Species identification of *Rhododendron* (Ericaceae) using the chloroplast deoxyribonucleic acid *PsbA-trnH* genetic marker

Yimei Liu, Lehua Zhang[1], Zhen Liu, Kun Luo[2], Shilin Chen[2], Keli Chen

Key Laboratory of Traditional Chinese Medicine Resource and Compound Prescription, Ministry of Education, Hubei University of Chinese Medicine, Wuhan, 430065, [1]The Chinese Academy of Sciences, Lushan Botanical Garden, Jiangxi, 332900, [2]Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing,100193, P.R. China

ABSTRACT

Background: *Rhododendron* is a group of famous landscape plants with high medicinal value. However, there is no simple or universal manner to discriminate the various species of this group. Deoxyribonucleic acid (DNA) barcoding technique is a new biological tool that can accurately and objectively identify species by using short and standard DNA regions. **Objective:** To choose a suitable DNA marker to authenticate the *Rhododendron* species. **Materials and Methods:** Four candidate DNA barcodes (*rbcL*, *matK*, *psbAtrnH*, and ITS2 intergenic spacer) were tested on 68 samples of 38 species. **Results:** The *psbAtrnH* candidate barcode yielded 86.8% sequencing efficiency. The highest interspecific divergence was provided by the *psbA-trnH* intergenic spacer, based on six parameters, and the Wilcoxon signed rank tests. Although there was not a clear barcoding gap, the Wilcoxon Two sample tests indicated that the interspecific divergence of the *psbA-trnH* intergenic spacer was significantly higher than the relevant intraspecific variation. The *psbA-trnH* DNA barcode possessed the highest species identification efficiency at 100% by the BLAST1 method. The present results showed that the *psbA-trnH* intergenic spacer was the most promising one of the four markers for barcoding the *Rhododendron* species. To further evaluate the ability of the *psbA-trnH* marker, to discriminate the closely related species, the samples were expanded to 94 samples of 53 species in the genus, and the rate of successful identification was 93.6%. The *psbA-trnH* region would be useful even for unidentified samples, as it could significantly narrow their possible taxa to a small area. **Conclusion:** The *psbA-trnH* intergenic region is a valuable DNA marker for identifying the Rhododendron species.

Key words: Deoxyribonucleic acid barcoding, *psbA-trnH*, *Rhododendron*, species identification

INTRODUCTION

Rhododendron is a very large genus in Ericaceae, with about 1000 known species in the world and more than 500 species in China.[1,2] Most species within this genus are widely cultivated in the temperate and sub-temperate regions as ornamentals.[3] Some *Rhododendron* species have been used in traditional Chinese medicine for treatment of various diseases. For example, the stems and leaves of *Rhododendron simsii* Planch and *Rhododendron anthopogonoides* Maxim have traditionally been used as folk medicines to treat chronic bronchitis.[4] Meanwhile, the fruit, flower,

and root of *Rhododendron molle* (Blum) G. Don relieve joint pains and have a remarkable therapeutic effect on rheumatoid arthritis (RA).[5] However, some species with no medicinal value are easily confused with the medicinal plants because of the similar morphological characteristics. Furthermore, in China, more than sixty species in this genus are poisonous, and some toxic medicinal plants can cause severe poisoning if they are confused with others.[4-6] Therefore, it is extremely important to accurately identify the *Rhododendron* species.

As the genus was established by Linnaeus, various methods have been constantly used to revise its classification system on the basis of morphology, cytology, chemotaxonomy, and molecular taxonomy. The contemporary classifications of *Rhododendron* are based on the seminal publication of

Address for correspondence:
Prof. Keli Chen, Hubei University of Chinese Medicine, Wuhan 430065, P.R. China. E-mail: kelichen@126.com

Sleumer.[7] Later, some researchers have conducted more morphological studies in infrageneric groupings and those studied have been integrated into the classification of Chamberlain *et al.*[1] Nowadays, this taxonomic system is generally accepted by *Rhododendron* specialists.[8] However, the *Rhododendron* genus still has some problems at various systematic classification levels and there is no simple or universal manner to discriminate the various species within the genus.

The DNA barcoding, based on a short DNA sequence to identify species has been proposed as a rapid, accurate, and convenient taxonomic tool.[9,10] The Consortium for the Barcode of Life (CBOL) Plant Working Group recommended the *rbcL* + *matK* combination as a barcode sequence in the plant kingdom, and they also suggested that ITS (ITS2) and *psbA-trnH* were good candidates for plant DNA barcoding, because of their fast evolution rates.[11] Chen *et al.* found that the ITS2 region possessed many advantages compared to the plastid loci, including the *rbcL* and *matK* regions. They also recommended that *psbA-trnH* could be a complementary barcode to ITS2 for a broad series of plants.[12] One of the problems for plant DNA barcoding was that the previous studies were mainly carried out on a large scale and rarely on a specific genus, with many closely related species, so some studies suggested that species identification using standard DNA sequences should be carried out within a narrow taxon (such as the genus).[13,14] In this study, we tried to assess the suitability of four potential DNA regions (*psbA-trnH*, *matK*, *rbcL,* and ITS2) as a DNA barcode, to identify species of *Rhododendron* across 68 samples belonging to 38 species. One of the challenges for any DNA barcode was its ability in discriminating closely related species (i.e., sister-species).[13,15] Furthermore, to evaluate the ability of the *psbA-trnH* region, the tested data were expanded to 94 samples belonging to 53 species, including 37 samples of 20 species within Subgenus *Hymenanthes* and 24 samples of 13 species within Subgenus Tsutsusi based on Chamberlain's classification system.[1]

