Scientia Horticulturae 192 (2015) 302–310

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti

Evaluation of chloroplast DNA markers for distinguishing *Phalaenopsis* species

Jhong-Yi Lin[a,1], Bo-Yen Lin[a,1], Ching-Dong Chang[b,1], Sin-Chung Liao[c,1], Yu-Chang Liu[a,1], Wen-Luan Wu[d,*], Ching-Chun Chang[a,*]

[a] Institute of Biotechnology, National Cheng Kung University, Tainan 701, Taiwan
[b] Department of Veterinary Medicine, National Pingtung University of Science and Technology, Pingtung 912, Taiwan
[c] Department of Biological Science and Technology, Meiho University, Pingtung 912, Taiwan
[d] Department of Life Sciences, National Cheng Kung University, Tainan 701, Taiwan

ARTICLE INFO

Article history:
Received 12 January 2015
Received in revised form 6 June 2015
Accepted 11 June 2015

Keywords:
Chloroplast DNA marker
DNA barcoding
Moth orchids

ABSTRACT

Comparative analysis of chloroplast DNA (cpDNA) between two endemic moth orchids in Taiwan, *Phalaenopsis aphrodite* subsp. *formosana* and *P. equestris*, previously revealed many evolutionary hotspots. In this study, we explored 27 pre-screened cpDNA regions to evaluate molecular markers for distinguishing *Phalaenopsis* species. Fifteen cpDNA markers were highly variable among moth orchids, with polymorphic information contents ≥8.0. The *rps16-trnQ* marker showed the best discrimatory power which 15 endemic moth orchids could be successfully separated into 13 groups. From the sequences of the seven selected cpDNA regions, the *rps16-trnQ* intergenic spacer was the best DNA barcode which 19 endemic moth orchids could be completely distinguished. The *petN-psbM* intergenic spacer could be better used as a DNA barcode to differentiate *P. aphrodite* subsp. *formosana* and *P. amabilis*. Additionally, the inheritance mode of cpDNA in moth orchids was maternally predominated. Phylogenetic analysis revealed non-monophyletic relationships below the *Phalaenopsis* genus level. In summary, we have revealed a set of cpDNA markers that could be used for identification and phylogenetic study of *Phalaenopsis* orchids.
© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The Orchidaceae comprise approximately 25,000 species and represent one of the largest and most diverse families in flowering plants (Dressler, 2005). The genus *Phalaenopsis* consists of approximately 66 endemic species; about 56 are extant worldwide (Christenson, 2001; Tsai et al., 2012). Among them, 17 endemic species (Supplemental Table 1) are the most commonly used as parental strains in commercial breeding programs. *P. aphrodite* subsp. *formosana* and *P. equestris* are the two endemic moth orchids in Taiwan; their genomic databases have been well established (Fu et al., 2011; Su et al., 2013). In addition, their chloroplast DNAs (cpDNAs) have been completely determined (Chang et al., 2006; Jheng et al., 2012). Comparative cpDNA study revealed many evolutionary hotspots that are potentially useful as molecular markers for phylogenetic classification among moth orchids (Jheng et al., 2012). Because moth orchids have many attractive commercial traits, such as numerous branches and spikes, along with colorful and long-lasting flowers, they are among the top-traded blooming potted plants in the world. Recently, the nuclear genome sequence of *P. equestris* was determined and contained rich simple sequence repeats (SSRs) (Cai et al., 2015; Hsu et al., 2011), which will greatly facilitate future marker-assisted breeding of the species.

More than 32,000 hybrid moth orchids have been registered in the database of Royal Horticultural Society (RHS). However, the erroneous registration of hybrid moth orchids resulted in conflicting genealogy and plastid genotypes (Tsai et al., 2012). Therefore, an efficient discriminatory system is urgently needed for correctly tracing the ancestors of hybrid orchids and for protecting intellectual property rights in commercial orchid breeding production. DNA-mediated marker technology is independent of morphology, developmental stage, and environmental factors (Agarwal et al., 2008) and is particularly useful for distinguishing moth orchids

Abbreviations: AFLP, amplified fragment length polymorphism; cpDNA, chloroplast DNA; cpSSR, chloroplast DNA simple sequence repeat; InDel, insertion and deletion; ML, maximum likelihood; NJ, neighbor-joining; nrITS, internal transcribed spacers of nuclear ribosomal DNA; PAGE, polyacrylamide gel electrophoresis; PIC, polymorphic information content; RAPD, random amplified polymorphic DNA; RHS, Royal Horticultural Society; SSR, simple sequence repeat; SNP, single nucleotide polymorphism; STR, sequence tandem repeat; UPGMA, unweighted pair-group method with arithmetic mean.

* Corresponding authors. Fax: +886 6 2766505.
E-mail addresses: wenluan2@mail.ncku.edu.tw (W.-L. Wu), chingcc@mail.ncku.edu.tw (C.-C. Chang).
[1] These five authors are contributed equally to this work.

http://dx.doi.org/10.1016/j.scienta.2015.06.019
0304-4238/© 2015 Elsevier B.V. All rights reserved.

because the orchids are morphologically similar in appearance and have a long vegetative growth stage. Various PCR-based approaches such as random amplified polymorphic DNA (RAPD) (Goh et al., 2005; Niknejad et al., 2009), amplified fragment length polymorphism (AFLP) (Chang and Veilleux, 2009), SSR marker (Fatimah and Sukma, 2011) and cpDNA marker (Jheng et al., 2012; Tsai et al., 2012) technologies have been used for the identification of moth orchids.

Alternatively, DNA barcoding, based on variable DNA sequences to identify species, is becoming a rapid, accurate, and convenient taxonomic tool. The mitochondrial *coxI* gene is well established for barcoding in animals (Hollingsworth et al., 2011). Although a universal DNA barcode in plants is still disputed, various combinations of four plastid coding genes (*rpoB*, *rpoC1*, *rbcL* and *matK*) and three intergenic spacers (*atpF-atpH*, *psbK-psbI* and *trnH-psbA*) have been selected as useful barcodes by the consortium for the barcode of life (CBOL) (Hollingsworth et al., 2011). Analysis of a large-scale samples found chloroplast *matK* or *trnH-psbA* as the most suitable region for plant DNA barcoding (Lahaye et al., 2008; Hollingsworth et al., 2011). Indeed, the *trnH-psbA* intergenic spacer could be used to distinguish various species in *Dendrobium*, *Rhododendron* or *Pteridophytes* and to differentiate these respective species from other adulterating species (Yao et al., 2009; Ma et al., 2010; Liu et al., 2012). DNA barcoding with *matK* alone could distinguish and reveal cryptic species of Mesoamerican orchids and was useful in identifying endangered species (Lahaye et al., 2008). However, in orchids, *ycf1* was found more variable than *matk* (Neubig et al., 2009; Jheng et al., 2012), so it might be a better barcode. Furthermore, the *atpF-atpH* intergenic spacer could be a beneficial DNA barcode for identifying duckweed at the species level (Wang et al., 2010). Additionally, the chloroplast *trnL* intron might be a potential barcode (Taberlet et al., 2007; Tsai et al., 2012). However, several studies have suggested that at low taxonomical levels, more than one barcode might be required to show enough variation for this technique to work well because the cpDNA evolves slowly relative to other genomes (Hollingsworth et al., 2011; Dong et al., 2012).

