Evaluation Of Three Candidate DNA Barcoding Loci In Selected *Ficus* L. (Moraceae)

Jay Edneil C. Olivar, Rochelle Y. Brillantes, Rosario R. Rubite, Grecebio Jonathan D. Alejandro

Abstract: The genus *Ficus* L. (Moraceae) is known for its diverse morphology and habit as well as economical, ecological, and pharmaceutical importance. In this study, six Philippine *Ficus* species were utilized to determine the best barcoding loci among cpDNA (rbcL, trnH-psbA) and nrDNA (ITS) markers on the basis of universality, discriminatory power, and resolution of species. Both trnH-psbA and ITS showed 100% PCR success rate and 67% sequencing success, whereas rbcL exhibited 67% PCR success rate but 100% sequencing success. The trnH-psbA marker performed best in terms of discriminatory power, showing the highest variable informative site (71.37%) and the highest mean interspecific distance (48.32% ± 19.89%). ITS ranked next to trnH-psbA since it was able to generate 100% resolution of species and a comparable mean intraspecific distance score with trnH-psbA. Meanwhile, rbcL failed to resolve any species correctly; thus, it has a 0% resolution of species. We initially recommend trnH-psbA and ITS as potential DNA barcodes for molecular authentication of *Ficus* species.

Index Terms— DNA barcoding, ITS, Philippine *Ficus*, *rbcL, trnH-psbA*

1 Introduction

DNA barcoding is a technique that makes use of one or more short gene sequences from a standardized region of a genome to provide quick and reliable identification of species among all forms of life [1]. Aside from introducing a revolutionary advancement in taxonomic and phylogenetic studies, DNA barcoding has many useful applications that may spur the interest from the general public into studies on plant science among various disciplines including ecology, food and drugs safety, conservation biology and environmental protection [2]. Since a DNA barcode is made up of short gene sequence (600 to 800 bp), it has a special advantage of being retrieved and analyzed even from fragments of degraded samples. In this connection, DNA barcoding can ensure consumers whether the constituent plant material which an herbal supplement claims to contain is authentic and not a mere adulterant [3], [4]. In animals, DNA barcoding has been successfully carried out using a universal barcoding locus, the mitochondrial cytochrome oxidase I (MT–COI) gene [5]. In contrast, no single universal barcoding locus has been identified for plants and the MT–COI gene is not applicable due to its slow mutation rate [5]. The Consortium for the Barcoding of Life (CBOL) Plant Working Group proposed the combination of rbcL and matK as the core barcode. However, difficulties in assessing phylogenetic relationships of plants using the proposed barcode led taxonomists to appraise a particular barcoding locus that best applies to a plant taxon [6]. Hence, the public consortium on DNA barcoding suggests that whenever plants are subjected to DNA barcoding protocols, different DNA barcoding loci must first be assessed for its applicability to a specific taxon. The species-rich *Ficus* L. (Moraceae), comprising about 755 species, is among the largest genera of angiosperms [6]. In the Philippines alone, there are 108 indigenous *Ficus* species, 28 of which are endemic according to L. Co's Nominal Checklist of Philippine seed plants. *Ficus* also exhibits vast diversity not only in terms of its morphological traits but also in the wide range of its phytochemical constituents [6]. Lansky et al. (2008) reported that *Ficus* extracts derived from bark, roots, leaves, fruits and latex exhibit cancer preventive, cancer therapeutic and anti-inflammatory activities [7]. Moreover, they found out that *Ficus* contains compounds like phenanthroindolizidine alkaloid, triterpenoid with C-18 carboxylic acid, flavonoids and lectin that are known to mediate anti-inflammatory, cytotoxic, and immune modulator activities [7]. With the rising information on the pharmaceutical importance of *Ficus* species, sooner or later, herbal supplements will be developed. Thus, to prevent fraud and to ensure consumer's health, the application of DNA barcoding to *Ficus* species proves to be a useful endeavor. This starts with the assignment of a DNA barcoding locus appropriate for the identification of species within the taxon. In this study, the applicability of cpDNA (rbcL, trnH–psbA) and nrDNA (internal transcribed spacer ITS) markers was evaluated in selected Philippine *Ficus* species based on universality (PCR and sequencing success) and discriminatory power (mean inter- and intraspecific divergence and resolution of species).

