MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2012) **12**, 783–790 doi: 10.1111/j.1755-0998.2012.03147.x

Evaluation of six candidate DNA barcoding loci in *Ficus* (Moraceae) of China

H.-Q. LI,[1] J.-Y. CHEN,[1] S. WANG and S.-Z. XIONG
School of Life Science, East China Normal University, Shanghai 200241, China

Abstract

***Ficus*, with about 755 species, diverse habits and complicated co-evolutionary history with fig wasps, is a notoriously difficult group in taxonomy. DNA barcoding is expected to bring light to the identification of *Ficus* but needs evaluation of candidate loci. Based on five plastid loci (*rbc*L, *mat*K, *trn*H-*psb*A, *psb*K-*psb*I, *atp*F-*atp*H) and a nuclear locus [internal transcribed spacer (ITS)], we calculated genetic distances and DNA barcoding gaps individually and in combination and constructed phylogenetic trees to test their ability to distinguish the species of the genus. A total of 228 samples representing 63 putative species in *Ficus* (Moraceae) of China were included in this study. The results demonstrated that ITS has the most variable sites, greater intra- and inter-specific divergences, the highest species discrimination rate (72%) and higher primer universality among the single loci. It is followed by *psb*K-*psb*I and *trn*H-*psb*A with moderate variation and considerably lower species discrimination rates (about 19%), whereas *mat*K, *rbc*L and *atp*F-*atp*H could not effectively separate the species. Among the possible combinations of loci, ITS + *trn*H-*psb*A performed best but only marginally improved species resolution over ITS alone (75% vs. 72%). Therefore, we recommend using ITS as a single DNA barcoding locus in *Ficus*.**

Keywords: DNA barcoding, *Ficus*, internal transcribed spacer, *psb*K-*psb*I, species discrimination, *trn*H-*psb*A

Received 26 November 2011; revision revised 22 February 2012; accepted 1 March 2012

Introduction

Ficus (Moraceae, $2n = 26$), comprising almost 755 species (Noort *et al.* 2007), is among the largest genera of angiosperms. With regard to habit, growth and life form, *Ficus* is also one of the most diverse genera and includes terrestrial trees, hemi-epiphytes, holo-epiphytes, climbers, creeping shrubs, rheophytes, etc. (Corner 1967; Berg & Wiebes 1992). Such diversity extends to phytochemical compounds, which makes the genus an important source of medicine (Fan *et al.* 2005). All *Ficus* species are recognized by a specialized syconia that exhibits one of two breeding systems: monoecy (*c.* 400 species) or gynodioecy (*c.* 350 species), and by an obligate pollination mutualism syndrome with Agaonidae wasps, which has long been considered a classic case of co-evolution and co-speciation (Janzen 1979; Wiebes 1979). However, a rapid and accurate method for *Ficus* identification is urgently needed to facilitate its use in biodiversity, food and pharmaceutical industry and co-evolution research (Harrison 2005; Solomon *et al.* 2006).

DNA barcoding is a well-known technique that aims to facilitate rapid species identification based on short, standardized DNA sequences in cases in which rapid taxonomic identification is not feasible (Tautz *et al.* 2002, 2003; Hebert *et al.* 2003; Kress *et al.* 2005; Savolainen *et al.* 2005). The Plant Working Group of the Consortium for the Barcode of Life (CBOL) proposed to use the combination of *rbc*L + *mat*K as a core plant barcode. However, the ability to distinguish closely related species is uncertain. As a consequence, more efforts are needed to avoid possibly overestimating the discriminatory power of the barcodes by considering an insufficient sample or focussing on a restricted geographical area of closely related species (Sass *et al.* 2007; Edwards *et al.* 2008; Lahaye *et al.* 2008b; Fazekas *et al.* 2009; Hollingsworth *et al.* 2009; Yan *et al.* 2011).

In the present study, we selected six commonly recommended DNA regions (*rbc*L, *mat*K, *trn*H-*psb*A, *psb*K-*psb*I, *atp*F-*atp*H and internal transcribed spacer (ITS); CBOL Plant Working Group, 2009) and tested their polymerase chain reaction (PCR) amplification universality and discriminatory power using an extensive sampling of *Ficus* species, including most species from China, to assess their suitability as DNA barcodes for *Ficus*.

Correspondence: Hong-Qing Li, Fax: (86)21-62233754;
E-mail: hqli@bio.ecnu.edu.cn

[1]These authors contributed equally to this work and should be considered co-first authors.

Materials and methods

Sampling

A total of 228 samples from 63 *Ficus* species were sampled. To eliminate biases associated with uneven sampling among taxa, all species to be included contain three to five (at least two) representatives from different populations when possible. Voucher specimens (deposited at HSNU), collection information, geographic location information and GenBank accession numbers are available as detailed in Table S1 (Supporting information). Because material of sect. *Pharmacosycea* is absent, we used *Humulus scandens* (Lour.) Merr. (Cannabaceae) and *Artocarpus altilis* (Park.) Fosb. (Moraceae) as outgroups for the phylogenetic analyses instead.

