Identification of the original plants of cultivated Bupleuri Radix based on DNA barcoding and chloroplast genome analysis

Gaixia Zhang¹, Hui Wang¹², Linchun Shi¹, Yang Liu¹, Ruyu Yao¹, Chun Sui¹, Chengmin Yang¹, Hongliang Ji¹, Qiuling Wang¹ and Jianhe Wei¹

¹Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
²Institute of Sericulture, Chengde Medical University, Chengde, China

ABSTRACT

Bupleuri Radix is the dry root of certain species of the genus Bupleurum and is commonly used in traditional Chinese medicine. The increasing global demand for Bupleuri Radix cannot be fulfilled with wild populations only. Therefore, cultivated Bupleurum is now the main commercial source of this medicinal product. Different species of Bupleurum show different medicinal properties and clinical effects, making reliable authentication and assignment of correct botanical origin for medicinal species critical. However, accurate identification of the cultivated Bupleurum species is difficult due to dramatic morphological variations resulting from cultivation. In this study, we sampled 56 cultivated Bupleurum populations of six different morphotypes (Types A-F) from the main production areas of China, and 10 wild populations of four species were used as reference materials. Conventional DNA barcoding was conducted to identify cultivated Bupleurum species. Additionally, verification based on complete chloroplast genomes was performed and new chloroplast markers were developed and evaluated. The combination of these methods resulted in the successful identification of all cultivated Bupleurum individuals. Three chloroplast regions are recommended as additional barcodes for the genus: ycf4_cemA, psaJ_rpl33, and ndhE_ndhG. This is a reliable and promising strategy that can be applied to the authentication of natural products and the identification of other medicinal plant species with similar taxonomic problems.

INTRODUCTION

As a plant with high diversity, Bupleuri Radix has been widely used as a herbal drug in Asia over the past 2,000 years and is famous for the treatment of various diseases such as typhoid fever, malaria, epidemic colds, hepatitis, menstrual irregularities, and pain from swollen breasts, uterine prolapse, and prolapse of the rectum (Chinese Pharmacopoeia Commission, 2015; Gorovoy, Ketrus & Grief, 1980; Pan, 2006; Young Hwa et al., 2012; Yuan, Li & Ma, 2012).
It has also been recommended to treat COVID-19 with other traditional Chinese medicine (Fan et al., 2020; Sun et al., 2020; Zhou & Wang, 2020). The original herbs of Bupleuri Radix belong to Bupleurum L., which is a primitive and large genera of the family Apiaceae (Umbelliferae) (Wang et al., 2008). Bupleurum L. is widely distributed throughout the North Temperate Zone and consists of around 180 species, of which 44 species, 17 varieties, and 7 forma have been reported in China (Pan et al., 2002; She & Watson, 2005). Some studies suggested that the Chinese Bupleurum species are divided into two major lineages and should be placed in Neves and Watson’s subgenus Bupleurum (Wang, Ma & He, 2011). Given the conflicting opinions about the taxonomic rank of most Bupleurum taxa and the dramatic morphological variations within species, the taxonomy of the group is complex, and several species remain difficult to identify (Chao et al., 2014; Pan, 2006; She & Watson, 2005; Wang, Ma & He, 2011; Wang et al., 2008). Furthermore, wild populations of Bupleurum are not sufficient to meet the current market demand for Bupleuri Radix. Thus, cultivated species are now the main commercial source of this medicinal herb because they produce a stable supply and considerable yield (Liang, 2012; Liang, Liu & Chao, 2012; Liu et al., 2011; Minami et al., 1997; Qin et al., 2012; Shimokawa & Oiishi, 1980; Shon, Haryanto & Yoshida, 1997; Zhao et al., 2011; Zhu et al., 2017).

In China, Bupleurum species are extensively cultivated in the provinces of Gansu, Shaanxi, Shanxi, Heilongjiang, and Hebei. In our previous study (Zhang et al., 2021), through observation and systematic comparison of living plants and specimens of Bupleurum spp., we found that the cultivated Bupleurum spp. could be preliminarily divided into six phenotypes (Types A-F) and ascribed to four species (B. chinense DC., B. scorzonerifolium Willd., B. falcatum Linneus, and B. marginatum var. stenophyllum (Wolff) Shan et Y. Li) based on 13 distinguishable morphological characteristics, including root color, growth pattern of basal leaf phyllotaxis and rhizome buds, roughness of stem surface, and fruit length. During the previous investigation, we found that morphological variations and the lack of reliable identification methods of cultivated Bupleurum populations have made it increasingly difficult to accurately delimit species. The same Bupleurum plant may be considered as different species due to morphological variations and/or the employment of mislabeled sequences into the identification analysis. For example, the same cultivated germplasm of Bupleuri Radix from Gansu Province was identified as three different species (B. chinense DC., B. yinchowense Shan et Y. Li, or B. marginatum Wall. ex DC.) in various studies (Geng et al., 2010; Guo et al., 2018; Liang, 2012; Qin et al., 2012; Yang et al., 2019; Zhu et al., 2017). Two morphotypes produced in Heilongjiang Province were temporarily treated as the same germplasm and its botanical origin could not be determined (Du et al., 2019). Accurate species identification within Bupleurum is critical because the pharmacological value and clinical properties vary between species and they should not be used interchangeably (Ashour et al., 2009; Huang et al., 2009; Lee et al., 2011; Li et al., 2007; Li et al., 2005; Liu et al., 2011; Qin et al., 2012). Though B. chinense and B. scorzonerifolium Willd. are regarded as standard medical plants of Bupleuri Radix in the Pharmacopeia of the People’s Republic of China (Chinese Pharmacopoeia Commission, 2015), many species recorded in the literature are regional substitutes for Bupleuri Radix or for other medicinal uses, some of them are even with
toxicity. For example, it has been reported that *B. longiradiatum* Turcz. is toxic and cannot be used as a source of Bupleuri Radix (*Ashour & Wink, 2011; Lin, Zhang & Su, 2016*). The safe use of this medicinal plant and its derived products urgently demand the development of specific and accurate methods to effectively determine the species that make Bupleuri Radix.