- *Jay Edneil Olivar is currently pursuing masters degree program in Biological Science in The Graduate School, University of Santo Tomas, España, Manila 1015 Philippines, E-mail: jayedneilolivar07@gmail.com*
- *Rochelle Brillantes and Rosario Rubite are currently faculty members of the Department of Biology, College of Arts and Sciences, University of the Philippines Manila, Philippines*
- *Grecebio Jonathan Alejandro is a faculty researcher at the Research Center for the Natural and Applied Sciences, University of Santo Tomas, España, Manila 1015 Philippines*

2 METHODOLOGY

2.1 Collection and Preparation of Plant Material

Six leaf samples of Philippine *Ficus* species (Fig. 1), five of which are endemic (*F. balete* Merr., *F. benjamina* L., *F. odorata* (Blanco) Merr., *F. pseudopalma* Blanco, *F. linearifolia* Elmer, and *F. ulmifolia* Lam.), were collected in Mts. Palay-palay-Mataas na Gulod National Park (Ternate, Cavite), Quezon Province National Park (Atimonan, Quezon), and Mt. Makiling Forest Reserve (Los Banos, Laguna). Leaves of two individuals from different populations were collected for every species and were placed in Ziplock™ bags with silica-gel beads (Chase & Hill, 1991) for DNA sequence analyses.

2.2 DNA Extraction, Amplification, Purification, and Sequencing

Nuclear ITS and two chloroplast genome regions (*rbcL* and tr*nH-psbA*) were evaluated. Genomic DNA was extracted from silica gel-dried leaves using the DNeasy Plant Mini Kit (Qiagen®, Germany) following the manufacturer's protocol. PCR amplifications were performed on Biometra T-gradient. The primers and reaction procedures are listed in Table 1. The PCR products were inspected on 1% TAE agarose gels and subsequently purified using the QIAquick PCR Purification Kit (Qiagen®, Germany). The purified PCR products were sent to MACROGEN, South Korea for bidirectional sequencing

2.3 Data Analysis

PCR success rate was evaluated by computing for the percentage of successfully amplified DNA.Sequences were assembled and edited with Codon Code Aligner v 4.2.5. Sequencing success was then evaluated by computing for the percentage of successfully assembled contigs. All sequences were aligned using ClustalW of MEGA 6.0 package and the alignment adjusted manually using MEGA 6.0 (Tamura et al., 2013).Inter- and intraspecific genetic divergences were calculated using p-distance with pairwise deletion options in MEGA 6.0. One-Way ANOVA was performed following Tukey's Post Hoc test to prove statistical significance between the mean inter- and intraspecific genetic divergences of the three barcoding loci. To further evaluate the effectiveness of barcoding candidates for species discrimination, a Maximum Likelihood (ML) Tree was constructed using Kimura 2-parameter (K2P) distances using MEGA 6.0. Resolution of species was then calculated based on the constructed tree.

