BIOLOGIA PLANTARUM 60 (1): 68-78, 2016 DOI: 10.1007/s10535-015-0566-2

The mitochondrial DNA markers for distinguishing *Phalaenopsis* species and revealing maternal phylogeny

B.-Y. LIN[1], C.-D. CHANG[2], L.L.H. HUANG[1], Y.-C. LIU[1], Y.-Y. SU[1], T.-C. CHEN[1], R.-H. LEE[3], H.-J. HUANG[4], W.-L. WU[4]*, and C.-C. CHANG[1]*

Institute of Biotechnology, National Cheng Kung University, Tainan 701, Taiwan[1]
Department of Veterinary Medicine, National Pingtung University of Science and Technology, Pingtung 912, Taiwan[2]
Institute of Tropical Plant Sciences, National Cheng Kung University, Tainan 701, Taiwan[3]
Department of Life Sciences, National Cheng Kung University, Tainan 701, Taiwan[4]

Abstract

Moth orchids (*Phalaenopsis*) are among the top-traded blooming potted plants in the world. To explore mitochondrial DNA (mtDNA) markers for species identification, we located simple sequence repeats in the mtDNA of *Phalaenopsis aphrodite* subsp. *formosana* and then pre-screened them for polymorphic markers by their comparison with corresponding mtDNA regions of *P. equestris*. The combination of 13 selected markers located in intergenic spacers could unambiguously distinguish 15 endemic moth orchids. Five most variable markers with polymorphic information content (PIC) ≥ 0.7 could be combined to classify 18 of 19 endemic moth orchids including parental strains most commonly used in breeding programs. The sequences of four selected mtDNA regions were highly variable, and one region (MT2) could be used to completely distinguish 19 endemic moth orchids. Though mitochondrial introns were highly conserved among moth orchids, evolutionary hotspots, such as variable simple sequence repeats and minisatellite repeats, were identified as useful markers. Furthermore, a marker technology was applied to reveal the maternal inheritance mode of mtDNA in the moth orchids. Moreover, phylogenetic analysis indicates that the mtDNA was non-monophyletic below the *Phalaenopsis* genus. In summary, we have revealed a set of mtDNA markers that could be used for identification and phylogenetic study of *Phalaenopsis* orchids.

Additional key words: minisatellite repeats, moth orchids, simple sequence repeats.

Introduction

Plant mitochondrial genes are highly conserved, except in the *Silene* genus, but evolved rapidly in terms of gene order and intergenic spacers even in closely related species (Barr *et al.* 2007, Sloan *et al.* 2012). Sequence tandem repeats (STRs), which include both minisatellites and microsatellites or simple sequence repeats (SSRs), are formed by iterations of DNA sequence motifs. Minisatellites are commonly found in the mitochondrial DNA (mtDNA) of higher plants; however, only polymorphic minisatellites in *Picea* (Sperisen *et al.* 2001, Bastien *et al.* 2003), *Pinus* (Godbout *et al.* 2005), *Beta* (Nishizawa *et al.* 2000, Cheng *et al.* 2011, Yoshida *et al.* 2012), *Oryza*, and *Brassica* (Honma *et al.* 2011) have been revealed and used in breeding, evaluation of germplasm resources, studies of the mode of maternal inheritance, and phytogeographic distribution. In addition, the plant mtDNA contains abundant SSRs. For instance, 2 528 SSRs most abundant in di-nucleotide

Submitted 7 April 2015, *last revision* 26 August 2015, *accepted* 28 August 2015.
Abbreviations: AFLP - amplified fragment length polymorphism; CAPS - cleaved amplified polymorphic sequence; cpDNA - chloroplast DNA; InDel - insertion and deletion; ML - maximum likelihood; mtDNA - mitochondrial DNA; NGS - next generation sequencing; nrITS - internal transcribed spacers of nuclear ribosomal DNA; PAGE - polyacrylamide gel electrophoresis; PCR - polymerase chain reaction; PIC - polymorphic information content; RAPD - random amplified polymorphic DNA; RHS - Royal Horticultural Society; SNP - single nucleotide polymorphism; SSR - simple sequence repeat; STR - sequence tandem repeat; UPGMA - unweighed pair-group method with arithmetic mean.
Acknowledgements: We thank Dr. Chi-Chu Tsai for sparing the moth orchids. We are also grateful to the anonymous reviewers for their comments that helped to improve the quality of this manuscript. This work was financially supported by the grant (MOST 102-2321-B-006-017 and 103-2321-B-006-009) from the Ministry of Science and Technology to C.-C. Chang.
* Corresponding authors; fax: (+886) 6 2766505; e-mails:chingcc@mail.ncku.edu.tw; wenluan2@mail.ncku.edu.tw

repeats were identified in the rice mtDNA (Rajendrakumar *et al.* 2007). Previously, organellar SSRs have been used for identifying cytoplasmic male sterility in cabbage (Wang *et al.* 2012). Recently, *MitoSatPlant*, a database of mitochondrial microsatellites in green plants has been developed; it may be a valuable resource for facilitating the screening of polymorphic SSRs usable in various applications (Kumar *et al.* 2014). Moreover, mtDNA markers have been developed for identifying date palm resistant to Bayoud disease (Quenzar *et al.* 2001), and mtDNA cleaved amplified polymorphic sequence (CAPS) markers can be a reliable tool for distinguishing *Panax* species and cultivars (Lee *et al.* 2012).

Orchidaceae is one of the largest and most diverse families within flowering plants (Atwood 1986). The genus *Phalaenopsis* (moth orchids) contains approximately 66 endemic species (Christenson 2001, Tsai *et al.* 2012). More than 32 000 hybrid moth orchids were bred and registered in the database of the Royal Horticultural Society (RHS), and they are among the top-traded blooming potted plants in the world. An efficient identification system is useful for correctly tracing the parental pedigree of the orchids and for plant variety protection. A DNA-mediated marker technology is particularly useful for distinguishing moth orchids because they are very similar in appearance and have a long vegetative period. Several PCR-based approaches, such as the random amplified polymorphic DNA (RAPD; Goh *et al.* 2005, Niknejad *et al.* 2009), amplified fragment length polymorphism (AFLP; Chang and Veilleux 2009), and SSR marker (Fatimah and Sukma 2011, Hsu *et al.* 2011) and chloroplast (cp) DNA marker technology (Jheng *et al.* 2012, Tsai *et al.* 2012) have been developed for species identification in moth orchids. However, no molecular marker technique is perfect with respect to important features such as genomic abundance, polymorphism level, specificity, reproducibility, technical requirements, and cost (Agarwal *et al.* 2008). To our best knowledge, no mtDNA-based marker has been applied for molecular identification of moth orchids.

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

Materials and methods

Plant material and DNA extraction: Plants used in this study are listed in Table 1 Suppl. The total DNA was isolated from leaves of moth orchids by use of a *Tri-plant* genomic DNA reagent kit (*Geneaid*, New Taipei, Taiwan).

