Identification of medical plants of 24 *Ardisia* species from China using the *matK* genetic marker

Yimei Liu, Ke Wang, Zhen Liu, Kun Luo[1], Shilin Chen[1], Keli Chen

Key Laboratory of Traditional Chinese Medicine Resource and Compound Prescription, Ministry of Education, Hubei University of Chinese Medicine, Wuhan, [1]Institute of Medicinal Plant Development Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, P.R. China

Submitted: 16-09-2012 Revised: 04-11-2012 Published: 07-09-2013

ABSTRACT

Background: *Ardisia* is a group of famous herbs in China, which has been used as medical plants for more than 900 years. However, the species from the genus are so analogous that it is difficult to discriminate them just by morphological characteristics. DNA barcoding is a new technique that uses a short and standard fragment of DNA sequences to identify species. **Objective:** Choose a suitable DNA marker to authenticate *Ardisia* species. **Materials and Methods:** Four markers (*psbA-trnH*, internal transcribed spacer 2 [ITS2], *rbcL, matK*) were tested on 54 samples of 24 species from genus *Ardisia*. The success rates of polymerase chain reaction amplification and sequencing, differential intra- and inter-specific divergences, DNA barcoding gap and identification efficiency were used to evaluate the discrimination ability. **Results**: The results indicate that *matK* has the highest interspecific divergence and significant differences between inter- and intra-specific divergences, whereas *psbA-trnH*, ITS2 and *rbcL* have much lower divergence values. *Matk* possessed the highest species identification efficiency at 98.1% by basic local alignment search tool 1 [BLAST1], method and 91.7% by the nearest distance method. **Conclusion:** The *matK* region is a promising DNA barcode for the genus *Ardisia*.

Key words: *Ardisia*, DNA barcoding, identification, *matK*

Access this article online

Website: www.phcog.com

DOI: 10.4103/0973-1296.117829

Quick Response Code:

INTRODUCTION

Ardisia genus is a group of flowering plants belonging to *Myrsinaceae* family, native to tropical American, Austronesia, India Peninsula, East and South Asian, minority spread over Oceania. The genus includes about 300 species in the world and 68 species in China, which is widely and commonly cultivated in south area of Yangtze River.[1] Most species of *Ardisia* are medicinal plants and a few of them are ornamental plants in China. Some of them are famous on medicinal value. For example, *Ardisia japonica* (Hornst.) Blume is commonly used for treating chronic bronchitis; *Ardisia crenata* Sims *var. crenata* is used as oxytocics and anti-pregnancy drugs. *Ardisia pusilla* A. de Candolle. is used to treat traumatic injuries.[2] But the species from the genus are so analogous that it is very difficult to discriminate them just by morphological characteristics and it is often taken place that many species of the genus are confused and used by other different species. So, it is very important to accurately identify these medical plants from *Ardisia*.

DNA barcoding, which was first proposed by Hebert *et al.*[3] is a new technique that uses a short and standardized fragment of DNA sequences to identify species, and recently it has become a hotspot of biodiversity research.[4] In subsequent research,[5-7] Hebert *et al.* found that the CO1 gene is a standard DNA barcode for animals. But the studies on plant barcodes are much more complicated than that of animals, because of the hybridization and reticulate evolutionary histories.[8,9] Recently, a number of single loci and combined loci have been suggested as candidate barcode sequences for plant identification,[10-12] but there was no consensus on universal DNA barcode for all plant species. For every concrete group of species, especially those which contain many closely related species, applicable loci have to be studied and choose. Some scholars have done DNA barcoding researches in related species and genera, but no one has evaluated feasibility of the method in plants of *Ardisia*.

In this context, we choose four regions intensively recommended (*psbA-trnH*, *matK*, *rbcL*, internal transcribed

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

spacer 2 [ITS2]) to test and evaluate the feasibility of these regions as candidate DNA barcodes to discriminate medicinal species in China from *Ardisia* and try to find a new a digital identification method for medicinal plants of *Ardisia*.

MATERIALS AND METHODS

Plant materials

The experimental samples were collected from (1) South China Botanical Garden, Guangdong Research Institute of Traditional Chinese Medicine, Guangdong province, and authenticated by Prof. Yuewen Cai of the Institute; (2) Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Yunnan province, and authenticated by senior Engineer Chunfen Xiao of the Garden; (3) Wuhan Botanical Garden, Chinese Academy of Sciences, Hubei province, and authenticated by Engineer Shouzhong Zhang of the Garden. All voucher images and specimens were deposited in the herbarium of Hubei University of Chinese Medicine. The information of 54 samples belonged to 27 species are given in Table 1.