MATERIALS AND METHODS

Plant materials
All experimental samples were collected and authenticated by Prof. Lehua Zhang (Lushan Botanical Garden, The Chinese Academy of Sciences), a specialist in taxonomy and cultivation of *Rhododendron* plants. The voucher samples were kept in the Lushan Botanical Garden and the voucher images were deposited in the herbarium of the Hubei University of Chinese Medicine. Information on the 68 samples belonging to the 38 species is given in Table 1 and the information on the expanded samples for

further evaluating the identification efficiency of *psbA-trnH* is given in Table 2. The subgenus of all the samples are based on Chamberlain's classification system.[1]

DNA extraction, amplification, and sequencing
Leaf tissues were first dried in silica gel. The total DNA was extracted as instructed by the Plant Genomic DNA Kit (Tiangen Biotech Co., China). The PCR reaction mixture consisted of 2 μL (~ 60 ng) DNA, 4 μL of 25 mM MgCl$_2$, 5 μL of 10 × PCR buffer, 2 U of Taq DNA polymerase, 4 μL of 2.5 mM dNTPs mix (Biocolor BioScience and Technology Co., China), and 2.0 μL of 2.5 μM primers (Synthesized by Sangon Co., China); and the final volume was 50 μL. Sequences of the universal primers for the tested DNA barcode, as well as general PCR reaction conditions, were obtained from previous studies.[12] The PCR products were examined with 1.5% agarose gel electrophoresis and purified using the Gel Band Purification Kit (Tiangen Biotech Co., China) and then sequenced on an ABI3730XL sequencer (Applied Biosystems, USA). The sequences were submitted to the GenBank [Tables 1 and 2].

Sequence editing and contig assembly were conducted by the CodonCode Aligner (CodonCode Co., Germany). The sequences were aligned using CLUSTALW and analyzed using the program MEGA 4.0. Average interspecific distances, theta prime, and the smallest interspecific distances were used to characterize the interspecific divergences.[12,16,17] Average intraspecific distances, theta, and coalescent depth were calculated to determine intraspecific variations, using K2P distances.[17] Wilcoxon signed rank tests were performed as described previously.[18,19] We calculated the barcoding gap using the TAXON DNA.[17,18] To estimate the reliability of species identification using DNA barcoding technique, the BLAST1 method was used.[20]

RESULTS

In order to be useful, a DNA barcode sequence must be easily PCR amplified with universal reaction conditions and primers, and then successfully sequenced. In our pilot study, we tested the efficiency of sequencing, by employing 68 samples, and the results showed that *rbcL, psbA-trnH,* and ITS2 candidate barcodes yielded 89.7, 86.8, and 50% success rates of sequencing, respectively [Table 3]. However, the efficiency of the *matK* region was very poor, hence *matK* was not included in the subsequent experiments. The sequence lengths, GC contents of the three regions based on the results of the CodonCode Aligner and Clustal W alignment are presented in Table 3.

An ideal DNA barcode should own significant interspecific

Table 1: The collection sites and GenBank accession of 68 samples of the *Rhododendron genus*