Although DNA barcoding is informative, it is costly and inconvenient when compared to gel-based DNA markers (Pereira et al., 2008). In a previous study, sequence tandem repeats (STRs) in the cpDNAs of angiosperm showed hypervariable length variation and provided a good source of polymorphic DNA marker to study interspecific and intraspecific variability (Tingey and Rafalski, 1995; Powell et al., 1995; Ebert and Peakall, 2009). For instance, in the Mediterranean orchid, *Anacamptis palustris*, the *trnL* intron of the cpDNA showed a highly variable minisatellite repeat and was used to reveal the inheritance mode of the cpDNA by single-seed PCR (Cafasso et al., 2005). Because the cpDNA of moth orchids is maternally inherited (Chang et al., 2000), cpDNA markers may be helpful in distinguishing endemic species of moth orchids and verifying the maternal ancestors among hybrid species. Previously, with the combination of three cpDNA markers, 11 endemic moth orchids were distinguished (Jheng et al., 2012). The incorporation of additional cpDNA markers may increase discriminatory power.

We used a simple, rapid, sensitive and cost-effective approach to identify polymorphic cpDNA markers for species identification in the genus *Phalaenopsis* and used the information to confirm the mode of cpDNA inheritance. In addition, we documented the nature of the observed interspecific polymorphisms by sequencing and used the cpDNA to study phylogenetic relationships.

2. Materials and methods

2.1. Plant materials and DNA extraction

The plant materials used in this study and their origin are in Supplemental Table 1. Total DNA was isolated from the leaves of moth orchids using the Tri-plant genomic DNA reagent kit (Geneaid, Taiwan).

2.2. Searching for insertion/deletion (InDel) markers, designing primers and PCR amplification

The nucleotide sequence of the coding genes, introns and intergenic spacers were extracted from the cpDNAs of *P. equestris* (NC_017609) and *P. aphrodite* (NC_007499). The corresponding regions of cpDNAs between the two moth orchids were aligned by use of Lasergene software (DNAStar, USA). The variable sites including single nucleotide polymorphism (SNP) and InDel regions between the two moth orchids were identified. Approximately 200 bp upstream or downstream regions from the indicated 27 InDel sites, which also possess STRs, for the two moth orchids were selected for further marker development (Supplemental Table 2). The primers for these InDel sites were designed with use of Primer Premier (http://www.premierbiosoft.com/primerdesign/). The size of the expected cpDNA markers ranged from 207 to 494 bp. The primer pairs (Supplemental Table 2) for evaluating cpDNA markers were synthesized (Genomics, Taiwan). PCR reactions consisted of 5-μl DNA (20 ng) solution, with 200 pmol of each specific primer, 200 nM of each dNTP, 5 units of *Taq* DNA polymerase, and 5 μl of 10x Taq DNA polymerase buffer in a 50-μl reaction volume. The amplification cycle parameters consisted of a 2-min incubation at 95 °C followed by 30 cycles of 30 s at 95 °C, 40 s at 50–55 °C, and 40 s at 72 °C, then one 5-min incubation at 72 °C. Each PCR sample was electrophoresed on a 6% polyacrylamide gel, and visualized by staining with ethidium bromide. The gel images were captured by use of a CCD camera module (Major Science, Taiwan). The size of each DNA fragment was estimated using Quantity One (Bio-Rad, USA) and was assigned a number or character according to their size. The molecular identity of each *Phalaenopsis* species consisted of the combination of a set of numbers or characters derived from PCR products. The NTSYSpc software (Exeter Software, USA) was used to perform clustering analysis of unweighted pair-group method with arithmetic mean (UPGMA) to estimate the distance. The polymorphic information content (PIC) was calculated for each SSR marker according to the following equation (He et al., 2003; Lee et al., 2004): $1-\sum_{i=1}^{n} Pi^2$, where P*i* is the frequency of the *i*th allele to the total number of alleles at a given locus and *n* is the total number of different alleles at the locus.

2.3. Determination of cpDNA inheritance mode by marker technology

Two different crosses of hybrid moth orchids, one population with *P.* Sogo Golden as the male parent and *P.* Yungho Gelb Canary as the female parent, the other population with *P.* Han–Ben's Girl as the male parent and *P.* Timothy Christopher as the female parent, and their progenies were used for determining the mode of cpDNA inheritance. The source of moth orchids and total DNA isolation were described previously. PCR was used to amplify the selected cpDNA marker regions as indicated. PCR products were resolved on 6% or 8% polyacrylamide gel electrophoresis (PAGE) as indicated for 5–6 h at 120 V, and then visualized by staining with ethidium bromide.

2.4. Phylogenetic analysis

PCR primers (Supplemental Table 2) were designed for conserved sequences located upstream and downstream of genes or exons in selected intergenic spacer or intron regions of cpDNAs from the two moth orchids. PCR was performed to amplify selected

Table 1
Polymorphic information contents for 24 cpDNA markers in moth orchids.

Locus[a,b]	No. of samples	Amplificable samples	Estimated size (bp)	No. of alleles	PIC
accD-psaI[a]	14	14	258–317	5	0.68
ccsA-ndhD	14	12	352–374	6	0.79
ndhI-rps15	14	8	388–412	4	0.72
rps4-trnT	15	15	267–287	5	0.76
rps15-ycf1[a]	14	12	167–336	9	0.88
rps16-trnQ[b]	15	15	384–473	13	0.92
petA-psbJ[a,b]	15	14	400–492	8	0.84
petN-psbM[a,b]	15	15	77–588	10	0.88
psaC-ndhE	14	10	397–421	4	0.70
psbA-trnK[a]	14	14	280–319	8	0.84
psbB-psbT	14	14	118–367	10	0.86
psbM-trnD	15	15	251–382	7	0.78
trnE-trnT[a]	14	14	265–464	9	0.86
trnF-ndhJ[a,b]	15	14	232–372	11	0.90
trnN-rpl32[a,b]	15	15	144–422	12	0.90
trnR-atpA	15	15	131–292	8	0.85
trnT-psbD[b]	15	15	236–388	9	0.85
trnT-trnL	14	12	321–361	8	0.82
atpF[a] intron	14	14	400–499	8	0.84
petD intron	14	13	230–255	4	0.70
rpl16 intron	14	14	264–343	9	0.88
rps16[b] intron	14	14	233–280	7	0.79
trnL intron	15	15	241–376	10	0.88
trnV intron	14	13	232–254	4	0.68

[a] The markers could be used for distinguishing *P. aphrodite* and *P. amabilis* based on length variation.
[b] The selected cpDNA regions were sequenced in this study.

cpDNA regions from 19 endemic moth orchids as described previously, and the PCR products were directly sequenced. The sequence data was aligned using Muscle in MEGA5.2 (Tamura et al., 2011), the sequences were concatenated and manually adjusted. *Oncidium* Gower Ramsey (GQ324949) was used as an out-group in the phylogenetic analysis. The phylogenetic relationship was inferred by the maximum likelihood (ML) method based on the Tamura–Nei model with all positions retained or the neighbor-joining (NJ) method with 50% deletion of gap/missing data in MEGA5.2. The bootstrap replication was set to 1000, and the tree with the highest log likelihood is shown.