Searching for SSR and InDel markers and designing primers: The mtDNA of *P. aphrodite* subsp. *formosana* was *de novo* sequenced and assembled into contigs by next generation sequencing (NGS) approaches (unpublished data). The mtDNA contigs were used to screen for SSRs through a SSR identification tool (SSRIT, http://www.gramene.org/gramene/searches/ssrtool) and *SSR Locator* (Da Maia *et al.* 2008) with mono-, di-, tri-, tetra-, and penta-nucleotides containing at least 12, 6, 4, 3, and 3 repeats, respectively, and hexa- to deca-nucelotides containing at least 2 repeats. Regions across the potential polymorphic SSR sites from *P. aphrodite* subsp. *formosana* were further *BLAST* searched against the mtDNA of *P. equestris* which was also *de novo* sequenced and assembled by an NGS method (unpublished data). Approximately 200 bp upstream or downstream regions from SSR sites carrying an insertion and deletion (InDel) between two moth orchids were selected for further marker development. The design of primers for these SSR and InDel sites involved the use of *Primer Premier* (http://www.premierbiosoft.com/primerdesign/). Expected mtDNA markers ranged from 105 to 354 bp. In total, 13 primer pairs for mtDNA markers (Table 2 Suppl.) were synthesized (*Genomics*, New Taipei, Taiwan).

Amplification by polymerase chain reaction, electrophoresis, and polymorphic information content: PCR was performed in a 5-mm^3 DNA solution (20 ng), with 200 pmol each specific primer and 12.5 mm^3 of a *Go-Taq* green master mix (*Promega*, Madison, USA) in a 25-mm^3 reaction volume. Amplification was started with 1 cycle at 95 °C for 2 min followed by 35 - 45 cycles at 94 °C for 30 s, at 55 - 60 °C for 30 s, and at 72 °C for 30 s, then 1 cycle at 72 °C for 5 min. Each PCR sample was separated by 8 % (m/v) polyacrylamide gel electrophoresis (PAGE) at 120 V for 5 h, and then visualized by staining with ethidium bromide. The gel images were captured by use of a CCD camera module (*Major Science*, New Taipei, Taiwan). The size of each DNA fragment was estimated by use of *Quantity One* (*BioRad*, Hercules, USA) and was grouped according to size. The molecular identity for each *Phalaenopsis* species consisted of the combination of a set of numbers or characters derived from PCR products. The unweighed pair-group method with arithmetic mean (UPGMA) clustering analysis was used to estimate a phylogenetic distance among moth orchids by use of *NTSYSpc* (*Exeter Software*, East Setauket, USA). In addition, a polymorphic information content (PIC) for each SSR marker was calculated according to an equation PIC = 1 - $\sum_{i=1}^{n} Pi^2$ (He *et al.* 2003, Lee *et al.* 2004), where *Pi* is the frequency of

the *i*th allele among the total number of alleles at a given locus, and *n* is the total number of different alleles at the locus.

Determination of mtDNA inheritance mode: Two different crosses of hybrid moth orchids were used to investigate the mtDNA inheritance mode, one with Sogo Golden and Yungho Gelb Canary as paternal and maternal lines, respectively, and the other with Han-Ben's Girl and Timothy Christopher as male and female parents, respectively, and 20 randomly selected progenies from each cross. The source of moth orchids (Table 1 Suppl.) and total DNA isolation were described above. Polymerase chain reaction was carried out to amplify the regions of *nad5* and *nad7* introns as well as MT2, MT3, MT7, and MT12 markers in intergenic spacer regions as indicated. Products of the PCR were separated by 8 % (m/v) PAGE and then stained with ethidium bromide. The genealogy of parental hybrid moth orchids was retrieved from the RHS database. Sequencing selected mtDNA regions from the parental hybrid orchids to trace their ancestor is described below.

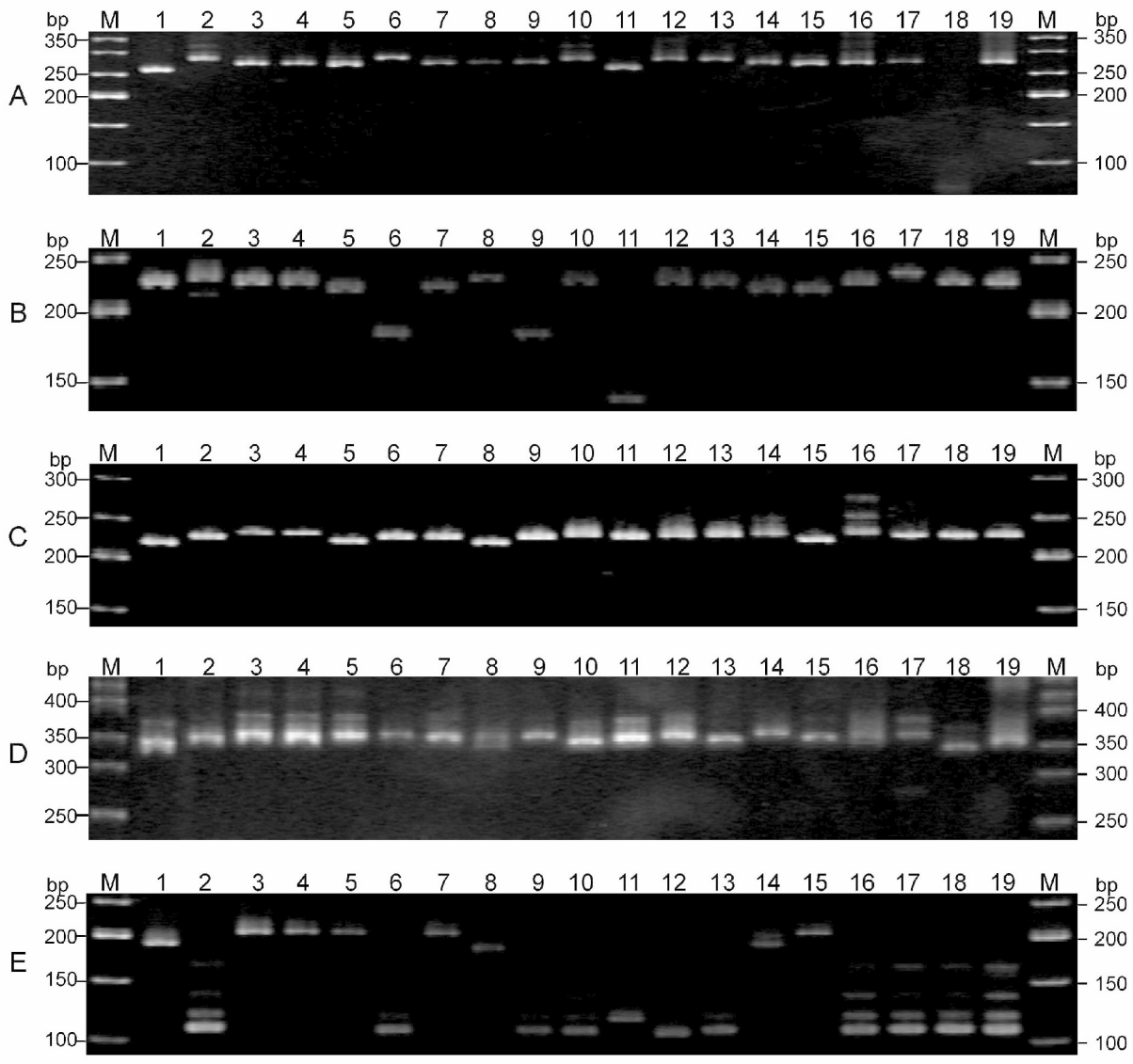
Sequencing mtDNA and phylogenetic analysis: To document the sequence information across the highly polymorphic markers, conserved PCR primers (Table 2 Suppl.) located further upstream and downstream of selected marker regions were designed by comparing the mtDNA sequences from the two moth orchids mentioned above. Polymerase chain reaction was used to amplify the mtDNA, and the PCR products were directly sequenced. In addition, eight introns from mitochondrial genes *nad2*, *nad5*, *nad7*, *rpl2*, *rps10,* and *ccmFc* were PCR amplified from the moth orchids with conserved forward and reverse primers located on the 3' end of uptream exons and the 5' end of downstream exons, respectively (Table 2 Suppl.). The amplified intron regions were directly sequenced. The corresponding marker regions or introns from the moth orchids were multiple aligned by use of *Muscle* in *MEGA5.2* (Tamura *et al.* 2011), concatenated, and then adjusted manually. Phylogenetic relationships were inferred by the maximum likelihood (ML) method based on the Tamura-Nei model with all positions retained in *MEGA5.2*. The bootstrap replication was set to 1 000 and a tree with the highest log likelihood was shown. Bootstrap support values are shown next to the branches. Branch lengths were measured in the number of substitutions per site.