Species name	Subgenus (Chamberlain *et al.*)	Collection sites	Voucher number	GenBank accession		
				ITS2	*psbA-trnH*	*rbcL*
Rhododendron championae	*Azaleastrum*	Jiangxi	LS0508MT01		HQ706957	HQ706893
Rhododendron championae	*Azaleastrum*	Yunnan	LS0508MT02	HQ707044		HQ706894
Rhododendron latoucheae	*Azaleastrum*	Jiangxi	LS0529MT01		HQ706992	HQ706917
Rhododendron moulmainense	*Azaleastrum*	Jiangxi	LS0541MT02		HQ707004	HQ706921
Rhododendron stamineum, var. stamineum	*Azaleastrum*	Jiangxi	LS0558MT01		HQ707025	
Rhododendron stamineum var. stamineum	*Azaleastrum*	Yunnan	LS0558MT02		HQ707026	
Rhododendron vialii	*Azaleastrum*	Jiangxi	LS0562MT01		HQ707031	HQ706939
Rhododendron vialii	*Azaleastrum*	Yunnan	LS0562MT02		HQ707032	HQ706940
Rhododendron agastum.	*Hymenanthes*	Jiangxi	LS0503MT01		HQ706949	HQ706887
Rhododendron agastum	*Hymenanthes*	Yunnan	LS0503MT02	HQ707042	HQ706950	HQ706888
Rhododendron auriculatum	*Hymenanthes*	Jiangxi	LS0505MT01		HQ706955	HQ706889
Rhododendron auriculatum	*Hymenanthes*	yunnan	LS0505MT02	HQ707043	HQ706956	HQ706890
Rhododendron chihsinianum	*Hymenanthes*	Jiangxi	LS0510MT01	HQ707045		HQ706895
Rhododendron chihsinianum	*Hymenanthes*	yunnan	LS0510MT02	HQ707046		HQ706896
Rhododendron decorum	*Hymenanthes*	Jiangxi	LS0511MT01		HQ706959	HQ706897
Rhododendron delavayi.	*Hymenanthes*	Jiangxi	LS0512MT01	HQ707047		HQ706898
Rhododendron delavayi	*Hymenanthes*	yunnan	LS0512MT02		HQ706961	
Rhododendron fortunei	*Hymenanthes*	Jiangxi	LS0518MT01		HQ706969	HQ706905
Rhododendron fortunei	*Hymenanthes*	yunnan	LS0518MT02		HQ706970	HQ706906
Rhododendron glanduliferum	*Hymenanthes*	Jiangxi	LS0520MT01	HQ707050	HQ706975	HQ706907
Rhododendron glanduliferum	*Hymenanthes*	yunnan	LS0520MT02	HQ707051	HQ706976	HQ706908
Rhododendron hemsleyanum	*Hymenanthes*	Jiangxi	LS0522MT01	HQ707052		HQ706909
Rhododendron hemsleyanum	*Hymenanthes*	yunnan	LS0522MT02	HQ707053		HQ706910
Rhododendron jinggangshanicum	*Hymenanthes*	Jiangxi	LS0527MT01	HQ707057	HQ706989	HQ706913
Rhododendron jinggangshanicum	*Hymenanthes*	yunnan	LS0527MT02	HQ707058	HQ706990	HQ706914
Rhododendron leptopeplum	*Hymenanthes*	Jiangxi	LS0530MT01		HQ706993	HQ706918
Rhododendron pachyphyllum	*Hymenanthes*	Jiangxi	LS0547MT01	HQ707064	HQ707010	HQ706927
Rhododendron simiarum	*Hymenanthes*	Jiangxi	LS0556MT01	HQ707070	HQ707020	HQ706935
Rhododendron simiarum	*Hymenanthes*	yunnan	LS0556MT02	HQ707071	HQ707021	HQ706936
Rhododendron vernicosum	*Hymenanthes*	Jiangxi	LS0561MT01	HQ707073	HQ707029	HQ706937
Rhododendron vernicosum	*Hymenanthes*	yunnan	LS0561MT02	HQ707074		HQ706938
Rhododendron williamsianum	*Hymenanthes*	Jiangxi	LS0564MT01		HQ707039	HQ706941
Rhododendron williamsianum	*Hymenanthes*	yunnan	LS0564MT02			HQ706942
Rhododendron zhangjiajieense	*Hymenanthes*	Jiangxi	LS0566MT01	HQ707075	HQ707040	HQ706943
Rhododendron zhangjiajieense	*Hymenanthes*	yunnan	LS0566MT02		HQ707041	HQ706944
Rhododendron molle	*Pentanthera*	Jiangxi	LS0567MT01	HQ707062	HQ707000	HQ706945
Rhododendron molle	*Pentanthera*	yunnan	LS0567MT02	HQ707063	HQ707001	HQ706946
Rhododendron molle	*Pentanthera*	Guangdong	LS0567MT03		HQ707002	HQ706947
Rhododendron mekongense	*Rhododendron*	Jiangxi	LS0506MT01		HQ706996	HQ706891
Rhododendron mekongense	*Rhododendron*	Jiangxi	LS0506MT02		HQ706997	HQ706892
Rhododendron edgeworthii	*Rhododendron*	Jiangxi	LS0515MT01		HQ706962	HQ706899
Rhododendron edgeworthii	*Rhododendron*	yunnan	LS0515MT02		HQ706963	HQ706900
Rhododendron excellens	*Rhododendron*	Jiangxi	LS0516MT01	HQ707048	HQ706964	HQ706901
Rhododendron excellens	*Rhododendron*	yunnan	LS0516MT02	HQ707049	HQ706965	HQ706902
Rhododendron kiangsiense	*Rhododendron*	Jiangxi	LS0528MT01	HQ707059	HQ706991	HQ706915
Rhododendron kiangsiense	*Rhododendron*	yunnan	LS0528MT02	HQ707060		HQ706916
Rhododendron micranthum	*Rhododendron*	Jiangxi	LS0536MT01		HQ706998	HQ706919
Rhododendron micranthum	*Rhododendron*	yunnan	LS0536MT02	HQ707061	HQ706999	HQ706920
Rhododendron rubiginosum	*Rhododendron*	Jiangxi	LS0554MT01		HQ707017	HQ706934
Rhododendron farrerae	*Tsutsusi*	Jiangxi	LS0517MT01		HQ706967	HQ706903
Rhododendron farrerae	*Tsutsusi*	yunnan	LS0517MT02		HQ706968	HQ706904
Rhododendron huanense	*Tsutsusi*	Jiangxi	LS0523MT01	HQ707054	HQ706978	
Rhododendron huanense	*Tsutsusi*	yunnan	LS0523MT02	HQ707055	HQ706979	
Rhododendron hypoblematosum	*Tsutsusi*	Jiangxi	LS0524MT01	HQ707056	HQ706980	HQ706911

Table 1: Contd...