3. Results

3.1. Identification of polymorphic cpDNA markers

Previously, comparative cpDNA studies of *P. aphrodite* subsp. *formosana* and *P. equestris* revealed many evolutionary hotspots; 64 cpDNA regions including 47 intergenic spacers, 12 introns, three protein-coding genes and two *ndh* pseudogenes carrying InDels resulted from the gain or loss of STRs between two moth orchids (Jheng et al., 2012). These cpDNA regions were potential candidates for marker development. In previous study, we evaluated three cpDNA intergenic spacers and revealed that 11 moth orchids could be separated based on combined length variation (Jheng et al., 2012).

To explore more variable regions for marker development in distinguishing moth orchids, we designed 27 pairs of conserved primers (Supplemental Table 2) flanking the variable regions of 21 intergenic spacers and six introns. All primer pairs were transferable among moth orchids, with at least 57% successful amplification among moth orchids except pairs located in the *ndhC-trnV*, *rps19-psbM* and *trnK-matK* intergenic spacers (Table 1). Each marker was polymorphic among moth orchids, and the estimated size of PCR products ranged from 77 to 588 bp (Table 1). In total, 188 polymorphic bands were amplified, with a mean of 7.8 bands per marker. The PIC for each marker was estimated to range from 0.68 to 0.92 (Table 1). Overall, 15 cpDNA markers were highly variable, with PIC $\geq$ 8.0, and each could classify 14 or 15 moth orchids into at least eight groups (Table 1). The three most informative markers were located in the *rps16-trnQ*, *trnF-ndhJ*, and *trnN-rpl32* intergenic spacers, and each could classify the 15 moth orchids into at least 11 groups (Fig. 1, Table 1). In particular, the *rps16-trnQ* markers showed the best discriminatory power which 15 endemic moth orchids could be successfully separated into 13 groups. Markers located in selected introns were highly variable, with the exception of *petD* and *trnV* introns; the marker targeting the *trnL* intron was the most informative (Fig. 1, Table 1). The *trnL* intron from most endemic moth orchids had been sequenced previously and found to be highly polymorphic (Tsai et al., 2012).

The six intergenic spacers (*petA-psbJ*, *petN-psbM*, *rps16-trnQ*, *trnF-ndhJ*, *trnN-rpl32*, *trnT-psbD*) as well as *rps16* intron were selected for sequencing to investigate the polymorphic status among 19 moth orchids including the 17 parental strains that are the most commonly used in commercial breeding programs (Table 2; Supplemental Table 1). All seven cpDNA regions were highly polymorphic, and each carried abundant SNPs (25–220) and InDels (4–29). In addition, some InDels resulted from the variable SSRs (Table 2). The 19 moth orchids could be classified into at least 15 groups on the basis of sequence variation of any cpDNA region. In particular, the *rps16-trnQ* intergenic spacer was the most variable and could completely distinguish the 19 endemic moth orchids on the basis of sequence variation (Supplemental Fig. 1; Table 2).

3.2. Polymorphic cpDNA markers for distinguishing P. aphrodite and P. amabilis

P. aphrodite subsp. *formosana* has been frequently confused with *P. amabilis* both in the literature and in hybrid registration (Christenson, 2001) because their morphological traits are very similar. The only morphological characteristic that can be used to differentiate them is the number of toothed calluses in flowers (Christenson, 2001). In this study, these two easily confused moth orchids could be distinguished using nine cpDNA markers (*trnE-trnT*, *trnN-rpl32*, *accD-psaI*, *rps15-ycf1*, *petN-psbM*, *petA-psbJ*, *trnF-ndhJ*, *psbA-trnK* and *atpF* intron) (Table 1). In addition, six of seven selected cpDNA regions, along with the previously iden-