Results

To identify potentially useful mtDNA markers in moth orchids, we searched for SSRs in mtDNA contigs of *P. aphrodite* subsp. *formosana,* whose mtDNA was *de novo* sequenced and assembled *via* NGS (unpublished data). We selected 1 093 SSRs, each with a repeat length of at least 12 bp (Table 3 Suppl.). To search for polymorphic markers, we retrieved flanking regions approximately 200 bp from both upstream and downstream SSRs from the mtDNA of *P. aphrodite* subsp. *formosana* for a *BLAST* search against the mtDNA contigs of *P. equestris*, also assembled by *de novo* NGS (unpublished data). Thirteen polymorphic markers, each carrying variable SSRs or InDels between the two moth orchids, were further selected for marker development (Table 1). Polymerase chain reaction was performed to amplify the 13 SSR markers from 15 endemic moth orchids to evaluate their transferability. All marker regions were amplifiable from each moth orchid, and the

Table 1. Polymorphic information content for 13 mitochondrial SSR markers among moth orchids. Numbers in parentheses indicate results from analysis of 19 moth orchids. PIC - polymorphic information content.

Locus	SSR	Number of samples	Size estimation [bp]	Number of alleles	PIC
MT1	$(C)_{12}$	15	193-197	2	0.12
MT2	$(A)_{12}$	15 (19)	259-291 (77-291)	7 (8)	0.72 (0.73)
MT3	$(TGCTGAG)_2$	15 (19)	137-312 (137-312)	8 (9)	0.85 (0.83)
MT4	$(GATT)_3$	15	190-202	3	0.48
MT5	$(AAGA)_3$	15	203-218	4	0.67
MT6	$(AGA)_4$	15 (19)	218-233 (220-233)	5 (5)	0.73 (0.75)
MT7	$(CT)_6$	15	162-186	4	0.68
MT8	$(ATA)_4$	15	235-269	3	0.56
MT9	$(A)_{12}$	15 (19)	334-362 (334-362)	5 (6)	0.7 (0.75)
MT10	$(CTTT)_3$	15	105-114	2	0.32
MT11	$(TC)_6$	15	148-156	2	0.48
MT12	$(A)_{12}$	15 (19)	106-208 (106-208)	6 (6)	0.75 (0.68)
MT13	$(TTTC)_3$	15	163-178	3	0.52



F

Identity	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Groups
A	b	f	d	d	d	h	de	e	de	g	c	g	g	de	de	de	de	a	e	8
B	e	g	ef	f	c	b	d	i	b	ef	a	f	f	de	cd	f	h	f	f	9
C	a	ab	e	e	ab	c	c	a	cd	c	c	cd	d	d	b	e	cd	d	d	5
D	a	c	ef	de	de	de	de	e	de	c	e	d	d	f	f	ef	f	b	e	6
E	e	b	f	f	f	b	f	d	b	b	c	a	b	e	f	b	b	b	b	6

Fig. 1. The identification of moth orchids by mitochondrial DNA markers. Total DNA was isolated from 19 endemic species and PCR was used to amplify regions of MT2 (*A*), MT3 (*B*), MT6 (*C*), MT9 (*D*), and MT12 (*E*) markers; then the products were separated by 8 % (m/v) polyacrylamide gel electrophoresis (PAGE) and stained with ethidium bromide. The species of moth orchids were: 1 - *Phalaenopsis 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,* 16 - *P. bellina*, 17 - *P. fuscata*, 18 - *P. pulchra*, 19 - *P. violacea*; M - 50 bp DNA ladder. *F* - The sizes of DNA fragments obtained by PAGE for each locus (*A* - *E*) were estimated by use of *Quantity One* and classified according to allele size differences > 2 bp in the same locus. *Lower-case letters* indicate different classification groups and *double letters* indicate indistinguishable ones between the two groups. The molecular identity of the moth orchids (1 - 19) was established by the combination of five sets of letters.

Table 2. Variation in 12 mitochondrial DNA regions among 19 moth orchids ([a] - regions located in intergenic spacers, [b] - regions located in introns, [c] - a number (i) after a gene name is the i*th* intron of the given gene, [d] - a number of variable simple sequence repeats (SSR) or minisatellites among the 19 moth orchids). InDel - insertion and deletion, SNP - single nucleotide polymorphism.

