 11:e0160611

22. Wang QZ, He XJ, Zhou SD, Wu YK, YuPang YL (2008) Phylogenetic inference of the genus *Bupleurum* (Apiaceae) in Hengduan Mountains based on chromosome counts and nuclear ribosomal DNA ITS sequences. J Syst Evol 46: 142-54.

23. Wang QZ, Zhou SD, Liu TY, Pang YL, Wu YK, He XJ (2008) Phylogeny and classification of Chinese *Bupleurum* based on nuclear ribosomal DNA internal transcribed spacer and *rps16*. Acta Biol Cracov Bot 50: 105–16.

24. Xiao PG (2005) Modern Chinese Materia Medica Beijing: Chemical Industry Press.

25. Yao RY, Chen XF, Zhang BL, Li ZF, Yang XW (2016) Germplasm resource status and breeding prospects of Bupleuri Radix in China. Chin Tradit Herbal Drugs, 44: 1349-53.

26. Yuan BC, Li WD, Ma YS, Zhou S, Zhu LF, et al. (2017) The molecular identification of *Bupleurum* medicinal species and the quality investigation of *Bupleuri Radix*. *Acta Pharmaceutica Sinica* 52: 162–71.

27. Yuan BS, Ma YS, Yang R, Zhou S, Lin RC, Liu Y (2016) Molecular identification of *Bupleurum chinense* DC. and *B. yinchowense* Shan et Y. Li based on ITS sequence. *Letters in Biotechnol* 27: 101–5.

28. Zhang YF (2020) Advances in pharmacological effects of saikosaponins from *Bupleurum chinense*. Clin J Chin Med 12: 120-1.

29. Zhang J, Chen M, Dong XY, Lin RZ, Fan JH, et al. (2015) Evaluation of four commonly used DNA barcoding loci for Chinese medicinal plants of the family Schisandraceae. PLoS One 10: e0125574.

30. Zheng TT, Sui C, Wei JH, Jin Y, Chu QL, et al. (2010) Breeding of new varieties “Zhongchai No. 2” and “Zhongchai No. 3” of *Bupleurum chinense*. Chin J Chin Materia Medica 35: 1931-4.

31. Zheng TT (2010) Evaluation of quality and genetic integrity of *Bupleurum* breeding lines and main cultivated germplasm. Beijing, Peking Union Medical College.