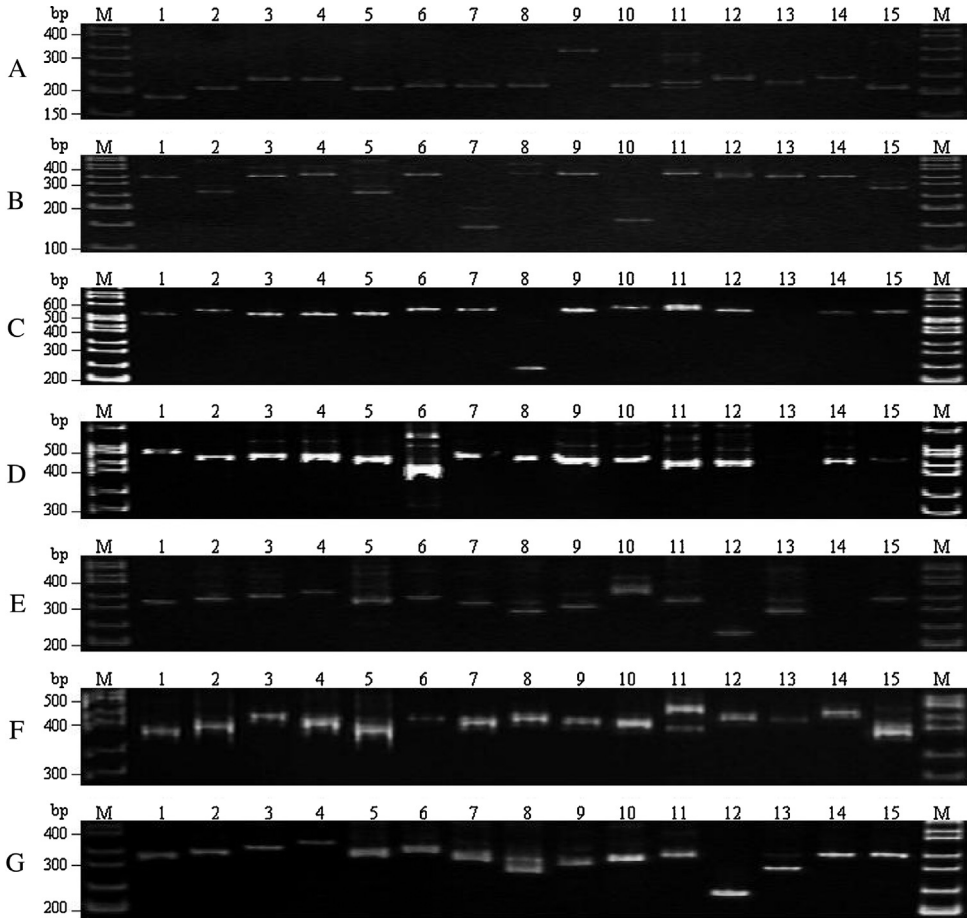

Fig. 1. Identification of 15 endemic species of moth orchids by cpDNA markers.
The selected cpDNA markers located in intergenic spacers *trnT-psbD* (A), *trnN-rpl32* (B), *petN-psbM* (C), *petA-psbJ* (D), *trnF-ndhJ* (E), *rps16-trnQ* (F) and *trnL* intron (G) were PCR amplified from 15 endemic moth orchids, respectively. The PCR products were separated on 6% PAGE, and then visualized by staining with ethidium bromide. The moth orchids were 1. *P. equestris*, 2. *P. aphrodite*, 3. *P. schilleriana*, 4. *P. stuartiana*, 5. *P. sanderiana*, 6. *P. lueddemanniana*, 7. *P. amboinensis*, 8. *P. pulcherrima*, 9. *P. fasciata*, 10. *P. venosa*, 11. *P. gigantea*, 12. *P. mannii*, 13. *P. javanica*, 14. *P. parishii,* 15. *P. amabilis*. M, 50 bp DNA ladder.

Table 2
Sequence variations of selected cpDNA regions among 19 moth orchids.

cpDNA regions	SNP	InDel	SSR[a]	Size range (bp)	Length variations	Sequence variations	Accession no.
petN-psbM	109	28	4	592–896	15	16	KP288923~41
petA-psbJ	220	29	8	955–1055	15	16	KP288904~22
trnT-psbD	57	14	6	965–1023	12	16	KP289018~36
trnF-ndhJ	47	4	1	263–276	9	16	KP288980~98
trnN-rpl32	25	13	3	450–661	16	16	KP288999~9017
rps16-trnQ	115	15	7	852–884	13	19	KP288961~79
rps16 intron	94	12	8	909–951	9	15	KP288942~60
trnL[b] intron	36	15	9	530–603	14	18	Tsai et al. (2012)

[a] The number of variable SSRs caused the InDel among 19 moth orchids.
[b] The sequences of trnL intron were from Tsai et al., 2012.

tified *trnL* intron (Tsai et al., 2012), showed sequence variation between the two moth orchids (Table 3). Though *rps16-trnQ* could be an ideal barcode for moth orchids as described above, it carried only a single nucleotide InDel resulting from mononucleotide repeats between *P. aphrodite* and *P. amabilis* (Table 3). Therefore, a second barcode might be required to complement *rps16-trnQ*. The minisatellite (19 bp) in *petN-psbM* intergenic spacer, which caused InDels between *P. aphrodite* and *P. amabilis*, and other signatures, could better serve as DNA barcoding markers along with *rps16-trnQ* to distinguish species (Fig. 2 and Supplemental Fig. 2). Alternatively, the *petN-psbM* intergenic spacer combined with the *trnT-psbD* intergenic spacer and/or *trnL* intron could provide substantial resolution for differentiation (Table 3).

Table 3
Sequence variations in cpDNA regions between *P. aphrodite* and *P. amabilis*.

cpDNA regions	SNP	InDel	SSR[a,b]	Minisatellites[a]	Size (bp)
petN-psbM	1	2	0	2	896 vs. 895
petA-psbJ	1	1	1	0	984 vs. 983
trnT-psbD	3	1	1	0	986 vs. 985
trnF-ndhJ	1	0	0	0	270 vs. 270
trnN-rpl32	0	1	1	0	537 vs. 538
rps16-trnQ	0	1	1	0	865 vs. 864
rps16 intron	0	0	0	0	914 vs. 914
trnL[c] intron	2	1	1	0	561 vs. 562

[a] The number of variable STRs caused the InDel between two moth orchids.
[b] The SSR was a mononucleotide repeat, and caused a single nucleotide InDel.
[c] The sequence for *trnL* intron was from Tsai et al., 2012.

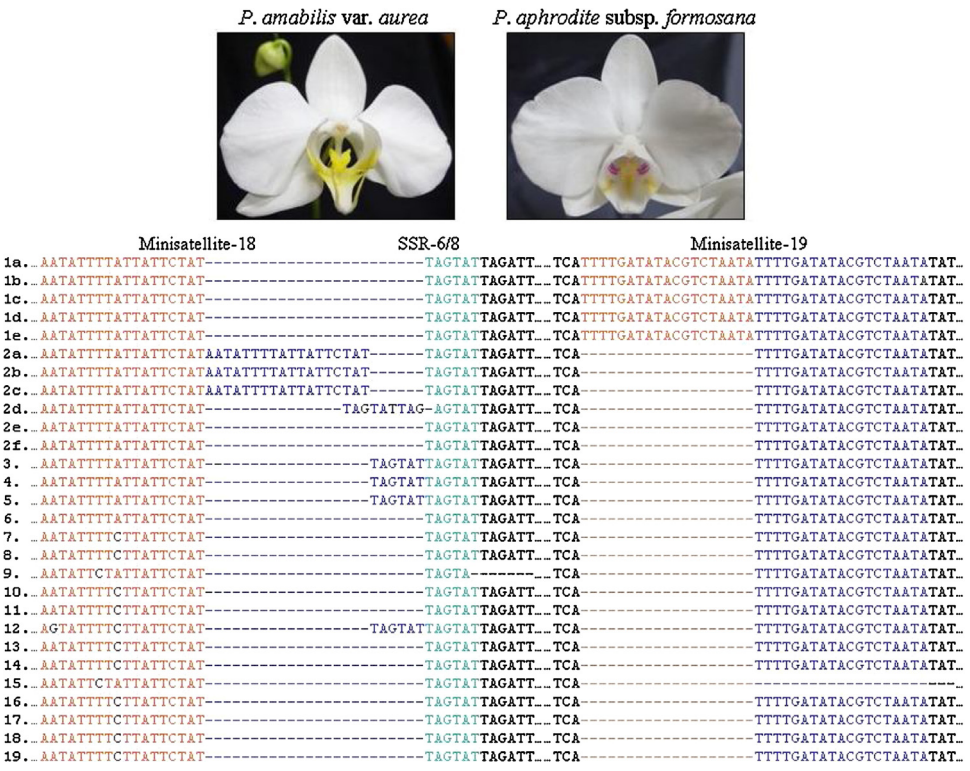

Fig. 2. The InDel marker in *petN-psbM* intergenic spacer for distinguishing *P. aphrodite* and *P. amabilis*.
The sequences of the *petN-psbM* intergenic spacer from 19 moth orchids were multiple aligned. The selected polymorphic SSR (6 or 8 bp) and minisatellite repeats (18 or 19 bp) in the *petN-psbM* intergenic spacer are shown. The repeat motif is shown as red, blue or green. The dot indicates that the upstream or downstream sequence was not shown. The dash indicates the nucleotide deletion. The moth orchids are 1. *P. aphrodite* (1a–e), 2. *P. amabilis* (2a–f). 3. *P. equestris*, 4. *P. schilleriana*, 5. *P. stuartiana*, 6. *P. sanderiana*, 7. *P. lueddemanniana*, 8. *P. amboinensis*, 9. *P. pulcherrima*, 10. *P. fasciata*, 11. *P. venosa*, 12. *P. gigantea*, 13. *P. mannii*, 14. *P. javanica*, 15. *P. parishii*. 16. *P. bellina*, 17. *P. fuscata*, 18. *P. pulchra*, 19. *P. violacea*. The floral morphology of *P. aphrodite* subps. *formosana* and *P. amabilis* var. *aurea* is shown on the top.