Here, the DNA barcoding technology was applied to further identify the cultivated germplasm of Bupleuri Radix at the species level. The phylogenetic analysis using chloroplast genomes and molecular markers developed from chloroplast genome sequences were introduced to evaluate and validate the identification result based on DNA barcodes. As an efficient tool for the authentication of medicinal plants and herbal materials, the DNA barcoding technology has been widely applied for species identification (*Chen et al., 2014; Song et al., 2009*). Chloroplast genomes are a useful tool for phylogenetic analyses and comparative studies, and as a source of alternative DNA markers because they are highly conserved with respect to their genome size, structure, and gene content (*Asaf et al., 2017; Chen et al., 2018; Jiang et al., 2017*). Additional DNA markers were harvested from complete chloroplast genomes and evaluated for species that could not be fully identified using DNA barcoding alone.

**MATERIALS & METHODS**

**Materials**

Leaves for all six morphotypes (Types A-F) were freshly collected in triplicates from 56 cultivated populations from the main Bupleuri Radix production areas of China. Four wild species (*B. falcatum*, *B. scorzonerifolium*, *B. marginatum* var. *stenophyllum* (Wolff) Shan et Y.Li, and *B. chinense*) corresponding to cultivated species or with ambiguous phylogenetic classifications to cultivated species were collected and used as the reference material for cultivated species identification (Table 1, Table S1). Herbarium vouchers for both cultivated and wild specimens were deposited in the Institute of Medicinal Plant Development (IMPLAD).

**DNA barcoding analysis**

Four conventional DNA barcodes (internal transcribed spacer - ITS, *psbA-trnH*, *rbcL*, and *matK*) were initially tested for their ability to discriminate cultivated and wild specimens. The primers used for amplification were as previously reported (*Chen, 2012*). A preliminary survey to assess barcode suitability was done with 63 samples from 47 cultivated populations representing six morphotypes. Meanwhile, 19 samples from seven wild populations were selected as reference material for the identification of cultivated species (Tables S1, S2). DNA extraction, PCR amplification, sequencing and sequence alignment were performed according to previously published procedures (*Han et al., 2013; Song et al., 2009*). Analysis of sequence variation among the cultivated species and the reference species was performed using the Molecular Evolutionary Genetics Analysis (MEGA) software (*Kumar et al., 2018*). A phylogenetic tree was constructed using the neighbor-joining algorithm (NJ tree) with 1,000 bootstrap replicates. Once the most suitable barcode (ITS) was selected, we expanded our sample set to further include 36 cultivated individuals (Tables S1, S2) and 11 wild
Table 1  Population information of *Bupleurum* used in this study.

| Serial No. of populations | Morphological type/species | Origin       | Producing area |
|---------------------------|----------------------------|--------------|----------------|
| HLC01-HLC02, HLC05        | A                          | Cultivated   | Heilongjiang   |
| HEC01-HEC03               | A                          | Cultivated   | Hebei          |
| HLC03-HLC04, HLC06        | B                          | Cultivated   | Heilongjiang   |
| GSC01-GSC05               | C                          | Cultivated   | Gansu          |
| GSC06-GSC18               | D                          | Cultivated   | Gansu          |
| HEC04-HEC06               | D                          | Cultivated   | Hebei          |
| SXC08                     | D                          | Cultivated   | Shanxi         |
| SXC01-SXC07, SXC09-SXC12  | E                          | Cultivated   | Shanxi         |
| SNC01-SNC06               | E                          | Cultivated   | Shaanxi        |
| SNC07-SNC14               | F                          | Cultivated   | Shaanxi        |
| GSW01-GSW02               | *B. chinense*              | Wild         | Gansu          |
| HLW01                     | *B. falcatum*              | Wild         | Heilongjiang   |
| HLW02                     | *B. scorzonerifolium*      | Wild         | Heilongjiang   |
| HEW01                     | *B. scorzonerifolium*      | Wild         | Hebei          |
| SXW01                     | *B. chinense*              | Wild         | Shanxi         |
| SNW01-SNW03               | *B. chinense*              | Wild         | Shaanxi        |
| XZW01                     | *B. marginatum var. stenophyllum* | Wild     | Xizang         |

individuals (Tables S1, S2). As a result, the ITS was examined in a total of 99 cultivated samples from 56 populations and 30 wild samples from 10 populations. ITS sequences of *Angelica sinensis* (JN704870) and *Hansenia forbesii* (JQ936553) were obtained from GenBank and used as outgroups for the NJ tree.

**Verification based on complete chloroplast genomes**

DNA extraction, sequencing, and annotation of chloroplast genomes were conducted as per *Chen et al. (2018)* and *Zhou et al. (2018)*. Chloroplast cluster analysis was performed on seven representative cultivated samples (6 cultivation morphotypes represented by HEC02-3, HLC04-3, GSC03-1, GSC06-1, SXC04-1, SNC10-1, and 1 adulterant germplasm of cultivated *B. scorzonerifolium* represented by HLC05-3). The chloroplast genomes of *B. falcatum* (HEC02-3, MT075714; HLC05-3, MT075716), *B. chinense* (GSC06-1, MT075713; SXC04-1, MT075710; SNC10-1, MT075709), and *B. scorzonerifolium* (HLC04-3, MT075715) are newly generated in this research, and the chloroplast genomes of GSC03-1(MT075712) was obtained from my previous research (*Zhang et al., 2021*). Similarities between the chloroplast genomes of the seven samples were calculated as described by *Park et al. (2018)*.

