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

**A R T I C L E   I N F O**

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

**A B S T R A C T**

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 discriminatory 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.
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) (Goel 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 cox1 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 discriminative 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 identified 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 mM of each dNTP, 5 units of Taq DNA polymerase, and 5 μl of 10× 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):

\[ PIC = \frac{1}{n} \sum_{i=1}^{n} p_i^2, \]

where Pi is the frequency of the ith 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–zen’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.

<table>
<thead>
<tr>
<th>Locus</th>
<th>No. of samples</th>
<th>Amplifiable samples</th>
<th>Estimated size (bp)</th>
<th>No. of alleles</th>
<th>PIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>accD-psal</td>
<td>14</td>
<td>14</td>
<td>258–317</td>
<td>5</td>
<td>0.68</td>
</tr>
<tr>
<td>ccsA-ndhC</td>
<td>14</td>
<td>12</td>
<td>352–374</td>
<td>6</td>
<td>0.79</td>
</tr>
<tr>
<td>ndhI-rps15</td>
<td>14</td>
<td>8</td>
<td>388–412</td>
<td>4</td>
<td>0.72</td>
</tr>
<tr>
<td>rps4-trnT</td>
<td>15</td>
<td>15</td>
<td>267–287</td>
<td>5</td>
<td>0.76</td>
</tr>
<tr>
<td>rps15-ycf1</td>
<td>14</td>
<td>12</td>
<td>167–136</td>
<td>9</td>
<td>0.88</td>
</tr>
<tr>
<td>rps16-trnQ</td>
<td>15</td>
<td>15</td>
<td>384–473</td>
<td>13</td>
<td>0.92</td>
</tr>
<tr>
<td>petA-psbH</td>
<td>15</td>
<td>14</td>
<td>400–492</td>
<td>8</td>
<td>0.84</td>
</tr>
<tr>
<td>petN-psbM</td>
<td>15</td>
<td>15</td>
<td>77–588</td>
<td>10</td>
<td>0.88</td>
</tr>
<tr>
<td>psbC-ndhE</td>
<td>14</td>
<td>10</td>
<td>397–421</td>
<td>4</td>
<td>0.70</td>
</tr>
<tr>
<td>psbA-trnK</td>
<td>14</td>
<td>14</td>
<td>280–319</td>
<td>8</td>
<td>0.84</td>
</tr>
<tr>
<td>psbB-psbT</td>
<td>14</td>
<td>14</td>
<td>118–367</td>
<td>10</td>
<td>0.86</td>
</tr>
<tr>
<td>psbM-trnD</td>
<td>15</td>
<td>15</td>
<td>251–382</td>
<td>7</td>
<td>0.78</td>
</tr>
<tr>
<td>trnE-trnT</td>
<td>14</td>
<td>14</td>
<td>265–254</td>
<td>9</td>
<td>0.86</td>
</tr>
<tr>
<td>trnF-ndhJ</td>
<td>15</td>
<td>15</td>
<td>232–372</td>
<td>11</td>
<td>0.90</td>
</tr>
<tr>
<td>trnN-rpl32</td>
<td>15</td>
<td>14</td>
<td>144–422</td>
<td>12</td>
<td>0.90</td>
</tr>
<tr>
<td>trnR-tpnA</td>
<td>15</td>
<td>15</td>
<td>131–292</td>
<td>8</td>
<td>0.85</td>
</tr>
<tr>
<td>trnT-psbD</td>
<td>15</td>
<td>15</td>
<td>236–388</td>
<td>9</td>
<td>0.85</td>
</tr>
<tr>
<td>trnT-trnL</td>
<td>14</td>
<td>12</td>
<td>321–361</td>
<td>8</td>
<td>0.82</td>
</tr>
<tr>
<td>atpD</td>
<td>14</td>
<td>14</td>
<td>400–499</td>
<td>8</td>
<td>0.84</td>
</tr>
<tr>
<td>petD</td>
<td>14</td>
<td>13</td>
<td>230–255</td>
<td>4</td>
<td>0.70</td>
</tr>
<tr>
<td>rpl16</td>
<td>14</td>
<td>14</td>
<td>264–343</td>
<td>9</td>
<td>0.88</td>
</tr>
<tr>
<td>rpu16</td>
<td>14</td>
<td>14</td>
<td>233–280</td>
<td>7</td>
<td>0.79</td>
</tr>
<tr>
<td>trnl</td>
<td>15</td>
<td>15</td>
<td>241–376</td>
<td>10</td>
<td>0.88</td>
</tr>
<tr>
<td>trnV</td>
<td>14</td>
<td>13</td>
<td>232–254</td>
<td>4</td>
<td>0.68</td>
</tr>
</tbody>
</table>

* The markers could be used for distinguishing *P. aphrodite* and *P. amabilis* based on length variation.

