

Article

# A Systematic Study on DNA Barcoding of Medicinally Important Genus *Epimedium* L. (Berberidaceae)

Mengyue Guo <sup>1</sup> , Yanqin Xu <sup>2</sup>, Li Ren <sup>1</sup>, Shunzhi He <sup>3</sup> and Xiaohui Pang <sup>1,\*</sup>

<sup>1</sup> Key Lab of Chinese Medicine Resources Conservation, State Administration of Traditional Chinese Medicine of the People's Republic of China, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China; guomy0908@hotmail.com (M.G.); renliwangyi@hotmail.com (L.R.)

<sup>2</sup> College of Pharmacy, Jiangxi University of Traditional Chinese Medicine, Nanchang 330004, China; 20081016@jxutcm.edu.cn

<sup>3</sup> Department of Pharmacy, Guiyang College of Traditional Chinese Medicine, Guiyang 550002, China; shunzhihe2018@hotmail.com

\* Correspondence: xhpang@implad.ac.cn; Tel.: +86-10-57833051

Received: 27 October 2018; Accepted: 10 December 2018; Published: 17 December 2018



**Abstract:** Genus *Epimedium* consists of approximately 50 species in China, and more than half of them possess medicinal properties. The high similarity of species' morphological characteristics complicates the identification accuracy, leading to potential risks in herbal efficacy and medical safety. In this study, we tested the applicability of four single loci, namely, *rbcL*, *psbA-trnH*, internal transcribed spacer (ITS), and ITS2, and their combinations as DNA barcodes to identify 37 *Epimedium* species on the basis of the analyses, including the success rates of PCR amplifications and sequencing, specific genetic divergence, distance-based method, and character-based method. Among them, character-based method showed the best applicability for identifying *Epimedium* species. As for the DNA barcodes, *psbA-trnH* showed the best performance among the four single loci with nine species being correctly differentiated. Moreover, *psbA-trnH* + ITS and *psbA-trnH* + ITS + *rbcL* exhibited the highest identification ability among all the multilocus combinations, and 17 species, of which 12 are medicinally used, could be efficiently discriminated. The DNA barcode data set developed in our study contributes valuable information to Chinese resources of *Epimedium*. It provides a new means for discrimination of the species within this medicinally important genus, thus guaranteeing correct and safe usage of *Herba Epimedii*.

**Keywords:** *Epimedium*; DNA barcode; identification; super barcode

## 1. Introduction

Over 200 million people suffer from osteoporosis around the world, and the prevalence of osteoporosis keeps rising with the increasing elderly population [1]. In 2005, the total fractures were more than 2 million, costing nearly \$17 billion in the United States. The rapid growth in the disease burden was also projected from 2005 to 2025 [2]. For a long time, Chinese herbal medicine has been used to treat fractures and joint diseases in China. *Herba Epimedii* (Yinyanghuo) is one of the most widely used herbs that are prescribed in formulas to treat osteoporosis [3], and its extract can reduce the occurrence of osteoporosis both in experimental models and clinical studies [4]. The term Yinyanghuo was first listed in Shen Nong Ben Cao Jing as a middle-grade herb during 200–300 B.C., and many species in genus *Epimedium* L. (Berberidaceae) have been used in traditional Chinese medicine (TCM) for a long time to nourish the kidney and reinforce the Yang [5,6]. At present, five species are recorded

in the Chinese pharmacopoeia. The dried leaves of *Epimedium brevicornu* Maxim, *Epimedium sagittatum* (Sieb. et Zucc) Maxim, *Epimedium pubescens* Maxim, and *Epimedium koreanum* Nakai are the sources of *Epimedium Folium* (Yinyanghuo), and the dried leaves of *E. wushanense* T.S. Ying are the resources of *Epimedium Wushanensis Folium* (Wushan Yinyanghuo) [7]. Simultaneously, *Epimedium* species are used medically in other Asian countries. For instance, *E. sagittatum* and *E. grandiflorum* have been used to treat impotence, prostermia, hyperdiuresis, and osteoporosis in Japan [6], and *E. koreanum* has been used in Korea as an aphrodisiac and hypotensive [8]. Up to now, more than 270 compounds were isolated from *Epimedium* species. The major active components are flavonoids and corresponding derivatives, with antitumor, antioxidative, antiosteoporosis, and protective effects [9–11].

In terms of the great medicinal value of *Epimedium* species, considerable scientific interest has been aroused in this genus. The completely accurate identification of the plant material is the basis for a scientific study. However, the classification and discrimination of *Epimedium* species have always been controversial. Genus *Epimedium* consists of approximately 58 species distributed from Japan in Asia to Algeria in North Africa [12]. Most of the *Epimedium* species exist in Eastern Asia and Mediterranean countries [12], and approximately 50 species have been reported in China [13–15]. The first *Epimedium* species, *E. alpinum* L., was recorded by Linnaeus whereas the most comprehensive classification system of this genus was established by Stearn [12]. In Stearn's monograph, genus *Epimedium* is divided into subgenera *Rhizophyllum* and *Epimedium* on the basis of the flower and leaf morphology, C-banding of the chromosomes [16], and geographical distribution [12]. Furthermore, subgenus *Epimedium* is divided into four sections, namely, *Epimedium*, *Polyphyllum*, *Macroceras*, and *Diphyllum*. According to the corolla characteristics, section *Diphyllum*, which has approximately 47 known species in China, is further subdivided into four series, namely, *Campanulatae* Stearn, *Davidianae* Stearn, *Dolichocerae* Stearn, and *Brachycerae* Stearn. Chinese sect. *Diphyllum* reaches the highest species diversity with more than 80% species in *Epimedium* [17], presenting numerous taxonomic controversies [18–21]. Morphological classification experts identify and publish new species according to their leaf and corolla characteristics [12]. Most species in this genus have similar leaf morphology (three branches from the stem and three leaves in every branch). Thus, the species identification on the basis of leaf morphology becomes difficult when the short flowering time has passed. For instance, *E. acuminata* and *E. pubescens* are difficult to differentiate due to their least variable leaf morphology [12]. Meanwhile, because of the similar shape among different species and intraspecific variation caused by geographical distribution, distinguishing closely related species becomes difficult. For instance, *E. sagittatum* and its related species, namely, *E. sagittatum* complex, present the most disputable questions in taxonomy due to the high morphological variations with extensive geographical distribution [20,22]. These taxonomic issues mentioned above complicate the species discrimination in this genus. At present, approximately 23 *Epimedium* species not recorded in the Chinese pharmacopoeia, such as *E. acuminatum*, *E. miryanthum*, and *E. leptorrhizum*, are commonly used as Yinyanghuo in particular localities of Guizhou, Sichuan, Chongqing, Jiangxi, Hunan, and Hubei provinces in China [23,24]. Considering differences in bioactive constituents among *Epimedium* species [6], incorrect identification of species is likely to result in potential risks in herbal efficacy and medical safety. Furthermore, *Herba Epimedium* is an herb that is harvested in the wild and is on the edge of extinction [25]. Improper or wrong harvest may induce the reduction of species abundance in genus *Epimedium*. Thus, effective methods for the accurate discrimination and sustainable utilization of the *Epimedium* resources are urgently needed.

