BMC
Plant Biology

RESEARCH ARTICLE　　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

CrossMark

Biogeography of the *Phalaenopsis amabilis* species complex inferred from nuclear and plastid DNAs

Chi-Chu Tsai[1,2†], Chang-Hung Chou[3,4†], Hao-Ven Wang[5†], Ya-Zhu Ko[3], Tzen-Yuh Chiang[5*] and Yu-Chung Chiang[3,6*]

Abstract

Background: *Phalaenopsis* is one of the important commercial orchids in the world. Members of the *P. amabilis* species complex represent invaluable germplasm for the breeding program. However, the phylogeny of the *P. amabilis* species complex is still uncertain. The *Phalaenopsis amabilis* species complex (Orchidaceae) consists of subspecies *amabilis*, *moluccana*, and *rosenstromii* of *P. amabilis*, as well as *P. aphrodite* ssp. *aphrodite*, *P. ap.* ssp. *formosana*, and *P. sanderiana*. The aims of this study were to reconstruct the phylogeny and biogeographcial patterns of the species complex using Neighbor Joining (NJ), Maximum Parsimony (MP), Bayesian Evolutionary Analysis Sampling Trees (BEAST) and Reconstruct Ancestral State in Phylogenies (RASP) analyses based on sequences of internal transcribed spacers 1 and 2 from the nuclear ribosomal DNA and the *trn*H-*psb*A spacer from the plastid DNA.

Results: A pattern of vicariance, dispersal, and vicariance + dispersal among disjunctly distributed taxa was uncovered based on RASP analysis. Although two subspecies of *P. aphrodite* could not be differentiated from each other in dispersal state, they were distinct from *P. amabilis* and *P. sanderiana*. Within *P. amabilis*, three subspecies were separated phylogenetically, in agreement with the vicariance or vicariance + dispersal scenario, with geographic subdivision along Huxley's, Wallace's and Lydekker's Lines. Molecular dating revealed such subdivisions among taxa of *P. amabilis* complex dating back to the late Pleistocene. Population-dynamic analyses using a Bayesian skyline plot suggested that the species complex experienced an *in situ* range expansion and population concentration during the late Last Glacial Maximum (LGM).

Conclusions: Taxa of the *P. amabilis* complex with disjunct distributions were differentiated due to vicariance or vicariance + dispersal, with events likely occurring in the late Pleistocene. Demographic growth associated with the climatic oscillations in the Würm glacial period followed the species splits. Nevertheless, a subsequent population slowdown occurred in the late LGM due to extinction of regional populations. The reduction of suitable habitats resulted in geographic fragmenttation of the remaining taxa.

Keywords: Biogeography, Demographic dynamics, *Phalaenopsis amabilis*, Phylogeny, Species complex, Vicariance

Background

According to the historical geology of Southeast Asia, most Philippine islands are relatively young, originating about five million years ago (Mya) [1, 2]. The older islands of the Philippines, including Palawan, Mindoro and Zamboanga, are located on the edges of the Eurasian

Plate and may have been moving away from the main landmass since the early Miocene (approximately 30 Mya). Until approximately 5–10 Mya, the crust of the older plate was a part of Borneo [3–5]. The Malay Peninsula, Borneo, Sumatra and Java together comprise the Sunda Shelf. When sea levels were low during the glacial maxima, the Malay Peninsula, Borneo, Sumatra, Java, Bali and the Philippines were interconnected via land bridges, making species migrations possible [6]. In addition, western Sulawesi had been a part of the Sunda Shelf in ancient times but slid away during the Palaeocene (approximately 50 Mya). The formation of the deep Makassar

* Correspondence: tychiang@mail.ncku.edu.tw; yuchung@mail.nsysu.edu.tw
†Equal contributors
[5]Department of Life Science, National Cheng Kung University, Tainan 701, Taiwan
[3]Department of Biological Sciences, National Sun Yat-sen University, Kaohsiung 804, Taiwan
Full list of author information is available at the end of the article

Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 2 of 16

Strait divided western Sulawesi from the Sunda Shelf, preventing further dispersal of Bornean species to Sulawesi. This history geology accounts for the high degree of endemism among the fauna and flora of Sulawesi [7] (Fig. 1).