Methods

DNA extraction, PCR amplification and sequencing. Nuclear ITS and five chloroplast genome region, including two coding genes (*mat*K, *rbc*L) and three intergenic spacers (*atp*F-*atp*H, *trn*H-*psb*A and *psb*K-*psb*I), were evaluated. Genomic DNA was extracted either from fresh or silica gel-dried leaves using the CTAB method (Doyle & Doyle 1987). PCR amplifications were performed on TAKARA TP600 thermocycler (Takara Bio, Inc., Otsu, Shiga, Japan). The primers and reaction procedures are listed in Table 1. The PCR products were inspected on 1% TAE agarose gels and then purified using the Tiangen Midi Purification Kit (Tiangen Biotech, Beijing, China) according to the manufacturer's protocol. The purified PCR products were bidirectionally sequenced by Shanghai Invitrogen Biotechnology (Shanghai, China).

Sequence analysis. Sequences were assembled and edited with SeqMan (DNA STAR package, Madison, WI, USA; Burland 2000). The cleaned sequences were aligned using ClustalW of MEGA 5 package (Tamura *et al.* 2011), and the alignment finally adjusted manually.

Inter- and intra-specific genetic divergences were calculated using Kimura 2-parameter (K2P) distances in MEGA 5.0 following the instruction of the CBOL for distance calculations (Tamura *et al.* 2007, 2011). Wilcoxon signed-rank tests were performed to compare intra- and inter-specific variability for every pair of barcodes (http://www.fon.hum.uva.nl/Service/Statistics/Signed_Rank_Test.html). Barcoding gaps were estimated by comparing the distributions of intra- and inter-specific divergences of each candidate locus using the program TaxonDNA (Meier *et al.* 2006).

To further evaluate the effectiveness of barcoding candidates for species discrimination, we conducted a tree-based analysis. A phylogenetic tree was estimated using maximum likelihood (ML) in MEGA 5.0 using K2P distances. Bootstrap values were calculated with 5000 random addition sequence replicates.

Results

PCR amplification and sequencing

The efficiency of PCR amplification and sequencing is an important index to evaluate the candidate barcoding. In the present study, 993 new sequences were generated, which included 199, 182, 178, 192, 121 and 121 sequences of ITS, *rbc*L, *mat*K, *trn*H-*psb*A, *atp*F-*atp*H and *psb*K-*psb*I, respectively (Table 2). All barcode loci but one (*matK*) exhibited high PCR success (>98%), which adequately indicated the universality of the primers (Table 2). The most successful PCR rate was observed for *atp*F-*atp*H (100%), followed by *trn*H-*psb*A (98.68%), *rbc*L (98.66%), *psb*K-*psb*I (98.39%) and ITS (98.20%), whereas *mat*K showed the worst performance at 88.31%. Sequencing success ranged from 98.64 to 100% for all candidate barcodes (Table 2).

Intra- and inter-specific diversities

For individual regions, aligned sequence lengths range from 541 bp for *psb*K-*psb*I to 838 bp for *mat*K (Table 2). Nuclear ITS was found to contain the most variable sites and parsimony-informative characters. Among plastid noncoding regions, the *psb*K-*psb*I and *trn*H-*psb*A had almost two, three and four times more variable characters than the *mat*K, *atp*F-*atp*H and *rbc*L regions, respectively (Table 2). Our results demonstrated that ITS exhibited the highest mean inter-specific and lower intra-specific divergence; *trn*H-*psb*A and *psb*K-*psb*I performed better than *atp*F-*atp*H, *mat*K and *rbc*L (Table 2). Similar results were obtained using Wilcoxon signed-rank tests (Tables S2 and S3, Supporting information).

DNA barcoding gap

Ideally, barcodes must exhibit a 'barcoding gap' between inter- vs. intra-specific divergences. To evaluate whether such gaps exist, we checked the distribution of divergences at classes of 0.005 distance units. As can be seen from the graph (Fig. 1), barcoding gaps were not found in any of the six candidate loci or the combination of ITS + *trn*H-*psb*A. Each distribution graph showed an overlap between intra- and inter-specific distances. ITS and the combination of ITS + *trn*H-*psb*A have the least overlap with their threshold values of intra- and inter-specific divergences located between 1.5% and 2.0%. Inter-specific divergences showed normal distribution (Fig. 1). Among the cpDNA loci, *psb*K-*psb*I and *trn*H-*psb*A