In order to resolve the identification problem, some researchers have attempted to use molecular methods to discriminate *Epimedium* species. For example, 5S ribosomal RNA (rRNA) gene spacer sequencing [26], expressed sequence tag (EST) dataset [27], and microsatellites [28–30] methods have been applied to identify medicinal *Epimedium* species. Some identification problems of *Epimedium* species were settled in these studies. However, a comprehensive discrimination study of medicinally important *Epimedium* species is absent at present. Recently, DNA barcoding technique has been widely used for species discrimination. Up to now, *matK*, *rbcL*, *psbA-trnH*, internal transcribed spacer (ITS), and ITS2 are the most commonly used regions for DNA barcoding in plants [31–34]. In this study,

we systematically evaluated the feasibility and identification efficiency of four generally acknowledged single loci, namely, *rbcl*, *psbA-trnH*, ITS, and ITS2, and their combinations to discriminate the *Epimedium* species distributed in China to identify the most suitable barcodes for genus *Epimedium* and provide effective information for the species identification of this genus.

## 2. Materials and Methods

### 2.1. Plant Materials

A total of 72 samples representing 37 *Epimedium* species, of which 20 are medicinally used, were collected from the Jiangxi, Guizhou, Hubei, and Jilin provinces in China. In addition to the widely distributed *Epimedium* species, some narrowly distributed species of the genus in China were collected as well. For instance, *E. dewuense* and *E. shuichengense* are endemic species to certain areas of Guizhou province. *E. zhushanense* only grows in Zhushan County in the Hubei province, and *E. tianmenshanensis* is merely distributed in Hunan province. Moreover, some rare species, including *E. pauciflorum*, *E. platypetalum*, *E. lishihchenii*, *E. franchetii*, *E. glandulosopilosum*, *E. truncatum*, *E. dolichostemon*, *E. pudingense*, *E. ilicifolium*, *E. borealiguizhouense*, and *E. qingchengshanense*, are mostly distributed in certain areas with extremely limited resources and survive with difficulty when transplanted. All samples were authenticated by Prof. Yanqin Xu of Jiangxi University of Traditional Chinese Medicine and Prof. Shunzhi He of Guiyang College of Traditional Chinese Medicine. The details of the samples are listed in Table S1. All corresponding voucher samples were deposited in the Herbarium of the Institute of Medicinal Plant Development, Chinese Academy of Medicinal Sciences & Peking Union Medical College, Beijing, China.

### 2.2. DNA Isolation, Amplification, and Sequencing

Approximately 30 mg fresh leaf from each sample was ground for 2 min at a frequency of 30 times/s in a FastPrep bead mill (Retsch MM400, Haan, Germany). The total genomic DNA was extracted using the Universal Plant Genome DNA Kit (Tiangen Biotech, Beijing, China) according to the manufacturer's instructions. Short fragments of the specific regions of the plastid (*rbcl*), noncoding (*psbA-trnH*), and nuclear DNA (ITS, including 18S, ITS1, 5.8 S, ITS2, and 28 S) sequences were amplified from the extracts. The ITS2 sequence was retrieved from the ITS region directly. Universal primers for candidate barcodes and reaction conditions were used as previously reported [35–37]. The primers and reaction conditions used to amplify these regions are listed in Table S2. PCR was performed in 25  $\mu$ L of the reaction system, containing 2  $\mu$ L of template DNA (approximately 30 ng), 12.5  $\mu$ L of 2 $\times$  Taq PCR Master Mix (Aidlab Biotechnologies Co., Ltd., Beijing, China), and 1  $\mu$ L of each primer (2.5  $\mu$ mol/L), and filled with double-distilled water. The purified PCR products were sequenced in both directions by using an ABI 3730XL sequencer (Applied Biosystems, Inc., Foster City, CA, USA) on the basis of the Sanger sequencing method by the Major Engineering laboratory of the Chinese Academy of Agricultural Sciences (Beijing, China). Altogether, 70 samples representing 36 species were sequenced successfully with all the three loci (*rbcl*, *psbA-trnH*, and ITS), and 210 new sequences were submitted to GenBank under the accession number MG837275–MG837475 and MH252065–MH252073 (Table S1). These sequences were used in the subsequent analysis.

