Korean J. Pl. Taxon.
43(4): 252-262 (2013)
http://dx.doi.org/10.11110/kjpt.2013.43.4.252

ISSN 1225-8318
Korean Journal of
Plant Taxonomy

Identification of host plant species of *Balanophora fungosa* var. *indica* from Phnom Bokor National Park of Cambodia using DNA barcoding technique

Joo Hwan Kim and Hyosig Won*

Department of Biological Science, Daegu University, Gyungsan, Gyungbuk 712-714, Korea
(Received 7 October 2013; Accepted 4 November 2013)

캄보디아 프놈보콜국립공원의 *Balanophora fungosa* var. *indica* 의 숙주식물에 대한 DNA barcoding 기법을 통한 동정

김주환 · 원효식*
대구대학교 생명과학과

ABSTRACT: During the floristic survey on Phnom Bokor National Park, Kampot, Cambodia, we encountered *Balanophora fungosa* var. *indica*, which is a tropical holoparasitic plant. To identify its host species, we collected host roots and trees nearby and tried to identify them using DNA barcoding approach. We applied plastid *rbcL* and *matK* gene regions as DNA barcode markers, and successfully amplified and sequenced the markers from 15 host roots and seven tree samples. Obtained host root sequences were identified as Primulaceae, Celastraceae, Myrtaceae, and Oleaceae, while trees nearby are Oleaceae, Myrtaceae, Sapindaceae, Rosaceae, Clusiaceae, Ericaceae, and Lauraceae. At genus level, host species are identified as *Myrsine*, *Euonymus*, *Syzygium*, and *Olea*, but failed in species discrimination. *Myrsine* (Primulaceae) and *Olea* (Oleaceae) are reported here as host species of *B. fungosa* var. *indica* for the first time. Further sampling and comparative work, and DNA barcoding will help recognize the biodiversity of the area and host species of *Balanophora*, together with their evolution.

Keywords: parasitic plants, host plant, *Balanophora fungosa* var. *indica*, DNA barcoding, Phnom Bokor National Park, Cambodia

적 요: 캄보디아 캄폿주 프놈보콜국립공원에 대한 식물상 조사 중 열대성 전기생식물인 *B. fungosa* var. *indica*를 발견하였다. 이들의 숙주를 확인하기 위해 숙주의 뿌리와 더불어 주변에 위치한 목본 식물을 채집하였으며, 이들을 DNA barcoding 방법을 사용하여 동정하였다. DNA barcode 마커로는 엽록체 *rbcL* 및 *matK* 유전자 구간을 적용하였으며, 15개의 숙주 뿌리와 7개의 주변 목본식물로부터 성공적으로 PCR 증폭 및 염기서열을 확보하였다. 숙주 뿌리로부터 얻어진 숙주의 염기서열은 앵초과, 노박덩굴과, 도금양과, 그리고 물푸레나무과로 식별되었으며, 주변의 목본식물은 물푸레나무과, 도금양과, 무환자나무과, 장미과, 물레나물과, 철쭉과와 녹나무과였다. 속 수준에서 앵초과는 *Myrsine*, 노박덩굴과는 *Euonymus*, 도금양과는 *Syzygium*, 물푸레나무과는 *Olea* 등으로 각각 식별되었으나, 종 수준에서의 동정은 불가능하였다. 앵초과 *Myrsine*와 물푸레나무과 *Olea*는 본 연구를 통해 최초로 *B. fungosa* var. *indica*의 숙주로 확인되었다. 추가적인 채집 조사 및 비교 연구, DNA barcoding을 통해 해당 지역의 생물다양성과 *Balanophora*속 식물의 숙주 식물 및 진화에 대해 좀더 명확하게 확인이 가능할 것으로 판단된다.

주요어: 기생식물, 숙주식물, *Balanophora fungosa* var. *indica*, DNA barcoding, 프놈보콜국립공원, 캄보디아

*Author for correspondence: wonhs@daegu.ac.kr

http://www.pltaxa.or.kr

Balanophoraceae are root holoparasites, characterized by underground tuber-like structure and fleshy, club-shaped inflorescences of fungoid appearance (Kujit, 1969). The color of plant ranges from pale yellows and browns through various hues of pink and red to deep purple, but no green due to lack of chlorophyll. Balanophoraceae is mostly tropical and consisted of 17 genera of ca. 50 species (Kujit, 1969; Huang and Murata, 2003). Balanophoraceae is placed in Santales of APG III (The Angiosperm Phylogeny Group, 2009), following the results of phylogenetic study (Nickrent et al., 2005)

Balanophora is mainly distributed in temperate and tropical Asia, and extends to Africa, Madagascar, Australia, and Pacific Islands, with about 19 species (Hansen, 1972). *Balanophora* species are either dioecious or monoecious, with male flowers are regularly arranged directly on the fleshy main axis and female flowers are inserted not only on the stalk of the spadicles (claviform bodies, according to Eberwein et al., 2009), but also around its base, on the main inflorescence axis. Female flower of *Balanophora* is an example of extreme reduction in floral structure, in that entire flower is reduced to a structure like the archegonium of a moss, without any trace of vasculation or perianth (Hansen, 1972). Vascular strands of the host reach far into the mature tuber of *Balanophora*, by forming composite bundles (Hsiao et al., 1995). Ants, cockroaches, and pyriad moths have been reported as visitors to the inflorescences of *B. kuroiwai* and *B. tobiracola* in subtropical Japan by Karakita and Kato (2002), although these two species basically undergo self-pollination through geitonogamy. Ridley (1930) suggested fruit-eating animals as seed disperser of *Balanophora* or wind-dispersal, while Kujit (1969) disagreed with the wind-dispersal.

We have been surveying seed plants of Cambodia from 2010, together with Forest Administration, Government of Cambodia. In January 2012, we discovered *B. fungosa* J.R. Forst. & G. Forst. subsp. *indica* (Arn.) B. Hansen var. *indica* B. Hansen from Bokor National Park, Kampot, Cambodia. Identification of *Balanophora* host is not easy task as stated by Hansen (1972): "It takes prolonged digging in the forest floor where an immense amount of roots are traveling. In two cases only I have succeeded in tracing the host root of *B. fungosa* ssp. *indica* directly to *Cissus* and a *Tetrastigma*. In other cases the host-roots broke." Because of the difficulty, Hansen (1972) anticipated revision of his Table 5, where the host species known to that time were compiled, and add-up of new host species. Host species of 13 *Balanophora* species are at least 70 species belonging to 35 families, and 12 species of 8 families for *B. fungosa* var. *indica* (Hansen, 1972).