Table 1 PCR primers of six candidate barcodes

DNA region	Primer name	Sequence (5′-3′)	Amplication protocol	Reference
*rbc*L	a_f	ATGTCACCACAAACAGAGACTAAAGC	95 °C 4 min; 94 °C 30 s, 55 °C 1 min, 72 °C 1 min, 35 cycles; 72 °C 10 min	Kress & Erickson (2007)
	a_r	CTTCTGCTACAAATAAGAATCGAT CTC		
*trn*H-*psb*A	*trn*H	CGCGCATGGTGGATTCACAATCC	95 °C 4 min; 94 °C 30 s, 55 °C 1 min	Tate & Simpson (2003)
	*psb*A	GTTATGCATGAACGTAATGCTC	72 °C 1 min, 35 cycles; 72 °C 10 min	Sang *et al.* (1997)
*mat*K	3F_KIM	CGTACAGTACTTTTGTGTTTA CGAG	94 °C 1 min; 94 °C 30 s, 52 °C 20 s	CBOL (http://barcoding.si.edu)
	IR_KIM	ACCCAGTCCATCTGGAAATCTTGGTTC	72 °C 50 s, 35 cycles; 72 °C 5 min	
ITS	ITS 3	GCATCGATGAAGAACGTAGC	94 °C 5 min; 94 °C 1 min, 50 °C 45 s, 72 °C 1 min, 30 cycles; 72 °C 5 min	White *et al.* (1990)
	ITS 4	TCCTCCGCTTATTGATATGC		
*atp*F-*atp*H	*atp*F	ACTCGCACACACTCCCTTTCC	94 °C 5 min; 94 °C 30 s, 50 °C 30 s, 72 °C 40 s, 35 cycles; 72 °C 5 min	Lahaye *et al.* (2008a)
	*atp*H	GCTTTTATGGAAGCTTTAACAAT		
*psb*K-*psb*I	*psb*K	TTAGCCTTTGTTTGGCAAG	94 °C 5 min; 94 °C 30 s, 50 °C 30 s, 72 °C 40 s, 35 cycles; 72 °C 5 min	Lahaye *et al.* (2008a)
	*psb*I	AGAGTTTGAGAGTAAGCAT		

Table 2 Properties of the six candidate barcoding loci (the outgroup sequences are not included in the calculation)

	*psb*K-*psb*I	*trn*H-*psb*A	*atp*F-*atp*H	*rbc*L	ITS	*mat*K
No. of samples/species	121/48	223/63	121/48	221/63	218/61	200/61
% PCR success	98.39	98.68	100	98.66	98.20	90.91
% Sequencing success	100	99.55	100	100	98.64	100
% Variable/informative sites	17.2/9.2	22.57/14.29	7.60/4.73	5.36/2.68	37.02/29.78	14.20/6.92
Aligned length (bp)	541	567	697	709	759	838
Mean inter-specific distance	0.012743 ± 0.004167	0.013619 ± 0.004786	0.005729 ± 0.002542	0.003277 ± 0.001743	0.049713 ± 0.008321	0.004398 ± 0.001882
Mean intra-specific distance	0.007236 ± 0.002377	0.004724 ± 0.002106	0.001675 ± 0.000989	0.001032 ± 0.000594	0.002863 ± 0.001319	0.001678 ± 0.000891
Resolution of species (%)	20.83	19.05	18.75	6.35	72.13	16.39

performed better than the other three candidate loci (*atp*F-*atp*H, *rbc*L and *mat*K) at threshold values of about 2% and 3%, respectively (Fig. 1).

Phylogenetic analyses

To evaluate whether species were recovered as monophyletic under each barcode, we constructed phylogenetic trees with single or combined candidate loci. Among single-locus analyses, only ITS showed relatively high levels of species discrimination, with a success rates of 72.13%, followed by *psb*K-*psb*I (20.83%), *trn*H-*psb*A (19.05%), *atp*F-*atp*H (18.75%) and *mat*K (16.39%), whereas *rbc*L had the lowest discrimination levels (6.35%). The core combination of *mat*K + *rbc*L was only in 27.87% of the cases able to discriminate species. Given that *trn*H-*psb*A and *psb*K-*psb*I performed better than the other candidates, we combined each with ITS, which lead to a resolution of 75% (Fig. S3, Supporting information) and 74.47% (tree not shown) species as exclusive lineages, respectively.