### 2.3. Statistical Analysis

Consensus sequences and contig generation were performed using the CodonCode Aligner V 7.1.1 (CodonCode Co., Dedham, MA, USA). The ITS sequences were subjected to hidden Markov model analysis to obtain the complete ITS2 sequences [38]. Sequences from each DNA region were aligned using MEGA 6.0 [39]. For species with more than three individuals, the average intraspecific distance, theta, and coalescent depth were calculated to evaluate the intraspecific variation based on the Kimura two-parameter (K2P) model; the average interspecific distance, minimum interspecific distance, and theta primer were used to assess the interspecific divergence using the K2P

model [40–42]. To assess the potential of each single locus and their combinations for accurate species discrimination, distance-based methods (TaxonDNA and neighbor-joining trees) and character-based approach (BLOG) were employed for species with more than one individual. The best match (BM) and best close match (BCM) tests in the Species Identifier 1.8 program of TaxonDNA were run [43]. The neighbor-joining (NJ) trees of each single locus and their combinations were constructed using MEGA 6.0, and bootstrap tests were performed with 1000 replicates [39]. The character-based DNA barcode method in BLOG (Barcoding with LOGic) 2.0 was applied to classify species with more than one individual [44]. Each single locus and their combinations were subjected to 100% slicing within species level with a maximum of 500 iterations (GRASPITER = 500) and a maximum time of 5 min for analysis (GRASPSECS = 300). The logic formula with the lowest false positive rate against the reference dataset was taken as discrimination basis. Moreover, species that could be identified were shown using Venn diagrams.

### 3. Results

#### 3.1. Efficiency of PCR Amplification and Sequencing

The three loci, *rbcL*, ITS, and *psbA-trnH*, all showed high PCR amplification and sequencing efficiency (100%). Moreover, the effective sequence ratio of *rbcL* and *psbA-trnH* were the same (100%), followed by ITS (97.2%). Detailed information regarding the PCR amplification and sequencing efficiency of the candidate barcodes is provided in Table S1.

#### 3.2. Genetic Divergence Determination

The length of the aligned sequence (base pairs) and variable sites for the *rbcL*, *psbA-trnH*, ITS, and ITS2 regions were 703/10, 572/59, 705/45, and 247/23, respectively. Six parameters were used to characterize inter- and intraspecific divergence. Results indicated that *psbA-trnH* had the highest interspecific divergence, followed by ITS2 and ITS. Meanwhile, *rbcL* exhibited a relatively lower interspecific divergence compared with the other regions (Table S3). At the intraspecific level, *rbcL* showed the lowest divergence, while *psbA-trnH* displayed the highest variation level (Table S3).

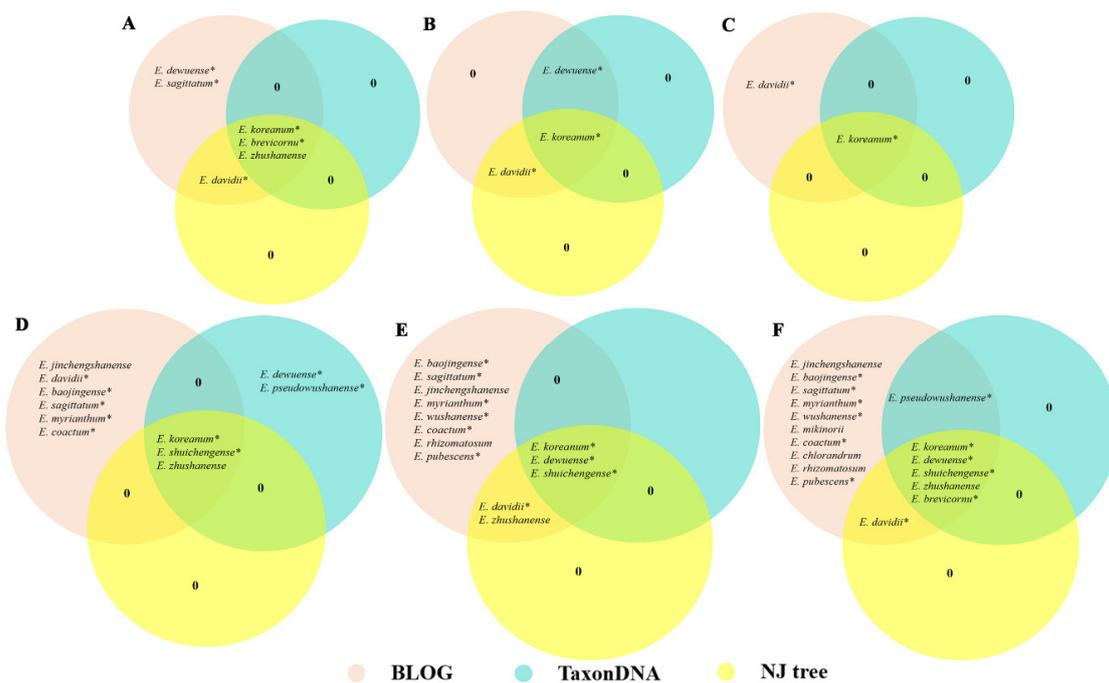
#### 3.3. Evaluation of the Identification Efficiency of the DNA Barcodes

In the present study, the distance-based method, namely, TaxonDNA, was used to assess the applicability of the different regions for species discrimination. Similar identification efficiency was obtained based on the BM and BCM methods (Table 1). For the single locus, *psbA-trnH* exhibited the highest species identification efficiency (29.62%), followed by the ITS (22.22%) and ITS2 regions (18.51%). Meanwhile, the *rbcL* region showed the lowest resolution rate of 3.70%. In the two loci combinations, *psbA-trnH* + ITS2 showed a higher resolution rate (38.88%) than *psbA-trnH* + ITS (37.03%). Moreover, compared with the two loci combinations, the resolution rate was not increased when three loci (*psbA-trnH* + ITS + *rbcL* and *psbA-trnH* + ITS2 + *rbcL*) were combined (37.03%). Results showed that only six species (*E. koreanum*, *E. dewuense*, *E. zhushanense*, *E. shuichengense*, *E. brevicornu*, and *E. pseudowushanense*) could be identified efficiently using single locus and their combinations (Figure 1).