To determine the phylogenetic relationship and genetic distance between morphotypes and each species, chloroplast genome sequences for *B. chinense* (NC_046774, MN893666), *B. scorzonerifolium* (MT239475), *B. falcatum* (NC_027834, MT821947), *B. marginatum* (MN968501), *B. latissimum* (NC_033346, MT821949), *Angelica sinensis* (MH430891), and *Hansenia forbesii* (NC_035054) (outgroup), were obtained from GenBank. Firstly, using MAFFT(v7.309) (*Katoh & Standley, 2013*), complete genome alignments were generated as well as with 74 genes shared by the 17 genome sequences. MrModeltest (v2.4)
Nylander, 2004 was then used to determine the best-fitting model based on the Akaike Information Criterion, and the optimal model, GTR+I+G, was selected for both datasets. Maximum likelihood (ML) analysis was performed using RaxML (v8.2.12) (Stamatakis, 2014) with 1000 bootstrap replicates. Bayesian Inference (BI) analysis was performed using MrBayes (v3.2.7) (Ronquist et al., 2012). Markov Chain Monte Carlo simulations for 2,000,000 generations were independently performed twice, sampling every 100th generation. Convergence was determined by examining the average standard deviation of split frequencies (<0.01). The first 25% of trees were discarded as burn-in, and the remaining trees were used to build a majority-rule consensus tree. Maximum parsimony (MP) analysis was run in Paup (v4.0b10) (Swofford, 2003), using heuristic search and tree bisection-reconnection (TBR) branch swapping with 1000 bootstrap replicates. Neighbor-Joining (NJ) analysis and genetic distance calculation were conducted using MEGA X (Kumar et al., 2018).

Development and validation of additional chloroplast DNA markers

In order to select a short and informative region with enough variation, nucleotide variability (Pi) was calculated for both coding and non-coding regions of the chloroplast using DnaSP version 5.1 (Librado & Rozas, 2009). Highly variable regions with Pi values greater than or equal to 0.015 and with high discriminatory power were screened as potential barcodes through extraction, aligned using MUSCLE, and analysed using the neighbor-joining algorithm (NJ tree) of sequences. Primers were designed using Primer Premier 5.0. PCR amplification was performed in a 25-µl reaction as follows: initial denaturation at 94 °C for 5 min; 40 cycles at 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 45 s; and final extension at 72 °C for 10 min. The PCR products were sequenced on an ABI 3730 sequence analyzer (Applied Biosystems Inc., CA, United States) with the same primers used for PCR amplification. The 21 samples used for marker verification are listed in Table S3. All primers for marker selection are shown in Table S4. The identification efficiency of potential markers was evaluated as described in ‘DNA Barcoding Analysis’.

RESULTS

DNA barcoding identification

Four conventional DNA barcodes (ITS, psbA-trnH, rbcL, and matK) were tested to evaluate their identification efficiency. Since psbA-trnH is a non-coding region, it is rich in long indels and poly (dA) and poly (dT), these sequence features will interfere with sequencing results (Fig. S1, Table S5). Therefore, psbA-trnH is not suitable for species identification. matK and rbcL were not variable enough to discriminate morphotypes or species (Figs. S2–S3, Tables S6–S7). ITS, on the other hand, showed effective discriminatory power and was selected to confirm species identification of the cultivated samples (Fig. 1, Table 2). A total of 129 ITS sequences were obtained: 30 wild specimens of B. chinense, B. scorzonerifolium, B. falcatum, and B. marginatum var. stenophyllum, and 99 cultivated samples of all six morphotypes (24 from Gansu Province, 18 from Heilongjiang Province, 12 from Hebei Province, 15 from Shanxi Province, and 30 from Shaanxi Province). Sequence length before alignment was 603–609 bp. No variability was observed within the phenotypes, with
the exception of Type F, which had one variable site (420 T-C) and was therefore divided into two haplotypes (Table 2). Types A, B, and C were identified as B. falcatum (HLW01), B. scorzonerifolium (HEW01; HLW02), and B. marginatum var. stenophyllum (XZW01), respectively, and Types D, E and F as B. chinense, Type D, Types E and F matched B. chinense (GSW01), B. chinense (SXW01, SNW01-1, SNW01-3; GSW02; SNW03, SNW02), respectively (Fig. 1, Table 2), and the three types were grouped into a single, separate clade (Fig. 1).
Table 2 Variable sites in the haplotypes of different cultivated *Bupleurum* types.

| Haplotypes | Type A                        | Type B                        | Type C                        | Type D                        | Type E                        | Type F1                        | Type F2*                       |
|------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
|            | AAAGTAGAGG                    | TCTAGCTACT                    | GCGAA.TGT.                    | ..T....                        | ..........                       | ..........                       | ..........                       |
|            | TCTAGCTACT                    | TCGGACCCCC                    | CTCT.TATTT                    | ..C....C.                      | ..........                       | ..........                       | ..........                       |
|            | TCGGAGCTAG                    | ACTTAAACTC                    | CCGAGACG.T                    | ..AA......                     | ..........                       | ..........                       | ..........                       |
|            | ATCTTAGATG                    | CCTCCCTACA                    | CTAATGT.TT                    | ..AA......                     | G...A.TG..                     | A...T.A...                     | ..........                       |
|            | GATG                          | GAT...T...                    | A.CTA.CG..                    | A...T.A...                     | G...A.TG..                     | A...T.A...                     | T...                           |
|            |                               |                               | ..C...A...                    | T...                           |                               |                               |                               |
|            |                               |                               |                               |                               |                               |                               |                               |

Notes.
* Type F has two haplotypes: SNC08-1, SNC08-3, SNC09-2, SNC09-3, and SNC11-1 belong to haplotype F2, and the remaining sequences are consistent with the dominant F1 haplotype.

Verification based on chloroplast genomes

A total of 79 protein-coding genes were annotated in the chloroplast genome of the studied *Bupleurum* species. Complete chloroplast genomes and all the 74 genes shared among 17 members of the genus *Bupleurum* and two other species within the family Umbelliferae (*A. sinensis* and *H. forbesii*) were analyzed. Bayesian inference (BI), maximum parsimony (MP), Neighbor-Joining (NJ), and maximum likelihood (ML) generated identical tree topologies for the main clades (Fig. 2 and Figs. S4–S11).