Regions	SNP	InDel	SSR[d]	Minisatellite[d]	Size range [bp]	Variation in length	Variation in sequence	Genebank acc. No.
MT2[a]	245	17	9	0	510-775	13	19	KP262225~43
MT3[a]	426	49	16	0	741-1013	13	17	KP262244~62
MT9[a]	94	10	8	0	1079-1100	11	16	KP262263~81
MT12[a]	45	3	3	0	920-996	6	9	KP262282~300
ccmFc[b]	13	5	4	1	1007-1156	9	13	KP262073~91
nad2-1[bc]	4	0	0	0	732	0	3	KP262092~110
nad2-3[bc]	3	1	0	0	1176-1177	2	5	KP262111~29
nad5-1[bc]	14	5	5	0	808-813	6	12	KP262130~48
nad5-4[bc]	3	0	0	0	661	0	4	KP262149~67
nad7-3[bc]	10	6	5	0	996-1014	6	13	KP262168~86
rpl2[b]	19	6	4	0	1558-1563	5	15	KP262187~205
rps10[b]	8	10	5	0	758-764	6	11	KP262206~24

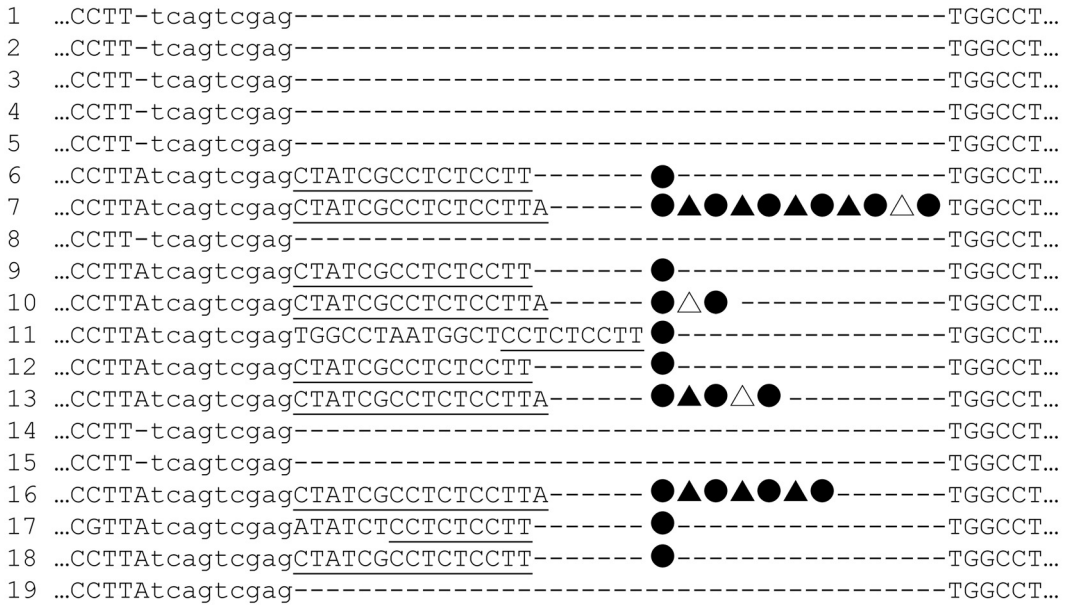


Fig. 2. Polymorphic minisatellite repeats (25 bp) in the *ccmFc* intron of moth orchids. They consisted of a perfect nonanucleotide (*lowercase letters*) and an imperfect hexadecanucleotide (*underlined*). *Black circles* indicate a nonanucleotide (TCAGTCGAG). *Black triangles* indicate a hexadecanucleotide (CTATCGCCTCTCCTTA). *Empty triangles* indicate a pentadecanucleotide (CTATCGCCTCTCCTT). *Dots* indicate not shown sequences. *Dashes* indicate nucleotide deletions. The species of moth orchids were: 1 - *Phalaenopsis 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,* 16 - *P. bellina*, 17 - *P. fuscata*, 18 - *P. pulchra*, and 19 - *P. violacea*.

PCR products ranged from 105 to 312 bp (Table 1). Every marker was polymorphic among the 15 moth orchids with the PIC ranging from 0.12 to 0.85 (Table 1). In total, 54 polymorphic bands were amplified from the 15 moth orchids with a mean of 4.15 bands per marker. The 15 moth orchids could be clustered into three major branches on the basis of UPGMA analysis and could be completely discriminated by the combination of 13 SSR markers (Fig. 1 Suppl.).

Five markers (MT2, MT3, MT6, MT9, and MT12) were highly variable, and each could be used to classify the 15 moth orchids into at least 5 groups. The five highly variable markers were further used to investigate polymorphic status among 19 moth orchids (Table 1 and Table 1 Suppl.) including 17 parental strains that are most commonly used in a commercial breeding program. All PCR primer pairs were transferable, and the PCR products ranged from 77 to 362 bp. The PIC value ranged from 0.7 (MT9) to 0.85 (MT3). The most polymorphic marker (MT3) could differentiate the 19 moth orchids