Second only to ITS region, *psbA-trnH* exhibited comparatively higher identification efficiency among the four single loci by using the NJ tree method (Figure 1). Three species, namely, *E. koreanum*, *E. zhushanense*, and *E. shuichengense*, could be identified using the locus *psbA-trnH*. Moreover, results suggested that the number of species that can be authenticated was increased when two or three loci were combined (Figure 1). Six species, namely, *E. koreanum*, *E. dewuense*, *E. zhushanense*, *E. shuichengense*, *E. davidii*, and *E. brevicornu*, could be discriminated by *psbA-trnH* + ITS and *psbA-trnH* + ITS + *rbcL*. Additionally, five *Epimedium* species, namely, *E. koreanum*, *E. dewuense*, *E. zhushanense*, *E. shuichengense*, and *E. davidii*, could be differentiated by *psbA-trnH* + ITS2 and *psbA-trnH* + ITS2 + *rbcL*. Detailed

information regarding the NJ trees of each single locus and their combinations is provided in Figures S1 and S2.

Compared with the distance-based method, the identification efficiency of each single locus and their combinations was significantly improved by using the character-based approach (Table 1). Among the four single loci, *psbA-trnH* displayed the highest resolution rate (61.11%) with nine species being identified. By contrast, *rbcL* showed the lowest (14.81%). Moreover, combination of barcodes increased the species identification efficiency significantly. Among the four multilocus combinations, *psbA-trnH* + ITS and *psbA-trnH* + ITS + *rbcL* both showed the highest resolution rate (92.59%). Aside from *E. acuminatum*, *E. leptorrhizum*, and *E. epsteinii*, a total of 17 species, of which 12 are medicinally used, could be distinguished (Figure 1). The logic formulas to identify the 17 species by using *psbA-trnH* + ITS region are listed in Table 2.



**Figure 1.** Species that can be identified based on three methods for each single locus and their combinations by using Venn diagrams. A, ITS; B, ITS2; C, *rbcL*; D, *psbA-trnH*; E, *psbA-trnH* + ITS2/*psbA-trnH* + ITS2 + *rbcL*; F, *psbA-trnH* + ITS/*psbA-trnH* + ITS + *rbcL*. "\*" represents medicinally used species.

**Table 1.** Identification efficiency of the four loci on the basis of the BM, BCM, and BLOG analytical methods.

Region	Best Match			Best Close Match (Threshold 3%)				BLOG		
	Correct	Ambiguous	Incorrect	Correct	Ambiguous	Incorrect	No Match	Correct	Not Classified	Wrong
<i>psbA-trnH</i>	29.62%	27.77%	42.59%	29.62%	27.77%	38.88%	3.70%	61.11%	38.89%	0.00%
ITS	22.22%	70.37%	7.40%	22.22%	70.37%	7.40%	0.00%	48.15%	51.85%	0.00%
ITS2	18.51%	74.07%	7.40%	18.51%	74.07%	7.40%	0.00%	24.07%	75.93%	0.00%
<i>rbcL</i>	3.70%	96.29%	0.00%	3.70%	96.29%	0.00%	0.00%	14.81%	85.19%	0.00%
<i>psbA-trnH</i> + ITS	37.03%	24.07%	38.88%	37.03%	24.07%	38.88%	0.00%	92.59%	7.41%	0.00%
<i>psbA-trnH</i> + ITS2	38.88%	25.92%	35.18%	38.88%	25.92%	33.33%	3.70%	77.78%	22.22%	0.00%
<i>psbA-trnH</i> + ITS + <i>rbcL</i>	37.03%	18.51%	44.44%	37.03%	18.51%	44.44%	0.00%	92.59%	7.41%	0.00%
<i>psbA-trnH</i> + ITS2 + <i>rbcL</i>	37.03%	24.07%	38.88%	37.03%	24.07%	38.88%	0.00%	81.48%	18.52%	0.00%

ITS, internal transcribed spacer; BLOG, Barcoding with LOGic.

**Table 2.** Diagnostic formulas of *psbA-trnH* + ITS region generated by BLOG for species identification. “\*” represents medicinally used species.

Species	Formulas	Score	Coverage
<i>E. davidii</i> *	pos1043 = T	0.136	1
<i>E. baojingense</i> *	pos314 = G OR pos1085 = C	0.174	1
<i>E. koreanum</i> *	pos201 = T	0.136	1
<i>E. dewuense</i> *	pos395 = C & pos1220 = G	0.136	1
<i>E. sagittatum</i> *	pos201 = G & pos320 = T & pos395 = T & pos411 = G & pos1052 = T OR pos91 = C OR pos113 = A	0.24	1
<i>E. jinchengshanense</i>	pos224 = C & pos411 = T & pos1220 = G OR pos411 = G & pos1052 = A	0.136	1
<i>E. chlorandrum</i>	pos90 = T & pos150 = G & pos201 = G & pos224 = A & pos395 = T & pos411 = T & pos536 = A & pos1043 = C & pos1052 = T & pos1220 = A & pos1252 = G OR pos1026 = C & pos1182 = T	0.174	1
<i>E. pseudowushanense</i> *	pos90 = T & pos91 = A & pos224 = C & pos411 = T & pos788 = C & pos1182 = C & pos1220 = A	0.136	1
<i>E. rhizomatosum</i>	pos411 = T & pos1052 = A	0.136	1
<i>E. pubescens</i> *	pos90 = T & pos150 = G & pos201 = G & pos201 = T & pos395 = T OR pos90 = G & pos224 = C & pos411 = T & pos1026 = C & pos1182 = C OR pos320 = A	0.174	1
<i>E. shuichengense</i> *	pos201 = A & pos1052 = T	0.136	1
<i>E. myrianthum</i> *	pos90 = G & pos411 = G OR pos113 = G & pos150 = A & pos314 = A & pos730 = A	0.208	1
<i>E. wushanense</i> *	pos164 = T & pos224 = A & pos411 = T & pos1085 = T & pos1220 = G OR pos90 = G & pos201 = G & pos224 = A & pos411 = T & pos1085 = T	0.174	1
<i>E. brevicornu</i> *	pos1252 = C	0.136	1
<i>E. coactum</i> *	pos505 = C & pos536 = C OR pos1026 = A	0.136	1
<i>E. mikinorii</i>	pos788 = T	0.136	1
<i>E. zhushanense</i>	OR pos90 = T & pos224 = C & pos411 = G & pos1220 = A pos164 = G	0.136	1

