

RESEARCH ARTICLE

# Evaluation of DNA barcodes in *Codonopsis* (Campanulaceae) and in some large angiosperm plant genera

De-Yi Wang<sup>1,2</sup>, Qiang Wang<sup>1</sup>, Ying-Li Wang<sup>3</sup>, Xiao-Guo Xiang<sup>1</sup>, Lu-Qi Huang<sup>4\*</sup>, Xiao-Hua Jin<sup>1,5\*</sup>

**1** State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China, **2** University of Chinese Academy of Sciences, Beijing, China, **3** Baotou Medical College, Donghe District, Baotou City, Inner Mongolia, China, **4** National Resource Centre for Chinese Materia Medica, China Academy of Chinese Medical Science, Beijing, China, **5** Southeast Asia Biodiversity Research Institute, Chinese Academy of Science (CAS-SEABRI), Menglun, Mengla, Yunnan, China

☯ These authors contributed equally to this work.

\* [xiaohuajin@ibcas.ac.cn](mailto:xiaohuajin@ibcas.ac.cn) (XHJ); [huangluqi01@163.com](mailto:huangluqi01@163.com) (LQH)



**OPEN ACCESS**

**Citation:** Wang D-Y, Wang Q, Wang Y-L, Xiang X-G, Huang L-Q, Jin X-H (2017) Evaluation of DNA barcodes in *Codonopsis* (Campanulaceae) and in some large angiosperm plant genera. PLoS ONE 12(2): e0170286. doi:10.1371/journal.pone.0170286

**Editor:** Zhong-Jian Liu, The National Orchid Conservation Center of China; The Orchid Conservation & Research Center of Shenzhen, CHINA

**Received:** June 7, 2016

**Accepted:** January 3, 2017

**Published:** February 9, 2017

**Copyright:** © 2017 Wang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data availability statement:** All sequences of *Codonopsis* are available from the GenBank database (accession numbers: *rbcl* sequences, KY070603 - KY070736; *psbA-trnH* sequences, KY070737 - KY070869; *matK* sequences, KY070870 - KY071002; ITS sequences, KY071003 - KY071130). All relevant data are within the paper and its Supporting Information files.

## Abstract

DNA barcoding is expected to be one of the most promising tools in biological taxonomy. However, there have been no agreements on which core barcode should be used in plants, especially in species-rich genera with wide geographical distributions. To evaluate their discriminatory power in large genera, four of the most widely used DNA barcodes, including three plastid regions (*matK*, *rbcl*, *trnH-psbA*) and nuclear internal transcribed spacer (nrITS), were tested in seven species-rich genera (*Ficus*, *Pedicularis*, *Rhodiola*, *Rhododendron*, *Viburnum*, *Dendrobium* and *Lysimachia*) and a moderate size genus, *Codonopsis*. All of the sequences from the aforementioned seven large genera were downloaded from NCBI. The related barcodes for *Codonopsis* were newly generated in this study. Genetics distances, DNA barcoding gaps and phylogenetic trees of the four single barcodes and their combinations were calculated and compared in the seven genera. As for single barcode, nrITS has the most variable sites, the clearest intra- and inter-specific divergences and the highest discrimination rates in the seven genera. Among the combinations of barcodes, ITS + *matK* performed better than all the single barcodes in most cases and even the three- and four-loci combinations in the seven genera. Therefore, we recommend ITS+*matK* as the core barcodes for large plant genera.

## Introduction

DNA barcoding, the use of a short gene sequence from a standardized region of the genome as a tool for species identification, provides new tools for use in biological taxonomy [1–5]. It has shown promise in providing a practical, standardized, species-level identification tool that can be used for taxonomic research, population genetics [6], phylogenetics [7], biodiversity assessment [8], and ecological studies [9–11]. An ideal DNA barcode should be variable enough to

**Funding:** This study was supported by grants from the Chinese Special Fund for Medicine Research in the Public Interest (201407003), and the National Natural Science Foundation of China (Grant No. 31470299).

**Competing interests:** The authors have declared that no competing interests exist.

resolve closely related species and short enough for easy experimental manipulation at low cost [12]. The cytochrome oxidase I (COI) for the zoological community appears to generally fulfill these criteria [11, 13–15]. In contrast, there is no universally accepted counterpart barcode for plants yet [16]. In the past decade, four loci widely used in plant molecular systematics, namely, ITS, *matK*, *rbcl* and *trnH-psbA*, have been extensively evaluated [3, 16–19]. The lack of resolving power for single barcodes has led to the transition from a single- to a multi-region barcoding system [3, 4, 20–24]. Specifically, the combined use of short segments of the chloroplast genes *matK* and *rbcl* was proposed by the Plant Working Group of the Consortium for Barcodes of Life [3, 14].

Despite the significant progress made in the DNA barcoding of higher plants, some obstacles still hinder its extensive application in plant taxonomy [8]. Firstly, rates of successful amplification and sequencing of candidate DNA makers are highly variable among plant taxa [23, 25]. Secondly, discriminating closely related or recently evolved species remains a challenge for DNA barcoding [8, 14, 26]. Furthermore, one of the biggest challenges is the lack of a broad sampling of well-dispersed species across all plant taxa. It has been proven that barcoding based solely on a limited number of DNA sequences was often inappropriate at the species-specific level [15, 26]. Finally, to date, species discrimination within a genus has been evaluated in a number of cases; however, the barcoding of species-rich genera is particularly difficult and has still not been sufficiently evaluated [21]. Recently, several studies of large plant genera [21, 23, 27–29], such as *Pedicularis* (328 samples representing 88 species), *Rhododendron* (531 samples representing 173 species), and *Dendrobium* (1698 accessions representing 184 species), have been conducted. Particularly, based on the comparative analyses of barcodes in five plant genera with range from large to moderate size (*Dendrobium*, *Ficus*, *Lysimachia*, *Paphiopedilum*, *Pedicularis*), Xu et al. [21] proposed the combination of *matK* + ITS as the core barcode for large flowering plant genera.

*Codonopsis* s.l. (Campanulaceae) consists of 42 species, mainly distributed in Central, East and South Asia [5, 30, 31]. Many *Codonopsis* species are widely used in traditional medicine and foods across their distributed regions. Fresh or dried caudices and roots of *Codonopsis pilosula* and *C. tangshen* have a long history of use as herbal medicines "Dangshen" in China [5, 30]. The roots and caudices of other *Codonopsis* species, including *C. tubulosa*, *C. subglobosa*, *C. clematidea* and *C. lanceolata*, are used as vegetables across several Asian countries [5, 32]. Because of the highly similar morphological appearance of the roots and caudices of *Codonopsis* species, DNA barcodes may be valuable for the accurately identifying *Codonopsis* material. According to some studies, *Codonopsis* is a very difficult genus to identify due to its rich and complex species composition, dynamic evolutionary history, and extensive plastid genome rearrangements during diversification [30, 31, 33]. Therefore, the genus *Codonopsis* provides an excellent opportunity to evaluate the application of DNA barcoding.

In this study, our aims were as follows: (1) to evaluate the performance of DNA barcodes in *Codonopsis*; and (2) the performance of barcodes for species identification in large plant genera.

## Materials and methods

### Plant materials, DNA extraction, PCR amplification, sequencing and sequence downloaded

A total of 140 individuals of 35 *Codonopsis* species were used in this study (Table A in S1 File). In the present study, healthy and fresh leaves of each plant were collected and dried immediately in silica gel for DNA extraction. Total genomic DNA was isolated from approximately 1 g of dried leaves following a modified cetyltrimethylammonium bromide (CTAB) protocol.