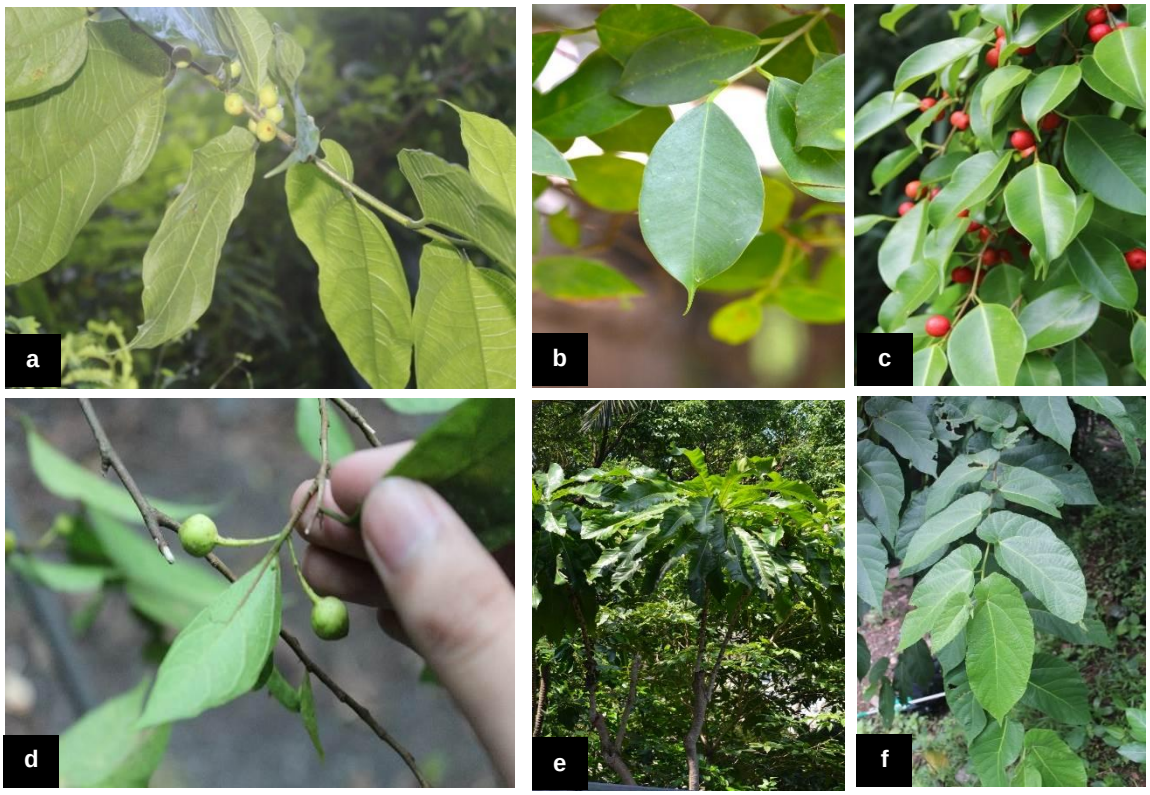

Fig. 1. Photographs of sampled *Ficus* species: (a) *F. ulmifolia* Lam.; (b) *F. balete* Merr.; (c) *F. benjamina* L.; (d) F. linearifolia Elmer (e) *F. pseudopalma* Blanco; and, (f) *F. odorata* (Blanco) Merr.

3 Results

3.1 PCR and Sequencing Success

The efficiency of PCR amplification and sequencing is an important index to evaluate the candidate barcoding locus. Both ITS and *trnH-psbA* exhibited 100% PCR success and 66.67% sequencing success while rbcL exhibited 66.67% PCR success and 100% sequencing success (Table 2). In terms of universality, no barcoding loci used in this study performed well.

3.2 Mean Inter- and Intraspecific Divergence

For individual regions, aligned sequence lengths range from 463 base pairs (bp) for *trnH-psbA* to 815 bp for *rbcL* (Table 2). The trnH-psbA locus contained the most variable sites (71.37%). Our results demonstrated that *trnH-psbA* exhibited the highest mean interspecific divergence (48.32% ± 19.89%). In terms of mean intraspecific divergence, both the ITS and the *trnH-psbA* loci can equally discriminate the same species as evidenced by the Tukey's Post Hoc result (p-value of 0.855 greater than 0.05) which suggests inconclusive statistical difference between the means of ITS and *trnH-psbA* (Table 2 and Figure 2). The *rbcL* locus exhibited an unsatisfactory quality of a barcoding loci-low interspecific divergence and high intraspecific divergence (Table 2).