Among the species having been identified by morphological characteristics and DNA barcodes, Type A samples from Heilongjiang (HLC05-3, MT075716) and Hebei Provinces (HEC02-3, MT075714) corresponded to the reference chloroplast genome of *B. falcatum* (NC_027834, MT821947). Type B samples (MT075715) corresponded to the reference Zhang et al. (2022), PeerJ, DOI 10.7717/peerj.13208
chloroplast genome of *B. scorzonerifolium* (MT239475). *B. marginatum* var. *stenophyllum* (MT075712) was closely clustered with *B. marginatum* (MN968501) and possessed a basal position sister to all the other *Bupleurum* species. Type D was clustered in the same clade with *B. chinense* (Type E, MT075710; Type F, MT075709) and corresponded to the reference chloroplast genome of *B. chinense* (NC_046774; MN893666). Support values of the species clades were high (100) and intraspecific support values varied between the phylogenetic tree constructed using the complete chloroplast genomes and that constructed using the shared genes (Figs. S4–S11). Furthermore, the maximum intraspecific genetic distance within each species was lower than the corresponding minimum interspecific genetic distance (Table S9), which confirmed the reliability of the identification results obtained from morphological characteristics and DNA barcodes.

**Highly Variable Chloroplast Regions for the Development of New DNA Markers**

Sequence divergence was further analyzed by extracting coding and non-coding regions from the chloroplast genomes sequences to calculate nucleotide variability (Pi) (Table S10–S11). Pi values ranged from 0 to 0.0433. Non-coding regions were more variable compared with the coding regions. Nineteen regions with nucleotide diversity >0.015 were selected and assessed through sequence variation analysis and phylogenetic analysis. The ideal DNA marker should be short enough for easy PCR amplification and sequencing, have sufficient interspecific variation but low intraspecific variation, and have conservative flanking sequences for easy primer design. Based on these considerations, three DNA markers (*ycf4_cemA*, *psaJ_rpl33*, and *ndhE_ndhG*), which were verified by conventional DNA barcoding methods to successfully discriminate cultivated *Bupleurum*, were selected and recommended as complementary barcodes for *Bupleurum* identification. Detailed results of the sequence variation and phylogenetic analysis are shown in Fig. 3.

**DISCUSSION**

**Accurate identification of cultivated *Bupleurum* species in China**

The conventional DNA barcoding technology was conducted to identify cultivated *Bupleurum* species, the complete chloroplast genomes were used to verify the identification results obtained with DNA barcodes, and DNA markers developed from chloroplast genome sequences were introduced to further evaluate and validate the results of previous identifications. The combination of these three methods successfully determined the species identity of cultivated *Bupleurum* in China, including *B. chinense*, produced mainly in Gansu Province, and *B. falcatum*, produced in Heilongjiang Province (Ding et al., 2016; Du et al., 2019; Geng et al., 2010; Guo et al., 2018; Qin et al., 2012; Yang et al., 2019; Yuan et al., 2017b; Zhu et al., 2017). Former studies were unable to determine the species identity for germplasm from Gansu Province. Three possible identifications have been proposed: *B. chinense*, *B. yinchowense*, and *B. marginatum* (Chao et al., 2014; Ding et al., 2016; Wang, Ma & He, 2011; Wang et al., 2008; Xie et al., 2009; Yang et al., 2007; Yuan et al., 2017b). Firstly, our previous morphological analysis supports the attribution of *B. chinense* based on morphological character descriptions published in Flora of China and the distinguishable
morphological characteristics from our analysis and summary. The verification results based on ITS sequences and chloroplast genome analysis and the newly developed markers in the present study all supported its attribution to _B. chinense_.

Previous studies have treated cultivated _B. falcatum_ as an adulterant of _B. scorzonerifolium_ produced in Heilongjiang Province, which was temporarily treated as a morphotype of _B. scorzonerifolium_ (Du et al., 2019). Our results indicate that _B. scorzonerifolium_ adulterants from the Heilongjiang and Hebei Provinces are the same species, and were identified as _B. falcatum_. Equal chromosome number and closer genome size are congruent with this conclusion (Du et al., 2019). _B. falcatum_ from China has been considered as the same species in Japan and Korea (Gorovoy, Ketrits & Grief, 1980; Jiang, Xu & Li, 2000; Jiang et al., 1994; Li et al., 1994; Matsumoto et al., 2004; Pan et al., 1995; Wang, 2011; Wang, Ma & He, 2011; Wang, Ma & He, 2013; Wang et al., 2016a), but it has not been included in the Flora of China. Considering its wide distribution and abundance (Jiang, Xu & Li, 2000), we suggest that _B. falcatum_ should be included in the Flora of China, which would facilitate and encourage its medicinal use.
Identification methods for cultivated *Bupleurum* and potential applications

Natural foods and medicines have become increasingly popular in recent years due to growing public awareness about nutrition and health issues (Phan, David & Sabaratnam, 2017; Xin et al., 2013; Yao et al., 2018). To ensure their appropriate, safe, and effective use, a precise and clear species identification of these products is paramount. Many plant species have similar taxonomic classification problems that result from domestication. For example, yams (*Dioscorea* spp.) are an important food crop with significant medicinal effects for spleen deficiency, reduced food intake, chronic diarrhea, etc. However, the taxonomy of the group is complex and remains unresolved because of the great variation resulting from domestication and artificial breeding (Gao et al., 2008; Wu, 2012). Similar issues are encountered in other medicinal crops such as mulberry (*Morus* spp.) and Goji (fruits of *Lycium barbarum* L. and *L. chinense* Mill.) (Gao et al., 2015; Xin et al., 2013; Yin, 2013; Zeng et al., 2015). Cultivated *Bupleurum* individuals were identified at the species level using DNA barcodes and further verified by phylogenetic analyses of complete chloroplast genomes and newly developed markers. The methods applied in this study provide a possible solution for these challenges and may serve as a powerful tool to solve taxonomic problems and ensure quality control of medicinal plants.