The *P. amabilis* species complex includes *P. amabilis* ssp. *amabilis*, *P. amabilis* ssp. *moluccana*, *P. amabilis* ssp. *rosenstromii*, *P. aphrodite* ssp. *aphrodite*, *P. aphrodite* ssp. *formosana* and *P. sanderiana*. These species and their subspecies were placed in the section *Phalaenopsis* [8, 9]. The species complex spans a broad geographic range, including southeastern Taiwan, the Philippines, Borneo, Sumatra (Mentawai Is.), Java, Sulawesi, Molucca Is., New Guinea, and northeastern Australia (Queensland) (Fig. 1), and has a complicated biogeography because these regions border two palaeo-continents [10]. These regions have long attracted the attention of biogeographers, and several phylogeographical breaks have been introduced (Fig. 1) [6, 7, 11]. Species/subspecies of the species complex shared most morphological characters, except for lip-midlobes and calli. The distributional patterns coincided with the geographic isolations via historical phylogeographic breaks between Borneo + Palawan and the Philippines (Huxley's Line), between Borneo and Sulawesi (Wallace's Line), between Sulawesi and Molucca Is. (Weber's Line), and between Sulawesi and New Guinea + Australia (Lydekker's Line) [6, 7, 11, 12].
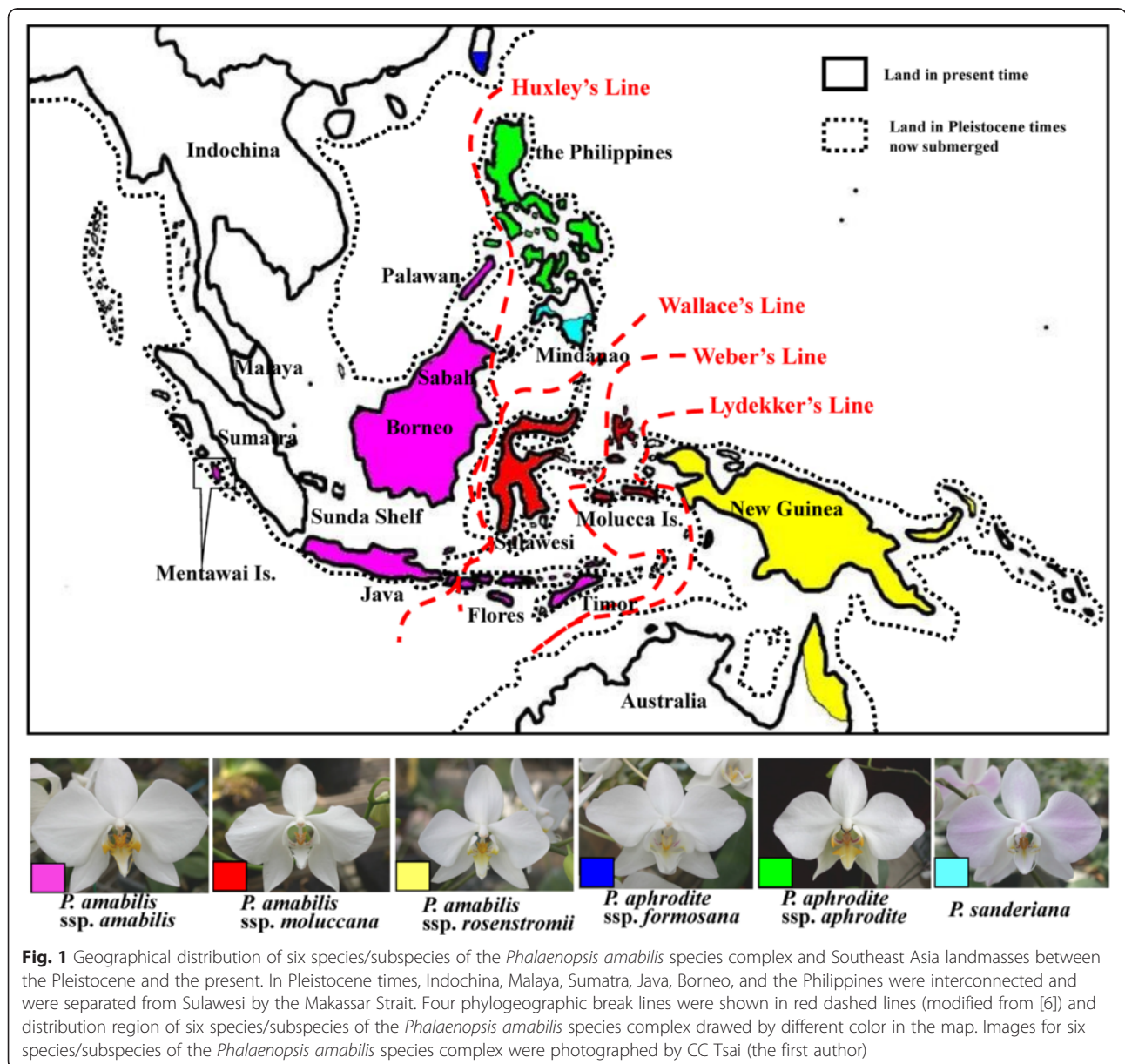