Three plastid barcodes (the coding genes *matK* and *rbcL*, and the spacer *trnH-psbA*) and a nuclear internal transcribed spacer (nrITS) were amplified and sequenced using universal primers (Table B in [S1 File](#)). Polymerase chain reaction (PCR) was used to amplify the selected DNA regions. The PCR mixture (25  $\mu$ L) contained approximately 10 ng (1–2  $\mu$ L) of template DNA, 12.5  $\mu$ L of 2 $\times$ PCR mix (0.005 units/ $\mu$ L Taq DNA polymerase; 4 mM MgCl<sub>2</sub>; and 0.4 mM dNTPs), 0.2  $\mu$ L of each primer and 6.5–7.5  $\mu$ L of ddH<sub>2</sub>O. The conditions of PCR were following Raskoti et al [34]. The sequencing reactions were performed using the Applied Biosystems Prism BigDye Terminator Cycle Sequencing Kit (FosterCity, CA).

Large genera (here considered as about 100 species or more for each genus) were chosen based on a literature survey using Web of Sciences (Accessed by Jan. 12, 2016). Seven genera, including *Dendrobium*, *Ficus*, *Lysimachia*, *Pedicularis*, *Rhodiola*, *Rhododendron*, and *Viburnum*, were chosen based on the number of species and the barcodes used [21, 23, 27, 28, 35–37]. Most of these genera have more than 200 species. Because the DNA barcoding results of *Dendrobium* [21] and *Lysimachia* [37] from prior studies could be used directly for comparison, here we focused on the other five genera. All sequences for the four extensively used barcodes (ITS, *matK*, *rbcL* and *trnH-psbA*) of the five species-rich genera were downloaded from NCBI. In order to improve accuracy of the evaluation of barcodes, the downloaded sequences were filtered and omitted if they met the following criteria: i) had a length less than 300 bp (length relative to ITS2, which has a length about 300 bp); ii) were with low quality (such as N/? in sequence); iii) lacked a voucher or were only identified to genera (voucher is important for the reliability of sequence). To save computational time, the representatives of each species were limited to fifteen samples. Conclusively, the number of species and individuals in these five genera are 858 sequences of 63 species in *Ficus*, 1306 sequences of 88 species in *Pedicularis*, 672 sequences of 47 species in *Rhodiola*, 1540 sequences of 130 species in *Rhododendron*, 694 sequences of 56 species in *Viburnum*. Although there are studies of the DNA barcoding on other three large genera, i.e., *Begonia* [38] and *Astragalus* [39], *Paphiopedilum* [40], very few species or different sampling strategy and/or barcodes had been included in the analyses. For example, there are about 3000 species in *Astragalus* [41], only eight species were included in the DNA barcoding research [39]. Therefore, these three genera were excluded from our further analyses. The taxa and GenBank accession numbers used in this study are shown in Tables C–G (see [S1 File](#)).

## Sequence alignment, genetic distance and barcoding gap

DNA barcodes were aligned with Clustal X V2.0 [42] and manually adjusted in BioEdit v7.2.5 [43]. Then, SequenceMatrix 1.7.7 [44] was used to combine matrixes of single marker into matrixes of multi-makers. The genetic distance of ITS (I), *matK* (M), *rbcL* (R) and *trnH-psbA* (T) and their combinations (I+M, I+R, I+T, M+R, M+T, R+T, I+M+R, I+M+T, I+R+T, M+R+T, I+M+R+T) were systematically analyzed and compared. The output data from BioEdit was processed to calculate the pairwise distance, between group distance and within group distance under the Kimura-2-Parameter (K2P) distances model [45] for each region using MEGA v6.0 [46]. Differences between intra- and inter-specific distances for each pair of four single barcodes were compared using IBM SPSS Statistics v19.0 with Wilcoxon signed-rank tests [47]. To compare barcoding gaps, the distributions of the pairwise intra- and inter-specific distances for each candidate barcode with 0.005 distance intervals were estimated in TaxonDNA with a 'pairwise summary' function [48].

## Species discrimination efficiency

The species discrimination efficiency of both single barcodes and their combinations was evaluated by two methods as described below. Firstly, 'Best Match' and 'Best Close Match'

functions in TaxonDNA were applied to calculate the accuracy of the barcode regions for species identification. To further evaluate the efficiency of candidate barcodes, a tree-based analysis was conducted to assess the monophyly of individuals representing the same species. The neighbor-joining (NJ) and the unweighted pair group method with arithmetic mean (UPGMA) trees were reconstructed by MEGA v6.0 with the K2P model, and node support was assessed by a bootstrap test with 1000 pseudo-replicates with the K2P distance options [49].

## Results

### Analyses of sequence characteristics

The sequence characteristics of the four regions (ITS, *matK*, *rbcL* and *trnH-psbA*) in *Codonopsis* were shown in Table I in [S1 File](#). The ITS had the highest percentage of variable sites (52.19%), whereas *rbcL* had the lowest (20.50%). *trnH-psbA* had the highest informative sites (30.85%), closely followed by ITS (30.05%) and *rbcL* had the lowest (11.71%).

The sequence characteristics of the four regions in the other five genera (*Ficus*, *Pedicularis*, *Rhodiola*, *Rhododendron*, and *Viburnum*) are summarized in Table J in [S1 File](#). Among the four single barcodes in *Ficus*, the *trnH-psbA* matrix showed the shortest length for the aligned sequences, and ITS provided the highest percentage of variable sites (38.63%) and the highest percentage of informative sites (32.64%), with *rbcL* having the lowest percentages of variable and informative sites (1.79% and 1.30%, respectively).

In *Pedicularis*, length variation exists in *matK* (728–788bp) and in *trnH-psbA* (388–882bp), whereas ITS (662bp) and *rbcL* (606bp) lengths were stable. The *trnH-psbA* provided the highest percentages of variable (67.35%) and informative (58.96%) sites, followed by those of ITS (54.59% and 48.87%, respectively) and *matK* (38.54% and 32.24%, respectively), with *rbcL* having the lowest percentages (14.69% and 14.03%, respectively).

In *Rhodiola*, the *rbcL* matrix had the longest length (1100 bp) of the aligned sequences, and the other matrixes exhibited variable lengths (630–671 bp for ITS, 726–737 bp for *matK*, and 366–381 bp for *trnH-psbA*). ITS had the highest percentages of variable and informative sites (39.94% and 31.59%, respectively).

In *Rhododendron*, the *trnH-psbA* matrix showed the shortest length (515 bp) followed by 701 bp for *rbcL*, 723 bp for ITS, and 765 bp for *matK*. *TrnH-psbA* also had the highest percentages of variable (21.17%) and informative (19.22%) sites followed by *matK* (12.42% and 11.11%, respectively).

In *Viburnum*, the barcode regions varied in length, among which *trnH-psbA* had the lowest range (405–471 bp). ITS had the highest percentages of variable and informative sites were the highest for ITS (29.98% and 23.82%, respectively), whereas *rbcL* had the lowest percentages of variable and informative sites (4.48% and 2.65%, respectively).

### DNA barcode intra- and inter-specific divergence

The mean intra- and inter-specific divergence of candidate barcodes and their combinations were different in the six large genera (Tables J, K in [S1 File](#)). Among the single barcodes, our results showed that ITS exhibited the highest mean inter-specific and lower intra-specific divergence in *Ficus*, *Rhodiola*, and *Viburnum*, whereas in *Codonopsis*, *Pedicularis* and *Rhododendron*, *trnH-psbA* had the highest mean inter-specific and lower intra-specific divergence, closely followed by ITS. To further test whether such barcoding gaps exist, the distribution of divergences of each barcodes for the six genera were drawn (Figure A-F in [S2 File](#)). Among the single barcodes, there was a clear separation for ITS in the six genera. For the combinations of

single barcodes, I+M, I+R, and I+T had less overlap of inter-specific and intra-specific divergence and performed better than the other combinations among the six genera.