Species name	Subgenus (Chamberlain *et al.*)	Collection sites	Voucher number	GenBank accession		
				ITS2	*psbA-trnH*	*rbcL*
Rhododendron hypoblematosum	*Tsutsusi*	yunnan	LS0524MT02		HQ706981	HQ706912
Rhododendron indicum	*Tsutsusi*	Jiangxi	LS0525MT01		HQ706982	
Rhododendron indicum	*Tsutsusi*	yunnan	LS0525MT02		HQ706983	
Rhododendron mucronatum	*Tsutsusi*	Jiangxi	LS0542MT02		HQ707005	HQ706922
Rhododendron obtusum	*Tsutsusi*	Jiangxi	LS0544MT01		HQ707006	HQ706923
Rhododendron obtusum	*Tsutsusi*	yunnan	LS0544MT02		HQ707007	HQ706924
Rhododendron oldhamii	*Tsutsusi*	Jiangxi	LS0545MT01		HQ707008	HQ706925
Rhododendron oldhamii	*Tsutsusi*	yunnan	LS0545MT02		HQ707009	HQ706926
Rhododendron strigosum	*Tsutsusi*	Jiangxi	LS0548MT02	HQ707072	HQ707027	HQ706928
Rhododendron pulchrum	*Tsutsusi*	Jiangxi	LS0549MT01	HQ707065	HQ707011	HQ706929
Rhododendron pulchrum	*Tsutsusi*	yunnan	LS0549MT02	HQ707066	HQ707012	HQ706930
Rhododendron rhuyuenense	*Tsutsusi*	Jiangxi	LS0551MT01	HQ707067	HQ707014	HQ706931
Rhododendron rhuyuenense	*Tsutsusi*	Yunnan	LS0551MT02	HQ707068	HQ707015	HQ706932

Table 2: The collection sites and GenBank accession of expanded samples of the *Rhododendron genus*

Samples name	Subgenus (Chamberlain *et al.*)	Sampling location	Voucher number	*psbA-trnH*
Rhododendron aganniphum	*Hymenanthes*	Xizang	LS0502MT01	HQ706948
Rhododendron annae	*Hymenanthes*	Jiangxi	LS0504MT01	HQ706951
Rhododendron annae	*Hymenanthes*	Yunnan	LS0504MT02	HQ706952
Rhododendron annae	*Hymenanthes*	Guangdong	LS0504MT03	HQ706953
Rhododendron anthosphaerum	*Hymenanthes*	Guangdong	LS0593MT01	HQ706954
Rhododendron delavayi	*Hymenanthes*	Guangdong	LS0512MT01	HQ706960
Rhododendron fortunei	*Hymenanthes*	Guangdong	LS0518MT03	HQ706971
Rhododendron fortunei	*Hymenanthes*	Guangdong	LS0518MT04	HQ706972
Rhododendron habrotrichum	*Hymenanthes*	Jiangxi	LS0597MT01	HQ706977
Rhododendron irroratum	*Hymenanthes*	Jiangxi	LS0526MT01	HQ706986
Rhododendron irroratum	*Hymenanthes*	Yunnan	LS0526MT02	HQ706987
Rhododendron irroratum	*Hymenanthes*	Guangdong	LS0526MT03	HQ706988
Rhododendron aberconwayi	*Hymenanthes*	Jiangxi	LS0514MT01	HQ707023
Rhododendron aberconwayi	*Hymenanthes*	Jiangxi	LS0514MT02	HQ707024
Rhododendron vernicosum .	*Hymenanthes*	yunnan	LS0561MT02	HQ707030
Rhododendron wardii	*Hymenanthes*	Xizang	LS0531MT01	HQ707037
Rhododendron wardii	*Hymenanthes*	Xizang	LS0531MT02	HQ707038
Rhododendron molle	*Pentanthera*	Guangdong	LS0567MT04	HQ707003
Rhododendron schlippenbachii	*Pentanthera*	Jiangxi	LS0586MT01	HQ707019
Rhododendron ciliatum	*Rhododendron*	Jiangxi	LS0507MT01	HQ706958
Rhododendron excellens	*Rhododendron*	Guangdong	LS0516MT03	HQ706966
Rhododendron liliiflorum	*Rhododendron*	yunnan	LS0588MT01	HQ706994
Rhododendron sargentianum	*Rhododendron*	Jiangxi	LS0513MT01	HQ707018
Rhododendron taronense	*Rhododendron*	Jiangxi	LS0519MT01	HQ707028
Rhododendron virgatum	*Rhododendron*	Jiangxi	LS0563MT01	HQ707033
Rhododendron virgatum	*Rhododendron*	yunnan	LS0563MT02	HQ707034
Rhododendron mariesii	*Tsususi*	Jiangxi	LS0535MT01	HQ706995
Rhododendron pulchrum	*Tsususi*	yunnan	LS0549MT03	HQ707013
Rhododendron indicum	*Tsutsusi*	Jiangxi	LS0525MT03	HQ706984
Rhododendron indicum	*Tsutsusi*	yunnan	LS0525MT04	HQ706985
Rhododendron simsii	*Tsutsusi*	Jiangxi	LS0583MT01	HQ707022
Rhododendron fragrans		yunnan	LS0509MT01	HQ706973
Rhododendron fragrans		yunnan	LS0509MT02	HQ706974
Rhododendron wanxia		yunnan	LS0521MT01	HQ707035
Rhododendron wanxia		yunnan	LS0521MT02	HQ707036