3.3 Resolution of Species

To evaluate whether species were recovered as monophyletic under each barcode, a ML tree was constructed using K2P distances. Both ITS and *trnH-psbA* were able to have a 100% species resolution (Table 2). The *rbcL* locus failed to resolve any of the samples correctly and therefore had a 0% resolution of species. Figure 3 presents the constructed trees and shows the resolution of species of the three barcoding loci.

TABLE 1

PCR primer sequences and amplification protocols of the DNA barcoding loci

DNA region	Primer Name	Sequence (5'→3')	Amplification Protocol
ITS	ITS 5	GGAAGTAAAAGTCGTAACAAGG	94°C 5 min; 94°C 1 min, 50°C 45 s, 72°C 1 min, 30 cycles; 72°C 5 min
	ITS 4	TCCTCCGCTTATTGATATGC	
trnH-psbA	trnH	CGCGCATGGTGGATTCACAATCC	95°C 4 min; 94°C 30 s, 55°C 1 min, 72°C 1 min, 35 cycles; 72°C 10 min
	psbA	GTTATGCATGAACGTAATGCTC	
rbcL	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
	a_r	CTTCTGCTACAAATAAGAATCGAT CTC	

TABLE 2

Properties of the three candidate DNA barcoding loci

Parameter	ITS	*trnH-psbA*	*rbcL*
PCR success (%)	100.00	100.00	66.67
Sequencing success (%)	66.67	66.67	100.00
Variable Informative Sites (%)	52.13	71.37	61.72
Aligned Length (bp)	681	463	815
Mean Interspecific distance (%)	30.05 ± 12.30	48.32 ± 19.89	33.50 ± 15.71
Mean Intraspecific distance (%)	1.02 ± 1.08	5.62 ± 4.49	40.08 ± 12.11
Resolution of Species (%)	100.00	100.00	0.00

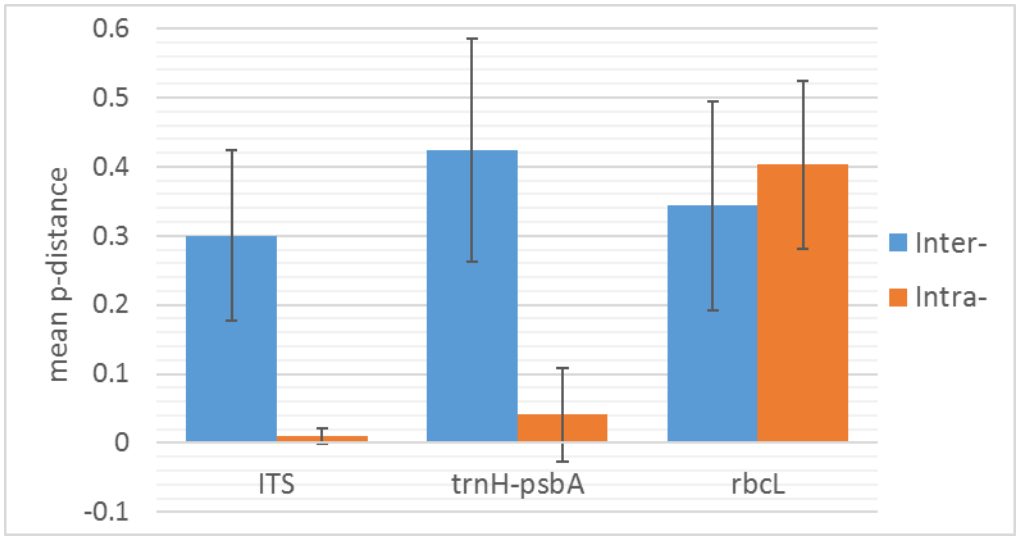

Fig. 2. Mean inter- and intraspecific genetic divergence with standard deviation error bars based on p-distance using MEGA6.0 with pairwise deletion options