ITS, internal transcribed spacer; BLOG, Barcoding with LOGic.

## 4. Discussion

### 4.1. Development of DNA Barcode Resources for Genus *Epimedium*

In our research, four universal candidate barcodes, namely, *rbcL*, *psbA-trnH*, ITS, and ITS2, were assessed for their applicability in identifying the *Epimedium* species. All three regions (*rbcL*, *psbA-trnH*, and ITS) showed high rates of sequencing recovery. The *psbA-trnH* intergenic spacer was reported as one of the most variable plastid regions in the angiosperms [35]. In our study, it demonstrated the highest interspecific divergence between species of *Epimedium*. Meanwhile, it exhibited better identification capability than the other regions for *Epimedium* species based on both distance (TaxonDNA and NJ tree) and character (BLOG) analyses. The BLOG method displayed the highest identification ability with nine species being differentiated, whereas five and three species can be respectively discriminated by TaxonDNA and NJ tree. Thus, *psbA-trnH* region may be used as a potential barcode to discriminate the *Epimedium* species. Second only to *psbA-trnH*, ITS region showed good identification ability with six species being differentiated. The China Plant BOL Group suggested that ITS/ITS2 should be incorporated into the core barcode for seed plants [45]. Here, ITS2 showed lower identification efficiency than ITS based on both distance- and character-based methods, which could be attributed to lower interspecific divergence of ITS2 for *Epimedium* species.

Considering the limited discriminatory power of the single-locus barcode, a multigene tiered approach for barcoding plants was recommended by Newmaster et al. [46]. In this study, the combination of two and three loci showed better identification efficiency than that of the single locus. The combination of two loci *psbA-trnH* and ITS, as well as the combination of three loci *psbA-trnH*, ITS, and *rbcL*, showed the highest discrimination power among all the multilocus combinations. As for the analytical methods, the character-based method exhibited the highest identification efficiency with 17 species being discriminated. The character-based method (BLOG), suggested to be efficient and precise, could provide diagnostic formulas listing the species-specific nucleotides to differentiate species from others [47]. On the contrary, the distance-based method showed lower resolution rate than the character-based method, and only six species could be discriminated. The intraspecific and interspecific genetic distances overlapped in our study (Table S3), making authentication of these species difficult by using BM test, BCM test, and NJ trees. Therefore, the character-based method is a more appropriate choice to identify *Epimedium* species.

This study contributes with DNA barcodes for more than 70% of the species in *Epimedium*, covering nearly all the commonly used medicinal plant species and rare species. For each species, information of the four universal barcodes, namely, *rbcL*, *psbA-trnH*, ITS, and ITS2, were provided. The dataset developed could provide assistance for accurate species identification, sustainable recourse utilization, and new medicinal resources development of this genus.

### 4.2. Chloroplast Genome-Based Super Barcode Has the Potential to Solve the Problem of Species Discrimination and Classification in Genus *Epimedium*

A number of taxonomic controversies exist in the Chinese sect. *Diphyllon* [18,19,48], and efficient methods for species identification and classification are lacking. In our study, the commonly used DNA barcodes were effective in identifying 17 *Epimedium* species, but the phylogenetic relationships of Chinese sect. *Diphyllon* were poorly resolved. The unresolved phylogeny of genus *Epimedium* might be ascribed to the following taxonomic issues. First, some species published earlier were only based on one or two specimens or even fragmentary specimens, resulting in the inaccurate description of the morphological characteristics. *Epimedium baojingensis* and *E. zhushanense* were believed to be the only species with unifoliolate leaves when they were published. However, the investigation of Zhang et al. [48] revealed that leaves of these species are usually trifoliolate and occasionally unifoliolate. Second, using the small variations as the reference to publish new species tends to ignore the possible connection between the new species and model specimens, which may lead to unnecessary publication to some extent and even cause problems for further study. Third, incomplete