DNA barcoding has been designed and applied to identify biological specimens, by amplifying and sequencing short gene sequence (Hebert et al., 2003; Kress et al., 2005). Recent studies have applied the method to ecological and community-level study, in addition to biodiversity assessment and species identification (Lahaye et al., 2008; Fazekas et al., 2008; Valentini et al., 2008; Kress et al., 2009; Yu et al., 2011b). Su et al. (2012) identified the host species of *B. japonica* and *B. yakushimensis* by amplifying chloroplast *matK* sequences from the connected root tissues. Here we are presenting the host species of *B. fungosa* var. *indica* from Cambodia, identified by DNA barcoding technique. This will broaden our knowledge on the host species diversity and evolution of the group.

Material and Methods

Collection sites: The collection sites are located at the Phnom Bokor National Park, which forms the part of the Elephant Mountain Range (Dâmrei Mountains) located on the southwestern coast of Cambodia. The highest elevation of the area is 1,081 m. The average rainfall is 3,800-5,000 mm per year on the western slopes facing the shores of the Gulf of Thailand, while only 1,020-1,520 mm on the eastern slopes (http://en.wikipedia.org/wiki/Dâmrei_Mountains). During the fieldwork at Phnom Bokor National Park, Kampot, Cambodia on January 13, 2012, first collection of *B. fungosa* was made (Table 1; Figs. 1 & 2A). As *B. fungosa* was in flowering, it could be easily recognized from the forest bed (Fig. 2A). The vegetation type of the collection site is a dwarf evergreen hill

Table 1. Collection information of *Balanophora fungosa* var. *indica* and its putative host plants from Phnom Bokor National Park, Kampot, Cambodia.

Site	GPS coordinates	Altitude	Coll. Date	Host plants collected		Voucher
				no. of branches	no. of roots	
1	N10° 38' 06.5" E104° 02' 14.6"	982 m	2012/01/13	–	–	*Won 6770* (DGU)
2	N10° 38' 19.3" E104° 00' 31.1"	1059 m	2013/01/06	6	5	*Won 9081* (DGU)
3	N10° 38' 19.9" E104° 00' 30.6"	1060 m	2013/01/07	7*	14	*Won 9203* (DGU)
Total				13	19	

*includes one tree traced from the root parasitized by *B. fungosa*.

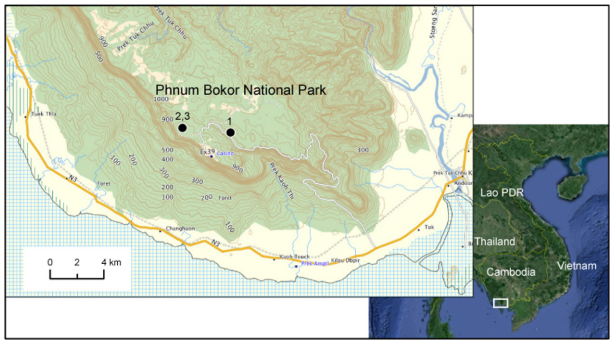

Fig. 1. Map showing study sites at Phnom Bokor National Park, Kampot, Cambodia. Black dots and numbers on them indicate collection sites of *Balanophora fungosa* var. *indica* and its putative host plants.

Fig. 2. *Balanophora fungosa* var. *indica*. A. Growth habit; B. Examples of excavated *Balanophora* samples; C. Female plant with cut tuber showing composite bundle spreading from the root of host plant; D. Tuber and root of the host plant attached. f, female plant; m, male plant; r, root of the host plant; t, tuber; br, bract; cb, composite bundle.

forest around 1,000 m high of which canopy is less than 15 m tall, according to the vegetation categories of Meng et al.(2000). The growth of the tree is limited due to the high wind blowing over the steep cliff from the ocean. The area often becomes misty, cloudy, or rainy from the moist provided from the ocean. The major rock formation of Bokor Plateau is sandstone and the soil is mainly composed of sand eroded from the sandstone. The second collection of *B. fungosa* and its putative host plant species was done at sites 2 and 3 between January 6 and 7, 2013 (Table 1; Figs. 1 & 2). The two sites are close-by, and about 3 km west from the first collection site of 2012. The vegetation type and the physical condition of the latter two sites are almost identical to those of the first collection sites, only slightly higher in elevation.

Materials: Whole plant body of *B. fungosa* was carefully dug out in the field, together with host roots attached to the tuber, and washed carefully with toothbrush under flowing water to remove dirt completely (Fig. 2). The host roots were easily recognized since the host roots end up inside the *B. fungosa* tuber and develop composite bundles (Fig. 2C), while *B. fungosa* has no roots. The diameter of the host roots ranged 2-5 mm. At least 5 cm long host roots were sampled per collection and stored in silica-gel until DNA extraction. The remaining tissues and *B. fungosa* were made into voucher specimen or liquid-collection, and stored at the Herbarium of Department of Biological Science, Daegu University (DGU). We tried to identify the host tree by tracing the whole length of the root attached, but only succeeded in one case after digging-up over 15 m.

In addition, branches of tree species, whose DBH (diameter at breast height) was over 5 cm and located less than ca. 5 m from the *B. fungosa* were collected as putative host species of *B. fungosa*. All the tree species sampled lacked flower and fruits, unfortunately, which rendered difficulties in identification.

In total, 19 host roots from ten male and nine female *B. fungosa* individuals and 13 branches of host candidates were obtained from the two sites (Table 1), with one host tree confirmed from the root tracing.