DNA extraction, amplification, and sequencing

First, leaf tissues were dried in silica gel. A total of 10 mg of each of the dried tissues was rubbed for 1 min at a frequency of 30 times/s in a FastPrep bead mill (Retsch MM400, Germany). Total DNA was extracted using the Plant Genomic DNA Kit (Tiangen Biotech Co., China). The polymerase chain reaction (PCR) reaction mixture consisted of 2 μL (~60 ng) DNA, 4 μL of 25 mM $MgCl_2$, 5 μL of 10 × PCR buffer, 2U of Taq DNA polymerase, 4 μL of 2.5 mM deoxy-ribonucleoside triphosphates [dNTPs] mix (Biocolor BioScience & Technology Co., China), 2.0 μL 2.5 μM of primers (Synthesized by Sangon Co., China), the final volume was 50 μL. The sequences of the universal primers for the DNA barcode to be tested and general PCR reaction conditions were obtained from previous studies by Chen *et al.*[13] PCR products were first examined with 1.5% agarose gel electrophoresis and purified using the Gel Band Purification Kit (Tiangen Biotech Co., China) and then sequenced in both directions with the primers used for PCR amplification on a 3730XL sequencer (Applied Biosystems, USA). The sequences were submitted to GenBank [Table 1].

Data analyses

The original forward and reverse sequences were assembled and edited using CodonCode Aligner 3.0 (CodonCode Co., USA) to estimate the quality of the generated sequence traces. Sequences alignment and checking were conducted by Clustal W. The ITS2 sequences were retrieved according to Keller *et al.*[14] and other sequences were retrieved using CodonCode Aligner. All the experimental materials were used to investigate the amplification efficiency of each sequence. The inter/intra-specific variation of the samples was calculated according to Luo *et al.*[15] and Zhu *et al.*[16] and Wilcoxon signed rank tests[17] were used to check the result. DNA barcoding gap was produced using Taxon DNA.[18] After the data from GenBank database were brought into, basic local alignment search tool 1 [BLAST1] and the nearest distance method were performed as described previously[19] to assess the identification efficiency of each candidate sequence.

RESULTS

PCR amplification efficiency and the success rate of sequencing

The efficiency of PCR amplification and the success rate of sequencing of the four candidates were compared. The result showed that the efficiency of PCR amplication of *rbcL*, *psbA-trnH*, ITS2 and *matK* region were 100%, 100%, 100% and 88.9%. And they were all successfully sequenced by 100% [Table 2]. The sequence length, Guanine and Cytosine [GC] content of the four regions based on the results of the CodonCode Aligner and Clustal W alignment were presented [Table 2].

The analysis of intra-specific variations and inter-specific divergences

An ideal barcode should show the low intra-specific variations and high inter-specific divergences in order to distinguish different species. Here, six parameters were used to characterize inter-specific versus intra-specific variation [Table 3]. Through comparison of interspecific genetic distances among congeneric species for four candidate barcode, ITS2 region exhibited the highest interspecific divergence with all four metrics, followed by *psbA-trnH* and *matK*, while *rbcL* provided the lowest [Table 3]. We also found that *rbcL* showed the lowest level of intraspecific variation with all four metrics, followed by *psbA-trnH* and *matK*, while ITS2 provided the highest [Table 3].

Validation of the different sequences' inter/intra-specific variation

The results of Wilcoxon signed rank tests confirmed that *matK* provided much higher inter-specific divergence among congeneric species [Table 4] and the higher variation between conspecific individuals [Table 5].