variation in DNA sequences, with a comparatively small variation between individuals, within a single species.[17,21,22] Therefore, six metrics were employed to characterize interspecific versus intraspecific variation. Through comparison of interspecific genetic distances among congeneric species for three candidate barcodes, the chloroplast non-coding region *psbA-trnH* exhibited the highest interspecific divergence with all three metrics, followed by ITS2, while *rbcL* provided the lowest [Table 4]. Moreover, the Wilcoxon signed rank tests confirmed that *psbA-trnH* provided the highest interspecific divergence between the congeneric species [Table 5]. We also found that *rbcL* showed the lowest level of intraspecific variation with all three parameters, followed by ITS2, while *psbA-trnH* provided the highest [Table 4]. The Wilcoxon signed rank tests showed that *rbcL* has the lowest variation between conspecific individuals, whereas, *psbA-trnH* showed the highest [Table 6].

The DNA barcode should exhibit a 'barcoding gap' between interspecific and intraspecific divergences.[17,18] Although there was no clear gap in the histogram between intraspecific variation and interspecific divergence in the distributions of the three loci (*rbcL*, TS2, *psbA-trnH* intergenic spacer) [Figure 1], the Wilcoxon two sample tests indicated that for the three loci the distribution of interspecifc divergences were higher than those of the corresponding intraspecific variations, with high significance [Table 7].

The BLAST1 method was used to test the applicability of different regions, for species identification.[20] The results indicated that the *psbA-trnH* intergenic spacer possessed the highest species identification efficiency at 100%, followed by *rbcL* at 59%, then ITS2 at 41.2% [Table 8]. To further evaluate the ability of the *psbA-trnH* region to identify the *Rhododendron* species with more closely related species in a wider range, 94 samples were tested. The rate of correct identification was 93.6% [Table 8], with six failed samples [Table 9].

DISCUSSION

In the present research, the feasibility of four potential DNA regions (*psbA-trnH*, *matK*, *rbcL*, ITS2) as a DNA barcode of the *Rhododendron* species was concretely tested. The *rbcL* sequence showed advantages of higher efficiency of PCR amplification and sequencing [Table 3]. However, the variation of the sequence in the species level was insufficient to discriminate the *Rhododendron* species, and the identification efficiency was only 59% [Table 8]. The *matK* showed lower sequencing efficiency and its successful identification rate of 131 samples from the GenBank database was 43.8%. At the Third International

Table 3: Success rate of sequencing, Length range, GC content

Markers	*psbA-trnH*	*rbcL*	ITS2
Number of samples / n	68	68	68
Success of Sequencing / n	59	61	34
Success rate of sequencing / %	86.8	89.7	50
Length range / bp	450 – 493	666	245 – 249
GC content / %	32.1	44	59.3

Table 4: Inter- and intraspecific genetic divergences of three candidate barcodes

Markers	ITS2 (34)	*psbA-trnH* (59)	*rbcL* (61)
All interspecific distance	0.0312 ± 0.0165	0.0728 ± 0.0237	0.0103 ± 0.0112
Theta prime	0.0293 ± 0.0109	0.0729 ± 0.0119	0.0101 ± 0.0103
Minimum interspecific distance	0.0079 ± 0.0123	0.0230 ± 0.0181	0.0045 ± 0.0110
All intraspecific distance	0.0043 ± 0.0136	0.0150 ± 0.0100	0.0002 ± 0.0007
Theta	0.0043 ± 0.0136	0.0151 ± 0.0105	0.0002 ± 0.0007
Coalescent depth	0.0043 ± 0.0136	0.0153 ± 0.0105	0.0002 ± 0.0007

Numbers in parentheses mean the whole sequenced samples of three candidate barcodes [Table 3].