DNA Barcoding & identification: For DNA extraction, PCR amplification of DNA barcode markers, and sequencing of the amplified barcode markers, we followed the methods used by Won (2009). As DNA barcode markers, we applied plastid *rbcL* and *matK* genes, since these are one of the most extensively sequenced and recommended barcoding markers by multiple sources (Kress et al., 2005; Newmaster et al., 2006; Fazekas et al., 2008; CBOL Plant Working Group, 2009; Li et al., 2011), although these genes are more or less conserved. For the PCR amplification of the barcode regions, we used matK_472F and matK_1248R (Yu et al., 2011a), or matK_3F and matK_3R primer combinations (Sang et al., 1997; correspond to base position 518-1268 from the starting position) for *matK* region, and rbcL_1F and rbcL_1460R primer combination (Fay et al., 1998) for *rbcL* region, respectively. We applied only forward primers, matK_472F, matK_3F, and rbcL_1F, for sequencing. When both the two *matK* primer combinations successful amplified, both PCR products were sequenced. Obtained chromatograms were checked and contiged in Sequencher (ver. 4.9; Gene Codes, USA). Obtained *rbcL* and *matK* gene sequences of host species were compared with sequences of GenBank by BLAST and also with BOLD System database for best matches. Sequences

of the top ten best matches and the five distinct genera from the BLAST search were analyzed by neighbor-joining to confirm the phylogenetic position of the host species sequences. Bootstrap support was obtained by 1000 re-sampling in PAUP* 4.0b10 (Swofford, 2002). Geographic distribution of the putative host genera was consulted with Mabberley (1997), Fu (2012), and sometimes with internet resources such as Tropicos (http://www.tropicos.org/), USDA Plants Database (http://plants.usda.gov/), Encyclopedia of Life website (http://www.eol.org/), and Wikipedia (http://www.wikipedia.org/).

Results

Identification of *B. fungosa*: Those *B. fungosa* individuals collected from the Phnom Bokor National Park, Kampot, Cambodia were identified as *B. fungosa* subsp. *indica* var. *indica*. Subsp. *indica* is distinguished from subsp. *fungosa* in having unisexual inflorescences, while subsp. *fungosa* has bisexual inflorescences with male flowers below female part. Within subsp. *indica*, Hansen (1972) recognized three varieties – var. *indica*, var. *globosa* (Jungh.) B. Hansen, and var. *minor* (Eichler) B. Hansen. Var. *globosa* has female inflorescence more than half of it is covered with leaves (no male plants have been observed yet), and extremely short stem. The tuber of var. *minor* is consisted of elongated, cylindrical, repeatedly branching parts forming an enlarged mass. On the contrary, var. *indica* has elongated stem with clearly exposed inflorescences and mass-form tuber. All the specimens collected from the Phnom Bokor National Park, Kampot, Cambodia matched well with the description of var. *indica* (Fig. 2). Hansen (1972, 1973) reported the distribution of *B. fungosa* subsp. *fungosa* from the Celebes and the Philippines eastwards up to Australia and Fiji, while subsp. *indica* mainly in the Indian and Indo-Chinese subcontinents which include India, Myanmar, Thailand, Lao PDR, Vietnam, Cambodia, Hainan, Malaya, Sumatra, and Australia, showing fairly separate distribution. Var. *minor* has been collected from India and Thailand, but rarely, and var. *globosa* only from Java, Indonesia. Hansen (1972, 1973) also reported collection of *B. fungosa* subsp. *indica* var. *indica* from the Phnom Bokor National Park, Kampot, Cambodia.

PCR amplification of barcode markers: We succeeded in amplifying plastid *rbcL* and *matK* gene regions from the host roots and the putative host trees nearby, but not all of them. We were able to amplify *rbcL* and *matK* regions from two roots collected from the site 2 and 13 from the site 3, respectively, with overall success rate of 78.9% (15 success out of 19). In case of *matK* region, primer combination matK_472F–matK_1248R didn't work for root samples later identified as Primulaceae (see below and Table 2), while primer combination matK_3F–matK_3R didn't work for those of Celastraceae, vice versa. We were not so successful in amplifying *rbcL* and *matK* regions from the putative host trees, probably due to bad quality of extracted DNA from leaf samples. The sampled branches of the putative host trees nearby mostly had only old leaves and troubled DNA extraction by excess of sticky material when ground in extraction buffer. We could only amplify seven samples for *rbcL* regions (53.8% success rate) and six samples for *matK* regions (46.2% success rate), respectively, out of thirteen.

Identification of host species: We obtained 15 *rbcL* and *matK* sequences each from the host roots, and seven *rbcL* and six *matK* sequences from the putative host trees nearby, respectively. The obtained *rbcL* sequences ranged 797-886 bp and *matK* 504-732 bp, respectively, after trimming, clearing and aligning in Sequencher program. All the *rbcL* sequences obtained from the host roots samples were grouped into four sequence types, and also the *matK* sequences were grouped into four sequence types, from at least three host root samples per type (Table 2). On the contrary, the seven *rbcL* sequences obtained from the nearby trees were categorized into six sequence types and the six *matK* sequences into five sequence types (Table 3). These obtained *rbcL* and *matK* sequence types were searched in GenBank and BOLD System database for best matches, and the results are summarized in Tables 2 and 3. Also, the neighbor-joining trees showing the phylogenetic position of the host species sequences are presented in Fig. 3. The four *rbcL* and *matK* sequence types obtained from the host roots were identified as members of Primulaceae, Celastraceae, Myrtaceae, and Oleaceae. However, the six *rbcL* and five *matK* sequence types obtained from the putative host trees nearby were identified as Oleaceae, Myrtaceae, Sapindaceae, Rosaceae, Clusiaceae, Ericaceae, and Lauraceae. The tree traced from the host root was identified as Oleaceae, identical with the result from the host root, and Myrtaceae was also common both to host roots and nearby trees. The 100% sequence similarities were observed from the two *matK* sequence types, which matched with *Syzygium* spp. (*S. buxifolium*, *S. hancei*, *S. rehderianum*, and *S. levinei*; Myrtaceae) and *Mischocarpus pentapetalus* (Sapindaceae). Except for the two cases, the maximum similarities ranged below 99.88%, which corresponds to at least one base pair difference to the whole length of the sequences obtained. All the host root *rbcL* and *matK* sequences obtained from the site 2, two and three, respectively, were identified as Primulaceae, while those from site 3 were Primulaceae, Celastraceae,

Table 2. List of host species of *Balanophora fungosa* var. *indica* identified by the DNA barcode sequences of the host roots. DNA barcode sequences were compared with GenBank BLAST and BOLD System. Numbers in the parentheses are maximum match scores and maximum similarities, respectively.