Discussion

Proposed barcoding loci

In general, several criteria should be applied to an ideal DNA barcode: (i) high inter-specific but low intra-specific divergence, so that they can be discriminated from one another and a clear threshold between intra- and inter-specific genetic variations; (ii) highly conserved flanking sites should be present for developing universal primers; (iii) an appropriately short sequence length (300–800 bp) suitable for DNA extraction, PCR amplification and sequencing with a pair of universal primers; and (iv) amenable to be directly sequenced and quickly aligned without manual editing (Cowan *et al.* 2006; CBOL Plant Working Group, 2009; Ford *et al.* 2009). The chloroplast coding region *rbc*L, as one of the core plant barcode, has been proven to be of universal primers, easy to be sequenced and aligned in ferns, mosses and angiosperms (Hasebe *et al.* 1995; Hollingsworth *et al.* 2009; Liu *et al.* 2011). However, previous studies have

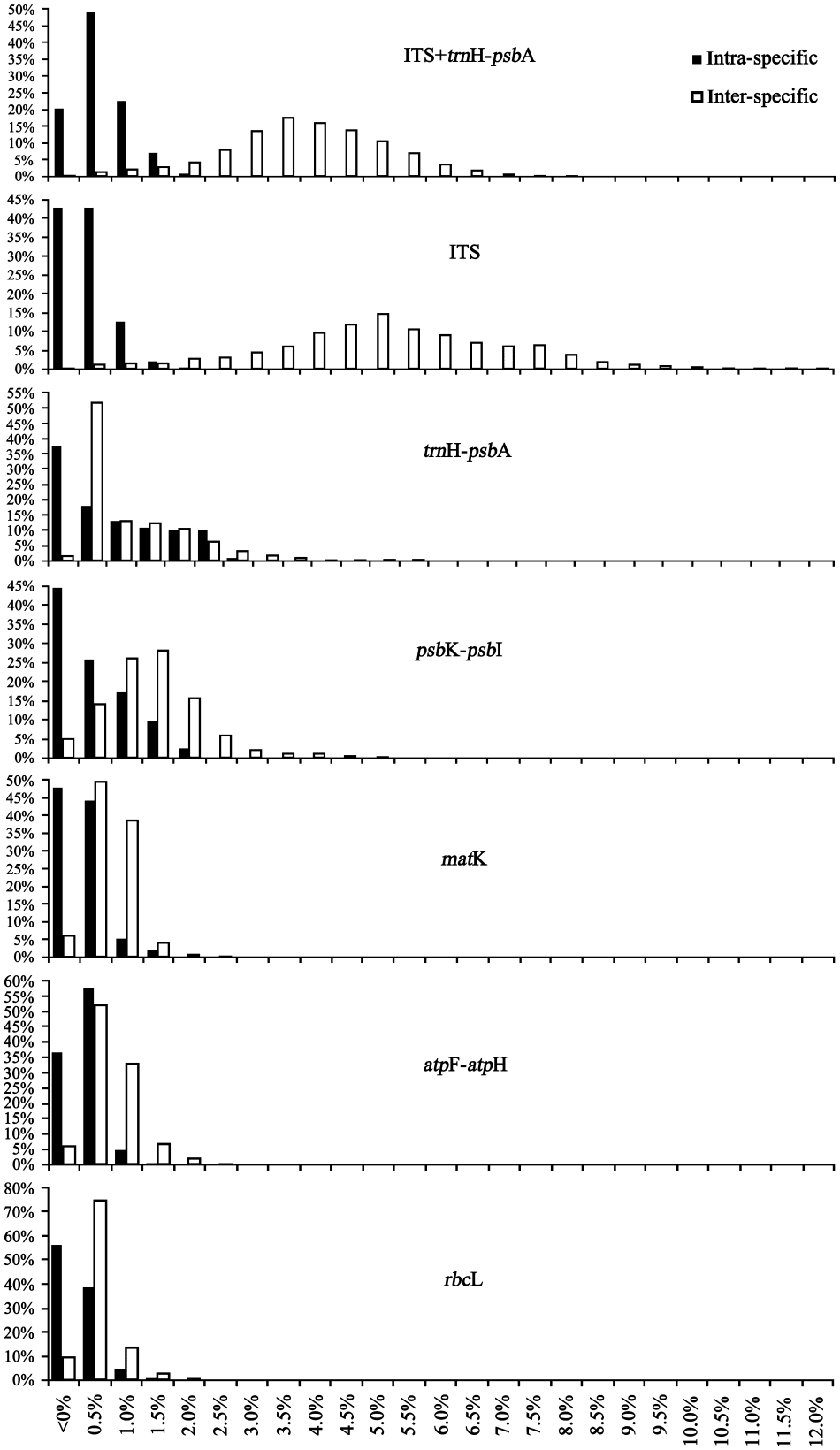

Fig. 1 Distribution of inter- and intra-specific Kimura 2-parameter (K2P) distances among all samples for the six candidate loci and the combination of ITS + *trn*H-*psb*A. *x*-axes: intra- and inter-specific K2P distances; *y*-axes: accumulative frequency.