3.3. Maternal inheritance of cpDNA in moth orchids

Although RFLP analysis revealed that the cpDNA of moth orchids was maternally inherited (Chang et al., 2000), we used cpDNA marker technology to further confirm the inheritance mode. Six cpDNA markers from the *psbM-trnD*, *rps4-trnT*, *trnE-atpA*, *trnF-ndhJ*, *trnN-rpl32* and *trnT-psbD* intergenic spacers were used to analyze the cpDNA inheritance mode from two different pedigrees of hybrid orchids. One pedigree consisted of hybrid orchids with *P.* Sogo Golden and *P.* Yungho Gelb Canary as male and female parents, respectively, and their random selected progeny. The other pedigree consisted of hybrid orchids with *P.* Han–Ben's Girl and *P.* Timothy Christopher as male and female parents, respectively, and their derived progeny. The PAGE (Fig. 3) and agarose (Supplemental Fig. 3) gel migration patterns of the PCR products from the progenies and female parents were identical. Therefore, our results further suggested the predominantly maternal inheritance mode of cpDNA in moth orchids.

3.4. Phylogenetic analysis

To investigate the phylogenetic relationship based on cpDNA, the sequences of six intergenic spacers and one intron from 19 endemic moth orchids (Table 2) were concatenated and subjected to phylogenetic analysis by the ML method using *Oncidium* Gower Ramsey as an out-group. The phylogenetic tree revealed two major clades. *P. aphrodite*, *P. saderiana* and *P. amabilis*, *P. stuartiana*, *P. schilleriana* and *P. equestris* formed a clade and *P. equestris* branched out from the others, and the former three and the latter two moth orchids could be further clustered together (Fig. 4A). The remaining 13 moth orchids formed the other clade (Fig. 4A). *P. parishii* and *P. pulcherrima* were clustered together and branched out from the other 11 moth orchids. *P. gigantea* and *P. mannii* were further branched out from others. *P. fuscata, P. fasciata*, *P. lueddemanniana* and *P. pulchra* formed a cluster, and *P. javanica*, *P. bellina*, *P. violacea*, *P. venosa* and *P. amboinensis* formed the other cluster; *P. bellina* and *P. violacea* were phylogenetically close with high bootstrap support. With the NJ method applied for phylogenetic analysis, similar results were observed, except that *P. equestris* was clustered with *P. stuartiana* and *P. schilleriana* (Fig. 4B). However, the molecular phylogeny based on the concatenated cpDNA sequence was not consistent with the taxonomical classification of species in the *Phalaenopsis* genus (Fig. 4). Our result showed a non-monophyletic relationship of cpDNA in moth orchids which may suggest the occurrence of interspecific hybridization during evolution.

4. Discussion

4.1. Identification of polymorphic cpDNA markers

The cpDNAs are maternally inherited in most angiosperms (Wicke et al., 2011). Although they evolve relatively slower than the nuclear genome, the noncoding regions of cpDNA are rich in SSRs, in particular, mononucleotide tandem repeats, and are significantly more variable than introns (Ebert and Peakall, 2009), so they can be useful molecular markers for species identification. Previously, the length variation and sequence complexity in noncoding regions of cpDNA was observed across the Orchidaceae, so developing a set of universal primers for cpDNA SSRs (cpSSRs) was difficult (Ebert and Peakall, 2009). However, the development of genus-specific cpSSR markers would be easier as shown in this study. DNA barcoding, based on DNA sequences to identify species, is a