Table 5: Wilcoxon signed rank test for interspecific variations

W+	W-	Inter relative ranks	Result
rbcL	ITS2	W+ = 4665.0, W- = 30846 (n = 266, *P* = 2.6958 × 10⁻¹¹)	ITS2 >> *rbcL*
psbA-trnH	ITS2	W+ = 35833, W- = 482.0 (n = 269, *P* = 3.1078 × 10⁻¹⁹)	*psbA-trnH* >> ITS2
rbcL	*psbA-trnH*	W+ = 160.0, W- = 36155 (n =269, *P* = 1.2216 × 10⁻¹⁹)	*psbA-trnH* >> *rbcL*

Table 6: Wilcoxon signed rank test for intraspecific variations

W+	W-	Intra relative ranks	Result
ITS2	*rbcL*	W+ = 1, W- = 0 (n = 1, *P* = 0.3171)	ITS2 = *rbcL*
psbA-trnH	ITS2	W+ = 21.0, W- = 7.0 (n = 7, *P* = 0.2367)	ITS2 = *psbA-trnH*
rbcL	psbA-trnH	W+ = 0, W- = 28 (n = 7, *P* = 0.01796)	*psbA-trnH* > *rbcL*

Barcoding Conference, the Plant Working Group of the Consortium for the Barcode of Life recommended the two-locus combination of *rbcL + matK,* for plant barcoding.[11] The two proposed regions were the most

Figure 1: The barcoding gap between interspecific and intraspecific divergences for three candidate barcodes. Histograms showing the relative distribution of pairwise (y-axes) intraspecific (blue bar) and interspecific (red bar) divergence distance estimates (x-axes) for *psbA-trnH*, ITS2, and *rbcL* intergenic spacers, respectively. The divergences were calculated using the Kimura 2-Parameter (K2P) model. Barcoding gaps were assessed by the Wilcoxon two-sample tests, and all were highly significant ($P < 0.0001$)

useful barcodes and provided a universal framework for land plants at and above the generic levels.[13] However, they showed a lower resolution rate to identify the species within a rapid evolutionary genus such as *Rhododendron*. In the meantime, many researchers have proposed the use of ITS2 as a suitable marker for taxonomic classification.[12,23,24] However, in our study, the success rate of sequencing with ITS2 was only 50% [Table 3], and the identification efficiency was only 41.2% [Table 8]. Above all, the results indicated that *matK, rbcL,* and ITS2 were not suitable as barcodes for the identification of the *Rhododendron* species.

The *psbAtrnH* intergenic spacer is among the most variable regions in the angiosperm chloroplast genome. It is a popular tool for plant population genetics and species level phylogenetics and has been proposed to be suitable for the DNA barcoding studies.[25,26] *Rhododendron* is a rapidly evolutionary genus within the angiosperms in
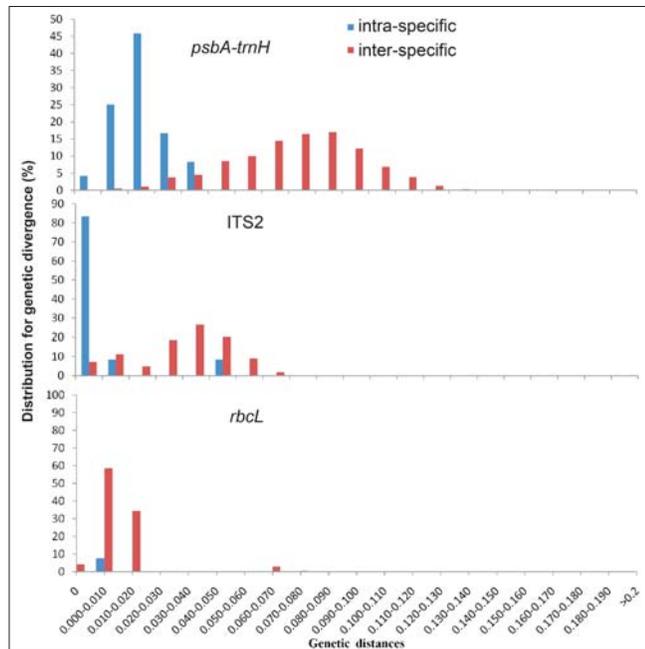
Table 7: Wicoxon two-sample tests for distribution of intra- versus interspecific divergences

Marker	No. of Interspecific distances	No. of Intraspecific distances	Wilcoxon W	P value
ITS2	549	12	795	8.5172×10^{-3}
psbA-trnH	1687	24	910	3.4910×10^{-7}
rbcL	1803	27	30	6.5088×10^{-8}

Table 8: Identification efficiency for potential deoxyribonucleic acid barcodes loci using the BLAST1 method

Marker	Number of samples (n)	Correct identification samples (n)	Correct identification efficiency (%)	Ambiguous identification samples (n)	Ambiguous identification efficiency (%)
ITS2	34	14	41.2	20	58.8
psbA-trnH	59	59	100	0	0
rbcL	61	36	59.0	25	41.0
psbA-trnH	94	88	93.6	6	6.4

Table 9: Unsuccessful identification sample pairs in BLAST1 based on Chamberlain's classification system