Assessment of barcoding gap

Barcodes should exhibit a "barcoding gap" between interspecific and intraspecific distances.[17,20] Although the histogram did not show a clear gap between intraspecific

Table 1: The collection sites and GenBank accession of 54 samples of the *Ardisia* genus

Species name[a]	Collection sites[b]	Voucher number	Gen Bank accession			
			ITS2	*psb A-trnH*	*rbcL*	*matK*
A. chinensis Benth.	1	KS0401MT01	JN252968	JN253070	JN253124	JN253022
A. dasyrhizomatica C. Y. Wu et C. Chen	1	KS0402MT01	JN252969	JN253071	JN253125	JN253023
A. dasyrhizomatica C. Y. Wu et C. Chen	1	KS0402MT02	JN252970	JN253072	JN253126	
A. squamulosa Presl.	1	KS0403MT01	JN252971	JN253073	JN253127	JN253024
A. squamulosa Presl.	1	KS0403MT02	JN252972	JN253074	JN253128	JN253025
A. squamulosa Presl.	1	KS0403MT03	JN252973	JN253075	JN253129	JN253026
A. squamulosa Presl.	2	KS0403MT04	JN252974	JN253076	JN253130	JN253027
A. curvula C. Y. Wu et C. Chen	1	KS0404MT01	JN252975	JN253077	JN253131	JN253028
A. corymbifera var. tuberifera C. Chen	1	KS0405MT01	JN252976	JN253078	JN253132	JN253029
A. humilis Vahl	1	KS0406MT01	JN252977	JN253079	JN253133	JN253030
A. humilis Vahl	1	KS0406MT02	JN252978	JN253080	JN253134	JN253031
A. humilis Vahl	3	KS0406MT03	JN252979	JN253081	JN253135	JN253032
A. humilis Vahl	3	KS0406MT04	JN252980	JN253082	JN253136	JN253033
A. virens Kurz *var. virens*	1	KS0407MT01	JN252981	JN253083	JN253137	JN253034
A. virens Kurz *var. virens*	1	KS0407MT02	JN252982	JN253084	JN253138	JN253035
A. virens Kurz *var. virens*	4	KS0407MT03	JN252983	JN253085	JN253139	
A. crispa (Thunb.) A. DC. *var. crispa*	1	KS0408MT01	JN252984	JN253086	JN253140	JN253036
A. crenata Sims *var. crenata*	5	KS0409MT01	JN252985	JN253087	JN253141	JN253037
A. crenata Sims *var. crenata*	5	KS0409MT02	JN252986	JN253088	JN253142	JN253038
A. crenata Sims *var. crenata*	1	KS0409MT03	JN252987	JN253089	JN253143	JN253039
A. crenata Sims *var. crenata*	1	KS0409MT04	JN252988	JN253090	JN253144	JN253040
A. crenata Sims *var. crenata*	3	KS0409MT05	JN252989	JN253091	JN253145	JN253041
A. crenata Sims *var. crenata*	3	KS0409MT06	JN252990	JN253092	JN253146	JN253042
A. virens Kurz *var. annamensis* Pitard	1	KS0410MT01	JN252991	JN253093	JN253147	JN253043
A. polycephala Wall. ex A. DC.	1	KS0411MT01	JN252992	JN253094	JN253148	JN253044
A. japonica (Thunb.) Blume	1	KS0412MT01	JN252993	JN253095	JN253149	JN253045
A. japonica (Thunb.) Blume	2	KS0412MT02	JN252994	JN253096	JN253150	JN253046
A. japonica (Thunb.) Blume	2	KS0412MT03	JN252995	JN253097	JN253151	JN253047
A. elegans Andr.	1	KS0413MT01	JN252996	JN253098	JN253152	JN253048
A. elegans Andr.	3	KS0413MT02	JN252997	JN253099	JN253153	JN253049
A. crenata Sims *var. bicolor* (Walker) C. Chen	1	KS0414MT01	JN252998	JN253100	JN253154	JN253050
A. pusilla A. DC.	3	KS0415MT02	JN252999	JN253101	JN253155	JN253051
A. pusilla A. DC.	4	KS0415MT03	JN253000	JN253102	JN253156	JN253052
A. pusilla A. DC.	4	KS0415MT04	JN253001	JN253103	JN253157	
A. fordii Hemsl.	1	KS0416MT01	JN253002	JN253104	JN253158	JN253053
A. quinquegona Bl. *var. quinquegona*	2	KS0417MT01	JN253003	JN253105	JN253159	JN253054
A. quinquegona Bl. *var. quinquegona*	2	KS0417MT02	JN253004	JN253106	JN253160	JN253055
A. quinquegona Bl. *var. quinquegona*	3	KS0417MT03	JN253005	JN253107	JN253161	JN253056
A. quinquegona Bl. *var. quinquegona*	3	KS0417MT04	JN253006	JN253108	JN253162	JN253057
A. punctata Lindl.	2	KS0418MT01	JN253007	JN253109	JN253163	JN253058
A. punctata Lindl.	2	KS0418MT02	JN253008	JN253110	JN253164	JN253059
A. ordinata Walker	2	KS0419MT01	JN253009	JN253111	JN253165	JN253060
A. mamillata Hance	3	KS0420MT01	JN253010	JN253112	JN253166	JN253061
A. densilepidotula Merr.	3	KS0421MT01	JN253011	JN253113	JN253167	JN253062
A. densilepidotula Merr.	3	KS0421MT02	JN253012	JN253114	JN253168	JN253063
A. hanceana Mez	3	KS0422MT01	JN253013	JN253115	JN253169	JN253064
A. hanceana Mez	3	KS0422MT02	JN253014	JN253116	JN253170	JN253065
A. triflora Hemsl.	4	KS0423MT01	JN253015	JN253117	JN253171	
A. triflora Hemsl.	4	KS0423MT02	JN253016	JN253118	JN253172	JN253066
A. corymbifera Mez *var. corymbifera*	4	KS0424MT01	JN253017	JN253119	JN253173	JN253067
A. corymbifera Mez *var. corymbifera*	4	KS0424MT02	JN253018	JN253120	JN253174	JN253068
A. primulaefolia Gardn. et Champ.	4	KS0425MT01	JN253019	JN253121	JN253175	JN253069
A. maculosa Mez *var. maculosa*	4	KS0426MT01	JN253020	JN253122	JN253176	
A. ensifolia Walker	4	KS0427MT01	JN253021	JN253123	JN253177	