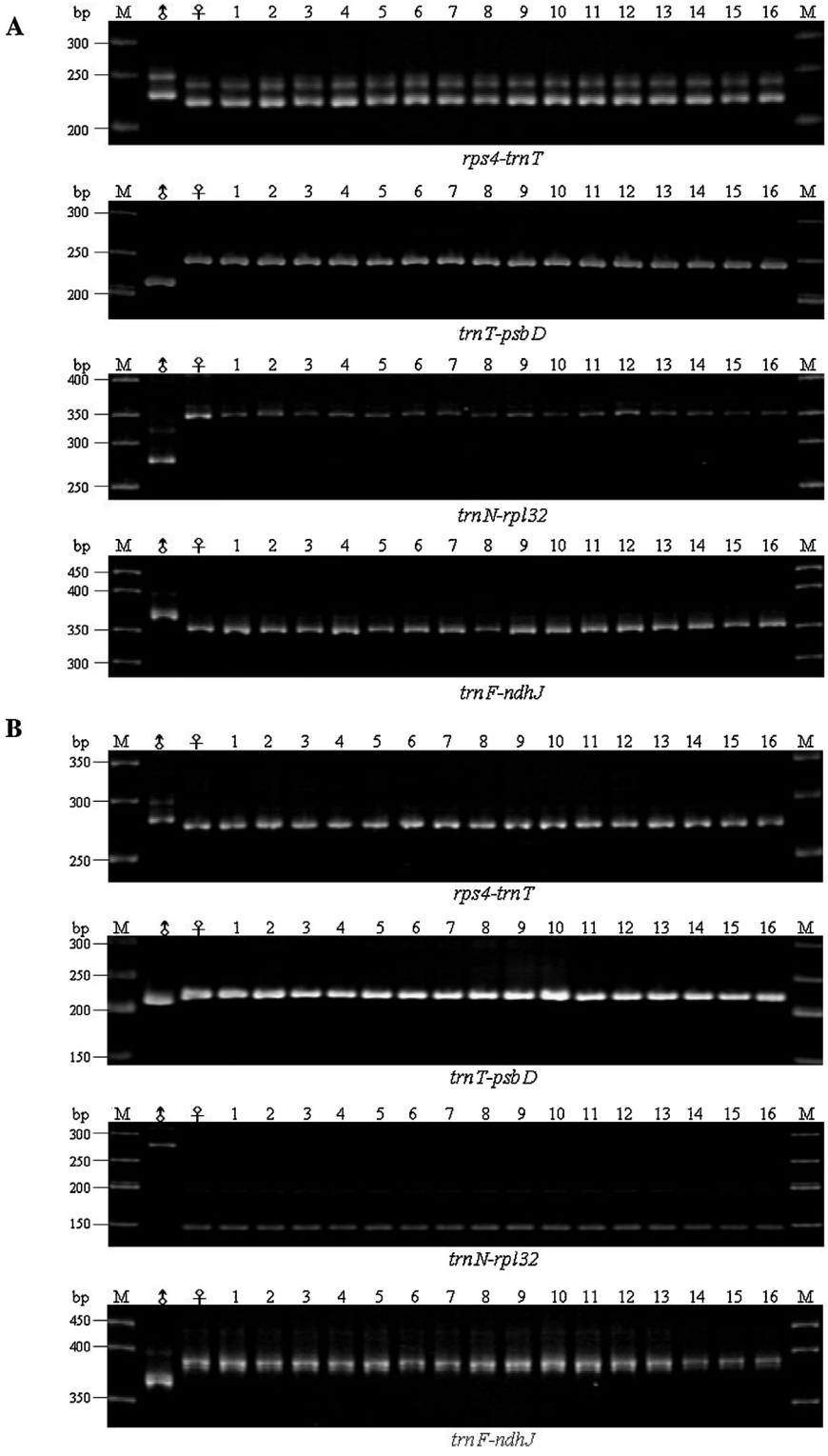

Fig. 3. Maternal inheritance mode of cpDNA in moth orchids.
Total DNA was isolated from two different pedigrees of male (♂) and female (♀) parents of hybrid moth orchids and their 16 progenies (1–16). PCR was carried out to amplify the indicated marker regions on the bottom from moth orchids. (A) The population with male and female parent is *P.* Sogo Golden and *P.* Yungho Gelb Canary, respectively. (B) The population with male and female parent is *P.* Han–Ben's Girl and *P.* Timothy Christopher, respectively. PCR products were separated on 6% PAGE with the exception of *rps4-trnT* marker in A which was resolved on 8% PAGE. Subsequently, the gels were stained with ethidium bromide. M, 50 bp DNA ladder.

useful taxonomic tool. Previously, the *psbA-trnH* intergenic spacer was found to be a valuable DNA marker for identifying species of *Dendrobium*, *Rhododendron* or *Pteridophytes* (Yao et al., 2009; Ma et al., 2010; Liu et al., 2012). A total of 13 of the 15 varieties in *Oncidiinae* could be successfully distinguished on the basis of two intergenic spacers (*psbA-trnH* and *trnF-ndhJ*) (Wu et al., 2010). In *Phalaenopsis*, the sequence of *trnL* intron was found to be highly polymorphic; however, the sequence of *trnL* intron alone was not sufficient to resolve all endemic species (Tsai et al., 2012). Comparison of the cpDNAs of two endemic moth orchids in Taiwan revealed

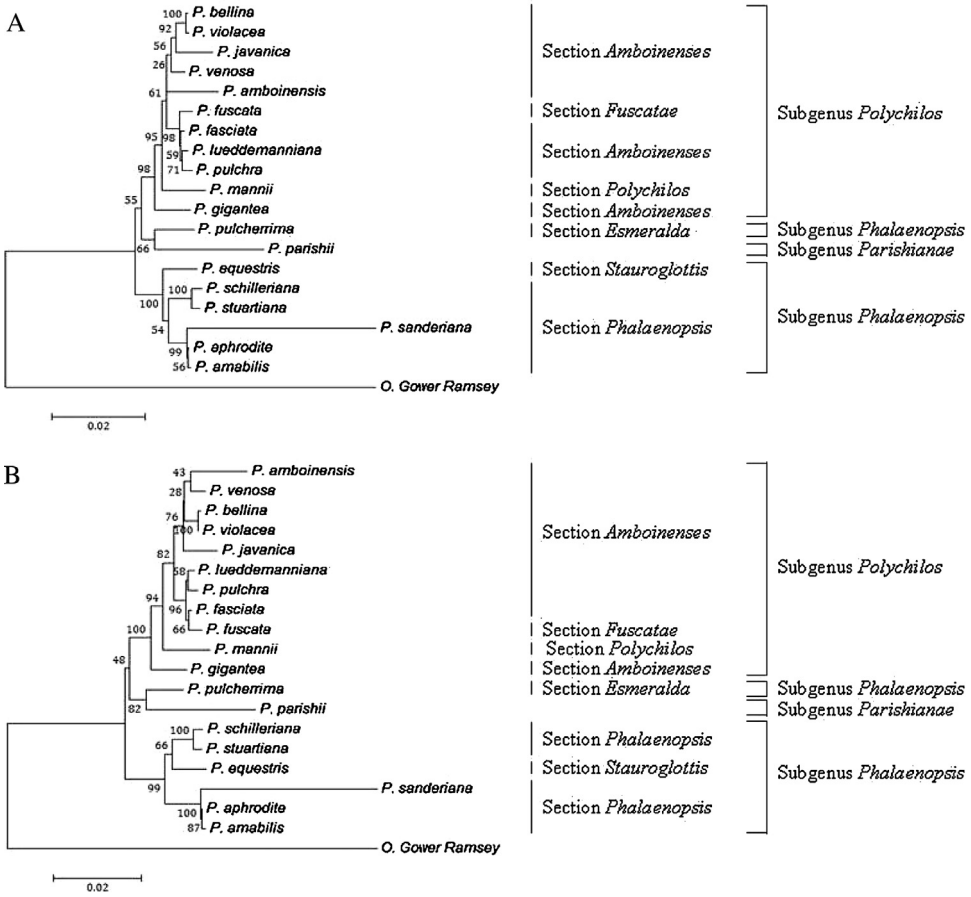

Fig. 4. Molecular phylogenetic analysis of *Phalaenopsis* species based on cpDNA.
The sequences from six intergenic spacers and one intron (Table 2) were multiple aligned by use of Muscle in MEGA5.2, then concatenated. The bootstrap replication was set to 1000. The bootstrap support value is shown beside branches. The ruler is the substitution rate. (A) Phylogenetic analysis by the maximum likelihood method with all sites retained. (B) Phylogenetic analysis by the neighbor-joining method with 50% deletion for gap/missing data.

64 candidate regions for marker development, particularly in intergenic spacers and intron regions (Jheng et al., 2012). The *atpF-atpH* and *psbK-psbI* intergenic spacers were proposed as the best plant barcoding regions (Hollingsworth et al., 2011). However, the variability of these two spacers between two endemic moth orchids in Taiwan is not among the highest (data not shown), so we did not include them for study here.

In this study, we tested 27 pre-screened regions consisting of 21 intergenic spacers and six introns and identified 15 highly variable cpDNA markers based on length variation, with PIC at least 8.0 (Table 1). Seven selected cpDNA regions for 19 endemic moth orchids, including the most commonly used parental species in commercially breeding programs, were further sequenced to reveal the polymorphic status (Table 2; Supplemental Table 1). From criteria for an ideal barcode marker, such as universality for ease of amplification and sequencing, sequence quality, and discriminatory power (Hollingsworth et al., 2011), the *rps16-trnQ* intergenic spacer was considered the best barcoding marker for identifying moth orchids (Table 2; Supplemental Fig. 1). As well, the variability of *rps16-trnQ* among the 19 moth orchids was higher than the previously reported *trnL* intron (Tsai et al., 2012; Table 2).