rbcL		*matK*	
GenBank BLAST	BOLD System	GenBank BLAST	BOLD System
Type 1 (n = 4; site 2, n = 2; site 3, n = 2; 797 bp)		Type 1 (n = 5; site 2, n = 3, site 3, n = 2; 673 bp)	
Primulaceae	**Primulaceae**	**Primulaceae**	**Primulaceae**
Rapanea sp. (1458/99.75%)	*Ardisia* sp. (773/98.49%)	*Myrsine* sp. (1238/99.85%)	*Rapanea* sp. (671/99.85%)
Myrsine sp. (1445/99.37%)		*Ardisia* sp. (1227/99.55%)	*Myrsine* sp. (671/99.85%)
Primula sp. (1411/98.62%)			
†*Stylogyne* sp. (1411/98.62%)			
Type 2 (n=3, site 3; 869 bp)		Type 2 (n=3, site 3; 729 bp)	
Celastraceae	**Celastraceae**	**Celastraceae**	**Celastraceae**
†*Paxistima* sp. (1583/99.54%)	†*Paxistima* sp. (861/99.54%)	*Euonymus* sp. (1336/99.73%)	*Euonymus* sp. (725/99.73%)
Euonymus sp. (1578/99.42%)	*Euonymus* sp. (859/99.42%)	*Glyptopetalum* sp. (1325/99.45%)	*Glyptopetalum* sp. (721/99.45%)
†*Maurocenia* sp. (1578/99.42%)	†*Maurocenia* sp. (859/99.42%)	†*Crossopetalum* sp. (1319/99.04%)	†*Crossopetalum* sp. (718/99.45%)
Maytenus sp. (1578/99.42%)	*Gymnosporia* sp. (859/99.42%)		*Gymnosporia* sp. (717/99.18%)
Type 3 (n=3, site 3; 886 bp)		Type 3 (n=3, site 3; 730 bp)	
Mytraceae	**Myrtaceae**	**Myrtaceae**	**Myrtaceae**
Syzygium sp. (1620/99.66%)	*Syzygium* sp. (880/99.66%)	*Syzygium* spp.* (1349/100%)	*Syzygium* sp. (726/99.73%)
†*Myrcianthes* sp. (1598/99.21%)		†*Piliocalyx* sp. (1321/99.32%)	
Eucalyptus sp. (1587/98.98%)			
†*Stockwellia* sp. (1581/98.87%)			
Type 4 (n=5, site 3; 875 bp)		Type 4 (n=5, site 3; 732 bp)	
Oleaceae	**Oleaceae**	**Oleaceae**	**Oleaceae**
Olea sp. (1605/99.77%)	*Olea* sp. (871/99.77%)	*Fraxinus* sp. (1325/99.32%)	*Olea* sp. (728/99.73%)
†*Forestiera* sp. (1594/99.54%)	†*Forestiera* sp. (867/99.54%)	*Osmanthus* sp. (1325/99.32%)	†*Nestegis* sp. (724/99.45%)
†*Picconia* sp. (1589/99.43%)	†*Picconia* sp. (865/99.43%)	†*Picconia* sp. (1319/99.18%)	†*Picconia* sp. (724/99.45%)
†*Comoranthus* sp. (1589/99.43%)	†*Comoranthus* sp. (865/99.43%)		†*Notelaea* sp. (724/99.45%)

*100% matches were observed for *S. buxifolium* (HQ427387), *S. hancei* (HQ415316), *S. rehderianum* (HQ415315), and *S. levinei* (HQ415313).

†indicates genus distributed outside of Indochina and SE Asia.

Table 3. List of putative host tree species of *Balanophora fungosa* var. *indica* identified by DNA barcoding approach. Numbers in the parentheses are maximum match scores and maximum similarities, respectively.

	rbcL		*matK*	
	GenBank BLAST	BOLD System	GenBank BLAST	BOLD System
Tree 1* (n=1, site 3; 852 bp)			(n=1, site 3; 723 bp)	
	Oleaceae	**Oleaceae**	**Oleaceae**	**Oleaceae**
	Olea sp. (1568/99.88%)	*Olea* sp. (850/99.88%)	*Fraxinus* sp. (1317/99.31%)	*Olea* sp. (721/99.73%)
	†*Forestiera* sp. (1557/99.65%)	†*Forestiera* sp. (846/99.65%)	*Osmanthus* sp. (1317/99.31%)	†*Nestegis* sp. (720/99.45%)
	†*Picconia* sp. (1552/99.53%)	†*Picconia* sp. (844/99.53%)	†*Picconia* sp. (1312/99.18%)	†*Picconia* sp. (720/99.45%)
	†*Comoranthus* sp. (1546/99.41%)	†*Comoranthus* sp. (842/99.41%)		†*Notelaea* sp. (720/99.45%)
Tree 2 (n=1, site 3; 872 bp)			(n=1, site 3; 720 bp)	
	Myrtaceae	**Myrtaceae**	**Myrtaceae**	**Myrtaceae**
	Syzygium sp. (1594/99.66%)	*Syzygium* sp. (866/99.66%)	*Syzygium* sp. (1325/99.86%)	*Syzygium* sp. (714/99.58%)
	Eucalyptus sp. (1570/99.2%)	*Eucalyptus* sp. (856/99.08%)		
	†*Myrcianthes* sp. (1572/99.2%)			
Tree 3 (n=1, site 2; 817 bp)			(n=1, site 2; 723 bp)	
	Sapindaceae	**Sapindaceae**	**Sapindaceae**	**Sapindaceae**
	Mischocarpus sp. (1498/99.76%)	*Mischocarpus* sp. (811/99.63%)	*Mischocarpus pentapetalus*‡ (1336/100%)	Mischocarpus *pentapetalus* (723/100%)
	Diploglottis sp. (1493/99.63%)	*Diploglottis* sp. (809/99.51%)	†*Mischarytera* sp. (1314/99.45%)	†*Mischarytera* sp. (715/99.45%)
	Guioa sp. (1411/99.51%)	*Cupaniopsis* sp. (807/99.39%)	*Diploglottis* sp. (1314/99.45%)	*Diploglottis* sp. (715/99.45%)
	†*Matayba* sp. (1411/99.51%)			
Tree 4 (n=1, site 3; 886 bp)			(n=2; site 2, n=1; site 3, n=1; 504 bp)	
	Rosaceae	**Rosaceae**	**Rosaceae**	**Rosaceae**
	Eriobotrya sp. (1626/99.77%)	*Eriobotrya* sp. (882/99.77%)	*Rhaphiolepis* sp. (909/99.21%)	*Rhaphiolepis* sp. (496/99.21%)
	Cotoneaster sp. (1615/99.66%)	*Cotoneaster* sp. (877/99.66%)	*Eriobotrya* sp. (898/98.81%)	
	Rhaphiolepis sp. (1615/99.55%)	*Rhaphiolepis* sp. (878/99.55%)	*Sorbus* sp. (898/98.81%)	
Tree 5 (n=2, site 2; 862 bp)				
	Clusiaceae	**Clusiaceae**	–	–
	Garcinia sp. (1565/99.42%)	*Garcinia* sp. (852/99.42%)		
	†*Allanblackia* sp. (1537/99.29%)			
	†*Symphonia* sp. (1526/98.61%)			