[a] *Ardisia*: *A.* ;[b]1Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Yunnan, China; [b]2 Guangdong Research Institute of Traditional Chinese Medicine, Guangdong,China; [b]3 South China Botanical Garden,Guangdong,China; [b]4 Wuhan Botanical Garden, Chinese Academy of Sciences, Hubei, China; [b]5 Kunming Institute of Botany, Yunnan, China; ITS2: Internal transcribed spacer 2

variation and interspecific divergence in the distributions of the four tested loci (*matK*, *rbcL*, ITS2, *psbA-trnH* intergenic spacer) [Figure 1], the results of Wilcoxon two-sample tests showed that the distribution of inter-specific divergences for the four barcodes were higher than that of intra-specific variations [Table 6]. All the four candidate sequences showed significant difference ($P < 0.05$).

Evaluation of identifying ability of barcodes

Two methods of species identification, including BLAST1 and the nearest distance method were used to test the applicability of using different regions for unique species identification. In the BLAST1 method, the results showed that the *matK* region identified correctly 98.1% of the samples at the species level. In contrast to *matK*, the identification efficiency of *psbA-trnH*, ITS2 and *rbcL* were much lower at the species level. The results confirmed that *matK* had the highest success rate at the species level identification with both two methods [Table 7].

Table 2: Success rate of sequencing, length range, GC content

Markers	*psb A-trnH*	ITS2	*rbcL*	*matK*
Number of samples /*n*	54	54	54	54
Success of sequencing /*n*	54	54	54	48
Success rate of sequencing/%	100	100	100	88.9
Length range/bp	439-494	219-225	717	928-966
GC content/%	0.292	0.610	0.430	0.332

ITS2: Internal transcribed spacer 2; GC: Guanine and Cytosine; Bp: Base pair

DISCUSSION

The screening of DNA barcode for the Ardisia genus

Optimal DNA barcode should meet following requirements: (1) Significant inter-species variance; (2) Sufficient small intra-species variance; (3) It should be amplified by single primer and have quality sequence by dual sequencing.[21] In this research, we tested four DNA regions (*psbA-trnH*, ITS2, *rbcL* and *matK*) using 55 plant samples belonging to 27 closely related species from the *Ardisia* genus.