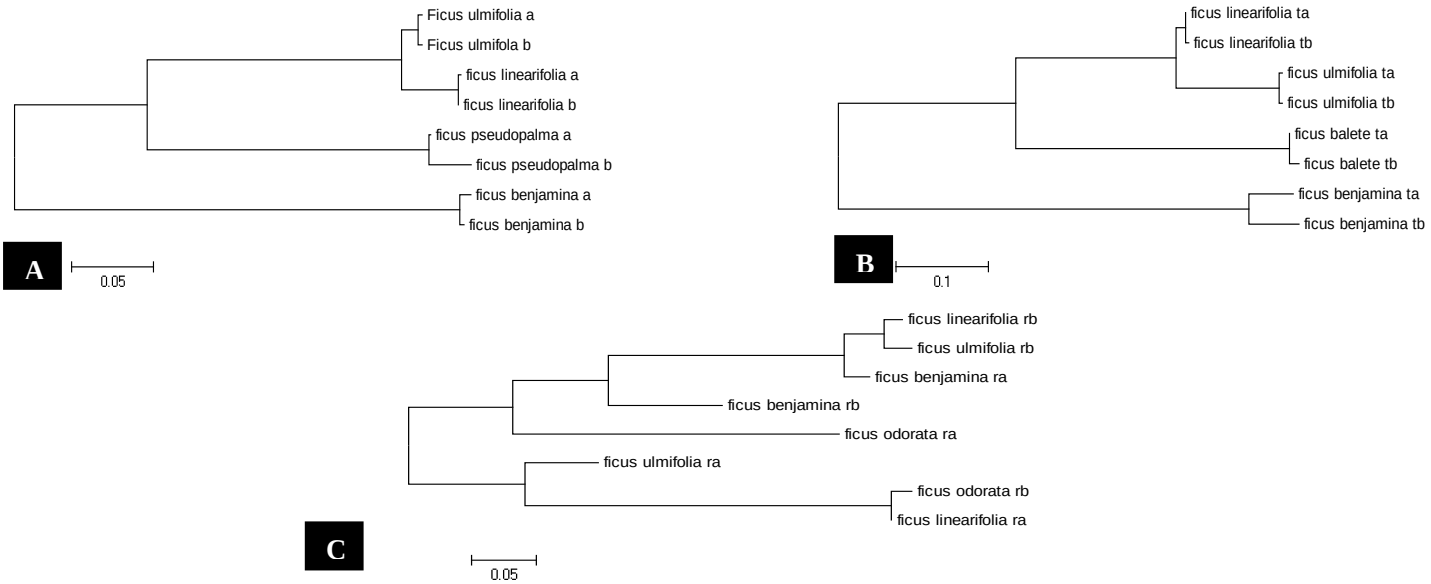

Fig. 3. Maximum likelihood trees constructed using K2P distances of: (A) ITS sequences showing 100% species resolution; (B) trnH-psbA sequences showing 100% species resolution; and, (C) rbcL sequences showing 0% resolution