Table 3. Continued.

	rbcL		*matK*	
	GenBank BLAST	BOLD System	GenBank BLAST	BOLD System
Tree 6 (n=1, site 3; 878 bp)				
	Ericaceae *Vaccinium* sp. (1561/98.75%)	**Ericaceae** *Vaccinium* sp. (856/98.75%)	–	–
Tree 7			(n=1, site 3; 727 bp)	
	–	–	**Lauraceae** *Beilschmiedia* sp. (1338/99.86%) *Endiandra* sp. (1315/99.31%) *Potameia* sp. (1314/99.17%)	**Lauraceae** *Beilschmiedia* sp. (725/99.86%) *Endiandra* sp. (717/99.31%)

*Tree 1 is the tree traced from the host root in site 3.
†indicates genus distributed outside of Indochina and SE Asia.
‡GenBank accession number is HQ415242.

Myrtaceae, and Oleaceae. The nearby trees of site 2 were identified as Sapindaceae, Rosaceae, and Clusiaceae, while those of site 3 were Oleaceae, Myrtaceae, Rosaceae, Ericaceae, and Lauraceae.

Discussion

The host species of *B. fungosa* var. *indica* identified by DNA barcoding approach broadly corresponds well with those compiled and suggested by Hansen (1972), but with newly identified host families and genera. Hansen (1972) compiled the host species of 13 *Balanophora* species to be at least 70 species belonging to 35 families, and 12 species of 8 families for *B. fungosa* var. *indica*: *Carissa carandas* L. (Apocynaceae), *Ilex wightiana* Wall. ex Wight (Aquifoliaceae), *Euonymus crenulatus* Wall. (Celastraceae), *Acacia melanoxylon* R. Br. (Fabaceae), *Albizzia lophantha* (Willd.) Benth. (Fabaceae), *Millettia* sp. (Fabaceae), *Pithecellobium* sp. (Fabaceae), *Ficus* sp. (Moraceae), *Barringtonia asiatica* (L.) Kurz (Lecythidaceae), *Syzygium cuminii* (L.) Skeels (Myrtaceae), *Cissus* sp. (Vitaceae), and *Tetrastigma* sp. (Vitaceae). In addition, Lauraceae, Rutaceae, Ebenaceae, Oleaceae, and Verbenaceae have been also reported as host of *B. fungosa*, as presented on the angiosperm phylogeny (Fig. 4). Primulaceae (including Myrsinaceae, sensu APG III) is the first report as a host of the genus *Balanophora*, and Oleaceae is the first as a host of *B. fungosa* var. *indica* since Oleaceae has only been reported for subsp. *fungosa*. The Primulaceae may be a member of *Myrsine* or *Ardisia*, as *Rapanea* Aubl. has been transferred to *Myrsine* L. by Chen and Pipoly (1996) and Ricketson and Pipoly (1997). Genus *Stylogyne* is unlikely host since it is distributed solely in North and South America (Table 2). In case of Oleaceae, the best matches were *Olea*, *Fraxinus*, or *Osmanthus*, since *Forestiera* (North America), *Picconia* (Europe and North America), *Comoranthus* (Africa), *Nestegis* (Australia and North America), and *Notelaea* (Australia) are unlikely. In case of Celastraceae and Myrtaceae, *Euonymus* and *Syzygium* correspond with Hansen (1972), respectively, as *E. crenulatus* and *S. cuminii* have already been reported. However, *Paxistima* (North and South America), *Maurocenia* (Africa), and *Crossopetalum* (North America) of Celastraceae, *Myrcianthes* (North and South America), *Stockwellia* (Australia), and *Piliocalyx* (Australia) of Myrtaceae are unlikely candidate for the host.

So far, we have not had chance to collect *Myrsine* species from Cambodia or have been unsuccessful in recognizing it, for Primulaceae, while seven *Ardisia* species, four from Bokor N.P. (*A. crenata*, *A. sanguinolenta*, *A. quinquegona*, and *A. smaragdina*), have been collected. In case of Celastraceae, we have collected *Euonymus cochinchinensis* at the foot hill of Bokor N.P., but not in the research site yet. In case of Myrtaceae, we have collected five *Syzygium* species, three from Bokor N. P. (*S. megacarpum*, *S. bokorense*, and *S. syzygioides*). In case of Oleaceae, only one species, *Olea macrophylla*, has been collected from Cambodia, but none yet from Bokor N.P. These suggest that there are still many unrecognized and un-