The results of the 'Best match' and 'Best close match' analyses indicated that the 'Best match' always performed better than the latter or resulted in equal individual identification rates (Table L in [S1 File](#)). All the six genera have varied individual identification rates by the 'Best match' ranging from 37.31% to 84.68% in *Codonopsis*, 2.73%–85.48% in *Ficus*, 51.07%–89.38% in *Pedicularis*, 19.04%–79.16% in *Rhodiola*, 11.94%–49.09% in *Rhododendron*, and 3.15%–67.53% in *Viburnum*. As a whole, the ability of DNA barcodes to discriminate between species was rather high in the six large genera with the exception of *Rhododendron*.

Among the single barcodes, ITS showed the highest individual identification rates among *Codonopsis* (62.07%), *Ficus* (77.88%), *Pedicularis* (83.79%), *Rhodiola* (68.45%), and *Viburnum* (45.34%) using the 'Best Match'; *matK* exhibited a higher discrimination rate in *Codonopsis* (61.65%) and *Pedicularis* (68.82%). Overall, all the single barcodes produced particularly low discrimination rates in *Rhododendron* with the highest rate (28.26% for *trnH-psbA*) generated by the 'Best Match'.

The 'Best Match' analyses of the combined barcodes indicated that they performed differently in the six genera (Table L in [S1 File](#)). In *Codonopsis*, I+M showed the highest individual identification rate (82.20%) among two-locus combinations, which was slightly lower than the rate from a four-loci combination (84.68% for I+M+R+T (ITS+*matK*+*rbcL*+*trnH-psbA*)). Furthermore, I+M also exhibited the highest individual identification rate (85.48%) among all the combined barcodes in *Ficus*. In *Pedicularis*, the highest discrimination was 89.80% for I+M+T (ITS+*matK*+ *trnH-psbA*), followed by 89.38% for I+M+R+T; the two loci barcodes I+M and I+T also performed well with rather high discrimination rates of 85.98% and 88.99%, respectively. In *Rhodiola*, the highest discrimination rate was 79.16% for both I+M+T and I+M+R+T; and the two loci barcodes I+M and I+T also performed well with discrimination rates of 76.19% and 77.38%, respectively. In *Rhododendron*, all the combined barcodes had low discrimination rates that were below 50%, and among which I+M+T had the highest (49.09%). In *Viburnum*, I+M+R+T showed the highest discrimination rate (67.53%), slightly more than that of I+R+T (67.05%).

## Tree-based method analyses

Our results indicate that the UPGMA tree provided the better indications of discriminatory power than the NJ tree (Table M in [S1 File](#), Figure A-Z, a-d in [S3 File](#), Figure A-Z, a-d in [S4 File](#), Figure A-Z, a-d in [S5 File](#), Figure A-Z, a-d in [S6 File](#), Figure A-Z, a-d in [S7 File](#), Figure A-Z, a-d in [S8 File](#)). The discrimination rate using two phylogenetic methods was high in the six genera with exception *Rhododendron*, for which the discrimination rates were below 50%. All the single barcodes showed lower discrimination rates that were below 50% in *Codonopsis*, among which *rbcL* had the lowest identification rate, and the other three barcodes (ITS, *matK*, and *trnH-psbA*) did not have distinctive discriminatory power with either the NJ and UPGMA tree. Among the other five genera, ITS had the highest discriminatory power with both the NJ tree and UPGMA tree, respectively, with discrimination rates of 68.85% and 72.13% in *Ficus*, 79.55% and 72.73% in *Pedicularis*, 53.19% and 51.06% in *Rhodiola*, and 42.22% and 40.00% in *Viburnum*. Additionally, *matK* had the highest discriminatory power (21.54% and 22.31%) among the single barcodes in *Rhododendron* using both phylogenetic methods.

Among the combined markers in *Codonopsis*, I+M and I+M+R+T showed the highest discrimination rate (66.67%) with both the NJ and UPGMA tree methods. Furthermore, I+M exhibited the highest discriminatory power with both the NJ tree and UPGMA trees, respectively, with discrimination rates of 74.58% and 69.49% in *Ficus*, and 83.91% and 82.76% in

*Pedicularis*, respectively. In *Rhodiola*, I+T had the highest discriminatory power (63.83% and 57.45%) with both the NJ tree and UPGMA tree methods, followed by 57.45% and 55.32% for I+M. In *Viburnum*, I+M+R provided the highest discrimination rates with values of 62.07% with the NJ tree and 48.28% with the UPGMA tree. Using combinations of barcodes, however, failed to improve the discriminatory power in *Rhododendron* merely with the highest values being 35.38% and 38.46% for the NJ tree UPGMA tree of I+M+R+T, respectively.

## Discussion

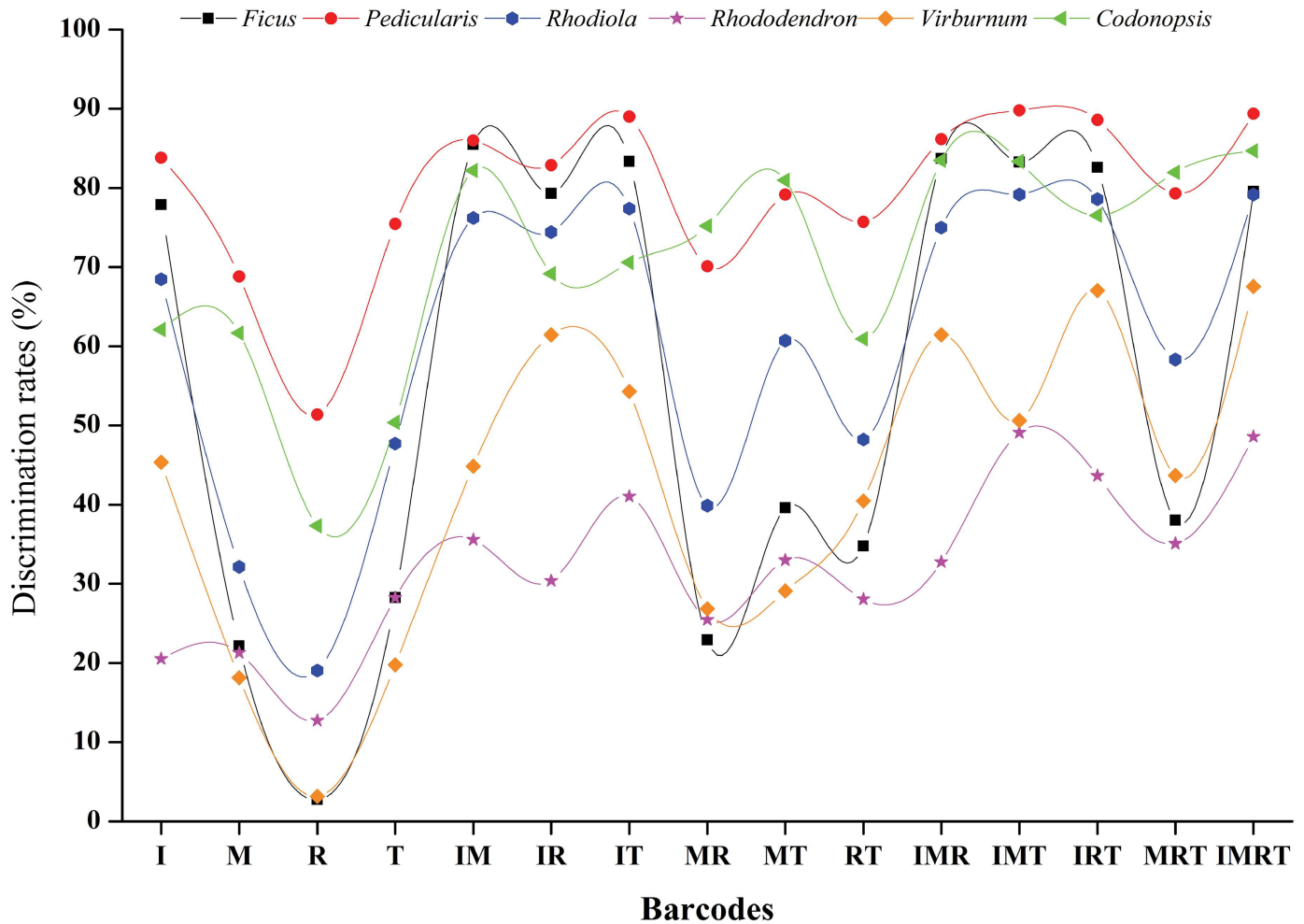
### Evaluation of barcodes for *Codonopsis* (Campanulaceae)