shown that *rbc*L possesses very low inter-specific variation (Kress *et al.* 2005), especially in closely related species (Newmaster *et al.* 2008). In our study, it contained the lowest number of informative sites and the lowest genetic variability and species discrimination rate (Table 2). Thus, the *rbc*L region is not suitable for DNA barcoding in *Ficus*. The coding region *mat*K was also recommended as one of the core plant barcode by CBOL, yet its universality was questioned in some groups (Cuénoud *et al.* 2002; Chase *et al.* 2007). In our study, we tried another primer pair (390F/1326R) suggested by Lahaye *et al.* (2008b) for the difficult samples. Regretfully, PCR success did not improve. Besides the relatively low amplification success, low genetic variability also excluded *mat*K from being a useful barcode in *Ficus*. The core combination of *rbc*L + *mat*K suggested by CBOL only provided 27.87% species discrimination rate; once again, it is incapable of functioning as a DNA barcode in *Ficus.*

In the present study, ITS exhibited the most variation, greater intra- and inter-specific divergences, the highest species discrimination rate, higher primer universality and suitable aligned sequence lengths (about 750 bp; Table 2, Fig. 2), which make it the best choice as DNA barcode in *Ficus*. In fact, ITS has long been used as a phylogenetic marker in plants because it was approximately three to four times more variable than plastid genes (Chase *et al.* 2007; Nieto & Rossello 2007; Roy *et al.* 2010) and has been successfully used in phylogenetic analyses in related studies in *Ficus* (Weiblen 2000; Jousselin *et al.* 2003; Rønsted *et al.* 2008). In the studies of plant DNA barcodes, this region has been demonstrated to be highly effective in discriminating closely related species in numerous genera and be suggested as a suitable barcode as well (Alvarez & Wendel 2003; Chen *et al.* 2010; China Plant BOL Group 2011). However, previous studies have also shown some obstacles to ITS as a universal barcode: the universal primers failed in gymnosperms, ferns and mosses (Kress & Erickson 2007); many species required an intermediate primer as their lengths were more than 1100 bp; the presence of homopolymers and multiple copies in the genome that led to difficulties with PCR, sequencing and sequence analysis (Sass *et al.* 2007); pseudogenes and contamination with fungi were found regularly (Alvarez & Wendel 2003). It goes without saying that some of these problems could also exist in *Ficus*. For example, poly-G and poly-C stretches made sequencing difficult at times. In addition, three samples of *F. racemosa* were contaminated by fungi; though, we repeated the experiment three times. Coincidently, the study of Xu *et al.* (2011) on this species showed the same result, although the samples were collected in different locations. But so far, such cases have only been detected in this particular species. For all of these reasons, ITS, even with its recognized limitations, is an effective candidate DNA barcode in *Ficus*.

The ends of the noncoding region *trn*H-*psb*A have a 75-bp highly conserved sequence, which was beneficial for designing universal primers (Shaw *et al.* 2005). However, indels are common in this region, which resulted in greater length variation and increased difficulty in alignment (Chase *et al.* 2007; Lahaye *et al.* 2008a; Hollingsworth *et al.* 2009). To a certain degree, the presence of poly-A/T decreased sequence quality (Shaw *et al.* 2005; Sass *et al.* 2007; Fazekas *et al.* 2008). Besides universal primers, relatively high variability and short lengths, Wilcoxon signed-rank tests demonstrated that *trn*H-*psb*A was outstanding in its inter-specific variation compared to *psb*K-*psb*I, *mat*K, *rbc*L or *atp*F-*atp*H and, thus, met most necessary criteria as DNA barcodes. But unfortunately, it managed only 19% species resolution (Fig. S1, Supporting information), and all these species can also be discriminated by ITS. To further explore the potential of *trn*H-*psb*A as candidate barcode in *Ficus*, we constructed trees of ITS + *trn*H-*psb*A, which provided 75% species resolution (Fig. S3, Supporting information). Considering that ITS alone provided up to 72% species resolution, it appears both impractical and unnecessary money and effort to add *trn*H-*psb*A just for the extra 3% resolution. Therefore, we suggest that ITS could be regarded as the only DNA barcode in *Ficus*.

Exploratory work for more valuable DNA barcoding in Ficus

Considering the recommend DNA barcode ITS obtained only 72% species resolution and some closely related species were still not effectively identified, it will be necessary to screen more informative regions. According to the proposal of Ki-Joong Kim at the Second International Barcode of Life Conference (September 2007), we evaluated the potential of the *psb*K-*psb*I- and *atp*F-*atp*H-regions as DNA barcodes in *Ficus*. The results indicate that *psb*K-*psb*I had more variable and informative sites, higher genetic distances and species discrimination rate than *atp*F-*atp*H (Table 2, Fig. S2, Supporting information). When *psb*K-*psb*I was compared to *rbc*L, *mat*K, *trn*H-*psb*A and ITS, its performance is very similar to that of *trn*H-*psb*A in all aspects (Table 2), but what is more important, unlike *trn*H-*psb*A, is that it has fewer indels and homopolymers, and the sequencing success rate and quality is much higher. However, it acquired only about 21% species resolution, which restricted it from being a suitable barcode in *Ficus*. Accordingly, if a second high species discrimination region should be explored, more effort should be concentrated on other regions, especially nuclear loci.