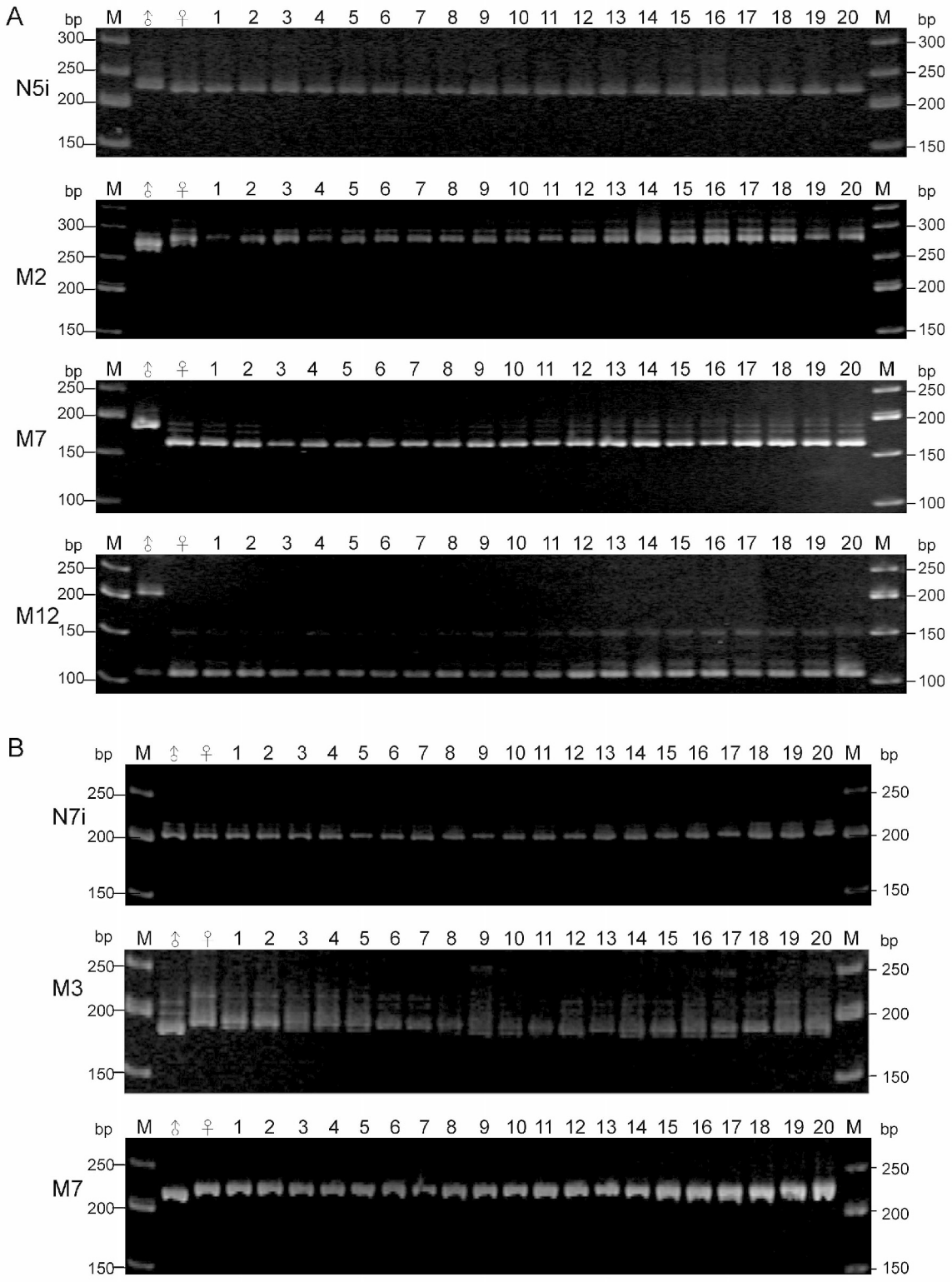


Fig. 3. The maternal inheritance of mitochondrial DNA in moth orchids. The total DNA was isolated from paternal (♂) and maternal (♀) lines of hybrids and their 20 progenies (1 - 20). Polymerase chain reaction was used to amplify indicated marker regions, and then the products were separated by 8 % (m/v) polyacrylamide gel electrophoresis and stained with ethidium bromide. *A* - The male and female parents are Sogo Golden and Yungho Gelb Canary, respectively. *B* - The male and female parents are Han-Ben's Girl and Timothy Christopher, respectively. N5i - *nad5* intron, N7i - *nad7* intron, M2 - MT2 marker, M3 - MT3 marker, M7 - MT7 marker, M12 - MT12 marker, M - 50 bp DNA ladder.

into 9 groups. A combination of the five highly variable markers could discriminate the 19 moth orchids into 18 groups (Fig. 1). In addition, to investigate the polymorphic sites of SSR markers, the regions across the markers MT2, MT3, MT9, and MT12 were subjected to sequencing. The MT2, MT3, and MT9 regions were highly polymorphic, and the 19 moth orchids could be completely distinguished or classified into 17 or 16 groups, respectively (Table 2). The polymorphisms among the 19 moth orchids were attributed to abundant single nucleotide polymorphisms (SNPs) and variable SSRs and InDels (Table 2).

To investigate the polymorphic status of intron regions for marker development in moth orchids, we sequenced eight introns from the mitochondrial genes *ccmFc*, *nad2*, *nad5*, *nad7*, *rpl2*, and *rps10* among the 19 moth orchids. The mitochondrial introns were highly conserved among the moth orchids (Table 2). For instance, intron 1 of *nad2* (*nad2*-1) and intron 4 of *nad5* (*nad5*-4) had only 4 and 3 SNPs, respectively. The other introns had also small-scale SNPs and InDels. Intron 3 of *nad7* (*nad7*-3) and *rpl2* intron were two most variable and could separate the 19 moth orchids into 13 and 15 groups, respectively. Although the introns were highly conserved, evolutionary hotspots could still be found, mainly caused by variations in SSRs and minisatellites (Table 2, Fig. 2 and Fig. 2 Suppl.). For instance, the *ccmFc* intron showed a highly polymorphic 25 bp minisatellite in which the repeat motif could iterate 0.4 to 6.4 times among the 19 moth orchids. With this variable minisatellite, the 19 moth orchids could be classified into at least 6 groups (Fig. 2).

Previous restriction fragment length polymorphism analysis demonstrated that the chloroplast DNA (cpDNA) of moth orchids is maternally inherited (Chang *et al.* 2000), however, the inheritance mode of mtDNA has not been reported yet. To investigate the mode of mtDNA inheritance in moth orchids, we used an SSR marker technology to analyze two different pedigrees of hybrid orchids (Fig. 3). We used four mtDNA SSR markers, three (MT2, MT7, and MT12) from intergenic regions and one from the *nad5*-1 intron, to analyze mtDNA inheritance. Polymerase chain reaction was carried out to amplify the marker regions from a pedigree of hybrid orchids with Sogo Golden and Yungho Gelb Canary used as male and female parents, respectively, and their 20 progenies. Using 8 % PAGE, the gel migration patterns of the PCR products amplified from all four markers were identical between the progeny and their female parent (Fig. 3*A*). Therefore, the mtDNA may be maternally inherited in moth orchids. Furthermore, three mtDNA markers, two (MT3 and MT7) located in intergenic regions and one located in the *nad7*-3 intron, were used to analyze mtDNA inheritance from the other pedigree of hybrid orchids with Han-Ben's Girl and Timothy Christopher used as male and female parents, respectively (Fig. 3*B*). The gel migration patterns of the PCR products amplified from two markers (MT3 and MT7) were also identical between progeny and their female parent except for that from the *nad7*-3 intron (Fig. 3*B*). Therefore, our results confirm the maternally inherited mode of mtDNA in moth orchids.

To investigate a phylogenetic relationship among moth orchids based on the mtDNA, concatenated sequences surrounding MT2, MT3, MT9, and MT12 marker regions (Table 2) were subjected to phylogenetic analysis by the ML method, which revealed two major clades branching from *P. pulchra* (Fig. 4 Suppl.). *P. fuscata* and *P. fasciata* formed a clade and the other moth orchids clustered as a clade. Furthermore, it appears that *P. venosa* and *P. javanica*, *P. bellina* and *P. violacea*, *P. aphrodite* and *P. amboinensis*, *P. schilleriana* and *P. stuartiana*, and *P. saderiana* and *P. amabilis* were more closely related to each other, respectively, with a high bootstrap support. In addition, the phylogenetic analysis of the concatenated eight introns (Table 2) reveals that the 19 moth orchids branched into three major clades (Fig. 4 Suppl.). *P. fuscata, P. fasciata*, *P. lueddemanniana, P. fasciata,* and *P. pulchra* formed a clade, with a high bootstrap support, branched from other moth orchids. In addition, *P. venosa*, *P. bellina*, and *P. violacea* formed a clade, and *P. bellina* and *P. violacea* were phylogenetically closer, with a high bootstrap support. The other moth orchids formed another clade. Among them, *P. aphrodite* and *P. equestris, P. javanica* and *P. parishii, P. sanderiana* and *P. amabilis* were phylogenetically closer, with a high bootstrap support (Fig. 4 Suppl.). Furthermore, the phylogenetic analysis of concatenated four intergenic spacers and eight introns (Table 2) revealed two major clades branching from *P. pulchra* (Fig. 4*C*). *P. fuscata* and *P. fasciata* formed a clade and the other moth orchids clustered as a branch which was highly similar to the phylogenetic tree constructed from the concatenation of intergenic spacer regions described above. Alternatively, the neighbor-joining method was used in phylogenetic analysis and it resulted in similar outcomes (data not shown). However, molecular phylogeny (Fig. 4 Suppl.) based on the mtDNA sequence was not consistent with the morphologically taxonomic classification of species in the *Phalaenopsis* genus (Christenson 2001, Tsai *et al.* 2010). This result suggests a non-monophyletic relationship of mtDNA in moth orchids during evolution.