Ideally, DNA barcodes should meet several critical criteria: (1) having high inter-specific but low intra-specific divergence so that they can be discriminated from one another; (2) having highly conserved flanking sites for developing universal primers; (3) having appropriately short length for DNA extraction, PCR amplification and sequencing; (4) easy alignment without manual editing [3, 9, 28, 50]. Although the use of DNA barcoding for identification and taxonomy has been controversial [16], a growing number of barcodes have been proposed for plant species identification. A list of barcodes have been proposed as universal barcodes for land plants, such as *rbcL* (easy to be sequenced and aligned in plants) [3]; *matK* (one of the most rapidly evolving plastid coding regions) [3, 16]; ITS (more variable sites and greater intra- and inter-specific divergences) [23, 28, 35, 51, 52]; *trnH-psbA* (variable sites to discriminate recently evolved species) [29, 53, 54]; the *trnL* intron [55] and *ycf1* [12], etc. In our study, *rbcL* contained the lowest percentages of informative and variable sites and the lowest discrimination ability among all studied genera (Fig 1). In contrast, ITS contained the highest percentage of variable sites, had greater intra- and inter-specific divergence, the highest discrimination rates and suitable alignment lengths in our study (Fig 1). Given its superior performance, ITS is considered to be an optional core barcode for species-level barcoding *Codonopsis*.

Specimens of *Codonopsis* are challenging for molecular taxonomy because of their complicated species composition, biparental inheritance of chloroplast, widespread distribution, dynamic evolutionary history, and extensive plastid genome rearrangements during diversification [5, 56, 57]. Our results indicated that the combination of ITS+*matK* performed well with a high resolution for over 80% of species discrimination (Figs 1 and 2). Because the identification success rates of three- or four-loci combinations were lower or slightly higher than the two-locus barcodes, ITS + *matK*, we recommended ITS+*matK* as the most suitable barcodes for large genera.

### Evaluation of combined barcodes for large plant genera

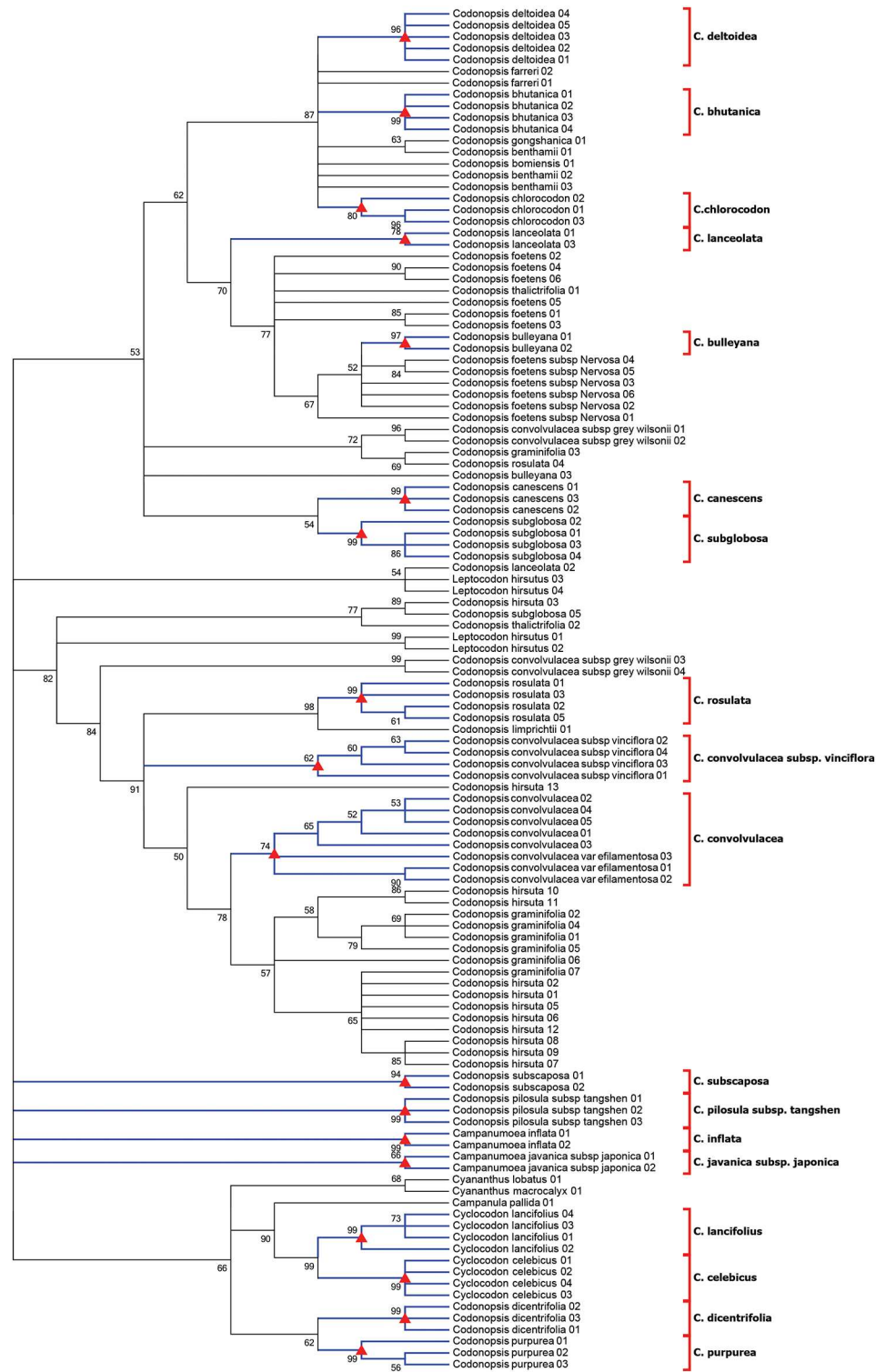
Because of the lower species discrimination rates and varied performance among different plant communities of single barcodes, the combined barcodes have been applied in recent studies [3]. Significant progress has been made in the past decades specifically, the CBOL has advocated *rbcL* + *matK* as the standard combination for combined barcodes [3]. However, several studies have demonstrated that *rbcL* + *matK* have poor identification abilities [58]. Additionally, the two-locus barcodes (ITS + *matK*, ITS + *trnH-psbA*, ITS + *rbcL*, *matK* + *rbcL*, *rbcL* + *trnH-psbA*) and three- or four-loci barcodes (ITS + *matK* + *rbcL*, ITS + *matK* + *trnH-psbA*, ITS + *matK* + *rbcL* + *trnH-psbA*) have been taken into consideration [3, 20, 22, 23, 27–29, 51, 59–62]. Xu *et al.* [21] have utilized the power of barcodes in the extraordinarily large genus *Dendrobium* based on 1,698 accessions of 184 species, and they found that the combination of ITS + *matK* performed best in *Dendrobium*, and they also verified the efficiency of ITS + *matK* in four other large genera including *Ficus*, *Lysimachia*, *Paphiopedilum*, and *Pedicularis*.