4 DISCUSSION

4.1 *rbcL*

A high PCR success score is rbcL's major strength which paved the way for its selection by CBOL Plant Working Group as the core barcode together with *matK* [10]. Previous literature reported that *rbcL* performed best in terms of PCR success with an average score not lower than 90%. In fact, Li et al. (2012) derived 98.66% PCR success score using *rbcL* [6]. Several other studies [(e.g., Hasebe, et al., (1995), Hollingsworth, et al., (2009), and Liu, et al., (2011)] proved that rbcL is easily sequenced and aligned in plant groups such as ferns, mosses and angiosperms [11], [12], [13]. In this study, findings on *rbcL*'s PCR success score (67%) proved contrary with Li et al. (2012). Findings might have changed if further PCR adjustments until results for unamplified samples become available. Since amplification of pure DNA was not successful for the three barcode markers (*rbcL, trnH-psbA* and ITS), dilution of DNA samples to 1:10 for second trial was carried out. ITS and trnH-psbA markers were able to yield 100% PCR success score using diluted DNA. To keep a standard level of comparison, the performance of three markers was based on second dilution amplification results. Perhaps *rbcL* could also have generated higher score if 1:100 dilution was tested. As a consequence of rbcL's low discriminatory power, it was not able to resolve any species correctly in the ML tree generated. Based on our results, *rbcL* has low interspecific but exceedingly high intraspecific divergence. Hence, *rbcL* was not able to grouped similar species in our samples. In Li (2011), plastid DNA barcoding loci may not be able to discriminate between closely related species within a genus [10]. They posited that the low performance

of *rbcL* in this parameter is attributed to the low mutation rate of bases in plastid DNA markers [10].

4.2 *trnH-psbA*

PCR success rate for *trnH-psbA* is highly satisfactory supporting previous literatures. In general, success rate for *trnH-psbA* does not exceed that of *rbcL*. It was reported by Shaw et al. (2005) that *trnH-psbA* has 75 -bp highly conserved sequences that are highly advantageous in developing universal primers [14]. However, sequencing problems encountered for *trnH-psbA* could be attributed to prevalence of indels, more variable lengths, and increased difficulty in alignment [11], [1], [10]. In addition, the presence of poly A/T sequences contributes to the low sequence quality [15]. Since *trnH-psbA* had a good discriminatory power based on mean inter- and intraspecific divergence score, it was able to generate a ML tree that correctly grouped similar species.

4.3 ITS

ITS exhibited moderate and in previous studies low PCR success rate. In some plant groups like gymnosperms, ferns and mosses, universal primers failed[5]. The use of primers ITS 3 and ITS 4 did not work for our samples, yet this primer set was recommended by Li H-Q et al [6]. Instead, ITS 4 and ITS 5 worked well and generated a perfect PCR success rate. The ITS marker also presented sequencing difficulties which could be attributed to prevalence of homopolymers and presence of multiple gene copies [15]. In addition, poly G and poly C sites in ITS sequences present themselves as sequencing difficulties [6]. The ITS marker has gained popularity as a barcoding locus in terms of its capacity to resolve species with the highest success rate. Even before DNA barcoding has gained popularity, botanists and taxonomists were already using ITS in phylogenetic studies since it can successfully discriminate different plant species [6]. In Li et al. (2012), ITS has the highest score at 72%. The high discriminatory power of ITS could be attributed to it possession of more variable sites than in plastid DNA [6].

5 CONCLUSION

Based on our results, trnH-psbA is the best performing locus in terms of discriminatory power considering its high variable informative sites and mean interspecific divergence score. Similar with trnH-psbA ITS can equally discriminate species. Hence, these two markers are initially recommended as potential DNA barcodes for molecular authentication of *Ficus* species. A need to increase sample size will develop more conclusive findings.

ACKNOWLEDGMENT

We thank the Thomasian Angiosperm Phylogeny and Barcoding Group (TAPBG) staff for assisting us in laboratory work; Research Center for the Natural & Applied Sciences, University of Santo Tomas for the facilities, Mr. Danny Tandang of the National Museum for plant identification. Mr. Li H-Q for generously sharing literature. GJDA thanks the DOST Philippine Council for Health, Research and Development for the funding.

REFERENCES

[1] W.J. Kress and D.L. Erickson, “DNA barcoding: a windfall for tropical biology?,” BIOTROPICA: The Journal of Tropical Biology and Conservation, vol. 40, no. 4, pp. 405-408, Mar. 2008.

[2] P.M. Hollingsworth, “DNA barcoding: potential user,” Genomics, Society and Policy, vol. 3, no. 2, pp. 44-47, 2007.