Our results confirmed that the relatively less sequence variations in conventional chloroplast barcodes (i.e., *rbcL*, *matK*, and *trnH-psbA*) among *Bupleurum* species might lead to incorrect identification result at the inter-generic level (Tables S5–S7, Figs. S1–S3). However, complete chloroplast genome analyses did provide enough discriminatory power to identify all species and morphotypes. Since the use of chloroplast genomes is not applicable to all sample types (e.g., degraded and processed samples with low DNA concentration and quality) and available to all research groups, we selected the three most variable chloroplast regions and recommend their use for species identification in *Bupleurum* to complement ITS: *ycf4_cemA*, *psaJ_rpl33*, and *ndhE_ndhG*. These markers can be used to streamline the identification of degraded and processed samples, and to facilitate and expedite the identification of *Bupleurum* species at a reduced cost: *ycf4_cemA*, *psaJ_rpl33*, and *ndhE_ndhG*. In future studies, we will include more species or samples to further exert the identification effectiveness of complete chloroplast genomes and expand the application of the developed markers on crude drugs of *Bupleurum* species as well as their products.

**CONCLUSIONS**

We presented an identification pipeline to accurately and specifically identify cultivated species of *Bupleurum* in China. This approach combines DNA barcoding, chloroplast genomes, and genus specific markers (*ycf4_cemA*, *psaJ_rpl33*, and *ndhE_ndhG*), and provides multiple and independent evidence to verify species identity. It also improves the efficiency and accuracy for the identification of cultivated *Bupleurum* species, which is critical for the development of resources that can be used in natural products, and for the safe and effective use of Bupleuri Radix. The combination of these methods could be equally successful to address similar taxonomic problems in other plant groups.
ACKNOWLEDGEMENTS

The authors wish to express their sincere thanks to Profs. Yaodong Qi and Baolin Guo (Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College), Prof. Yuan Chen (Agronomy College, Gansu Agricultural University), Prof. Tianlin Zhou (College of Life Science and Technology, Longdong University), Bingdong Nan, Xin Li (Qingyang Academy of Agricultural Sciences), Jinwen Zhang (Longxi Agricultural Technology Promotion Center), Shengjian Dong (Lintao Agricultural School of Dingxi City in Gansu Province), Zhongsen Ma (Longxi Baohetang Pharmaceutical Co. Ltd.), Shengli Hui (Shaanxi huayuan pharmaceutical group shenbao Co. Ltd.), Xunhui Teng (Shanxi Medicinal Animals and Plants Cultivation Research Base), Xianlin He (Agriculture and Animal Husbandry Bureau of Shexian County), Genlai Kou (Anguo Science and Technology Bureau) for their encouragement and the help with sample collection, and Prof. Yulin Lin (Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College), Prof. Chunsheng Liu (School of Chinese Materia Medica, Beijing University of Chinese Medicine) for the help with morphological identification.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding
This work was supported by CAMS Innovation Fund for Medical Sciences (2016-I2M-2-003), National Ten Thousand Talents Program (99950534), National TCM production technology service platform project, Ministry of Industry and Information Technology of the People’s Republic of China (0714-EMTC-02-00195), China Agriculture Research System (CARS-21), and National Key R&D Program of China (2019YFC1710800). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures
The following grant information was disclosed by the authors:
CAMS Innovation Fund for Medical Sciences: 2016-I2M-2-003.
National Ten Thousand Talents Program: 99950534.
National TCM production technology service platform project.
Ministry of Industry and Information Technology of the People’s Republic of China: 0714-EMTC-02-00195.
China Agriculture Research System: CARS-21.
National Key R&D Program of China: 2019YFC1710800.

Competing Interests
The authors declare there are no competing interests.
Author Contributions

- Gaixia Zhang conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Hui Wang and Hongliang Ji performed the experiments, prepared figures and/or tables, and approved the final draft.
- Linchun Shi and Ruyu Yao analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Yang Liu performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Chun Sui, Qiuling Wang and Jianhe Wei conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Chengmin Yang conceived and designed the experiments, prepared figures and/or tables, and approved the final draft.

Field Study Permissions

The following information was supplied relating to field study approvals (i.e., approving body and any reference numbers):

The field materials were belong to individual farmers, the materials used in this study were obtained with verbal permission from local farmers of every collection point. For farmers who are not willing to disclose their personal information, we have written the informations of the guides. The relevant informations have been provided as a confidential supplemental file named “Field Permit”. Thank you.

Data Availability

The following information was supplied regarding data availability:

The raw data are available in the Supplemental Files and at Genbank: ITS (MT074485–MT074610, MT074614–MT074616); psbA-trnH (MT085836–MT085848, MT085850–MT085866, MT085869–MT085880, MT085883–MT085886, MT085888–MT085902, MT085905–MT0859217, MT085920–MT085925); matK (MT085926–MT085974, MT085976–MT085989, MT085991–MT085996, MT085998–MT086011); rbcL (MT086018–MT086021, MT086023, MT086025–MT086056, MT086058–MT086063, MT086065–MT086071, MT086074–MT086077, MT086079–MT086085, MT086089–MT086096, MT086098–MT086110); Chloroplast genome (MT075709, MT075710, MT075712–MT075716).

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.13208#supplemental-information.

REFERENCES

Asaf S, Waqas M, Khan AL, Khan MA, Kang SM, Imran QM, Shahzad R, Bilal S, Yun BW, Lee IJ. 2017. The complete chloroplast genome of wild rice (Oryza
minuta) and its comparison to related species. *Frontiers in Plant Science* 8:304 DOI 10.3389/fpls.2017.00304.