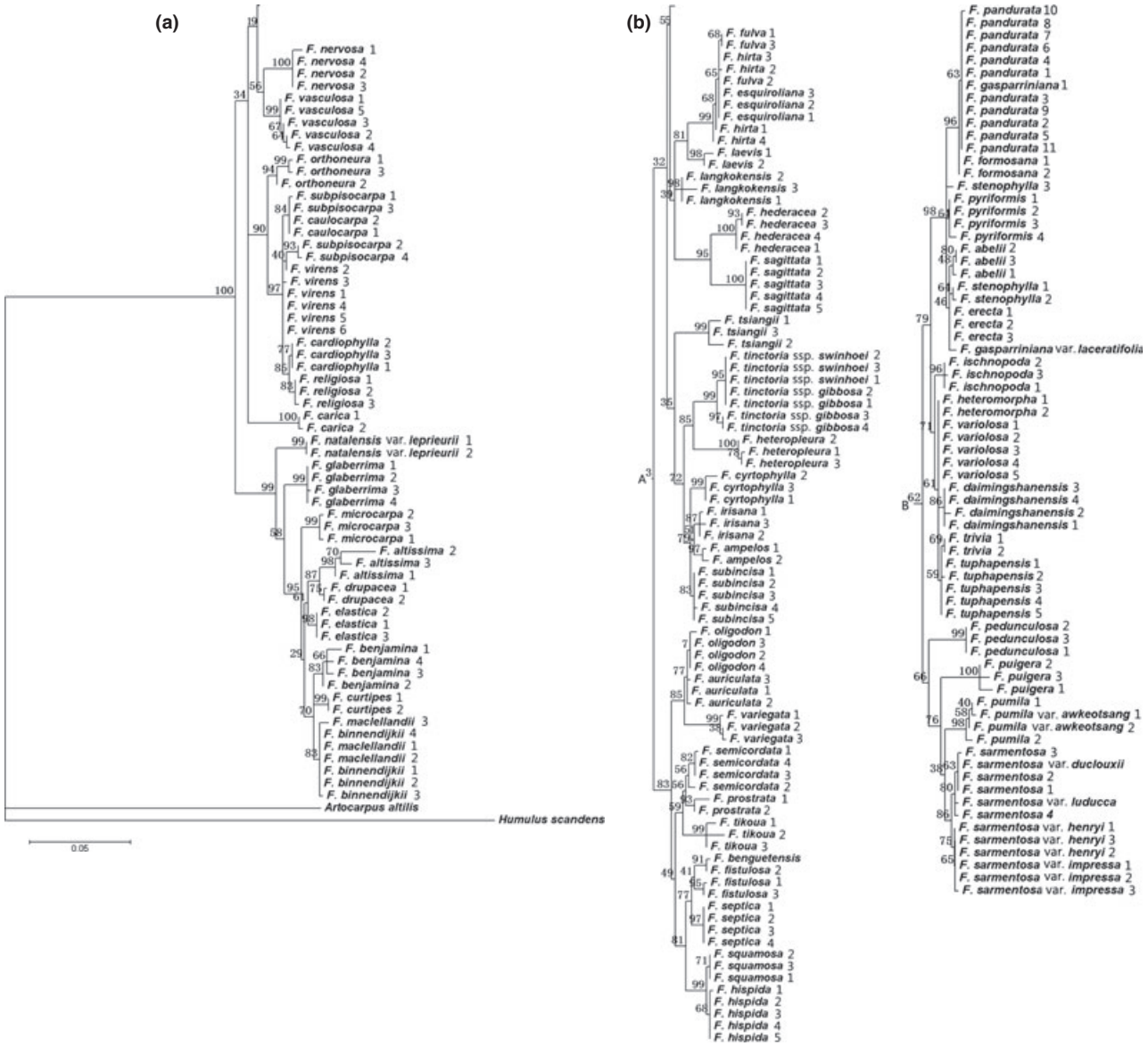

Fig. 2 A tree of ITS sequences generated using maximum likelihood (ML) methods with Kimura 2-parameter distances of *Ficus*. Bootstrap values are shown above the relevant branches. The tree was divided into three successive parts.