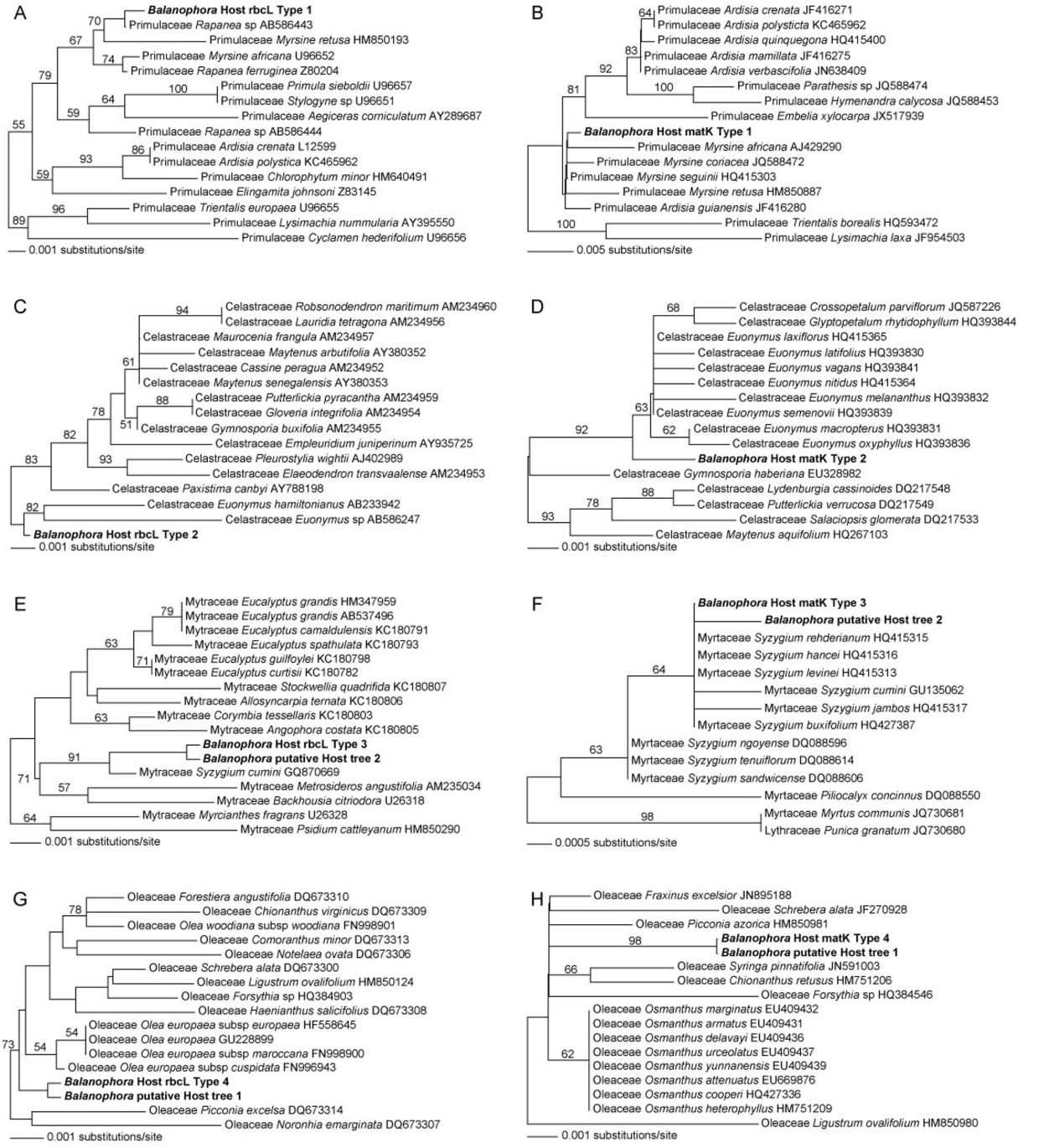

Fig. 3. Neighbor-joining trees of the best match sequences obtained from the GenBank BLAST search of the chloroplast *rbcL* and *matK* sequences of the host roots. Numbers on the branches indicate bootstrap supports. A, C, E, G: cp *rbcL* sequence trees; B, D, F, H: cp *matK* sequence trees.

reported species in Bokor N.P. and further floristic work and taxonomic study, also with DNA barcoding, will help identify the host species of *Balanophora* and enlighten the biodiversity of the area.

By the way, those trees collected nearby the *Balanophora* were identified as Oleaceae, Myrtaceae, Sapindaceae, Rosaceae, Clusiaceae, Ericaceae, and Lauraceae, by DNA barcoding. Except for Oleaceae and Myrtaceae, which were also identified as root hosts, the other tree species lack evidence as hosts of *B. fungosa* var. *indica*. Tree 7, identified as Lauraceae, although Lauraceae was enlisted as host of *B. fungosa* var. *minor*, *Beilschmiedia* was not in the list of Hansen

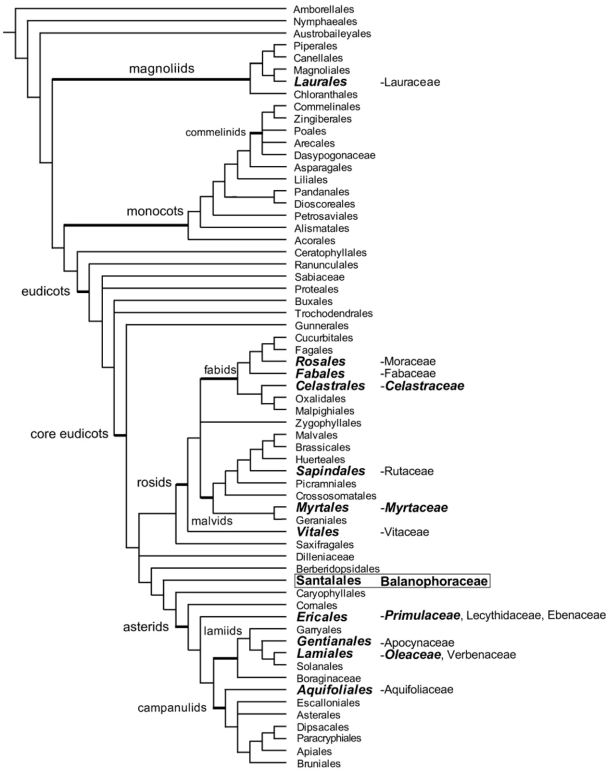

Fig. 4. Taxonomic distribution of host species of *Balanophora fungosa*, presented on the angiosperm phylogeny modified from the APGIII (The Angiosperm Phylogeny Group, 2009) at ordinal level. Families of host species of *B. fungosa*, compiled by Hansen (1972), are listed next to the corresponding order. The four host families of *B. fungosa* var. *indica* identified from the Phnom Bokor National Park, Kampot, Cambodia are specified as bold italic characters.

(1972). Likewise in Table 2, *Picconia* (Europe and North America), *Comoranthus* (Africa), *Nestegis* (Australia and North America), and *Notelaea* (Australia) of Oleaceae, *Myrcianthes* (North and South America) of Myrtaceae, *Matayba* (North and South America, *Mischarytera* (Australia) of Sapindaceae, *Allanblackia* (Africa) and *Symphonia* (Africa and South America) of Clusiaceae are unlikely host of *B. fungosa* var. *indica*, since they simply don't occur in the area. However, *Olea* (Oleaceae), *Syzygium* (Myrtaceae), *Eriobotrya* (Rosaceae), *Rhaphiolepis* (Rosaceae), *Garcinia* (Clusiaceae), and *Vaccinium* (Ericaceae) have been collected from Cambodia.