Ashour ML, El-Readi M, Youns M, Mulyaningsih S, Sporer F, Efferth T, Wink M. 2009. Chemical composition and biological activity of the essential oil obtained from *Bupleurum marginatum* (Apiaceae). *Journal of Pharmacy and Pharmacology* 61(8):1079–1087 DOI 10.1211/jpp/61.08.00xx.

Ashour ML, Wink M. 2011. Genus *Bupleurum*: a review of its phytochemistry, pharmacology and modes of action. *Journal of Pharmacy and Pharmacology* 63(3):305–321 DOI 10.1111/j.2042-7158.2010.01170.x.

Chao Z, Zeng WP, Liao J, Liu L, Liang ZB, Li XL. 2014. DNA barcoding Chinese medicinal *Bupleurum*. *Phytomedicine* 21:1767–1773 DOI 10.1016/j.phymed.2014.09.001.

Chen S. 2012. *Traditional Chinese medicine molecular identification by DNA barcode*. Beijing: People's Medical Publishing House.

Chen SL, Pang XH, Song JY, Shi LC, Yao H, Han JP, Leon C. 2014. A renaissance in herbal medicine identification: from morphology to DNA. *Biotechnol Adv* 32:1237–1244 DOI 10.1016/j.biotechadv.2014.07.004.

Chen XL, Zhou JG, Cui YX, Wang Y, Duan BZ, Yao H. 2018. Identification of *Ligularia* herbs using the complete chloroplast genome as a super-barcode. *Frontiers in Pharmacology* 9:695 DOI 10.3389/fphar.2018.00695.

Chinese Pharmacopoeia Commission. 2015. *Pharmacopoeia of the People's Republic of China*. Beijing: China Medical Science Press.

Ding C, Xu Y, Ma XX, Wang ZP, Mao WW, Zhang LL. 2016. Identification on five kinds of easily confused *Bupleurum* medicinal materials. *Journal of Chinese Medicinal Materials* 39:1975–1981 DOI 10.13863/j.issn1001-4454.2016.09.012.

Du M, Xu J, Zhu C, Wang L, Wei J. 2019. Chromosome number determination and genome size estimation of three *Bupleurum* species. *Jiangsu Agricultural Sciences* 47:191–193 DOI 10.15889/j.issn1002-1302.2019.11.043.

Fan YP, Wang YP, Zhang HM, Wang YY. 2020. Analysis on the treatment of new coronavirus pneumonia (COVID-19) from the cold epidemic treatment. *Journal of Traditional Chinese Medicine* 61:369–374 DOI 10.13288/j.jl.11-2166/.2020.05.001.

Gao X, Zhu YP, Wu BC, Zhao YM, Chen JQ, Hang YY. 2008. Phylogeny of *Dioscorea* sect. Stenophora based on chloroplast *matK, rbcL* and *trnL-F* sequences. *Journal of Systematics and Evolution* 46:315–321 DOI 10.3724/SP.J.1002.2008.08007.

Gao Y, Yu XZ, Xu LR, Wang N, Zhang R. 2015. A novel approach to discriminate *Lycium barbarum* from the Zhongning area using FT-IR spectroscopy and chemometrics. *Analytical Methods* 7:9108–9113 DOI 10.1039/c5ay02346g.

Geng ML, Song PP, Xu ZL, Wang NH. 2010. Comparison on HPLC analysis of aerial part of Bupleurums from different germplasm resources. *China Journal of Chinese Materia Medica* 35:574–578+582 DOI 10.4268/cjcm20100507.

Gorovoy PG, Ketrits LM, Grief VG. 1980. A study of East Asian *Bupleurum falcatum*. *Feddes Repertorium* 91:57–62 DOI 10.1002/fedr.19800910112.
Guo JQ, Yang YJ, Fang W, Guo BL. 2018. Saikosaponins content in *Bupleurum marginatum* var. *stenophyllum* and cultivated *B. chinese* and *B. marginatum*. *Modern Chinese Medicine* 20:34–38 DOI 10.13313/j.issn.1673-4890.20170524003.

Han JP, Zhu YJ, Chen XL, Liao BS, Yao H, Song JY, Chen SL, Meng FY. 2013. The short ITS2 sequence serves as an efficient taxonomic sequence tag in comparison with the full-length ITS. *BioMed Research International* 2013:1–7 DOI 10.1155/2013/741476.

Huang HQ, Zhang X, Xu ZX, Su J, Yan SK, Zhang WD. 2009. Fast determination of saikosaponins in *Bupleurum* by rapid resolution liquid chromatography with evaporative light scattering detection. *Journal of Pharmaceutical and Biomedical Analysis* 49:1048–1055 DOI 10.1016/j.jpba.2009.01.011.

Young Hwa K, Kim ES, Ko BS, Uddin MR, Park JY, Lee MY. 2012. Internal transcribed spacer-based identification of *Bupleurum* species used as sources of medicinal herbs. *Journal of Medicinal Plants Research* 6:841–848 DOI 10.5897/jmpr11.1484.

Jiang CM, Xu N, Li B. 2000. A new discovery in *Bupleurum scorzonerifolium* investigation. *Bulletin of Botanical Research* 20:162–167 DOI 10.3969/j.issn.1673-5102.2000.02.007.

Jiang CM, Xu N, Wang HY, Li RJ, Liu MY. 1994. Cytotaxonomical studies on *Bupleurum* L. in northeast China I. karyology analysis of six taxa and their significance of taxonomy. *Bulletin of Botanical Research* 14:267–272 +333-334.

Jiang D, Zhao ZY, Zhang T, Zhong WH, Liu CS, Yuan QJ, Huang LQ. 2017. The chloroplast genome sequence of *Scutellaria baicalensis* provides insight into intraspecific and interspecific chloroplast genome diversity in *Scutellaria*. *Genes* 8:227–240 DOI 10.3390/genes8090227.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30:772–780 DOI 10.1093/molbev/mst010.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35:1547–1549 DOI 10.1093/molbev/msy096.