**Fig 1. Discrimination rates of DNA barcoding in six genera in the analyses of 'Best Match' in TaxonDNA.** The labels in X-axis representing all the single barcodes and their combinations used in this study. I: ITS, M: *matK*, R: *rbcl*, T: *trnH-psbA*. IM: ITS + *matK*; IR, ITS + *rbcl*; IT, ITS + *trnH-psbA*; MR, *matK* + *rbcl*; MT, *matK* + *trnH-psbA*; RT, *rbcl* + *trnH-psbA*; IMR, ITS + *matK* + *rbcl*; IMT, ITS + *matK* + *trnH-psbA*; IRT, ITS + *rbcl* + *trnH-psbA*; MRT, *matK* + *rbcl* + *trnH-psbA*; IMRT, ITS + *matK* + *rbcl* + *trnH-psbA*.

doi:10.1371/journal.pone.0170286.g001

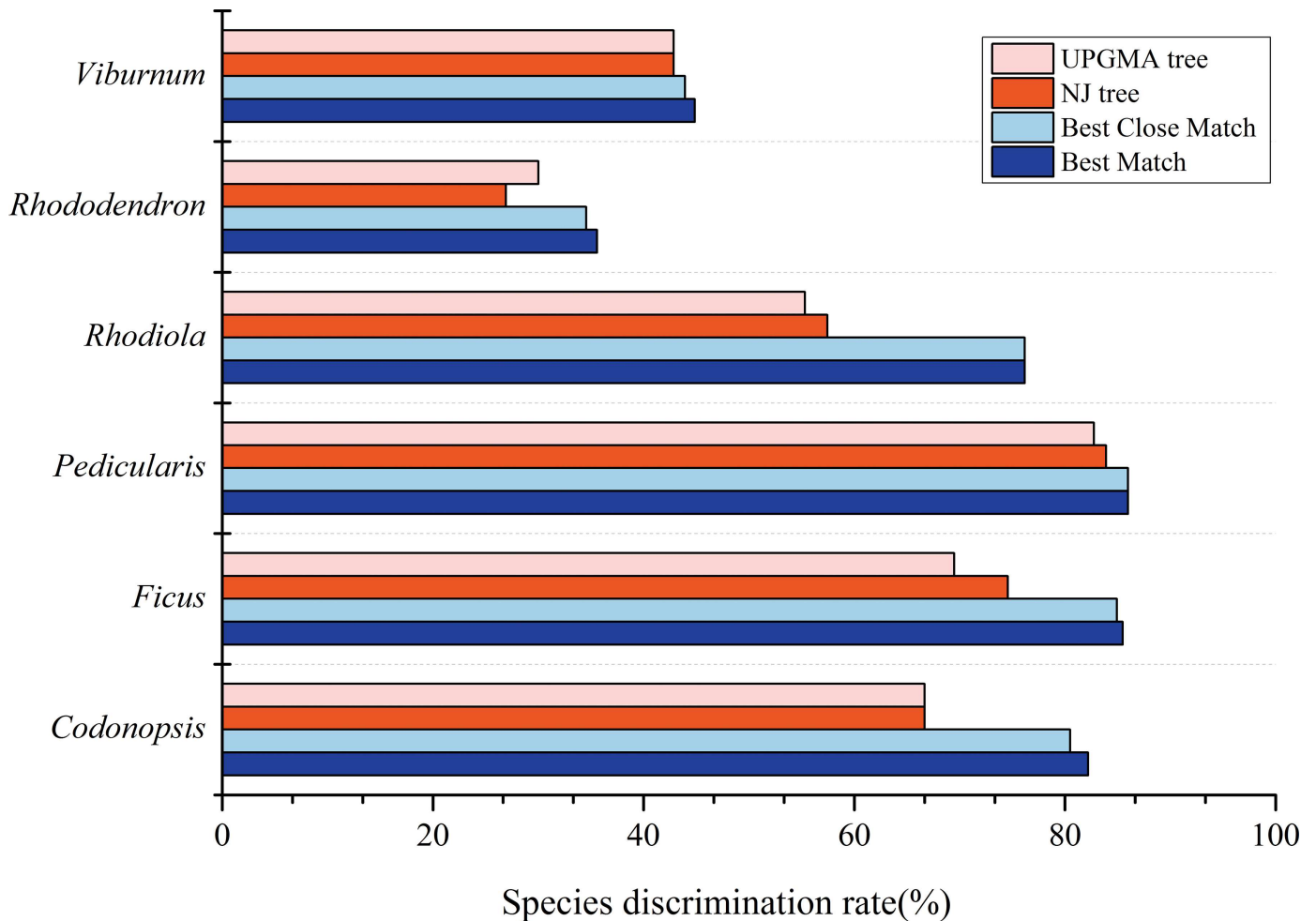
Our results indicated that *rbcl* showed the lowest sequence variation and performed poorly for species identification in the studied genera. Little to no improvement of species resolution was obtained even if *rbcl* was combined with other barcodes (Figure A-F in [S2 File](#), Figure A-Z, a-d in [S3 File](#), Figure A-Z, a-d in [S4 File](#), Figure A-Z, a-d in [S5 File](#), Figure A-Z, a-d in [S6 File](#), Figure A-Z, a-d in [S7 File](#), Figure A-Z, a-d in [S8 File](#)). Therefore, combinations with *rbcl* are not suitable for species identification in large genera. In contrast, combinations with *matK* or ITS showed significantly increased in discriminatory power (Fig 1). Specifically, ITS + *matK* performed well in almost all of the studied genera compared to the other single candidates or combinations. Thus, *matK* and ITS are recommended as core barcodes for large genera in our study. Similar results have been found in previous studies [16, 51, 60, 62]. Although other combinations, such as I+T, I+M+R, I+M+T, I+M+R+T also had high discriminatory ability in some genera, however, these barcodes performed more poorly than ITS+*matK* in some of the genera tested (Fig 1). The identification success rates of three- or four-loci combinations were lower or slightly higher than the two-locus barcodes, ITS + *matK*, which suggests



**Fig 2. The NJ tree of *Codonopsis* based on the two-locus barcodes 'ITS+matK'.** Numbers at nodes, bootstrap values with 1000 replicates (only values>50 were shown). Blue species name, resolved species. Black species name, unresolved species.

doi:10.1371/journal.pone.0170286.g002





**Fig 3. The comparison of discrimination power of ITS + *matK* in large genera using different methods.** Four methods ('Best Match', 'Best Close Match', 'NJ tree', and 'UPGMA tree') used to evaluate the discrimination power of 'ITS+*matK*' in six genera, *Codonopsis*, *Ficus*, *Pedicularis*, *Rhodiola*, *Rhododendron*, and *Viburnum*.

doi:10.1371/journal.pone.0170286.g003

that the identification success rates did not increase with an increase in the number of barcodes. Although different methods generated different results, one consistent result was that ITS+*matK* showed better overall performance with multiple evaluation methods (Fig 3, Tables J, K, L, M in S1 File).