4.2. Application of cpDNA markers in orchid breeding

RFLP analysis showed the cpDNA of *Phalaenopsis* to be maternally inherited (Chang et al., 2000), and we further confirmed the predominant inheritance mode by using cpDNA markers (Fig. 3; Supplemental Fig. 3). More than 32,000 hybrid species of moth orchids are registered in the RHS; however, the genealogy of some registered hybrid orchids is erroneous because their plastid genotypes are not consistent with that of the female ancestor (Tsai et al., 2012). The registration of hybrid moth orchids began a century ago and the taxonomic classification of endemic moth orchids has been revised (Christenson, 2001). Therefore, the RHS database may contain erroneous genealogies. In light of the maternal inheritance mode of the cpDNA of *Phalaenopsis*, with the development of more cpDNA markers, clarifying the source of maternal ancestor among hybrid species would be helpful. Breeding a new variety of moth orchid and commercializing it is time-consuming and labor-intensive. However, intellectual property rights might be easily infringed upon with tissue culture technology to mass-propagate moth orchids for the market. Therefore, a powerful discriminatory system for early, fast and reliable identification of moth orchids is urgently needed and is important to protect novel commercial breeding varieties in the orchid industry. DNA-based markers such as RAPD (Goh et al., 2005; Niknejad et al., 2009), AFLP (Chang and Veilleux, 2009), SSR markers (Fatimah and Sukma, 2011), and cpDNA markers (Jheng et al., 2012; Tsai et al., 2012) have been used for species identification of moth orchids. The organellar DNA markers can complement nuclear markers by providing additional pedigree information (Petit et al., 2005). In this study, at least nine cpDNA markers (Table 1) could be used to distinguish the frequently confused moth orchids, *P. aphrodite* subsp. *formosana* and *P. amabilis*. In addition, the *petN-psbM* intergenic spacer could be better used as a DNA barcode to differentiate these two orchids (Fig. 2; Supplemental Fig. 2), and in combination with other

barcodes could provide substantial resolution for the differentiation (Table 3).

4.3. Phylogenetic analysis of moth orchids

The genus *Phalaenopsis* has been classified into five subgenera, *Proboscidioides*, *Aphyllae*, *Parishianae*, *Polychilos*, and *Phalaenopsis*, mainly by plant size and floral morphology (Christenson, 2001). The subgenus *Polychilos* can be subdivided into four sections (*Polychilos*, *Fuscatae*, *Amboinenses*, and *Zebrinae*), and the subgenus *Phalaenopsis* can be further subdivided into four sections (*Phalaenopsis*, *Deliciosae*, *Esmeralda*, and *Stauroglottis*). Previously, monophyly of *Phalaenopsis* at the genus level but non-monophyly below the genus level was supported by internal transcribed spacers of nuclear ribosomal DNA (nrITS), cpDNA and the concatenation of both data matrixes of about 2886 bp (Tsai et al., 2006, 2010). In our study, we analyzed the phylogenetic relationship among 19 moth orchids based on concatenation of six cpDNA intergenic spacers and one intron with data matrixes from 4986 to 5746 bp (Table 2, Fig. 4). Our results also suggested that the cpDNA of *Phalaenopsis* was non-monophyletic below the genus level. Within the *Polychilos* subgenus, species in *Amboinenses* section were non-monophyletic, which agrees with previous findings (Tsai et al., 2010). Within the *Phalaenopsis* subgenus, species in the *Phalaenopsis* section were monophyletic based on the ML method (Fig. 4A), which also agrees with previous findings (Tsai et al., 2010). However, with the NJ method, species in *Phalaenopsis* section were non-monophyletic (Fig. 4B). The different results might reflect the parameter settings used in phylogenetic analysis, with all data retained with the ML method and 50% deletion of gap/missing data with the NJ method. The non-monophyletic relationship of cpDNA in moth orchids might indicate the ancient natural occurrence of interspecific hybridization in some moth orchids during evolution.

5. Conclusions

Our comparative cpDNAs analysis between the two endemic moth orchids in Taiwan, *P. aphrodite* subsp. *formosana* and *P. equestris*, previously identified many InDels resulting from STRs, which might provide potential targets for further evaluation of polymorphic markers. We evaluated the pre-screened 27 cpDNA regions to uncover molecular markers for distinguishing *Phalaenopsiss* pecies. More than 62% of cpDNA markers were highly variable, each with PIC > 8.0 among moth orchids. This set of cpDNA markers will be useful for future interspecific and/or intraspecific identification of moth orchids. In particular, the *rps16-trnQ* marker showed the best discrimatory power on the basis of length variation as well as sequence variation. Additionally, the variability of *rps16-trnQ* was higher than the previously reported *trnL* intron among tested moth orchids. Therefore, our results suggested that *rps16-trnQ* might serve as the best barcode in moth orchids at species level. However, in combination with other barcodes could significantly increase discriminatory power. For instance, the InDel marker of *petN-petM* is more informative for the distinguishing between *P. aphrodite* and *P. amabilis.* By taking advantage of marker technology, we suggested the predominantly maternal inheritance mode of cpDNA in moth orchids, which is consistent with previous RFLP analysis. Furthermore, phylogenetic analysis based on the concatenation of six intergenic spacers and one intron of cpDNA regions revealed the non-monophyletic relationship of moth orchids below the *Phalaenopsis* genus level. This result might suggest the possible occurrence of ancient interspecific hybridization in some endemic moth orchids during evolution.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We thank Dr. Chi-Chu Tsai for sparing the moth orchids. We are also grateful to the anonymous reviewers for their comments that help to improve the quality of this manuscript. This work was financially supported by the grants (MOST 102-2321-B-006-017 and 103-2321-B-006-009) to C.-C. Chang from the Ministry of Science and Technology.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.scienta.2015.06.019