Discussion

Plant mitochondrial genomes are highly conserved in coding regions but are highly variable in intergenic spacers even in closely related taxa. For instance, mitochondrial STRs, most present in intergenic spacer regions, are common in plants, but accumulation of STR hotspots appears to be taxa-specific (Jaramillo-Correa *et al.* 2013). Recently, a plant mitochondrial SSR database was established to facilitate the screening of polymorphic SSRs (Kumar *et al.* 2014); however, no species from the *Orchidaceae* family was included. In this study, we searched for available mtDNA contigs of *P. aphrodite* (unpublished data) and found 4 340 SSRs most abundant in di-nucleotide repeats followed by mono-nucleotide repeats (Table 3 Suppl.). In addition, by comparative mtDNA analysis of two endemic moth orchids in Taiwan, we prescreened potentially useful polymorphic SSRs and identified five highly variable SSR markers (Table 1, Fig. 1). These five markers were used to study the alternative source of the same 19 endemic moth orchids, which also showed a high discriminating power (data not shown). Although plant mitochondrial introns are highly conserved, evolutionary hotspots can be found. Previously, the phylogenetic relationship of *Orchidaceae* species was well established at the subfamily level based on variations in the *nad1* intron (Freudenstein and Chase 2001). In this study, eight mitochondrial introns of moth orchids were highly conserved; however, we found several evolutionarily variable regions that might be useful as markers such as the polymorphic minisatellite repeat of the *ccmFc* intron (Table 2, Fig. 2 and Fig. 2 Suppl.). Though no single intron could completely identify 19 moth orchids, combining multiple introns could provide enough discriminatory power for a complete resolution (Table 2, Fig. 2, and Fig. 2 Suppl.). Our results suggest that the variable SSRs and minisatellite repeat in introns along with five other highly polymorphic SSR markers in intergenic spacers will be useful for future application.

P. aphrodite subsp. *formosana* has been frequently confused with *P. amabilis* both in literature and in hybrid registration (Christenson 2001) because their morphological traits are similar. The only morphological characteristic that can distinguish *P. aphrodite* from *P. amabilis* is the number of toothed calluses in flowers (Christenson 2001). In this study, these two easily confused moth orchids could be distinguished by six mtDNA SSRs (MT2, 3, 5, 7, 9, and 12) based on length variations (Fig. 1, and unpublished data). In addition, the sequences surrounding four marker regions (MT2, 3, 9, and 12) were variable (Table 2), mainly in SSRs and InDels between two moth orchids, which will provide useful information for DNA barcoding to distinguish them.

Although the maternal inheritance mode of organellar DNA is the dominant form in plants, some exceptions reported include paternal inheritance of cpDNA in most conifers (Neale and Sederoff 1989) and mtDNA in cucumber (Harvey 1997) as well as biparental inheritance of mtDNA in *Pelargonium* (Apitz *et al.* 2013). In the Mediterranean orchid *Anacamptis palustris*, a cpDNA minisatellite repeat marker was used to demonstrate the maternal inheritance mode of cpDNA (Cafasso *et al.* 2005). The mtDNA markers have also been used to reveal maternal phylogeny in conifers (Jaramillo-Correa *et al.* 2003) and citrus (Froelicher *et al.* 2011). Previous study showed that in moth orchids, the cpDNA was maternally inherited (Chang *et al.* 2000). In this study, we used multiple mtDNA markers to investigate two different pedigrees of hybrid orchids and their derived progeny and concluded the maternal inheritance mode of mtDNA in the moth orchids (Fig. 3).

According to the registered genealogy in the RHS database, the female ancestors of hybrid orchids Sogo Golden, Yungho Gelb Canary, Han-Ben's Girl, and Timothy Christopher, used for investigating the inheritance mode of mtDNA in this study, originated from *P. amabilis*, *P. amboinensis*, *P. amabilis*, and *P. equestris*, respectively (Fig. 3 Suppl.). However, from our sequencing results, the *nad7*-3 intron revealed 10 bp differences between *P. amabilis* and *P. equestris* (data not shown) with no difference in length between the two hybrid parents Han-Ben's Girl and Timothy Christopher using PAGE analysis (Fig. 3*B*). Therefore, we further sequenced the *nad7*-3 intron and MT3 marker regions of Han-Ben's Girl and Timothy Christopher, respectively, and compared them with the corresponding regions of 19 endemic moth orchids. The result shows that Han-Ben's Girl (acc. Nos. KP262331 and KP262333) and Timothy Christopher (acc. Nos. KP262334 and KP262332) were identical to *P. amabilis* and *P. stuartiana* (Table 2), respectively (data not shown). Our results strongly suggest that the original female ancestor Timothy Christopher was from *P. stuartiana*. This finding could explain why no length variation was observed in the *nad7*-3 intron among hybrid parents and progeny (Fig. 3*B*). Therefore, the erroneous genealogy in the RHS might be the female parent (Cassandra) of Timothy Christopher (Fig. 3 Suppl.). Previously, based on the *trnL* intron and *atpB-rbcL* intergenic spacer of cpDNA, a conflict in plastid genotype in Timothy Christopher and its female parent Cassandra was discovered, which probably resulted from the erroneous switched registration of the male and female parents of Cassandra (Tsai *et al.* 2012). The registration of hybrid moth orchids began a century ago and the taxonomic classification of endemic moth orchids was disputed and revised (Christenson 2001), so the RHS database may have some erroneous registrations. Since we determined that the mtDNA of moth orchids was maternally inherited, the development of mtDNA markers may be

helpful in clarifying the source of maternal ancestor among hybrid species.