PsbA-trnH fragment has one of the biggest evolution rate among chloroplast compartment and flanked with approximate 75 bp conservative sequences at two ends, which can be used for designing universal primer.[8,11,22]

Table 3: Analysis of inter-specific divergence between congeneric species and intra-specific variation for the whole sample

Markers	*psbA-trnH*	ITS2	*matK*	*rbcL*
All inter-specific distance	0.0162±0.0098	0.0368±0.0159	0.0159±0.0074	0.0044±0.0028
Theta prime	0.0142±0.0065	0.0358±0.0133	0.0158±0.0040	0.0041±0.0017
Minimum inter-specific distance	0.0033±0.0056	0.0154±0.0157	0.0048±0.0035	0.0014±0.0018
All intra-specific distance	0.0053±0.0064	0.0119±0.0155	0.0056±0.0052	0.0009±0.0017
Theta	0.0069±0.0084	0.0095±0.0138	0.0119±0.0069	0.0009±0.0017
Coalescent depth	0.0089±0.0090	0.0142±0.0196	0.0148±0.0069	0.0012±0.0023

ITS2: Internal transcribed spacer 2

Table 4: Wilcoxon signed rank test for interspecific variations

W+	*W–*	Inter relative ranks, *n*, *P* value	Result
matK	*psbA-trnH*	*W+*= 306877.0, *W–*= 271473.0, *n* = 1075, *P* = 0.0821	*P*>0.05, *matK*=*psbA-trnH*
matK	ITS2	*W+*= 21701.0, *W–*= 556649.0, *n* = 1075, *P* = 4.2474E-152	*P*<10-10, *matK*>>ITS2
matK	*rbcL*	*W+*= 575250.0, *W–*= 3100.00, *n* = 1075, *P* = 1.0756E-173	*P*<10-10, *matK*>>*rbcL*
psbA-trnH	ITS2	*W+*= 21847.0, *W–*= 554354.0, *n* = 1073, *P* = 1.4515E-151	*P*<10-10, *psbA-trnH*>>ITS2
psbA-trnH	*rbcL*	*W+*= 549304.0, *W–*= 19407.0, *n* = 1066, *P* = 4.8884E-153	*P*<10-10, *psbA-trnH*>>*rbcL*
ITS2	*rbcL*	*W+*= 573079.0, *W–*= 977.0, *n* = 1071, *P* = 1.3494E-175	*P*<10-10, ITS2> >*rbcL*

Table 5: Wilcoxon signed rank test for intraspecific variations

W+	*W–*	Intra relative ranks, *n*, *P* value	Result
matK	*psbA-trnH*	*W+*= 1225.0, *W–*= 206.0, *n* = 53, *P* = 6.4680E-6	*P*<0.05, *matK*>*psbA-trnH*
matK	ITS2	*W+*= 735.0, *W–*= 696.0, *n* = 53, *P* = 0.8629	*P*>0.05, *matK*=ITS2
matK	*rbcL*	*W+*= 1431.0, *W–*= 0.0, *n* = 53, *P* = 2.3865E-10	*P*<0.05, *matK*>*rbcL*
psbA-trnH	ITS2	*W+*= 254.0, *W–*= 827.0, *n* = 46, *P* = 0.0017	*P*<0.05, *psbA-trnH*>ITS2
psbA-trnH	*rbcL*	*W+*= 865.0, *W–*= 38.0, *n* = 42, *P* = 2.3263E-7	*P*<0.05, *psbA-trnH*>*rbcL*
ITS2	*rbcL*	*W+*= 521.0, *W–*= 7.0, *n* = 32, *P* = 1.5153E-6	*P*<0.05, ITS2>*rbcL*