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 ≥ 0.8, 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-psbH, 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-psal, rps15-ycf1, petN-psbM, petA-psbH, trnF-ndhJ, psbA-trnK and atpD intron*) (Table 1). In addition, six of seven selected cpDNA regions, along with the previously ident-

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

<table>
<thead>
<tr>
<th>cpDNA regions</th>
<th>SNP</th>
<th>InDel</th>
<th>SSR</th>
<th>Size range (bp)</th>
<th>Length variations</th>
<th>Sequence variations</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>petN-psbM</td>
<td>109</td>
<td>28</td>
<td>4</td>
<td>592–896</td>
<td>15</td>
<td>16</td>
<td>KP288923–41</td>
</tr>
<tr>
<td>petA-psbf</td>
<td>220</td>
<td>29</td>
<td>8</td>
<td>955–1055</td>
<td>15</td>
<td>16</td>
<td>KP288904–22</td>
</tr>
<tr>
<td>trnT-psbD</td>
<td>57</td>
<td>14</td>
<td>6</td>
<td>965–1023</td>
<td>12</td>
<td>16</td>
<td>KP289018–36</td>
</tr>
<tr>
<td>trnF-ndfj</td>
<td>47</td>
<td>4</td>
<td>1</td>
<td>263–276</td>
<td>9</td>
<td>16</td>
<td>KP288980–98</td>
</tr>
<tr>
<td>trnN-petL2</td>
<td>25</td>
<td>13</td>
<td>3</td>
<td>450–661</td>
<td>16</td>
<td>16</td>
<td>KP288995–9017</td>
</tr>
<tr>
<td>rps16-trnQ</td>
<td>115</td>
<td>15</td>
<td>7</td>
<td>852–884</td>
<td>13</td>
<td>19</td>
<td>KP288961–79</td>
</tr>
<tr>
<td>rps16 intron</td>
<td>94</td>
<td>12</td>
<td>8</td>
<td>909–951</td>
<td>9</td>
<td>15</td>
<td>KP288942–60</td>
</tr>
<tr>
<td>trnL intron</td>
<td>36</td>
<td>15</td>
<td>9</td>
<td>530–603</td>
<td>14</td>
<td>18</td>
<td>Tsai et al. (2012)</td>
</tr>
</tbody>
</table>

* The number of variable SSRs caused the InDel among 19 moth orchids.

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

<table>
<thead>
<tr>
<th>cpDNA regions</th>
<th>SNP</th>
<th>InDel</th>
<th>SSR</th>
<th>Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>petN-psbM</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>petA-psbf</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>trnT-psbD</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>trnN-petL2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rps16-trnQ</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>rps16 intron</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>trnL intron</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

* The number of variable STRs caused the InDel between two moth orchids.

**Fig. 1.** Identification of 15 endemic species of moth orchids by cpDNA markers.

The identified trn L 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).
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-apaA, 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. manii were further branched out from others. P. fusca, 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 \textit{psbA-trnH} intergenic spacer was found to be a valuable DNA marker for identifying species of \textit{Dendrobium}, \textit{Rhododendron} or \textit{Pteridophytes} (Yao et al., 2009; Ma et al., 2010; Liu et al., 2012). A total of 13 of the 15 varieties in \textit{Oncidinae} could be successfully distinguished on the basis of two intergenic spacers (\textit{psbA-trnH} and \textit{trnF-ndhJ}) (Wu et al., 2010). In \textit{Phalaenopsis}, the sequence of \textit{trnL} intron was found to be highly polymorphic; however, the sequence of \textit{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
64 candidate regions for marker development, particularly in intergenic spacers and intron regions (Jheng et al., 2012). The atmF-atpH and psbK-psbl 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 2). 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 Zebriniae), and the subgenus Phalaenopsis can be further subdivided into four sections (Phalaenopsis, Deliciosae, Esmeralda, and Stauroglautes). 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 matrices 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 matrices 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. *formsana* 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 *Phalaenopsis* species. 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 discriminatory 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**