### Conclusion

The synthetic analyses of identification ability for all barcodes in the seven species-rich genera (*Codonopsis*, *Dendrobium*, *Ficus*, *Pedicularis*, *Rhodiola*, *Rhododendron*, and *Viburnum*) agreed with previous studies that 'ITS+*matK*' may be the best core barcode combination for large genera in angiosperms [21, 51, 61, 62]. The ITS and *matK* exhibited more variable and informative sites for species identification. The combination of ITS and *matK* performs much better than other single barcode, and was almost equal to the discriminatory power of the three- or four-locus barcodes. Therefore, we propose the combined 'ITS + *matK*' as the core barcode for large plant genera.

## Supporting information

### S1 File. Supporting tables.

Table A. The sampling information of *Codonopsis* used in this study.

Table B. A list of primers used for PCR and sequence in this study.

Table C. The GenBank accession numbers of *Ficus* used in this study.

Table D. The GenBank accession numbers of *Pedicularis* used in this study.

Table E. The GenBank accession numbers of *Rhodiola* used in this study.

Table F. The GenBank accession numbers of *Rhododendron* used in this study.

Table G. The GenBank accession numbers of *Viburnum* used in this study.

Table H. Basic information of the candidate DNA markers in *Codonopsis*.

Table I. Basic information of the candidate DNA markers in *Ficus*, *Pedicularis*, *Rhodiola*, *Rhododendron*, *Viburnum*.

Table J. Summary of the pairwise intra- and inter-specific distances in *Ficus*, *Pedicularis*, *Rhodiola*, *Rhododendron*, *Viburnum*.

Table K. Wilcoxon signed-rank tests of intra- and inter-specific divergence among single barcodes.

Table L. Identification success rates generated by TaxonDNA in the six genera.

Table M. Identification success rates computed by NJ and UPGMA trees in the six genera. (RAR)

### S2 File. Barcoding gaps of all genera in this study.

(RAR)

**S3 File. 50% consensus NJ and UPGMA trees based on each barcodes for *Ficus*.** Numbers on branches represent NJ or UPGMA support values.

(RAR)

**S4 File. 50% consensus NJ and UPGMA trees based on each barcodes for *Pedicularis*.** Numbers on branches represent NJ or UPGMA support values.

(RAR)

**S5 File. 50% consensus NJ and UPGMA trees based on each barcodes for *Rhodiola*.** Numbers on branches represent NJ or UPGMA support values.

(RAR)

**S6 File. 50% consensus NJ and UPGMA trees based on each barcodes for *Rhododendron*.**

Numbers on branches represent NJ or UPGMA support values.

(RAR)

**S7 File. 50% consensus NJ and UPGMA trees based on each barcodes for *Viburnum*.** Numbers on branches represent NJ or UPGMA support values.

(RAR)

**S8 File. 50% consensus NJ and UPGMA trees based on each barcodes for *Codonopsis*.** Numbers on branches represent NJ or UPGMA support values.

(RAR)

## Author contributions

**Conceptualization:** XHJ QW.

**Data curation:** YLW DYW XGX QW.

**Formal analysis:** YLW DYW XGX.

**Funding acquisition:** XHJ.

**Investigation:** YLW QW.

**Methodology:** YLW DYW XGX.

**Project administration:** XHJ.

**Resources:** QW.

**Software:** DYW XGX.

**Supervision:** XHJ.

**Validation:** DYW.

**Visualization:** DYW.

**Writing – original draft:** DYW.

**Writing – review & editing:** XHJ QW XGX LQH.

## References

1. Hebert PDN, Cywinska A, Ball SL, DeWaard JR. Biological identifications through DNA barcodes. *P Roy Soc B-Biol Sci.* 2003; 270(1512):313–21.
2. Hebert PDN, Gregory TR. The promise of DNA barcoding for taxonomy. *Systematic biology.* 2005; 54(5):852–9. doi: [10.1080/10635150500354886](https://doi.org/10.1080/10635150500354886) PMID: [16243770](https://pubmed.ncbi.nlm.nih.gov/16243770/)
3. Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, et al. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America.* 2009; 106(31):12794–7. Epub 2009/08/12. doi: [10.1073/pnas.0905845106](https://doi.org/10.1073/pnas.0905845106) PMID: [19666622](https://pubmed.ncbi.nlm.nih.gov/19666622/)
4. Hollingsworth PM, Graham SW, Little DP. Choosing and using a plant DNA barcode. *PloS one.* 2011; 6:e19254. doi: [10.1371/journal.pone.0019254](https://doi.org/10.1371/journal.pone.0019254) PMID: [21637336](https://pubmed.ncbi.nlm.nih.gov/21637336/)
5. Van der Niet T, Pirie MD, Shuttleworth A, Johnson SD, Midgley JJ. Do pollinator distributions underlie the evolution of pollination ecotypes in the Cape shrub *Erica plukenetii*? *Annals of Botany.* 2014; 113(2):301–15. doi: [10.1093/aob/mct193](https://doi.org/10.1093/aob/mct193) PMID: [24071499](https://pubmed.ncbi.nlm.nih.gov/24071499/)
6. Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics.* 2007; 23(4):167–72. doi: [10.1016/j.tig.2007.02.001](https://doi.org/10.1016/j.tig.2007.02.001) PMID: [17316886](https://pubmed.ncbi.nlm.nih.gov/17316886/)
7. Lahaye R, Van Der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, et al. DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences.* 2008; 105(8):2923–8.
8. Kress WJ, Erickson DL, Jones FA, Swenson NG, Perez R, Sanjurjo O, et al. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proc Natl Acad Sci U S A.* 2009; 106(44):18621–6. Epub 2009/10/21. doi: [10.1073/pnas.0909820106](https://doi.org/10.1073/pnas.0909820106) PMID: [19841276](https://pubmed.ncbi.nlm.nih.gov/19841276/)
9. Valentini A, Pompanon F, Taberlet P. DNA barcoding for ecologists. *Trends Ecol Evol.* 2009; 24(2):110–7. doi: [10.1016/j.tree.2008.09.011](https://doi.org/10.1016/j.tree.2008.09.011) PMID: [19100655](https://pubmed.ncbi.nlm.nih.gov/19100655/)
10. Joly S, Davies TJ, Archambault A, Bruneau A, Derry A, Kembel SW, et al. Ecology in the age of DNA barcoding: the resource, the promise and the challenges ahead. *Mol Ecol Resour.* 2014; 14(2):221–32. doi: [10.1111/1755-0998.12173](https://doi.org/10.1111/1755-0998.12173) PMID: [24118947](https://pubmed.ncbi.nlm.nih.gov/24118947/)
11. Kress WJ, Erickson DL. DNA barcodes: genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences of the United States of America.* 2008; 105(8):2761–2. doi: [10.1073/pnas.0800476105](https://doi.org/10.1073/pnas.0800476105) PMID: [18287050](https://pubmed.ncbi.nlm.nih.gov/18287050/)
12. Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, et al. *ycf1*, the most promising plastid DNA barcode of land plants. *Scientific reports.* 2015; 5:8348. doi: [10.1038/srep08348](https://doi.org/10.1038/srep08348) PMID: [25672218](https://pubmed.ncbi.nlm.nih.gov/25672218/)
13. Meusnier I, Singer GA, Landry JF, Hickey DA, Hebert PD, Hajibabaei M. A universal DNA mini-barcode for biodiversity analysis. *BMC genomics.* 2008; 9:214. doi: [10.1186/1471-2164-9-214](https://doi.org/10.1186/1471-2164-9-214) PMID: [18474098](https://pubmed.ncbi.nlm.nih.gov/18474098/)
14. Chase MW, Fay MF. Barcoding of plants and fungi. *Science.* 2009; 325(5941):682–3. doi: [10.1126/science.1176906](https://doi.org/10.1126/science.1176906) PMID: [19644072](https://pubmed.ncbi.nlm.nih.gov/19644072/)