ITS2: Internal transcribed spacer 2

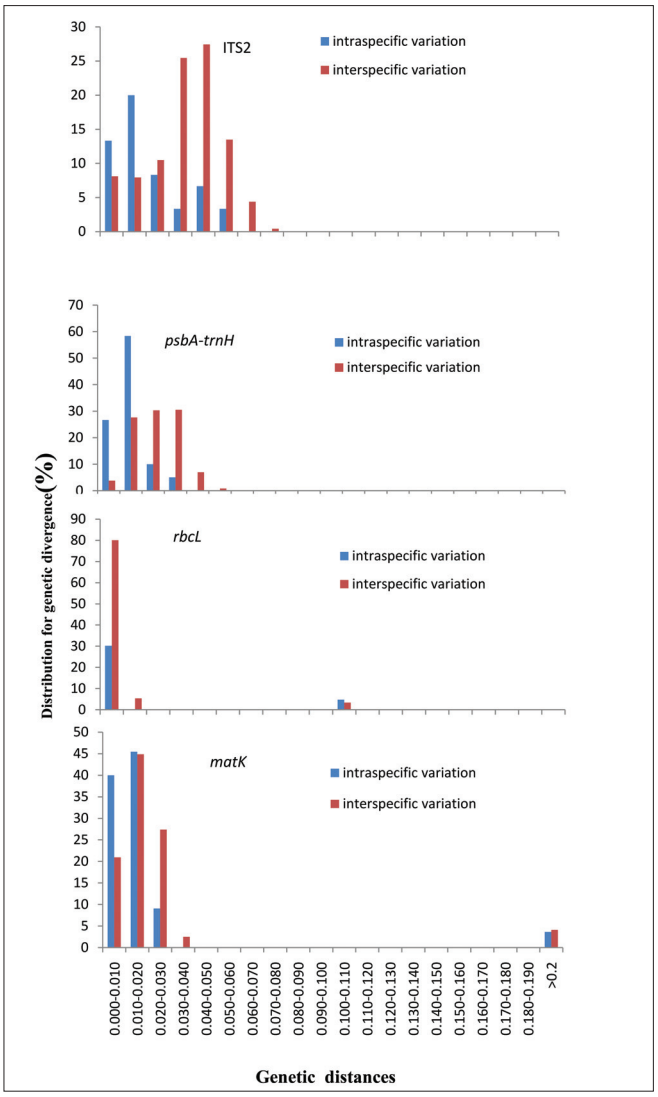

Figure 1: The barcoding gap between interspecific and intraspecific divergences for four candidate barcodes. Histograms showing the relative distribution of pairwise (y-axis) intraspecific (blue bar) and interspecific (red bar) divergence distance estimates (x-axis) for internal transcribed spacer 2(ITS2), psbA-trnH, rbcL and matK intergenic spacers, respectively. The divergences were calculated using the Kimura 2-parameter model. Barcoding gaps were assessed by the Wilcoxon two-sample tests, and all were significant ($P < 0.05$)

Yao *et al.* found it universal with high success rates of amplification, which is highly recommended in barcode research[23,24] In our study, we found *psbA-trnH* sequence has a successful identification rate of 70.4%. Although there is a significant difference at intra- and inter-species levels, it has low identification efficiency. Therefore, it is not suitable as the *Ardisia* barcode sequence.

Many researchers have proposed the ITS2 region as a suitable marker for taxonomic classification.[13,25] However in our study, the identification efficiency with ITS2 is only 51.9%. So ITS2 is also not suitable as a barcode for the identification of *Ardisia* species.

RbcL and *matK* are recommended as plant barcode sequence in the latest Consortium for the Barcode of Life [CBOL] Research.[26] There are large amount of data for *rbcL* in Genbank, which is universal, being easily amplified and compared, but its variance mainly exist in intra-species rather than inter-species.[17,27] As described before, *rbcL* fragment was chosen as plant barcode candidate by Kress *et al.*[8] However, there is no significant difference between the intra-species and inter-species in the research, moreover, the efficiencies of identification by BLAST1 and nearest distance are only 29.1% respectively, therefore, *rbcL* is not proper as the *Ardisia* DNA barcode sequence.

The *matK* fragment is emerging as a gene with potential contribution to plant molecular systematic and evolution.[28-31]

Table 6: Wilcoxon two-sample tests for distribution of intra- versus inter-specific divergences

Marker	Number of inter-specific distances	Number of intra-specific distances	Wilcoxon	*P* value
ITS2	1371.0	60.0	72.1	7.18×10^{-4}
psbA-trnH	1371.0	60.0	86.9	3.12×10^{-3}
matK	1075.0	53.0	103.6	5.64×10^{-3}
rbcL	1371.0	60.0	75.1	7.16×10^{-4}

ITS2: Internal transcribed spacer 2

Table 7: Comparison of identification efficiency for potential DNA barcodes loci using different methods of species identification

Marker	Method of species identification	Number of species	Number of samples	Correct identification	Incorrect identification	Ambiguous identification
				Species level %	Species level	Species level %
psbA-trnH	BLAST 1	24	54	70.4	0	29.6
	Distance	24	54	44.4	0	55.6
ITS2	BLAST 1	24	54	51.9	0	48.1
	Distance	24	54	51.9	0	48.1
matK	BLAST 1	22	48	98.1	0	1.9
	Distance	22	48	91.7	0	8.3
rbcL	BLAST 1	24	54	27.8	0	72.2
	Distance	24	54	27.8	0	72.2

BLAST: Basic local alignment search tool; ITS2: Internal transcribed spacer 2; DNA: Deoxyribonucleic acid

The fragment has a quicker evolution compared with other fragments. In the research, the *matK* region had the highest identification success rate at the species level; meanwhile, it exhibited well in PCR amplification and sequencing efficiency, differential intra- and inter-specific divergences and DNA barcoding gap. Therefore, we suggested *matK* region as the DNA barcode for the genus *Ardisia*.