Pairs	Samples name	Voucher number	Subgenus	Section	Subsection
1	*Rhododendron irroratum*	LS0526MT01	*Hymenanthes*	*Ponticum*	*Irrorata*
	Rhododendron anna	LS0504MT01	*Hymenanthes*	*Ponticum*	*Irrorata*
2	*Rhododendron irroratum*	LS0526MT02	*Hymenanthes*	*Ponticum*	*Irrorata*
	Rhododendron aberconwayi	LS0514MT02	*Hymenanthes*	*Ponticum*	*Irrorata*
3	*Rhododendron excellens*	LS0516MT01	*Rhododendron*	*Rhododendron*	*Maddenia*
	Rhododendron virgatum	LS0563MT01	*Rhododendron*	*Rhododendron*	*Virgata*

recent years with many closely related species and there are many artificial and natural hybrids.[27] The *psbA-trnH* region is one of the most variable non-coding regions of the plastid genome in the angiosperms, because of the highest percentages of variable sites.[28-30]

Moreover this variation indicated that this inter-genic spacer could offer high levels of species discrimination.[27,28] In our study, first we found that the average length of the *psbA-trnH* intergenic spacer was rather short at 450 – 493 base pairs. The *psbA-trnH* sequences were relatively easy to be amplified using one pair of universal primers. Second, examination of the genetic divergences using six parameters and statistical tests confirmed that the *psbA-trnH* intergenic spacer possessed high interspecific divergence. Analyses of the DNA barcoding gap and the Wilcoxon twosample tests supported the notion that the mean interspecific divergence of the psbAtrnH intergenic spacer was significantly higher than its mean intraspecific variation. Third, according to the BLAST1 method, the identification accuracy using the *psbAtrnH* intergenic spacer was 100%, and it could identify all the species that could be identified by ITS2 or *rbcL*. Therefore, it was quite clear that among the four sequences, *psbA-trnH* was the most promising one for barcoding the species within the rapid evolutionary genus.

One of the challenges for any DNA barcode is its utility in discriminating closely related species.[13,15] In this study, to further evaluate the ability of the *psbA-trnH* region, to identify the closely related species in a wider range, the samples were expanded to 94 samples belonging to 53 species. The result showed that the *psbA-trnH* region steadily kept a higher identification efficiency. Furthermore, we specifically tested the identification ability of *psbA-trnH* in two subgenera, and it showed that the success rate of identification was 100% for 24 samples of 13 species from Subgenus Tsutsusi and 89.2% for 37 samples of 20 species from Subgenus *Hymenanthes*. Therefore, *psbA-trnH* was confirmed as a useful marker for differentiating closely related species within *Rhododendron*.

Meanwhile, we noted that there were three pairs of samples which could not be accurately identified [Table 9]. The first and second pairs from three species, *Rhododendron annae, Rhododendron irroratum,* and *Rhododendron aberconwayi,* belonged to the same subsection *Irrorata,* as they shared exceedingly similar morphological characteristics of the corolla shape, leaves shape, glabrous petiole, and pedicel.[2] The third pair, *R. excellens* and *R. virgatum,* was classified by Chamberlain as the same subgenus *Rhododendron* and the same section *Rhododendron,* because of their similar morphological characters: Both of them have termina

inflorescence buds, vary rarely axillary from lower leaves and the whole plant of them is densely covered with peltate scales.[1,2] The failure of *psbA-trnH* in discriminating these species indicated that some morphologically similar species had no sufficient interspecific variation in the *psbA-trnH* region. In spite of this, the *psbA-trnH* region would still be significant for those unidentified samples as it could narrow their possible taxa to a small area, one subgenus, one section, or even to one subsection [Table 9].

CONCLUSION

The *psbA-trnH* intergenic region is a potential DNA barcoding sequence for identifying the *Rhododendron* species. Furthermore, it would still be useful, even for those unidentified species, because it could significantly narrow the possible taxa to a small area.

ACKNOWLEDGEMENTS

REFERENCES

1. Chamberlain DF, Hyam R, Argent G, Fairweather G, Walter KS. The genus Rhododendron: Its Classification and Synonymy. Edinburgh: Royal Botanic Gardens; 1996.

2. Fang MY, Fang RC, He MY, Hu LZ, Yang HB. Rhododendron Linnaeus. Flora China 2005;14:260-455.

3. Craven LA, Goetsch LA, Hall BD, Brown GK. Classification of the Vireya group of Rhododendron (Ericaceae). Blumea 2008;53:435-42.

4. Wang MJ. Study on Antitussive, Antiasthmatic Effects and Toxicity of Total Flavanoid of *Rhododendron simsii* Planch. Lishizhen Med Mater Med Res 2006;17:1678-9.

5. Xiang YN, Zhang CG, Zheng YJ. Studies on the Chemical Constituents of the roots of *Rhododendron molle* G. DON. J Huazhong Univ Sci Technol 2004;24:202-4.

6. Zhong GH, Hu MY. Advances in Studies on Activity Constituents of Ericaceae and Their Mechanisms. J Wuhan Bot Res 2000;18:509-14.