breeding isolation between the species and existence of natural hybrids may also complicate the relationships of the Chinese *Epimedium* species, thereby increasing the difficulty of morphological identification and taxonomy research [12]. Thus, a more effective method is needed to further research the discrimination and classification of this genus. Recently, the chloroplast (cp) genomes of plants have been applied as a useful tool for phylogenetic studies and species identification [49–51]. The cp genome can significantly increase the resolution at low taxonomic levels in plant phylogeny; it was proposed as a species-level DNA barcode [52] and has been used as a plant barcode to discriminate closely related species [52–54]. The complete cp genome with more variation information than single or multiple DNA barcodes was also suggested as a super barcode [55], which may be a solution for the identification and classification problems in the *Epimedium* species. Zhang et al. [13] sequenced the complete cp genomes of five *Epimedium* species and found that the phylogenetic relationships among these cp genomes were consistent with the updated system. However, the cp genome study of 90% of the approximately 50 *Epimedium* species distributed in China still has not been conducted. Additionally, only one sample was sequenced for each species, which was not adequate to demonstrate the evolutionary relationships and divisions within the section *Diphyllon* because some *Epimedium* species have large intraspecific variations. Therefore, more systematic and in-depth classification and identification investigation on the basis of the complete cp genomes of a large sample size are urgently needed. Thus, the cp genome-based super barcode will be utilized to study the identification and evolution relationships of all Chinese *Epimedium* species in our further research. Further research is expected to solve the complex problems in taxonomy and species discrimination and to guarantee the medical safety of species in genus *Epimedium*.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2073-4425/9/12/637/s1>, **Figure S1.** NJ trees of *Epimedium* species on the basis of the *psbA-trnH* (A), *rbcL* (B), ITS (C), and ITS2 (D) regions. The bootstrap scores (1000 replicates) are shown ( $\geq 50\%$ ) for each branch. **Figure S2.** NJ trees of *Epimedium* species on the basis of four multilocus combinations. The bootstrap scores (1000 replicates) are shown ( $\geq 50\%$ ) for each branch. A, *psbA-trnH* + ITS; B, *psbA-trnH* + ITS2; C, *psbA-trnH* + ITS2 + *rbcL*; D, *psbA-trnH* + ITS + *rbcL*. **Table S1.** Voucher information, PCR amplification and sequencing efficiency of the candidate barcodes, and GenBank accession numbers for *Epimedium* plant samples in this study. **Table S2.** List of universal primers and reaction conditions for candidate barcodes used in this study. **Table S3.** Inter- and intraspecific genetic divergences of the four loci.

**Author Contributions:** Conceptualization, X.P.; Data curation, M.G. and Y.X.; Formal analysis, M.G. and X.P.; Funding acquisition, X.P.; Investigation, L.R.; Methodology, M.G.; Resources, Y.X., L.R., and S.H.; Validation, X.P.; Writing—original draft, M.G.; Writing—review & editing, X.P.

**Funding:** This work was supported by grants from the National Natural Science Foundation of China (No. 81573541), CAMS Innovation Fund for Medical Sciences (CIFMS) (No. 2017-I2M-1-013), and the China Scholarship Fund (No. 201808110108).

**Acknowledgments:** We would like to acknowledge Professor Baolin Guo (the Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences) for providing experimental samples (RM01, JY01, WS01, and WS03).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Reginster, J.Y.; Burlet, N. Osteoporosis: A still increasing prevalence. *Bone* **2006**, *38*, 4–9. [[CrossRef](#)] [[PubMed](#)]
2. Burge, R.; Dawson-Hughes, B.; Solomon, D.H.; Wong, J.B.; King, A.; Tosteson, A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005–2025. *J. Bone Miner. Res.* **2007**, *22*, 465–475. [[CrossRef](#)] [[PubMed](#)]
3. Hsieh, T.P.; Sheu, S.Y.; Sun, J.S.; Chen, M.H.; Liu, M.H. Icariin isolated from *Epimedium pubescens* regulates osteoblasts anabolism through BMP-2, SMAD4, and Cbfa1 expression. *Phytomedicine* **2010**, *17*, 414–423. [[CrossRef](#)] [[PubMed](#)]
4. Qian, G.; Zhang, X.; Lu, L.; Wu, X.; Li, S.; Meng, J. Regulation of Cbfa1 expression by total flavonoids of *Herba epimedii*. *Endocr. J.* **2006**, *53*, 87–94. [[CrossRef](#)] [[PubMed](#)]