Discussion on samples with unsuccessful identification

In our study, the *matK* sequence was chosen as a DNA barcode to identify the species of Ardisia genus. Among the 48 samples, which were successfully sequenced, there was one sample (*A. japonica*) that could not be distinguished from *A. pusilla*. These two species are sister species both attached to the group of Sect. Bladhia. They show little differences in morphology and closely relationship between them and that may be the reason that they were difficult to differentiate from each other.

The present research found that of the four candidate loci (*psbA-trnH*, ITS2, *matK*, *rbcL*), *matK* produced the highest rate of successful identification in 91.7% at the species level and it can correctly discriminate 22 Chinese medicinal species from *Ardisia* according to the nearest distance method. Therefore, it is proposed that the *matK* region can be used as a DNA barcode to identify these medicinal plants from *Ardisia*. Collection of more samples and deep researches for those species of ambiguous identification are necessary to provide more effective information about phyletic evolution and more reliable method for the identification of genus *Ardisia*.

Measuring the success rates of identification methods

CBOL recommended *rbcL* and *matK* together as plant barcode sequence, but needs enough matching data from the experiment, which could increase cost, therefore, we just focus on the probability of single sequence, BLAST1/and the nearest distance methods are employed. Meanwhile, identification efficiency was measured in order to display the ability for all sequences. BLAST method compares sample's DNA sequence with total sequence in terms of base, which was ranked by base difference; the advantage of this method is high velocity and accuracy. Nearest distance method compares sample's DNA with the "Kimura 2-parameter (K2P)"distance of total sequence, which is based on overall comparison.[19] It can quantitate difference in single sequence with low velocity, meanwhile lost locus and variable locus are processed equally, which easily leads to the slight difference between data and facts, which is the reason why these two verifications are not uniform.

We will measure how the total data change when each sequence exist or not during the process of efficiency identification. When the data are abnormal, we blast the suspicious sequence with GenBank in order to exclude "false positive" data. The same as other authors, we define "inter-species variance" as the variance among different species under a genus without breaking the genus. It might get smaller results than real fact by using "inter-species" in above extent. Layhaye *et al*.[17] also got the same conclusion as we have, therefore, we will use some new method of identification, e.g., probability of correct identification [PCI], in order to exclude man-made disturbance.

DNA barcode technique has already been used in animal research and increasing used in plant research, which will assist non-systematic scholars to quickly and accurately identify different species. DNA barcode cannot replace traditional taxonomy, but it is accurate, abundant and unique with high repeats as digital DNA sequence, leading to a useful tool for taxonomists.[26,27,32] This research explored the application of DNA barcode technique and provided a new method and insight for molecular identification and relationship. As limits of sampling condition in this research, some species had no duplicates; some nearest sibling species were included under a genus. There should be more effective information and reliable method when more samples are included in further research in the future.

ACKNOWLEDGMENTS

Thanks to Prof. Yuewen Cai, Mrs. Chunfen Xiao and Mr. Shouzhong Zhang for helping with collection and authentication of experimental materials. This work was funded by the Key Projects in the National Science and Technology Pillar Program (2011BAI07B08) and the Special Funding of the Ministry of Health (200802043).

REFERENCES

1. Chen J. Myrsinaceae Flora Republicea Popularis Sinicae. Vol. 58. Beijing: Beijing Science Press; 1979. p. 35-42.
2. Jiang XM, Ye JS, Xin WR. Introduction to medical and horticultural values and research progress of *Ardisia* Species. Jiangxi For Sci Tech 2003;5:30-3.
3. Hebert PD, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. Proc Biol Sci 2003;270:313-21.
4. Gregory TR. DNA barcoding does not compete with taxonomy. Nature 2005;434:1067.
5. Vences M, Thomas M, Bonett RM, Vieites DR. Deciphering amphibian diversity through DNA barcoding: Chances and challenges. Philos Trans R Soc Lond B Biol Sci 2005;360: 1859-68.
6. Janzen DH, Hajibabaei M, Burns JM, Hallwachs W, Remigio E, Hebert PD. Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. Philos Trans R Soc Lond B Biol Sci 2005;360:1835-45.

7. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD. DNA barcoding Australia's fish species. Philos Trans R Soc Lond B Biol Sci 2005;360:1847-57.
8. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci U S A 2005;102:8369-74.
9. Newmaster SG, Fazekas AJ, Ragupathy S. DNA barcoding in land plants: evaluation of rbcL in a multigene tiered approach. Can J Bot 2006;84:335-41.
10. Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madrinan S, Petersen G, *et al*. A proposal for a standardised protocol to barcode all land plants. Taxon 2007;56:295-9.
11. 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.
12. Pennisi E. Taxonomy. Wanted: A barcode for plants. Science 2007;318:190-1.
13. Chen S, Yao H, Han J, Liu C, Song J, Shi L, *et al*. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 2010;5:e8613.
14. Keller A, Schleicher T, Schultz J, Müller T, Dandekar T, Wolf M. 5.8S–28S rRNA interaction and HMM-based ITS2 annotation. Gene 2009;430:50-7.
15. Luo K, Chen S, Chen K, Song J, Yao H, Ma X, *et al*. Assessment of candidate plant DNA barcodes using the Rutaceae family. Sci China Life Ser C 2010;40:342-51.
16. Zhu YJ, Chen SL, Yao H, Tan R, Song JY, Luo K, *et al*. DNA barcoding the medicinal plants of the genus Paris. Acta Pharm Sin 2010;45:376-82.
17. Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, *et al*. DNA barcoding the floras of biodiversity hotspots. Proc Natl Acad Sci U S A 2008;105:2923-8.
18. 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.
19. Ross HA, Murugan S, Li WL. Testing the reliability of genetic methods of species identification via simulation. Syst Biol 2008;57:216-30.
20. Meyer CP, Paulay G. DNA barcoding: Error rates based on comprehensive sampling. PLoS Biol 2005;3:2229-38.
21. Song J, Yao H, Li Y, Li X, Lin Y, Liu C, *et al*. Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique. J Ethnopharmacol 2009;124:434-9.
22. Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, *et al*. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. PLoS One 2008;3:e2802.
23. 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.
24. Liu Y, Zhang L, Liu Z, Luo K, Chen S, Chen K. Species identification of *Rhododendron* (Ericaceae) using the chloroplast deoxyribonucleic acid *PsbA-trn*H genetic marker. Pharmacogn Mag 2012;8:29-36.
25. Miao M, Warren A, Song W, Wang S, Shang H, Chen Z. 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.
26. CBOL Plant Working Group. A DNA barcode for land plants. Proc Natl Acad Sci U S A 2009;106:12794-7.
27. Ning SP, Yan HF, Hao G, Ge XJ. Current advances of DNA barcoding study in plants. Biodivers Sci 2008;16:417-25.
28. Johnson LA, Soltis DE. *Matk* DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. Syst Bot 1994;19:143-56.
29. Johnson LA, Soltis DE. Phylogenetic inference in saxifragaceae sensu stricto and gilia (Polemoniaceae) using *matK* sequences. Ann Mo Bot Gard 1995;82:149-75.
30. Steele KP, Vilgalys R. Phylogenetic analyses of Polemoniaceae using nucleotide sequences of the plastid gene *matK*. Syst Bot 1994;19:126-42.
31. Liang HP, Hilu KW. Application of the *matK* gene sequences to grass systematics. Can J Bot 1996;74:125-34.
32. Chen SL, Song JY, Yao H, Shi LC, Luo K, Han JP. Strategy and key technique of identification of Chinese herbal medicine using DNA barcoding. Chin J Nat Med 2009;7:322-7.

Cite this article as: Liu Y, Wang K, Liu Z, Luo K, Chen S, Chen K. Identification of medical plants of 24 Ardisia species from China using the matK genetic marker. Phcog Mag 2013;9:331-7.

Source of Support: Nil., **Conflict of Interest:** None declared.