Lee J, Yang DH, Suh JH, Kim U, Eom HY, Kim J, Lee MY, Kim J, Han SB. 2011. Species discrimination of Radix Bupleuri through the simultaneous determination of ten saikosaponins by high performance liquid chromatography with evaporative light scattering detection and electrospray ionization mass spectrometry. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences* 879:3887–3895 DOI 10.1016/j.jchromb.2011.10.040.

Li RJ, Jiang CM, Wang HY, Liu MY. 1994. Chromosomes of so-called *Bupleurum falcatum* in Northeast China. *Cytologia* 59:365–368 DOI 10.1508/cytologia.59.365.

Li XQ, He ZG, Bi KS, Song ZH, Xu L. 2007. Essential oil analyses of the root oils of 10 *Bupleurum* species from China. *Journal of Essential Oil Research* 19:234–238 DOI 10.1080/10412905.2007.9699268.

Li XQ, Jia Y, Song AH, Chen XH, Bi KS. 2005. Analysis of the essential oil from Radix Bupleuri using capillary gas chromatography. *Yakugaku Zasshi* 125:815–819 DOI 10.1248/yakushi.125.815.
Liang ZB. 2012. Resource investigation and metabolomics-based studies on quality evaluation of Bupleurum medicinal plants. Master thesis, Southern medical university.

Liang ZB, Liu L, Chao Z. 2012. Investigation of medicinal Bupleurum resources and current situation of Chaihu production. Lishizhen Medicine and Materia Medica Research 23:2011–2013 DOI 10.3969/j.issn.1008-0805.2012.08.075.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452 DOI 10.1093/bioinformatics/btp187.

Lin MY, Zhang WD, Su J. 2016. Toxic polyacetylenes in the genus Bupleurum (Apiaceae)- Distribution, toxicity, molecular mechanism and analysis. Journal of Ethnopharmacology 193:566–573 DOI 10.1016/j.jep.2016.09.052.

Liu XJ, Hu J, Li ZY, Qin XM, Zhang LZ, Guo XQ. 2011. Species classification and quality assessment of Chaihu (Radix Bupleuri) based on high-performance liquid chromatographic fingerprint and combined chemometrics methods. Archives of Pharmacal Research 34:961–969 DOI 10.1007/s12272-011-0613-2.

Matsumoto H, Ohta S, Yuan CQ, Zhu YC, Okada M, Miyamoto KI. 2004. Phylogenetic relationships among subgroups in Bupleurum falcatum L. sensu lato (Umbelliferae) based on restriction site variation of chloroplast DNA. Journal of Japanese Botany 79:79–90.

Minami M, Sugino M, Hasegawa C, Ohe C, Ashida K, Ogaki K. 1997. Seasonal variation on growth and saikosaponins content of Bupleurum falcatum (II) with special reference to non-bolting plant (Rosette plant) in first year. Natural Medicines 51:30–36.

Nylander JAA. 2004. MrModeltest Version 2. Evolutionary biology centre. Uppsala: Uppsala University.

Pan SL. 2006. Bupleurum species: scientific evaluation and clinical applications. Boca Raton: CRC Press.

Pan SL, Shun QS, Bo QM, Bao XS. 2002. The coloured atlas of the medicinal plants from genus Bupleurum in China. Shanghai: Shanghai Science and Technology Literature Publishing House.

Pan ZH, Zhuang TD, Zhou XL, Yao XM. 1995. Karyotype analysis of four taxa of Bupleurum used in Chinese drugs. Journal of Plant Resources and Environment 4:41–45.

Park I, Yang S, Kim WJ, Noh P, Lee HO, Moon BC. 2018. The complete chloroplast genomes of six Ipomoea species and indel marker development for the discrimination of authentic pharbitidis semen (seeds of I. nil or I. purpurea). Frontiers in Plant Science 9:965 DOI 10.3389/fpls.2018.00965.

Phan CW, David P, Sabaratnam V. 2017. Edible and medicinal mushrooms: emerging brain food for the mitigation of neurodegenerative diseases. Journal of Medicinal Food 20:1–10 DOI 10.1089/jmf.2016.3740.

Qin XM, Dai YT, Liu NQ, Li ZY, Liu XJ, Hu J, Young HC, Verpoorte R. 2012. Metabolic fingerprinting by 1HNMR for discrimination of the two species used as Radix Bupleuri. Planta Medica 78:926–933 DOI 10.1055/s-0031-1298496.
Ronquist F, Teslenko M, Van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**:539–542 DOI 10.1093/sysbio/sys029.

She ML, Watson MF. 2005. Bupleurum Linnaeus. In: *Flora of China, Apiaceae through Ericaceae*. Beijing: Science Press.

Shimokawa Y, Oiiashi H. 1980. Cultivation and breeding of *Bupleurum falcatum* relation among cultivation years, root growth and saikosaponin content. *Shoyakugaku Zasshi* 34:235–238.

Shon TK, Haryanto TAD, Yoshida T. 1997. Variation and distribution and saikosaponin in *Bupleurum falcatum* L. *Journal of the Faculty of Agriculture, Kyushu University* 42:17–22 DOI 10.5109/24185.

Song JY, Yao H, Li Y, Li XW, Lin YL, Liu C, Han JP, Xie CX, Chen SL. 2009. Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique. *Journal of Ethnopharmacology* **124**:434–439 DOI 10.1016/j.jep.2009.05.042.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**:1312–1313 DOI 10.1093/bioinformatics/btu033.

Sun KB, Zhang XY, Liu J, Sun R. 2020. Network pharmacological analysis and mechanism prediction of Xiaochaihu Decoction in treatment of COVID-19 with syndrome of pathogenic heat lingering in lung and obstructive cardinalate. *Chinese Traditional and Herbal Drugs* **51**:1750–1760 DOI 10.7501/j.issn.0253-2670.2020.07.009.