5. Wu, H.; Lien, E.J.; Lien, L.L. Chemical and pharmacological investigations of *Epimedium* species: A survey. *Prog. Drug Res.* **2003**, *60*, 1–57.
6. Ma, H.; He, X.; Yang, Y.; Li, M.; Hao, D.; Jia, Z. The genus *Epimedium*: An ethnopharmacological and phytochemical review. *J. Ethnopharmacol.* **2011**, *134*, 519–541. [[CrossRef](#)] [[PubMed](#)]
7. Chinese Pharmacopoeia Commission. *Pharmacopoeia of the People's Republic of China*; Chemical Industry Press: Beijing, China, 2015; Volume I.
8. Cho, W.-K.; Kim, H.; Choi, Y.J.; Yim, N.H.; Yang, H.J.; Ma, J.Y. *Epimedium koreanum* Nakai Water Extract Exhibits Antiviral Activity against Porcine Epidermic Diarrhea Virus In Vitro and In Vivo. *Evid.-Based Complement. Altern.* **2012**, *2012*, 1–10.
9. Yuan, D.; Wang, H.; He, H.; Jia, L.; He, Y.; Wang, T.; Zeng, X.; Li, Y.; Li, S.; Zhang, C. Protective effects of total flavonoids from *Epimedium* on the male mouse reproductive system against cyclophosphamide-induced oxidative injury by up-regulating the expressions of SOD3 and GPX1. *Phytother. Res.* **2014**, *28*, 88–97. [[CrossRef](#)]
10. Jiang, J.; Song, J.; Jia, X.B. Phytochemistry and ethnopharmacology of *Epimedium* L. species. *Chin. Herb. Med.* **2015**, *7*, 204–222. [[CrossRef](#)]
11. Telang, N.T.; Li, G.; Katdare, M.; Sepkovic, D.W.; Bradlow, H.L.; Wong, G.Y.C. The nutritional herb *Epimedium grandiflorum* inhibits the growth in a model for the Luminal A molecular subtype of breast cancer. *Oncol. Lett.* **2017**, *13*, 2477–2482. [[CrossRef](#)]
12. Stearn, W.T. *The genus Epimedium and other herbaceous Berberidaceae*; Timber Press: Portland, OR, USA, 2002.
13. Zhang, Y.; Du, L.; Liu, A.; Chen, J.; Wu, L.; Hu, W.; Zhang, W.; Kim, K.; Lee, S.; Yang, T.; et al. The Complete Chloroplast Genome Sequences of Five *Epimedium* Species: Lights into Phylogenetic and Taxonomic Analyses. *Front. Plant Sci.* **2016**, *7*, 306. [[CrossRef](#)] [[PubMed](#)]
14. Wang, Y.Y.; Xu, W.F.; He, S.Z. *Epimedium muhuangense* (Berberidaceae), a new species from China. *Phytotaxa* **2017**, *319*, 277–282. [[CrossRef](#)]
15. Wei, N.; Zhang, Y.J.; Xu, Z.; Kamande, E.M.; Ngumbau, V.M.; Hu, G.W. *Epimedium zhaotongense* (Berberidaceae), a New Species from Yunnan Province, China. *Phytotaxa* **2017**, *296*, 88. [[CrossRef](#)]
16. Takahashi, C. Karyomorphological studies on speciation of *Epimedium* and its allied *Vancouveria* with special reference to C-bands. *J. Sci. Hiroshima Univ. B* **1989**, *2*, 159–269.
17. Ying, J. Petal evolution and distribution patterns of *Epimedium* L. (Berberidaceae). *Acta Phytotaxon. Sin.* **2002**, *40*, 481–489.
18. De Smet, Y.; Goetghebeur, P.; Wanke, S.; Asselman, P.; Samain, M.S. Additional evidence for recent divergence of Chinese *Epimedium* (Berberidaceae) derived from AFLP, chloroplast and nuclear data supplemented with characterisation of leaflet pubescence. *Plant Ecol. Evol.* **2012**, *145*, 73–87. [[CrossRef](#)]
19. Guo, B.L.; Pei, L.K.; Xiao, P.G. Further research on taxonomic significance of flavonoids in *Epimedium* (Berberidaceae). *J. Syst. Evol.* **2008**, *46*, 874–885.
20. Ying, T.S.; Boufford, D.E.; Brach, A.R. *Epimedium* L. In *Flora of China*; Science Press: St. Louis, MO, USA; Botanical Garden Press: Beijing, China, 2011; Volume 19, pp. 787–799.
21. Govindaraghavan, S.; Hennell, J.R.; Sucher, N.J. From classical taxonomy to genome and metabolome: Towards comprehensive quality standards for medicinal herb raw materials and extracts. *Fitoterapia* **2012**, *83*, 979–988. [[CrossRef](#)]
22. Xu, Y.Q.; Cai, W.Z.; Hu, S.F.; Huang, X.H.; Ge, F.; Wang, Y. Morphological variation of non-glandular hairs in cultivated *Epimedium sagittatum* (Berberidaceae) populations and implications for taxonomy. *Biodivers. Sci.* **2013**, *21*, 185–196.
23. He, S.Z. *The Genus Epimedium of China in Colour*; Guizhou Publishing Group: Guizhou, China, 2014.
24. Xie, P.S.; Yan, Y.Z.; Guo, B.L.; Lam, C.W.K.; Chui, S.H.; Yu, Q.X. Chemical pattern-aided classification to simplify the intricacy of morphological taxonomy of *Epimedium* species using chromatographic fingerprinting. *J. Pharmaceut. Biomed.* **2010**, *52*, 452–460. [[CrossRef](#)]
25. Zhang, H.F.; Yang, X.H. Asian medicine: Protect rare plants. *Nature* **2012**, *482*, 35. [[CrossRef](#)] [[PubMed](#)]
26. Sun, Y.; Fung, K.P.; Leung, P.C.; Shi, D.; Shaw, P.C. Characterization of medicinal *Epimedium* species by 5S rRNA gene spacer sequencing. *Planta Med.* **2004**, *70*, 287–288. [[CrossRef](#)] [[PubMed](#)]
27. Zeng, S.; Xiao, G.; Guo, J.; Fei, Z.; Xu, Y.; Roe, B.A.; Wang, Y. Development of a EST dataset and characterization of EST-SSRs in a traditional Chinese medicinal plant, *Epimedium sagittatum* (Sieb. Et Zucc.) Maxim. *BMC Genom.* **2010**, *11*, 94. [[CrossRef](#)] [[PubMed](#)]