15. Frezal L, Leblois R. Four years of DNA barcoding: Current advances and prospects. *Infection Genetics and Evolution*. 2008; 8(5):727–36.
16. Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, et al. DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105(8):2923–8. Epub 2008/02/09. doi: [10.1073/pnas.0709936105](https://doi.org/10.1073/pnas.0709936105) PMID: [18258745](https://pubmed.ncbi.nlm.nih.gov/18258745/)
17. Li DZ, Liu JQ, Chen ZD, Wang H, Ge XJ, Zhou SL, et al. Plant DNA barcoding in China. *Journal of Systematics and Evolution*. 2011; 49(3):165–8.
18. Vassou SL, Kusuma G, Parani M. DNA barcoding for species identification from dried and powdered plant parts: a case study with authentication of the raw drug market samples of *Sida cordifolia*. *Gene*. 2015; 559(1):86–93. Epub 2015/01/18. doi: [10.1016/j.gene.2015.01.025](https://doi.org/10.1016/j.gene.2015.01.025) PMID: [25596347](https://pubmed.ncbi.nlm.nih.gov/25596347/)
19. Tang YL, Wu YS, Huang RS, Chao NX, Liu Y, Xu P, et al. Molecular identification of *Uncaria* (*Gouteng*) through DNA barcoding. *Chinese medicine*. 2016; 11:3. Epub 2016/02/05. doi: [10.1186/s13020-015-0072-7](https://doi.org/10.1186/s13020-015-0072-7) PMID: [26843891](https://pubmed.ncbi.nlm.nih.gov/26843891/)
20. Kress WJ, Erickson DL. 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(6):e508. Epub 2007/06/07. doi: [10.1371/journal.pone.0000508](https://doi.org/10.1371/journal.pone.0000508) PMID: [17551588](https://pubmed.ncbi.nlm.nih.gov/17551588/)
21. Xu SZ, Li DZ, Li JW, Xiang XG, Jin WT, Huang WC, et al. Evaluation of the DNA Barcodes in *Dendrobium* (Orchidaceae) from Mainland Asia. *PloS one*. 2015; 10(1):e0115168. doi: [10.1371/journal.pone.0115168](https://doi.org/10.1371/journal.pone.0115168) PMID: [25602282](https://pubmed.ncbi.nlm.nih.gov/25602282/)
22. Yan LJ, Liu J, Moller M, Zhang L, Zhang XM, Li DZ, et al. DNA barcoding of *Rhododendron* (Ericaceae), the largest Chinese plant genus in biodiversity hotspots of the Himalaya-Hengduan Mountains. *Molecular ecology resources*. 2015; 15(4):932–44. doi: [10.1111/1755-0998.12353](https://doi.org/10.1111/1755-0998.12353) PMID: [25469426](https://pubmed.ncbi.nlm.nih.gov/25469426/)
23. Zhang JQ, Meng SY, Wen J, Rao GY. DNA barcoding of *Rhodiola* (Crassulaceae): a case study on a group of recently diversified medicinal plants from the Qinghai-Tibetan Plateau. *PloS one*. 2015; 10(3):e0119921. doi: [10.1371/journal.pone.0119921](https://doi.org/10.1371/journal.pone.0119921) PMID: [25774915](https://pubmed.ncbi.nlm.nih.gov/25774915/)
24. Von Raab-Straube E, Raus T. Euro plus Med-Checklist Notulae, 3. *Willdenowia*. 2014; 44(2):287–99.
25. Smith SA, Donoghue MJ. Rates of molecular evolution are linked to life history in flowering plants. *Science*. 2008; 322(5898):86–9. doi: [10.1126/science.1163197](https://doi.org/10.1126/science.1163197) PMID: [18832643](https://pubmed.ncbi.nlm.nih.gov/18832643/)
26. Spooner DM. DNA barcoding will frequently fail in complicated groups: An example in wild potatoes. *American journal of botany*. 2009; 96(6):1177–89. doi: [10.3732/ajb.0800246](https://doi.org/10.3732/ajb.0800246) PMID: [21628268](https://pubmed.ncbi.nlm.nih.gov/21628268/)
27. Zhao Y, Yin J, Guo H, Zhang Y, Xiao W, Sun C, et al. The complete chloroplast genome provides insight into the evolution and polymorphism of *Panax ginseng*. *Frontiers in Plant Science*. 2015; 5.
28. Li HQ, Chen JY, Wang S, Xiong SZ. Evaluation of six candidate DNA barcoding loci in *Ficus* (Moraceae) of China. *Molecular ecology resources*. 2012; 12(5):783–90. doi: [10.1111/j.1755-0998.2012.03147.x](https://doi.org/10.1111/j.1755-0998.2012.03147.x) PMID: [22537273](https://pubmed.ncbi.nlm.nih.gov/22537273/)
29. Clement WL, Donoghue MJ. Barcoding success as a function of phylogenetic relatedness in *Viburnum*, a clade of woody angiosperms. *BMC evolutionary biology*. 2012; 12:73. <http://www.biomedcentral.com/1471-2148/12/73>. PMID: [22646220](https://pubmed.ncbi.nlm.nih.gov/22646220/)
30. Wang Q, Ma XT, Hong DY. Phylogenetic analyses reveal three new genera of the Campanulaceae. *Journal of Systematics and Evolution*. 2014; 52:541–50.
31. Cosner ME, Raubeson LA, Jansen RK. Chloroplast DNA rearrangements in Campanulaceae: phylogenetic utility of highly rearranged genomes. *BMC evolutionary biology*. 2004; 4:27. doi: [10.1186/1471-2148-4-27](https://doi.org/10.1186/1471-2148-4-27) PMID: [15324459](https://pubmed.ncbi.nlm.nih.gov/15324459/)
32. Hossen MJ, Kim MY, Kim JH, Cho JY. *Codonopsis lanceolata*: A review of its therapeutic potentials. *Phytotherapy Research*. 2016; 30:347–56. doi: [10.1002/ptr.5553](https://doi.org/10.1002/ptr.5553) PMID: [26931614](https://pubmed.ncbi.nlm.nih.gov/26931614/)
33. Knox EB. The dynamic history of plastid genomes in the Campanulaceae *sensu lato* is unique among angiosperms. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111(30):11097–102. doi: [10.1073/pnas.1403363111](https://doi.org/10.1073/pnas.1403363111) PMID: [25024223](https://pubmed.ncbi.nlm.nih.gov/25024223/)
34. Raskoti BB, Jin W-T, Xiang X-G, Schuiteman A, Li D-Z, Li J-W, et al. A phylogenetic analysis of molecular and morphological characters of *Herminium* (Orchidaceae, Orchidaceae): evolutionary relationships, taxonomy, and patterns of character evolution. *Cladistics*. 2015.
35. Yu WB, Huang PH, Ree RH, Liu ML, Li DZ, Wang H. DNA barcoding of *Pedicularis* L. (Orobanchaceae): Evaluating four universal barcode loci in a large and hemiparasitic genus. *Journal of Systematics and Evolution*. 2011; 49(5):425–37.
36. Averyanov LV, Ormerod PA, Nong Van D, Tran Van T, Chen T, Zhang D-X. *Bidoupia phongii*, new orchid genus and species (Orchidaceae, Orchidoideae, Goodyerinae) from southern Vietnam. *Phytotaxa*. 2016; 266(4):289–94.