Breeding a new cultivar of moth orchid is time-consuming since it takes nearly two to three years from germination to flowering. However, infringing on intellectual property rights by tissue culture technology to commercialize it to the market might be easy. Therefore, a powerful marker system for early, rapid, and sensitive identification of cultivars of moth orchids is urgently needed and important for protecting novel commercial cultivars in the orchid industry. Recently, the nuclear genome sequence of *P. equestris* was determined and it contains rich SSRs (Hsu *et al.* 2011, Cai *et al.* 2015), which will greatly facilitate a future marker-assisted breeding program. Organellar DNA markers can complement nuclear markers by providing additional pedigree information (Petit *et al.* 2005). Previously, 11 moth orchids could be separated on the basis of length variation from a combination of three cpDNA markers which was selected by comparative cpDNA analysis (Chang *et al.* 2006, Jheng *et al.* 2012). In addition, the *trnL* intron of cpDNA was highly polymorphic among moth orchids (Tsai *et al.* 2012) and was suggested as good target for barcoding. In this study, we identifed 13 variable mtDNA markers which were also selected by the comparative mtDNA analysis (Table 1). The more mtDNA markers are tested, the more moth orchids could be distinguished at a higher resolution. Alternatively, DNA barcoding, which is based on DNA sequencing to identify species, is convenient taxonomic tool. Previously, on the basis of a highly variable intron sequence, mitochondrial *nad1* and *nad2* were established as novel marker for phylogenetic analysis among orchids (Freudenstein and Chase 2001) and mosses (Beckert *et al.* 2001). Based on the criteria of an ideal barcode marker, such as universality for ease of amplification and sequencing, sequence quality, and discriminatory power (Hollingsworth *et al.* 2011), the MT2 region could be proposed as the best barcoding marker for identifying moth orchids (Table 2).

The genus *Phalaenopsis* can be divided into five subgenera, *Proboscidioides*, *Aphyllae*, *Parishianae*, *Polychilos*, and *Phalaenopsis*, mainly based on a 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, the monophyly of *Phalaenopsis* was supported by internal transcribed spacers of nuclear ribosomal DNA (nrITS), cpDNA, and the concatenation of both data matrixes of about 2 886 bp (Tsai *et al.* 2010). However, the subgenus *Phalaenopsis* was not monophyletic, but species in the section *Phalaenopsis* were shown to be monophyletic (Tsai *et al.* 2010). On the contrary, the subgenus *Polychilos* was shown to be monophyletic according to molecular evidence, but species in the section *Amboinenses* were not monophyletic (Tsai *et al.* 2010). In this study, we analyzed the phylogenetic relationship of 19 moth orchids based on mtDNA intergenic spacer regions (3 250 ~ 3 884 bp), introns (7 696 ~ 7 880 bp) and the concatenation of both data matrixes. Our results suggest that the mtDNA of *Phalaenopsis* was non-monophyletic below the genus level (Fig. 4 Suppl.). Within the *Polychilos* subgenus, species in the *Amboinenses* section were non-monophyletic, which agrees with previous findings (Tsai *et al.* 2010). However, within the *Phalaenopsis* subgenus, species in the *Phalaenopsis* section were non-monophyletic, which does not agree with previous findings (Tsai *et al.* 2010). Since the maternal inheritance mode of organellar DNA was confirmed in moth orchids (Chang *et al.* 2000; this study), the inconsistency in phylogenetic relationships based on the cpDNA and mtDNA might result from the limited matrix datasets used for analysis. Alternatively, the differential paternal leakage of the organellar DNA through ancient hybridization between moth orchids might explain the inconsistency and the polyphyletic relationship at or below the subgenus level. The paternal leakage of the mitochondrial genome has been previously observed in seed plants and might play an important role in mtDNA diversity (McCauley 2013).

References

Agarwal, M., Shrivastava, N., Padh, H.: Advances in molecular marker techniques and their applications in plant sciences. - Plant Cell Rep. **27**: 617-631, 2008.

Apitz, J., Weihe, A., Pohlheim, F., Borner, T.: Biparental inheritance of organelles in *Pelargonium*: evidence for intergenomic recombination of mitochondrial DNA. - Planta **237**: 509-515, 2013

Atwood, J.T.: The size of the *Orchidaceae* and the systematic distribution of epiphytic orchids. - Selbyana **9**: 171-186, 1986.

Barr, C.M., Keller, S.R., Ingvarsson, P.K., Sloan, D.B., Taylor, D.R.: Variation in mutation rate and polymorphism among mitochondrial genes of *Silene vulgaris*. - Mol. Biol. Evol. **24**: 1783-1791, 2007.

Bastien, D., Favre, J.M., Collignon, A.M., Sperisen, C., Jeandroz, S.: Characterization of a mosaic minisatellite locus in the mitochondrial DNA of Norway spruce [*Picea abies* (L.) Karst.]. - Theor. appl. Genet. **107**: 574-580, 2003.

Beckert, S., Muhle, H., Pruchner, D., Knoop, V.: The mitochondrial *nad2* gene as a novel marker locus for phylogenetic analysis of early land plants: a comparative analysis in mosses. - Mol. Phylogenet. Evol. **18**: 117-126, 2001.

Cafasso, D., Widmer, A., Cozzolino, S.: Chloroplast DNA

inheritance in the orchid *Anacamptis palustris* using single-seed polymerase chain reaction. - J. Hered. **96**: 66-70, 2005.

Cai, J., Liu, X., Vanneste, K., Proost, S., Tsai, W.C., Liu, K.W., Chen, L.J., He, Y., Xu, Q., Bian, C., Zheng, Z., Sun, F., Liu, W., Hsiao, Y.Y., Pan, Z.J., Hsu, C.C., Yang, Y.P., Hsu, Y.C., Chuang, Y.C., Dievart, A., Dufayard, J.F., Xu, X., Wang, J.Y., Wang, J., Xiao, X.J., Zhao, X.M., Du, R., Zhang, G.Q., Wang, M., Su, Y.Y., Xie, G.C., Liu, G.H., Li, L.Q., Huang, L.Q., Luo, Y.B., Chen, H.H., Van de Peer, Y., Liu, Z.J.: The genome sequence of the orchid *Phalaenopsis equestris*. - Nat. Genet. **47**: 65-72, 2015.

Chang, S.B., Chen, W.H., Chen, H.H., Fu, Y.M., Lin, Y.S.: RFLP and inheritance patterns of chloroplast DNA in intergeneric hybrids of *Phalaenopsis* and *Doritis.* - Bot. Bull. Acad. sin. **41**: 219-223, 2000.

Chang, C.C., Lin, H.C., Lin, I.P., Chow, T.Y., Chen, H.H., Chen, W.H., Cheng, C.H., Lin, C.Y., Liu, S.M., Chang, C.C., Chaw, S.M.: 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, 2006.

Chang, Y.K., Veilleux, R.E.: Analysis of genetic variability among *Phalaenopsis* species and hybrids using amplified fragment length polymorphism. - J. amer. Soc. hort. Sci. **134**: 58-66, 2009.