Although the number of samples and research sites are limited, current study is comparable with other researches on *Balanophora* hosts. For example, Su et al. (2012) identified the host species of *B. japonica* and *B. yakushimensis* by amplifying chloroplast *matK* sequences from the connected root tissues. They found that there exist differences in host between the two *Balanophora* species depends on species and geographic distribution; *B. japonica* parasitizes on *Symplocos* (Symplocaceae) species both in Japan and Taiwan; *B. yakushimensis* parasitizes on *Distylium racemosum* (Hamamelidaceae) in Japan, while on *Schima superba* (Theaceae) in Taiwan population. Also, Azukawa (1982) reported the host species of *B. tobiracola* are *Pittosporum tobira* (Pittosporaceae), *Ligustrum japonicum* (Oleaceae), and *Rhaphiolepis indica* var. *umbellata* (Rosaceae), and those of *B. kuroiwai* are *Pongamia pinnata* (Fabaceae) and *Macaranga tanarius* (Euphorbiaceae). These suggests that there may exists host specificity depends on the species and geographic distribution, even though host species of *B. fungosa* is quite diverse and geographic distribution of it is second widest of all *Balanophora*, following *B. abbreviata*. As shown in Fig. 4, the host species of *B. fungosa* are widely distributed in core eudicots, except Lauraceae in magnoliids. The current study also indicates that the hosts of *B. fungosa* var. *indica* in fabids (Celastraceae), malvids (Myrtaceae), basal asterids (Primulaceae), and lamiids (Oleaceae). This suggests that the host specificity of *B. fungosa* is rather broad, but still requires comprehensive sampling and comparative study to verify if there exists any correlation with geographic distribution and host specificity, especially with DNA barcoding approach. Hansen (1972, 1973) also reported the collection of *B. fungosa* var. *indica* from Phnom Samkos, in addition to Bokor N.P., and *B. latisepala* from Bokor N.P. There are several places in Cambodia with similar vegetation type and climate conditions to Bokor N.P., such as Phnom Samkos, Central Cardamom mountain ranges, and Phnom Aoral. Further sampling and comparative work will help understand the evolution of host specificity and evolution of the group.

The DNA barcoding markers applied in this study, the *rbcL+matK* combination, can only identify two cases with 100% sequence similarity at genus/species level. CBOL Plant Working Group (2009) presented that the *rbcL+matK* combination has successfully discriminated 72% of cases at specific level, with remaining species being matched to groups of congeneric species with 100% success. However, China Plant BOL Group et al. (2011) presented only 49.7% species discrimination using the *rbcL+matK* combination. Plastid *rbcL* gene is the one of the best characterized gene and can be amplified and sequenced across all land plants with ease. Newmaster et al. (2006) could discriminate approximately 85% congeneric species using *rbcL* sequences in GenBank. On the contrary, plastid *matK* gene is one of the most rapidly evolving plastid coding regions and consistently show high levels of discrimination among angiosperm species, though *matK* alone provided 66% species discrimination (CBOL Plant Working Group, 2009). Because of this, more rapidly evolving barcode markers have been recommended in addition to the two markers

for species level identification through DNA barcoding (Kress and Erickson, 2007; CBOL Plant Working Group, 2009; Chen et al., 2010; China Plant BOL Group et al., 2011). Therefore, the lack of 100% identity of DNA barcodes may have been originated from the nature of the two barcodes markers, and/or from incomplete taxon sampled in the GenBank and BOLD System database. For angiosperms, at least 250,000 species are recognized, but only about 34,319 species (13.7%) have *rbcL* sequence record in GenBank as of October 2013. Also, scarcity of floristic survey and molecular systematic works done for Cambodian plants may have caused lack of matching sequence data in the database. Considering several genera solely distributed in America, Africa, and Australia were included as highest matches altogether with Asian genera (Tables 2 & 3), more taxonomic representation of those area's flora in the barcode sequence database seems obvious. As Ekrem et al. (2007) pointed out, for correct identification of unknown taxon to species, genus, family or even order level, a comprehensive DNA sequence library is essential. Thus, to increase the resolution of the DNA barcodes, extensive taxonomic sampling and DNA barcode generation are necessary for the Cambodian flora.

In this study, we found that mature leaves are not good choice for DNA extraction due to existence of secondary compounds and possible contamination from the epiphytic mosses in tropics. Rather, cambium tissues and/or root tissues are better for DNA extraction and following PCR amplification. Fast dividing meristem tissue is a first choice and young shoots are recommended whenever available.

Acknowledgement

Authors thank two anonymous reviewers for their helpful comments on this manuscript, and Chang Kun Lim, Sunah Choi, Seung Hoon Baek at Daegu University, and Phourin Chhang at Forestry Administration of Cambodia for help with field work. Field work was supported by the SE Asia Biodiversity Survey Project of National Institute of Biological Resources, Ministry of Environment, Republic of Korea. This work was financially supported by a 2010 research grant of Daegu University.

Literature Cited.

Azukawa, E. 1982. Balanophoraceae. *In* Wild Flowers of Japan, vol. 2. Herbaceous Plants. Satake, Y. et al. (eds.), Heibonsha, Tokyo, Japan. Pp. 12-13.

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

Chen, J. and J. J. Pipoly III. 1996. Myrsinaceae. *In* Flora of China. Vol. 15. Wu, C. Y., P. H. Raven, and D. Y. Hong (eds.), Science Press & Missouri Botanical Garden Press, Beijing & St. Louis. Pp. 1-38.

Chen, S., H. Yao, J. Han, C. Liu, J. Song, L. Shi, Y. Zhu, X. Ma, T. Gao, X. Pang, K. Luo, Y. Li, X. Li, X. Jia, Y. Lin, and C. Leon. 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE 5: e8613.

China Plant BOL Group, D.-Z. Li, L.-M. Gao, H.-T. Li, H. Wang, X.-J. Ge, J.-Q. Liu, Z-D. Chen, S.-L. Zhou, S.-L. Chen, J.-B. Yang, C.-X. Fu, C.-X. Zeng, H.-F. Yan, Y.-J. Zhu, Y.-S. Sun, S.-Y. Chen, L. Zhao, K. Wang,, T. Yang, and G.-W. Duan. 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 of the United States of America 108: 19641-19646.