28. Zhou, J.; Xu, Y.; Huang, H.; Wang, Y. PRIMER NOTE: Identification of microsatellite loci from *Epimedium koreanum* and cross-species amplification in four species of *Epimedium* (Berberidaceae). *Mol. Ecol. Notes* **2006**, *7*, 467–470. [[CrossRef](#)]
29. Xu, Y.; Li, Z.; Wang, Y. Fourteen microsatellite loci for the Chinese medicinal plant *Epimedium sagittatum* and cross-species application in other medicinal species. *Mol. Ecol. Resour.* **2008**, *8*, 640–642. [[CrossRef](#)] [[PubMed](#)]
30. Yousaf, Z.; Hu, W.; Zhang, Y.; Zeng, S.; Wang, Y. Systematic validation of medicinally important genus *epimedium* species based on microsatellite markers. *Pak. J. Bot.* **2015**, *47*, 477–484.
31. Chen, S.; Pang, X.; Song, J.; Shi, L.; Yao, H.; Han, J.; Leon, C. A renaissance in herbal medicine identification: From morphology to DNA. *Biotechnol. Adv.* **2014**, *32*, 1237–1244. [[CrossRef](#)]
32. Hollingsworth, P.M.; Graham, S.W.; Little, D.P. Choosing and using a plant DNA barcode. *PLoS ONE* **2011**, *6*, e19254. [[CrossRef](#)]
33. Hollingsworth, P.M.; Li, D.Z.; van der Bank, M.; Twyford, A.D. Telling plant species apart with DNA: From barcodes to genomes. *Philos. Trans. R. Soc. B* **2016**, *371*, 20150338. [[CrossRef](#)]
34. Guo, M.; Ren, L.; Pang, X. Inspecting the true identity of herbal materials from *Cynanchum* using ITS2 barcode. *Front. Plant Sci.* **2017**, *8*, 1945. [[CrossRef](#)]
35. Kress, W.J.; Wurdack, K.J.; Zimmer, E.A.; Weigt, L.A.; Janzen, D.H. Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8369–8374. [[CrossRef](#)] [[PubMed](#)]
36. Kress, W.J.; Erickson, D.L. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE* **2007**, *2*, e508. [[CrossRef](#)] [[PubMed](#)]
37. Sass, C.; Little, D.P.; Stevenson, D.W.; Specht, C.D. DNA barcoding in the cycadales: Testing the potential of proposed barcoding markers for species identification of cycads. *PLoS ONE* **2007**, *2*, e1154. [[CrossRef](#)] [[PubMed](#)]
38. Keller, A.; Schleicher, T.; Schultz, J.; Muller, T.; Dandekar, T.; Wolf, M. 5.8S–28S rRNA interaction and HMM-based ITS2 annotation. *Gene* **2009**, *430*, 50–57. [[CrossRef](#)] [[PubMed](#)]
39. Tamura, K.; Stecher, G.; Peterson, D.; Filipowski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [[CrossRef](#)] [[PubMed](#)]
40. Chen, S.; Yao, H.; Han, J.; Liu, C.; Song, J.; Shi, L.; Zhu, Y.; Ma, X.; Gao, T.; Pang, X.; et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE* **2010**, *5*, e8613. [[CrossRef](#)] [[PubMed](#)]
41. Meyer, C.P.; Paulay, G. DNA barcoding: Error rates based on comprehensive sampling. *PLoS Biol.* **2005**, *3*, e422. [[CrossRef](#)] [[PubMed](#)]
42. Meier, R.; Zhang, G.; Ali, F. The use of mean instead of smallest interspecific distances exaggerates the size of the “barcoding gap” and leads to misidentification. *Syst. Biol.* **2008**, *57*, 809–813. [[CrossRef](#)]
43. Meier, R.; Shiyang, K.; Vaidya, G.; Ng, P.K. DNA barcoding and taxonomy in Diptera: A tale of high intraspecific variability and low identification success. *Syst. Biol.* **2006**, *55*, 715–728. [[CrossRef](#)]
44. Weitschek, E.; van Velzen, R.; Felici, G.; Bertolazzi, P. BLOG 2.0: A software system for character-based species classification with DNA Barcode sequences. What it does, how to use it. *Mol. Ecol. Resour.* **2013**, *13*, 1043–1046. [[CrossRef](#)]
45. China Plant BOL Group. Comparative analysis of a large dataset indicates that ITS should be incorporated into the core barcode for seed plants. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 19641–19646. [[CrossRef](#)] [[PubMed](#)]
46. Newmaster, S.G.; Fazekas, A.J.; Ragupathy, S. DNA barcoding in land plants: Evaluation of *rbcL* in a multigene tiered approach. *Botany* **2006**, *84*, 335–341. [[CrossRef](#)]
47. Goldstein, P.Z.; DeSalle, R. Integrating DNA barcode data and taxonomic practice: Determination, discovery, and description. *Bioessays* **2011**, *33*, 135–147. [[CrossRef](#)] [[PubMed](#)]
48. Zhang, Y.J.; Dang, H.S.; Wang, Y.; Li, X.W.; Li, J.Q. A taxonomic revision of unifoliolate Chinese *Epimedium* L. (Berberidaceae). *Kew Bull.* **2011**, *66*, 253. [[CrossRef](#)]
49. Wicke, S.; Schneeweiss, G.M.; dePamphilis, C.W.; Muller, K.F.; Quandt, D. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol. Biol.* **2011**, *76*, 273–297. [[CrossRef](#)]
50. Nystedt, B.; Street, N.R.; Wetterbom, A.; Zuccolo, A.; Lin, Y.-C.; Scofield, D.G.; Vezzi, F.; Delhomme, N.; Giacomello, S.; Alexeyenko, A.; et al. The Norway spruce genome sequence and conifer genome evolution. *Nature* **2013**, *497*, 579–584. [[CrossRef](#)]

51. Nguyen, P.A.T.; Kim, J.S.; Kim, J.H. The complete chloroplast genome of colchicine plants (*Colchicum autumnale* L. and *Gloriosa superba* L.) and its application for identifying the genus. *Planta* **2015**, *242*, 223–237. [[CrossRef](#)]
52. Parks, M.; Cronn, R.; Liston, A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biol.* **2009**, *7*, 84. [[CrossRef](#)]
53. Nock, C.J.; Waters, D.L.E.; Edwards, M.A.; Bowen, S.G.; Rice, N.; Cordeiro, G.M.; Henry, R.J. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnol. J.* **2011**, *9*, 328–333. [[CrossRef](#)]
54. Kane, N.; Sveinsson, S.; Dempewolf, H.; Yang, J.Y.; Zhang, D.; Engels, J.M.; Cronk, Q. Ultra-barcoding in cacao (*Theobroma* spp.; Malvaceae) using whole chloroplast genomes and nuclear ribosomal DNA. *Am. J. Bot.* **2012**, *99*, 320–329. [[CrossRef](#)]
55. Li, X.; Yang, Y.; Henry, R.J.; Rossetto, M.; Wang, Y.; Chen, S. Plant DNA barcoding: From gene to genome. *Biol. Rev. Camb. Philos. Soc.* **2015**, *90*, 157–166. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).