Swofford DL. 2003. PAUP. Phylogenetic analysis using parsimony (and other methods). Version 4. Sunderland: Sinauer Associates Press.

Wang CB. 2011. Systematic and taxonomical studies on Chinese *Bupleurum* L. (Apiaceae). PhD thesis, Sichuan University.

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. *Journal of Systematics and Evolution* **49**:558–589 DOI 10.1111/j.1759-6831.2011.00157.x.

Wang CB, Ma XG, He XJ. 2013. Rerecognize *Bupleurum falcatum* L. sensu lato (Apiaceae) in East Asia and evaluate some questionable ITS sequences in GenBank. *Journal of Jiangsu Normal University(Natural Science Edition)* **31**:64–73.

Wang CB, Wang XF, Zhang HB, Yang HS, Zhang HJ, Cheng HT. 2016a. Studies on biological and ecological features of medical plant *Bupleurum falcatum* complex. *Mod Chin Med* **18**:1153–1158 +1184 DOI 10.13313/j.issn.1673-4890.2016.9.015.

Wang QZ, Zhou SD, Liu TY, Pang YL, Wu Y, He XJ. 2008. Phylogeny and classification of chinese *Bupleurum* based on nuclear ribosomal DNA internal transcribed spacer and *rps16*. *Journal of Systematics and Evolution* **50**:105–116 DOI 10.1186/1471-2229-8-1.

Wu ZG. 2012. Study on genetic relationship and crude drugs quality of yam germplasm resources. PhD thesis, Guangzhou University of Chinese Medicine.
Xie H, Huo KK, Chao Z, Pan SL. 2009. Identification of crude drugs from Chinese medicinal plants of the genus Bupleurum using ribosomal DNA ITS sequences. *Planta Medica* 75:89–93 DOI 10.1055/s-0028-1088334.

Xin TY, Yao H, Gao HH, Zhou XZ, Ma XC, Xu CQ, Chen J, Han JP, Pang XH, Xu R, Song JY, Chen SL. 2013. Super food *Lycium barbarum* (Solanaceae) traceability via an internal transcribed spacer 2 barcode. *Food Research International* 54:1699–1704 DOI 10.1016/j.foodres.2013.10.007.

Yang YJ, Zheng W, Guo JQ, Fang W, Ma BP, Guo BL, Peng YX, Zhang DJ. 2019. Comparative analysis on chemical constituents in *Bupleurum chinense*, *B. marginatum*, *B. marginatum* var. *stenophyllum* and *B. smithii* var. *parvifolium*. *China Journal of Chinese Materia Medica* 44:332–337.

Yang ZY, Chao Z, Huo KK, Xie H, Tian ZP, Pan SL. 2007. ITS sequence analysis used for molecular identification of the *Bupleurum* species from northwestern China. *Phytomedicine* 14:416–422 DOI 10.1016/j.phymed.2007.04.009.

Yuan BC, Li WD, Ma YS. 2017a. The molecular identification of *Bupleurum* medicinal species and the quality investigation of Bupleuri Radix. *Acta Pharmaceutica Sinica* 52:162–171 DOI 10.1016/j.aps.2016.05.001.

Yuan BC, Li WD, Ma YS, Zhou S, Zhu LF, Lin RC, Liu Y. 2017b. The molecular identification of *Bupleurum* medicinal species and the quality investigation of Bupleuri Radix. *Acta Pharmaceutica Sinica* 52:162–171 DOI 10.1016/j.aps.2016.05.001.

Zeng QW, Chen HY, Zhang C, Han MJ, Li T, Qi XW, Xiang ZH, He NJ. 2015. Definition of eight mulberry species in the genus *Morus* by internal transcribed spacer-based phylogeny. *PLOS ONE* 10(8):e0135411 DOI 10.1371/journal.pone.0135411.

Zhang GX, Wang H, Jiang JM, Wang QL, Li BL, Shi LC, Wei JH. 2021. The complete chloroplast genome of *Bupleurum marginatum* var. *stenophyllum* (H. Wolff) Shan & Yin Li (Apiaceae), a new substitution for Chinese medicinal material, Bupleuri Radix (Chai hu). *Mitochondrial DNA Part B* 6(2):441–443 DOI 10.1080/23802359.2020.1870899.

Zhang GX, Wang H, Liu Y, Yao RY, Jiang JM, Wang QL, Wei JH. 2021. Survey and analysis of cultivated *Bupleurum* spp. germplasm resources in China. *Modern Chinese Medicine* 23(5):772–799 DOI 10.13313/j.issn.1673-4890.20200905002.

Zhao LC, Li W, He YF, Li XH, Wang Z, Liu WC, Zheng YN, Liang J. 2011. Fast and sensitive LC-DAD-ESI/MS method for analysis of saikosaponins c, a, and d from the roots of *Bupleurum falcatum* (Sandaochaihu). *Molecules* 16:1533–1543 DOI 10.3390/molecules16021533.

Zhou JG, Cui YX, Chen XL, Li Y, Xu ZC, Duan BZ, Li YH, Song JY, Yao H. 2018. Complete chloroplast genomes of *Papaver rhoeas* and *Papaver orientale*: molecular
structures, comparative analysis, and phylogenetic analysis. *Molecules* **23:**437 DOI 10.3390/molecules23020437.

**Zhou MX, Wang M. 2020.** TCM prevention and treatment strategies and prescription screening of COVID-19. *Acta Chinese Medicine* **35:**458–463 DOI 10.16368/j.issn.1674-8999.2020.03.102.

**Zhu L, Liang ZT, Yi T, Ma Y, Zhao ZZ, Guo BL, Zhang JY, Chen HB. 2017.** Comparison of chemical profiles between the root and aerial parts from three *Bupleurum* species based on a UHPLC-QTOF-MS metabolomics approach. *BMC Complementary and Alternative Medicine* **17:**305 DOI 10.1186/s12906-017-1816-y.