[3] N.J.C. Zerega, S. Mori, C. Lindqvist, Q. Zheng, and T.J. Motley, “Using amplified fragment length polymorphisms (AFLP) to identify black cohosh (Actaea racemosa),” Economic Botany, vol. 56, no. 2, pp. 154-164, 2002.

[4] I.B. Schnell, M. Fraser, E. Willerserv, and T.P. Gilbert, “Characterization of insect and plant origins using DNA extracted from small volumes of bee honey,” Arthropod-Plant Interactions, vol. 4, pp. 107-116, Apr. 2010.

[5] D. Rubinoff, S. Cameron, and K. Will, “Are plant DNA barcodes a search for the Holy Grail?,” Trends in Ecology and Evolution, vol. 21, no. 1, pp. 1-2, Jan. 2006.

[6] H.Q. Li, J.Y. Chen, S. Wang, and S.Z. Xiong, “Evaluation of six candidate DNA barcoding loci in Ficus (Moraceae) of China,” Molecular Ecology Resources, vol. 12, no. 5, pp. 783-790 Sep. 2012.

[7] E.P. Lansky, H.M. Paavilainen, A.D. Pawlus, and R.A. Newman, “Ficus spp (fig): Ethnobotany and potential as anticancer and anti-inflammatory agents,” Journal of Ethnopharmacology, vol. 119, no. 2, pp. 195-213, Sep. 2008.

[8] M.W. Chase and H.H. Hills, “Silica gel: An ideal material for field preservation of leaf samples for DNA studies,” Taxon, vol. 40, no. 2, pp. 215-220, May 1991.

[9] K. Tamura, G. Stecher, D. Peterson, A. Filipski, and S. Kumar, “MEGA6: Molecular Evolutionary Genetics Analysis version 6.0,” Molecular Biology and Evolution, vol. 30, no. 12, pp. 2725-2729, Dec. 2013.

[10] D.Z. Li, “Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants,” Proceedings of the National Academy of Sciences of the United States of America, 2011.

[11] M. Hasebe, P.G. Wolf, K.M. Pryer, K. Ueda, M. Ito, R. Sano, G.J. Gastony, J. Yokoyama, J.R. Manhart, N. Murakami, E.H. Crane, C.H. Haufler, and W.D. Hauk, “Fern phylogeny based on rbcL nucleotide sequences,” American Fern Journal, vol. 85, no. 4, pp. 134 -135, Dec. 1995.

[12] M.L. Hollingsworth, A.A. Clark, L.L. Forrest, J. Richardson, R.T. Pennington, D.G. Long, R. Cowan, M.W. Chase, M. Gaudeul, and P.M. Hollingsworth,

"Selecting barcoding loci for plants: evaluation of seven candidate loci with species–level sampling in three divergent groups of land plants," Molecular Ecology Resources, vol. 9, no. 2, pp. 439 – 457, Mar. 2009.

[13] J. Liu, M. Moller, L.M. Gao, D.Q. Zhang, D.Z. Li, "DNA barcoding for the discrimination of Eurasian yews (Taxus L., Taxaceae), and the discovery of cryptic species," Molecular Ecology Resources, vol. 11, no. 1, pp. 89 -100, Jan. 2011.

[14] J. Shaw, E.B. Lickey, J.T. Beck, S.B. Farmer, W. Liu, J. Miller, K.J. Siripun, C.T. Winder, E.E. Schilling, and R.L. Small, "The tortoise and the hare II: relative utility of 21 non-coding chloroplast DNA sequences for phylogenetic analysis," American Journal of Botany, vol. 92, no. 1, pp. 142–166, Jan. 2005.

[15] C. Sass, D.P. Little, D.W. Stevenson, and C.D. Specht, "DNA barcoding in the Cycadales: testing the potential of proposed barcoding markers for species identification of Cycads," PLoS ONE, vol. 2, no. 11, Nov. 2007.