Cheng, D., Yoshida, Y., Kitazaki, K., Negoro, S., Takahashi, H., Xu, D., Mikami, T., Kubo, T.: Mitochondrial genome diversity in *Beta vulgaris* L. ssp. *vulgaris* (leaf and garden beet groups) and its implications concerning the dissemination of the crop. - Genet. Resour. Crop Evol. **58**: 553-560, 2011.

Christenson E.A.: *Phalaenopsis*: a Monograph. - Timber Press, Portland 2001.

Da Maia, L.C., Palmieri, D.A., De Souza, V.Q., Kopp M.M., De Carvalho, F.I., De Costa, O.A.: SSR locator: tool for simple sequence repeat discovery integrated with primer design and PCR simulation. - Int. J. Plant Genomics **2008**: 412696, 2008.

Fatimah, D., Sukma, D.: Development of sequence-based microsatellite marker for *Phalaenopsis* orchid. - HAYATI J. Bios. **18**: 71-76, 2011.

Freudenstein, J.V., Chase, M.W.: Analysis of mitochondrial *nad1*b-c intron sequences in *Orchidaceae*: utility and coding of length-change characters. - Syst. Bot. **26**: 643-657, 2001.

Froelicher, Y., Mouhaya, W., Bassene, J.B., Costantino, G., Kamiri, M., Luro, F., Morillon, R., Ollitrault, P.: New universal mitochondrial PCR markers reveal new information on maternal citrus phylogeny. - Tree Genet. Genomes **7**: 49-61, 2011.

Godbout, J., Jaramillo-Correa, J.P., Beaulieu, J., Bousquet, J.: A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. - Mol. Ecol. **14**: 3497-3512, 2005.

Goh, M.W.K., Kumar, P.P., Lim, S.H., Tan, H.T.W.: Random amplified polymorphic DNA analysis of the moth orchids, *Phalaenopsis* (*Epidendroideae*: *Orchidaceae*). - Euphytica **141**: 11-22, 2005.

Hollingsworth, P.M., Graham, S.W., Little, D.P.: Choosing and using a plant DNA barcode. - PLoS ONE **6**: e19254, 2011.

Harvey, M.J.: Predominant paternal transmission of the mitochondrial genome in cucumber. - J. Hered. **88**: 232-235, 1997.

He, C., Poysa, V., Yu, K.: 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, 2003.

Honma, Y., Yoshida, Y., Terachi, T., Toriyama, K., Mikami, T., Kubo, T.: Polymorphic minisatellites in the mitochondrial DNAs of *Oryza* and *Brassica*. - Curr. Genet. **57**: 261-270, 2011.

Hsu, C.C., Chung, Y.L., Chen, T.C., Lee, Y.L., Kuo, Y.T., Tsai, W.C.,Hsiao, Y.Y., Chen, Y.W., Wu, W.L., Chen, H.H.: An overview of the *Phalaenopsis* orchid genome through BAC end sequence analysis. - BMC Plant Biol. **11**: 3-13, 2011.

Jaramillo-Correa, J.P., Bousquet, J., Beaulieu, J., Isabel, N., Perron, M., Bouill, M.: Cross-species amplification of mitochondrial DNA sequence-tagged-site markers in conifers: the nature of polymorphism and variation within and among species in *Picea*. - Theor. appl. Genet. **106**: 1353-1367, 2003.

Jaramillo-Correa, J.P., Aguirre-Planter, E., Eguiarte, L.E., Khasa, D.P., Bousquet, J.: Evolution of an ancient microsatellite hotspot in the conifer mitochondrial genome and comparison with other plants. - J. mol. Evol. **76**: 146-157, 2013.

Jheng, C.F., Chen, T.C., Lin, J.Y., Chen, T.C., Wu, W.L., Chang, C.C.: The comparative chloroplast genomic analysis of photosynthetic orchids and developing DNA markers to distinguish *Phalaenopsis* orchids. - Plant Sci. **190**: 62-73, 2012.

Kumar, M., Kapil, A., Shanker, A.: MitoSatPlant: mitochondrial microsatellites database of viridiplantae. - Mitochondrion **19**: 334-337, 2014.

Lee, J.W., Bang, K.H., Kim, Y.C., Seo, A.Y., Jo, I.H., Lee, J.H., Kim, O.T., Hyun, D.Y., Cha, S.W., Cho, J.H.: CAPS markers using mitochondrial consensus primers for molecular identification of *Panax* species and Korean ginseng cultivars (*Panax ginseng* C. A. Meyer). - Mol. Biol. Rep. **39**: 729-736, 2012.

Lee, J.M., Nahm, S.H., Kim, Y.M., Kim, B.D.: Characterization and molecular genetic mapping of microsatellite loci in pepper. - Theor. appl. Genet. **108**: 619-627, 2004.

McCauley, D.E.: Paternal leakage, heteroplasmy, and the evolution of plant mitochondrial genomes. - New Phytol. **200**: 966-977, 2013.

Neale, D.B., Sederoff, R.R.: Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. - Theor. appl. Genet. **77**: 212-216, 1989.

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

Nishizawa, S., Kubo, T., Mikami, T.: Variable number of tandem repeat loci in the mitochondrial genomes of beets. - Curr. Genet. **37**: 34-38, 2000.

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

Quenzar, B., Trifi, M., Bouachrine, B., Hartmann, C., Marrakchi, M., Benslimane, A.A., Rode, A.: A mitochondrial molecular marker of resistance to Bayoud disease in date palm. - Theor. appl. Genet. **103**: 366-370, 2001.

Rajendrakumar, P., Biswal, A.K., Balachandran, S.M.,

Srinivasarao, K., Sundaram, R.M.: Simple sequence repeats in organellar genomes of rice: frequency and distribution in genic and intergenic regions. - Bioinformatics **23**: 1-4, 2007.

Sloan, D.B., Alverson, A.J., Chuckalovcak, J.P., Wu, M., McCauley, D.E., Palmer, J.D., Taylor, D.R..: Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. - PLoS Biol. **10**: e1001241, 2012.

Sperisen, C., Büchler, U., Gugerli, F., Mátyás, G., Geburek, T., Vendramin, G.G.: Tandem repeats in plant mitochondrial genomes: application to the analysis of population differentiation in the conifer Norway spruce. - Mol. Ecol. **10**: 257-263, 2001.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S.: MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. - Mol. Biol. Evol. **28**: 2731-2739, 2011.

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

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

Wang, Q., Zhang, Y., Fang, Z., Liu, Y., Yang, L., Zhuang, M.: Chloroplast and mitochondrial SSR help to distinguish allo-cytoplasmic male sterile types in cabbage (*Brassica oleracea* L. var. capitata). - Mol. Breed. **30**: 709-716, 2012.

Yoshida, Y., Matsunaga, M., Cheng, D., Xu, D., Honm, A Y., Mikami, T., Kubo, T.: Mitochondrial minisatellite polymorphisms in fodder and sugar beets reveal genetic bottlenecks associated with domestication. - Biol. Plant **56**: 369-372, 2012.