37. Zhang CY, Wang FY, Yan HF, Hao G, Hu CM, Ge XJ. Testing DNA barcoding in closely related groups of *Lysimachia* L. (Myrsinaceae). *Molecular Ecology Resources*. 2012; 12(1):98–108. doi: [10.1111/j.1755-0998.2011.03076.x](https://doi.org/10.1111/j.1755-0998.2011.03076.x) PMID: [21967641](https://pubmed.ncbi.nlm.nih.gov/21967641/)
38. Li-Juan Jiao Y-M S. Evaluating Candidate DNA Barcodes among Chinese Begonia (Begoniaceae) Species. *Plant Diversity and Resources*. 2013; 35(6):715–24.
39. Zheng SH, Liu DW, Ren WG, Fu J, Huang LF, Chen SL. Integrated Analysis for Identifying Radix *Astragali* and Its Adulterants Based on DNA Barcoding. *Evidence-Based Complementary and Alternative Medicine*. 2014.
40. Guo YY, Huang LQ, Liu ZJ, Wang XQ. Promise and Challenge of DNA Barcoding in Venus Slipper (*Paphiopedilum*). *Plos One*. 2016; 11(1).
41. Podlech XLD. Fabaceae. In: Zhengyi Wu R PH, Hong Deyuan, editor. *Flora of China*. 25. Science Press, Beijing; Missouri Botanical Garden Press, St. Louis 2010. p. 328–453.
42. Thompson JD G T, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*. 1997; 25:4876–82. PMID: [9396791](https://pubmed.ncbi.nlm.nih.gov/9396791/)
43. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*. 1999; 41:95–8.
44. Vaidya G, Lohman D, Meier R. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*. 2011; 27(2):171–80.
45. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*. 1980; 16:111–20. PMID: [7463489](https://pubmed.ncbi.nlm.nih.gov/7463489/)
46. Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution*. 2013; 30:2725–9. doi: [10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197) PMID: [24132122](https://pubmed.ncbi.nlm.nih.gov/24132122/)
47. IBM Corp. IBM SPSS Statistics for windows, Version 19.0. Armonk, NY: IBM Corp. 2010.
48. Meier R, Shiyang K, Vaidya G, Ng PK. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic biology*. 2006; 55(5):715–28. doi: [10.1080/10635150600969864](https://doi.org/10.1080/10635150600969864) PMID: [17060194](https://pubmed.ncbi.nlm.nih.gov/17060194/)
49. Felsenstein J. Phylogenies from molecular sequences: inference and reliability. *Annu Rev Genet*. 1988; 22:521–65. doi: [10.1146/annurev.ge.22.120188.002513](https://doi.org/10.1146/annurev.ge.22.120188.002513) PMID: [3071258](https://pubmed.ncbi.nlm.nih.gov/3071258/)
50. Cristescu ME. From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. *Trends in ecology & evolution*. 2014; 29(10):566–71.
51. Liu J, Yan HF, Ge XJ. The use of DNA barcoding on recently diverged species in the genus *Gentiana* (Gentianaceae) in China. *PloS one*. 2016; 11(4):e0153008. doi: [10.1371/journal.pone.0153008](https://doi.org/10.1371/journal.pone.0153008) PMID: [27050315](https://pubmed.ncbi.nlm.nih.gov/27050315/)
52. China Plant BOLG, Li DZ, Gao LM, Li HT, Wang H, Ge XJ, et al. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108(49):19641–6. doi: [10.1073/pnas.1104551108](https://doi.org/10.1073/pnas.1104551108) PMID: [22100737](https://pubmed.ncbi.nlm.nih.gov/22100737/)
53. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(23):8369–74. doi: [10.1073/pnas.0503123102](https://doi.org/10.1073/pnas.0503123102) PMID: [15928076](https://pubmed.ncbi.nlm.nih.gov/15928076/)
54. Ojeda DI, Santos-Guerra A, Oliva-Tejera F, Jaen-Molina R, Caujape-Castells J, Marrero-Rodriguez A, et al. DNA barcodes successfully identified Macaronesian *Lotus* (Leguminosae) species within early diverged lineages of Cape Verde and mainland Africa. *AoB PLANTS*. 2014; 6.
55. Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, et al. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Res*. 2007; 35(3):e14. doi: [10.1093/nar/gkl938](https://doi.org/10.1093/nar/gkl938) PMID: [17169982](https://pubmed.ncbi.nlm.nih.gov/17169982/)
56. Morris KE, Lammers TG. Circumscription of *Codonopsis* and the allied genera *Campanumoea* and *Lep-tocodon* (Campanulaceae: Campanuloideae). I. Palynological data. *Botanical Bulletin of Academia Sinica*. 1997; 38:277–84.
57. Li CY, Xu HX, Han QB, Wu TS. Quality assessment of radix *Codonopsis* by quantitative nuclear magnetic resonance. *Journal of Chromatography A*. 2009; 1216:2124–9. doi: [10.1016/j.chroma.2008.10.080](https://doi.org/10.1016/j.chroma.2008.10.080) PMID: [19004445](https://pubmed.ncbi.nlm.nih.gov/19004445/)
58. Roy S, Tyagi A, Shukla V, Kumar A, Singh UM, Chaudhary LB, et al. Universal plant DNA barcode loci may not work in complex groups: a case study with Indian berberis species. *PloS one*. 2010; 5(10): e13674. Epub 2010/11/10. doi: [10.1371/journal.pone.0013674](https://doi.org/10.1371/journal.pone.0013674) PMID: [21060687](https://pubmed.ncbi.nlm.nih.gov/21060687/)

59. Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haider N, et al. Land plants and DNA barcodes: short-term and long-term goals. *Philos T R Soc B*. 2005; 360(1462):1889–95.
60. Ashfaq M, Asif M, Anjum ZI, Zafar Y. Evaluating the capacity of plant DNA barcodes to discriminate species of cotton (*Gossypium*: Malvaceae). *Molecular ecology resources*. 2013; 13(4):573–82. doi: [10.1111/1755-0998.12089](https://doi.org/10.1111/1755-0998.12089) PMID: [23480447](https://pubmed.ncbi.nlm.nih.gov/23480447/)
61. Alves LST, Chauveau O, Eggers L, de Souza-Chies TT. Species discrimination in *Sisyrinchium* (Iridaceae): assessment of DNA barcodes in a taxonomically challenging genus. *Molecular ecology resources*. 2014; 14(2):324–35. doi: [10.1111/1755-0998.12182](https://doi.org/10.1111/1755-0998.12182) PMID: [24119215](https://pubmed.ncbi.nlm.nih.gov/24119215/)
62. Xiang XG, Hu H, Wang W, Jin XH. DNA barcoding of the recently evolved genus *Holcoglossum* (Orchidaceae: Aeridinae): a test of DNA barcode candidates. *Molecular ecology resources*. 2011; 11(6):1012–21. doi: [10.1111/j.1755-0998.2011.03044.x](https://doi.org/10.1111/j.1755-0998.2011.03044.x) PMID: [21722327](https://pubmed.ncbi.nlm.nih.gov/21722327/)