Acknowledgements

This study was supported by the Large-Scale Scientific Facilities Research Project (2009-LSFGBOWS-01). Our work was accomplished in the laboratory of School of Life Science, East China Normal University. The authors sincerely thank C. C. Berg (Bergen Museum, University of Bergen) for his valuable suggestions and reference and Prof. Dirk C. Albach (Institut für Spezielle Botanik, Johannes Gutenberg-Universität Mainz) for his critical comments on our manuscript. We also thank Dr. Chau-Chin Lin, Fu-Shan Chou and Sheng-Shan Lu (Taiwan Forestry Research Institute) for their help in gathering specimens. The Herbaria of Kunming Institute of Botany, South China Botanical Garden and Guangxi Institute of Botany, and the Chinese Academy of Sciences are thanked for providing support for checking specimens.

References

Alvarez I, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, **29**, 417–434.

Berg CC, Wiebes JT (1992) *African Fig Trees and Fig Wasps*, p. 11. North-Holland, Amsterdam, New York.

Burland TG (2000) DNASTAR's Lasergene sequence analysis software. *Methods in Molecular Biology*, **132**, 71–91.

CBOL Plant Working Group (2009) A DNA barcode for land plants. *Proceedings of the National Academy Sciences of the United States of America*, **106**, 12794–12797.

Chase MW, Cowan RS, Hollingsworth PM *et al.* (2007) A proposal for a standardised protocol to barcode all land plants. *Taxon*, **56**, 295–299.

Chen SL, Yao H, Han JP *et al.* (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One*, **5**, e8613.

China Plant BOL Group (2011) 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*, **108**, 19641–19646.

Corner EJH (1967) *Ficus* in the Solomon Islands and its bearing on the post-Jurassic history of Melanesia. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **253**, 23–159.

Cowan RS, Chase MW, Kress WJ, Savolainen V (2006) 300000 species to identify: problems, progress, and prospects in DNA barcoding of land plants. *Taxon*, **55**, 611–616.

Cuénoud P, Savolainen V, Chatrou LW *et al.* (2002) Molecular phylogenetics of Caryophyllales based on nuclear 18s rDNA and plastid *rbc*L, *atp*B, and *mat*K DNA sequences. *American Journal of Botany*, **89**, 132–144.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19**, 11–15.

Edwards D, Horn A, Taylor D, Savolainen V, Hawkins JA (2008) DNA barcoding of a large genus, *Aspalathus* L. (Fabaceae). *Taxon*, **57**, 1317–1327.

Fan MS, Guan YE, Huang CG (2005) The advances of chemistry and pharmacological study of *Ficus* genus. *Nature Product Research and Development*, **17**, 497–504.

Fazekas AJ, Burgess KS, Kesanakurti PR *et al.* (2008) Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One*, **3**, e2802.

Fazekas AJ, Kesanakurti PR, Burgess KS *et al.* (2009) Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? *Molecular Ecology Resources*, **9**, 130–139.

Ford CS, Ayres KL, Toomey N *et al.* (2009) Selection of candidate coding DNA barcoding regions for use on land Plants. *Botanical Journal of the Linnean Society*, **159**, 1–11.

Harrison RD (2005) Figs and the diversity of tropical rainforests. *BioScience*, **55**, 1053–1064.

Hasebe M, Wolfe PG, Pryer KM (1995) Fern phylogeny based on *rbc*L nucleotide sequences. *American Fern Journal*, **85**, 134–181.

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **270**, 313–321.

Hollingsworth ML, Clark AA, Forrest LL *et al.* (2009) Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. *Molecular Ecology Resources*, **9**, 439–457.

Janzen DH (1979) How to be a fig. *Annual Review of Ecology and Systematics*, **10**, 13–51.

Jousselin E, Rasplus JY, Kjellberg F (2003) Convergence and coevolution in a mutualism: evidence from a molecular phylogeny of *Ficus*. *Evolution*, **57**, 1255–1269.

Kress WJ, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbc*L gene complements the non-coding *psb*A-*trn*H spacer region. *PLoS One*, **2**, e508.

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 8369–8374.

Lahaye R, Savolainen V, Duthoit S, Maurin O, van der Bank M (2008a) A test of *psb*K-*psb*I and *atp*F-*atp*H as potential plant DNA barcodes using the flora of the Kruger National Park as a model system (South Africa). Available from Nature Precedings, http://hdl.handle.net/10101/npre.2008.1896.1 (accessed on 16 May 2008).

Lahaye R, van der Bank M, Bogarin D *et al.* (2008b) DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 2923–2928.

Liu J, Möller M, Gao LM, Zhang DQ, Li DZ (2011) DNA barcoding for the discrimination of Eurasian yews (*Taxus* L.,Taxaceae), and the discovery of cryptic species. *Molecular Ecology Resources*, **11**, 89–100.

Meier R, Shiyang K, Vaidya G, Ng PKL (2006) DNA barcoding and taxonomy in Diptera: a tail of high intraspecific variability and low identification success. *Systematic Biology*, **55**, 715–728.