Eberwein, R., D. L. Nickrent, and A. Weber. 2009. Development and morphology of flowers and inflorescences in *Balanophora papuana* and *B. elongata* (Balanophoraceae). American Journal of Botany 96: 1055-1067.

Ekrem, T., E. Willassen, and E. Stur. 2007. A comprehensive DNA sequence library is essential for identification with DNA barcodes. Molecular Phylogenetics and Evolution 43: 530-542.

Fay, M. F., C. Bayer, W. S. Alverson, A. Y. de Bruijn, and M. W. Chase. 1998. Plastid *rbcL* sequence data indicate a close affinity between *Diegodendron* and *Bixa*. Taxon 47: 43–50.

Fazekas, A. J., K. S. Burgess, P. R. Kesanakurti, S. W. Graham, S. G. Newmaster, B. C. Husband, D. M. Percy, M. Hajibabaei, and S. C. H. Barrett. 2008. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. PLoS ONE 3: e2802.

Fu, D.-Z. 2012. The Plant-Dictionary. Qingdao Publishing House, Qingdao, China. Pp. 1520.

Hansen, B. 1972. The genus *Balanophora* J. R. & G. Forster - a taxonomic monograph. Dansk Botanisk Arkiv 28: 1-189.

Hansen, B. 1973. Balanophoraceae. *In* Flore du Cambodge, du Laos, et du Vietnam. Vol. 14. Aubreville, A.and J. F. Leroy J.-F. (eds.), Muséum National D'Histoire Naturelle, Paris. Pp. 49-58.

Hebert, P. D. N., A. Cywinska, S. L. Ball, J. R. deWaard. 2003. Biological identification through DNA barcodes. Proceedings of the Royal Society, Biological Science Series B. 270: 313-321.

Hooker, J. D. 1856. On the structure and affinities of Balanopho-

raceae. Transactions of the Linnean Society of London 22: 1-68.

Hsiao, S.-C., J. . Mauseth., and C.-I. Peng. 1995. Composite bundle, the host/parasite interface in the holoparasitic angiosperms *Langsdorffia* and *Balanophora* (Balanophoraceae). American Journal of Botany 82: 81-91.

Huang, S.-M. and J. Murata. 2003. Balanophoraceae *In* Flora of China. Vol. 5. We, C. Y., P. H. Raven, and D. Y. Hong (eds.), Science Press & Missouri Botanical Garden Press, Beijing & St. Louis. Pp. 272-276.

Kawakita, A. and M. Kato. 2002. Floral biology and unique pollination system of root holoparasites, *Balanophora kuroiwa* and *B. tobiracola* (Balanophoraceae). American Journal of Botany 89: 1164-1170.

Kress, W. J. and D. L. Erickson. 2007. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. PLoS ONE 2: e508.

Kress, W. J., D. L. Erickson, F. A. Jones, N. G. Swenson, R. Perez, O. Sanjur, and E. Bermingham. 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America 106: 18621-18626.

Kress, W. J., K. J. Wurdack, E. A. Zimmer, L. A. Weigt, and D. H. Janzen. 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.

Kujit, J. 1969. The Biology of Parasitic Flowering Plants. University of California Press, Berkeley and Los Angeles, California, USA. Pp. 246.

Lahaye, R., M. van der Bank, D. Bogarin, J. Warner, F. Pupulin, G. Gigot, O. Maurin, S. Duthoit, T. G. Barraclough, and V. Savolainen. 2008. DNA barcoding the floras of biodiversity hotspots. Proceedings of the National Academy of Sciences of the United States of America 105: 2923-2928.

Li, F.-W., L.-Y. Kuo, C. J. Rothfels, A. Ebihara, W.-L. Chiou, M. D. Windham, and K. M. Pryer. 2011. *rbcL* and *matK* earn two thumbs up as the core DNA barcode for ferns. PLoS One 6: e26597.

Mabberley, D. J. 1997. The Plant Book: A Portable Dictionary of the Vascular Plants, 2nd ed., Cambridge University Press, Cambridge, UK. Pp. 858.

Meng, M., K. E. Hourt, R. Bansok, E. Thol, and D. Ashwell. 2000. Plants. *In* Cardamom Mountains Biodiversity Survey 2000. Daltry, J. C. and F. Momberg (eds.). Flora & Fauna International, Cambridge, UK. Pp. 31-47.

Newmaster, S. G., A. J. Fazekas, and S. Ragupathy. 2006. DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. Canadian Journal of Botany 84: 335–441.

Nickrent, D. L., J. P. Der, and F. E. Anderson. 2005. Discovery of the photosynthetic relatives of the "Maltese mushroom" *Cynomorium*. BMC Evolutionary Biology 5: 38.

Ricketson, J. M. & I. J. J. Pipoly. 1997. Nomenclatural notes and a synopsis of the genus *Myrsine* (Myrsinaceae). Sida 17: 579–589.

Ridley, H. N. 1930. The Dispersal of Plants throughout the World. Kent, Ashford, UK.

Sang, T., D. J. Crawford, and T. F. Stuessy. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). American Journal of Botany 84: 1120-1136.

Su, H.-J., J. Murata, and J.-M. Hu. 2012. Morphology and phylogenetics of two holoparasitic plants, *Balanophora japonica* and *Balanophora yakushimensis* (Balanophoraceae), and their hosts in Taiwan and Japan. Journal of Plant Research 125: 317-326.

Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4. Sinauer Associates, Sunderland. Massachusetts.

The Angiosperm Phylogeny Group. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APGIII. Botanical Journal of the Linnean Society 161: 105-121.

Valentini, A., F. Pompanon, and P. Taberlet. 2008. DNA barcoding for ecologists. Trends in Ecology and Evolution 24: 110-117.

Won. H. 2009. Phylogenetic position of *Corchoropsis* Siebold & Zucc. (Malvaceae *s.l.*) inferred from plastid DNA sequences. Journal of Plant Biology 52: 411-416.

Yu, J., J.-H. Xue, and S.-L. Zhou. 2011a. New universal *matK* primers for DNA barcoding angiosperms. Journal of Systematics and Evolution 49: 176-181.

Yu, W.-B., P.-H. Huang, R. H. Ree, M.-L. Liu, D.-Z. Li, and H. Wang. 2011b. DNA barcoding of *Pedicularis* L. (Orobanchaceae): evaluating four universal barcode loci in a large and hemiparasitic genus. Journal of Systematics and Evolution 49: 425-437.