References

Agarwal, M., Shrivastava, N., Padh, H., 2008. Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep. 27, 617–631.
Cafasso, D., Widmer, A., Cozzolino, S., 2005. Chloroplast DNA inheritance in the orchid *Anacamptis palustris* using single-seed polymerase chain reaction. J. Hered. 96, 66–70.
Cai, J., Liu, X., Vanneste, K., Proost, S., Tsai, W.C., Liu, K.W., Chen, L.J., He, Y., et al., 2015. The genome sequence of the orchid *Phalaenopsis equestris*. Nat. Genet. 47, 65–72.
Chang, S.B., Chen, W.H., Chen, H.H., Fu, Y.M., Lin, Y.S., 2000. RFLP and inheritance patterns of chloroplast DNA in intergeneric hybrids of *Phalaenopsis* and *Doritis*. Bot. Bull. Acad. Sin. 41, 219–223.
Chang, C.C., Lin, H.C., Lin, I.P., Chow, T.Y., Chen, H.H., Chen, W.H., et al., 2006. The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae): comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. Mol. Biol. Evol. 23, 279–291.
Chang, Y.K., Veilleux, R.E., 2009. Analysis of genetic variability among *Phalaenopsis* species and hybrids using amplified fragment length polymorphism. J. Am. Soc. Hortic. Sci. 134, 58–66.
Christenson, E.A., 2001. Phalaenopsis: A Monograph. Timber Press, Portland, Oregon.
Dong, W., Liu, J., Yu, J., Wang, L., Zhou, S., 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. PLoS One 7, e35071.
Dressler, R.L., 2005. How many orchid species? Selbyana 26, 155–158.
Ebert, D., Peakall, R., 2009. Chloroplast simple sequence repeats (cpSSRs): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. Mol. Ecol. Resour. 9, 673–690.
Fatimah, Sukma, D., 2011. Development of sequence-based microsatellite marker for *Phalaenopsis* orchid. HAYATI J. Biosci. 18, 71–76.
Fu, C.H., Chen, Y.W., Hsiao, Y.Y., Pan, Z.J., Liu, Z.J., Huang, Y.M., et al., 2011. OrchidBase: a collection of sequences of the transcriptome derived from orchids. Plant Cell Physiol. 52, 238–243.
Goh, M.W.K., Kumar, P.P., Lim, S.H., Tan, H.T.W., 2005. Random amplified polymorphic DNA analysis of the moth orchids, *Phalaenopsis* (Epidendroideae: Orchidaceae). Euphytica 141, 11–22.
He, C., Poysa, V., Yu, K., 2003. Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. Theor. Appl. Genet. 106, 363–373.
Hollingsworth, P.M., Graham, S.W., Damon, P., Little, D.P., 2011. Choosing and using a plant DNA barcode. PLoS One 6, e19254.
Hsu, C.C., Chung, Y.L., Chen, T.C., Lee, Y.L., Kuo, Y.T., Tsai, W.C., et al., 2011. An overview of the *Phalaenopsis* orchid genome through BAC end sequence analysis. BMC Plant Biol. 11, 3–13.
Jheng, C.F., Chen, T.C., Lin, J.Y., Chen, T.C., Wu, W.L., Chang, C.C., 2012. The comparative chloroplast genomic analysis of photosynthetic orchids and developing DNA markers to distinguish *Phalaenopsis* orchids. Plant Sci. 190, 62–73.
Lahaye, R., van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., et al., 2008. DNA barcoding the floras of biodiversity hotspots. Proc. Natl. Acad. Sci. U. S. A. 105, 2923–2928.
Lee, J.M., Nahm, S.H., Kim, Y.M., Kim, B.D., 2004. Characterization and molecular genetic mapping of microsatellite loci in pepper. Theor. Appl. Genet. 108, 619–627.
Liu, Y., Zhang, L., Liu, Z., Luo, K., Chen, S., Chen, K., 2012. Species identification of Rhododendron (Ericaceae) using the chloroplast deoxyribonucleic acid *psbA-trnH* genetic marker. Pharmacogn. Mag. 8 (29), 29–36.

Ma, X.Y., Xie, C.Y., Liu, C., Song, J.Y., Yao, H., Luo, K., et al., 2010. Species identification of medicinal Pteridophytes by a DNA barcode marker, the chloroplast *psbA-trnH* intergenic region. Biol. Pharm. Bull. 33, 1919–1924.

Neubig, K.M., Whitten, W.M., Carlsward, B.S., Blanco, M.A., Endara, L., Williams, N.H., Moore, M., 2009. Phylogenetic utility of *ycf1* in orchids: a plastid gene more variable than *matK*. Plant Syst. Evol. 277, 75–84.

Niknejad, A., Kadir, M.A., Kadzimin, S.B., Abdullah, N.A.P., Sorkheh, K., 2009. Molecular characterization and phylogenetic relationships among and within species of *Phalaenopsis* (Epidendroideae: Orchidaceae) based on RAPD analysis. Afr. J. Biotechnol. 8, 5225–5240.

Pereira, F., Carneiro, J., Amorim, A., 2008. Identification of species with DNA-based technology: current progress and challenges. Recent Pat. DNA Gene Seq. 2, 187–200.

Petit, R.J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D., Vendramin, G.G., 2005. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Mol. Ecol. 14, 689–701.

Powell, W., Morgante, M., Andre, C., McNicol, J.W., Machray, G.C., Doyle, J.J., Tingey, S.V., Rafalski, J.A., 1995. Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. Curr. Biol. 5, 1023–1029.

Su, C.L., Chao, Y.T., Yen, S.H., Chen, C.Y., Chen, W.C., Chang, Y.C., Shih, M.C., 2013. Orchidstra: an integrated orchid functional genomics database. Plant Cell Physiol. 54, e11.

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., Willerslev, E., 2007. Power and limitations of the chloroplast *trnL*(UAA) intron for plant DNA barcoding. Nucleic Acids Res. 35, 1–8.

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. Mol. Biol. Evol. 28, 2731–2739.

Tingey, S.V., Rafalski, J.A., 1995. Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. Curr. Biol. 5, 1023–1029.

Tsai, C.C., Huang, S.C., Chou, C.H., 2006. Molecular phylogeny of *Phalaenopsis* Blume (Orchidaceae) based on the internal transcribed spacer of the nuclear ribosomal DNA. Plant Syst. Evol. 256, 1–16.

Tsai, C.C., Chiang, Y.C., Huang, S.C., Chen, C.H., Chou, C.H., 2010. Molecular phylogeny of *Phalaenopsis Blume* (Orchidaceae) on the basis of plastid and nuclear DNA. Plant Syst. Evol. 288, 77–98.

Tsai, C.C., Chiang, Y.C., Lin, Y.S., Liu, W.L., Chou, C.H., 2012. Plastid *trnL* intron polymorphisms among *Phalaenopsis* species used for identifying the plastid genome type of *Phalaenopsis* hybrids. Sci. Hortic. 142, 84–91.

Wang, W., Wu, Y., Yan, Y., Ermakova, M., Kerstetter, R., Messing, J., 2010. DNA barcoding of the Lemnaceae, a family of aquatic monocots. BMC Plant Biol. 10, 205.

Wicke, S., Schneeweiss, G.M., dePamphilis, C.W., Muller, K.F., Quandt, D., 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol. Biol. 76, 273–297.

Wu, F.H., Chan, M.T., Liao, D.C., Hsu, C.T., Lee, Y.W., Daniell, H., Duvall, M.R., Lin, C.S., 2010. Complete chloroplast genome of *Oncidium* Gower Ramsey and evaluation of molecular markers for identification and breeding in Oncidiinae. BMC Plant Biol 10, 68.

Yao, H., Song, J.Y., Ma, X.Y., Liu, C., Li, Y., Xu, H.X., et al., 2009. Identification of *Dendrobium* species by a candidate DNA barcode sequence: the chloroplast *psbA-trnH* intergenic region. Planta Med 75, 667–669.