Newmaster SG, Fazekas AJ, Steeves RAD, Janovec J (2008) Testing candidate plant barcode regions in the Myristicaeac. *Molecular Ecology Resources*, **8**, 480–490.

Nieto FG, Rossello JA (2007) Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants *Molecular Phylogenetics and Evolution*, **44**, 911–919.

Noort SV, Gardiner AJ, Tolley KA (2007) New records of *Ficus* (Moraceae) species emphasize the conservation significance of inselbergs in Mozambique. *South African Journal of Botany*, **73**, 642–649.

Rønsted N, Weiblen GD, Savolainen V, Cook JM (2008) Phylogeny, biogeography, and ecology of *Ficus* section *Malvanthera* (Moraceae). *Molecular Phylogenetics and Evolution*, **48**, 12–22.

Roy S, Tyagi A, Shukla V *et al.* (2010) Universal plant DNA barcode loci may not work in complex groups: a Case study with Indian *Berberis* species. *PLoS One*, **5**, e13674.

Sang T, Crawford DJ, Stuessy TF (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany*, **84**, 1120–1136.

Sass C, Little DP, Stevenson DW, Specht CD (2007) DNA barcoding in the Cycadales: testing the potential of proposed barcoding markers for species identification of *Cycads*. *PLoS One*, **2**, e1154.

Savolainen V, Cowan RS, Vogler AP, Roderick GK, Lane R (2005) Towards writing the encyclopaedia of life: an introduction to DNA barcoding. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **360**, 1805–1811.

Shaw J, Lickey EB, Beck JT *et al.* (2005) The tortoise and the hare II: comparison of the relative utility of 21 non-coding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, **92**, 142–166.

Solomon A, Golubowicz S, Yablowicz Z *et al.* (2006) Antioxidant activities and anthocyanin content of fresh fruits of common fig (*Ficus carica* L.). *Journal of Agricultural and Food Chemistry*, **54**, 7717–7723.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.

Tate JA, Simpson BB (2003) Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Biology*, **28**, 723–737.

Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP (2002) DNA points the way ahead of taxonomy – in assessing new approaches, it's time for DNA's unique contribution to take a central role. *Nature*, **418**, 479.

Tautz D, Arctander P, Thomas RH, Vogler AP (2003) A plea for DNA taxonomy. *Trends in Ecology and Evolution*, **18**, 70–74.

Weiblen GD (2000) Phylogenetic relationships of functionally dioecious *Ficus* (Moraceae) based on ribosomal DNA sequences and morphology. *American Journal of Botany*, **87**, 1342–1357.

White JT, Bruns T, Lee S, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Applications* (eds Innis MA, Gelfand DH, Sninsky JJ & White TJ), pp. 315–322. Academic Press, Inc., New York.

Wiebes JT (1979) Co-evolution of Figs and their insect pollinators. *Annual Review of Ecology and Systematics*, **10**, 1–12.

Xu L, Harrison RD, Yang P, Yang DR (2011) New insight into the phylogenetic and biogeographic history of genus Ficus: vicariance played a relatively minor role compared with ecological opportunity and dispersal. *Journal of Systematics and Evolution*, **49**, 546–557.

Yan HF, Hao G, Hu CM, Ge XJ (2011) DNA barcoding in closely related species: a case study of *Primula* L. Proliferae Pax (Primulaceae) in China sect. *Journal of Systematics and Evolution*, **49**, 225–236.

H.-Q. L., and J.-Y. C. conceived and designed the study. All four authors collected the specimens and H.-Q. L. identified them. J.-Y. C., and S. W. performed the molecular experiments. H.-Q. L., and J.-Y. C. analyzed the data and wrote the paper.

Data Accessibility

DNA Sequences: GenBank accessions JQ773465–JQ774462.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Fig. S1 A tree of *trn*H-*psb*A sequences generated using maximum likelihood (ML) methods with Kimura 2-parameter distances of *Ficus*. Bootstrap values are shown above the relevant branches. The tree was divided into three successive parts (For A, please see Fig. S1a; For B, please see Fig. S1b).

Fig. S2 A tree of *psb*K-*psb*I sequences generated using maximum likelihood (ML) methods with Kimura 2-parameter distances of *Ficus*. Bootstrap values are shown above the relevant branches. The tree was divided into two successive parts (Fig. S2 and S2a).

Fig. S3 A tree of the combined ITS + *trn*H-*psb*A data sets generated using maximum likelihood (ML) methods with Kimura 2-parameter distances of *Ficus*. Bootstrap values are shown above the relevant branches. The tree was divided into three successive parts (For A, please see Fig. S3a; For B, please see Fig. S3b).

Table S1 Voucher information and GenBank accession numbers for species examined in this study.

Table S2 Wilcoxon signed-rank tests of intraspecific divergence among loci

Table S3 Wilcoxon signed-rank tests of interspecific divergence among loci

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.