7. Sleumer H. Ein system der gattung *Rhododendron* L. Bot Jahrb Syst 1949;74:511-53.

8. Cox PA, Cox KN. Encyclopedia of Rhododendron species. Perth: Glendoick Publishing; 1997.

9. Hebert PD, Cywinska A, Ball SL, de Waard JR. Biological identifications through DNA barcodes. Proc Biol Sci 2003;270:313-21.

10. Han JP, Song JY, Liu C, Chen J, Qian J, Zhu YJ, *et al*. Identification of Cistanche species (Orobanchaceae) based on sequences of the plastid psbA-trnH intergenic region. Acta Pharmaceutica Sinica 2010;45:126-30.

11. CBOL Plant Working Group. A DNA barcode for land plants. Proc Natl Acad Sci USA 2009;106:12794-7.

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

13. Ren BQ, Chen ZD. DNA Barcoding Plant Life. Chin Bull Bot 2010;45:1-12.

14. Nitta JH. Exploring the utility of three plastid loci for biocoding the filmy ferns (Hymenophyllaceae) of Moorea. Taxon 2008;57: 725-36.

15. Newmaster SG, Fazekas AJ, Steeves RA, Janovec J. Testing candidate plant barcode regions in the Myristicaceae. Mol Ecol Resour 2008;8:480-90.

16. Meier R, Zhang GY, 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;57:809-13.

17. Meyer CP, Paulay G. DNA barcoding: Error rates based on comprehensive sampling. PLoS Biol 2005;3:2229-38.

18. Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G. DNA barcoding the floras of biodiversity hotspots. Proc Natl Acad Sci USA 2008;105:2923-8.

19. 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:e508.

20. Ross HA, Murugan S, Li WL. Testing the reliability of genetic methods of species identification via simulation. Syst Biol 2008;57:216-30.

21. Moritz C, Cicero C. DNA barcoding: Promises and pitfalls. PLoS Biol 2004;2:1529-34.

22. Slabbinck B, Dawyndt P, Martens M, De Vos P, De Baets B. TaxonGap: A visualization tool for intra- and inter-species variation among individual biomarkers. Bioinformatics 2008;24:866-7.

23. Miao M, Warren A, Song WB, Wang S, Shang HM, Chen ZG. Analysis of the internal transcribed spacer 2 (ITS2) region of scuticociliates and related taxa (Ciliophora, Oligohymenophorea) to infer their evolution and phylogeny. Protist 2008;159:519-33.

24. Schultz J, Maisel S, Gerlach D, Müller T, Wolf M. A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. RNA 2005;11:361-4.

25. Storchová H, Olson MS. The architecture of the chloroplast *psbA-trnH* non-coding region in angiosperms. Plant Syst Evol 2007;268:235-56.

26. Ma YP, Zhang CQ, Zhang JL, Yang JB. Natural Hybridization between Rhododendron delavayi and R. cyanocarpum (Ericaceae), from Morphological, Molecular and Reproductive Evidence. J Integr Plant Biol 2010;52:844-51.

27. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci USA 2005;102:8369-74.

28. Shaw J, Lickey EB, Schilling EE, Small RL. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms. The tortoise and the hare III. Am J Bot 2007;94:275-88.

29. Song JY, Yao H, Li Y, Li XW, Liu C, Han JP, *et al*. Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique. J Ethnopharmacol 2009;124:434-9.

30. Yao H, Song JY, Ma XY, Liu C, Li Y, Xu HX, *et al*. Identification of Dendrobium species by a candidate DNA barcode sequence: The chloroplast *psbA-trnH* intergenic region. Planta Med 2009;75:667-9.

Author Help: Online submission of the manuscripts

Articles can be submitted online from http://www.journalonweb.com. For online submission, the articles should be prepared in two files (first page file and article file). Images should be submitted separately.

1) **First Page File:**
Prepare the title page, covering letter, acknowledgement etc. using a word processor program. All information related to your identity should be included here. Use text/rtf/doc/pdf files. Do not zip the files.

2) **Article File:**
The main text of the article, beginning with the Abstract to References (including tables) should be in this file. Do not include any information (such as acknowledgement, your names in page headers etc.) in this file. Use text/rtf/doc/pdf files. Do not zip the files. Limit the file size to 1 MB. Do not incorporate images in the file. If file size is large, graphs can be submitted separately as images, without their being incorporated in the article file. This will reduce the size of the file.

3) **Images:**
Submit good quality color images. Each image should be less than **4096 kb (4 MB)** in size. The size of the image can be reduced by decreasing the actual height and width of the images (keep up to about 6 inches and up to about 1200 pixels) or by reducing the quality of image. JPEG is the most suitable file format. The image quality should be good enough to judge the scientific value of the image. For the purpose of printing, always retain a good quality, high resolution image. This high resolution image should be sent to the editorial office at the time of sending a revised article.

4) **Legends:**
Legends for the figures/images should be included at the end of the article file.