Fig. 1 Geographical distribution of six species/subspecies of the *Phalaenopsis amabilis* species complex and Southeast Asia landmasses between the Pleistocene and the present. In Pleistocene times, Indochina, Malaya, Sumatra, Java, Borneo, and the Philippines were interconnected and were separated from Sulawesi by the Makassar Strait. Four phylogeographic break lines were shown in red dashed lines (modified from [6]) and distribution region of six species/subspecies of the *Phalaenopsis amabilis* species complex drawn by different color in the map. Images for six species/subspecies of the *Phalaenopsis amabilis* species complex were photographed by CC Tsai (the first author)

Tsai et al. BMC Plant Biology (2015) 15:202

Page 3 of 16

Phalaenopsis amabilis ssp. amabilis is widespread, extending from the southern Philippines (Palawan) to Borneo, Sumatra and Java. Another subspecies, P. amabilis ssp. moluccana, is strictly distributed in the Molucca Islands (Seram, Buru) and Sulawesi. There is a phylogeographic break (Wallace's Line) between these two subspecies (Fig. 1) [6, 11]. Phalaenopsis amabilis ssp. moluccana is easily recognised by the shape of the midlobe of the lip, which is linear-oblong with a slight dilation toward the base instead of being cruciform [8] (Table 1). P. amabilis ssp. rosenstromii is distributed in New Guinea (Kaiser Wilhelms Land), New Ireland (Neumecklenburg) and Australia (northeastern Queensland). This subspecies differs from P. amabilis ssp. amabilis and P. amabilis ssp. moluccana in having a lip with a triangular midlobe, as in P. aphrodite and P. sanderiana (Table 1). Phalaenopsis amabilis ssp. moluccana and P. amabilis ssp. rosenstromii are separated by a phylogeographic break at the Lydekker's Line [12].

Given high resemblance between P. aphrodite and P. amabilis, P. aphrodite was ever once placed under P. amabilis, i.e., P. amabilis var. aphrodite [8, 9]. In contrast, Sweet [13] recognised it as a distinct species based on characteristics of the callus and midlobe shape, and the geographical subdivision by the Lydekker's Line [12]. P. aphrodite ssp. aphrodite is distributed throughout the Philippines with exceptions of Mindanao and Palawan [13]. Its calli have a posterior edge dividing into four teeth, unlike the two-toothed calli of P. amabilis [9] (Fig. 1). Another subspecies, P. aphrodite ssp. formosana, is distributed only in southeastern Taiwan and can be distinguished from P. aphrodite ssp. aphrodite by its apple-green leaves with no trace of anthocyanin pigments, somewhat smaller flowers, and many-branched panicles [9] (Table 1).

Another phyloegnetically related species [14], P. sanderiana, has a restricted distribution in the southern Philippines (including Mindanao Island, Igat Island and Balut and Sarangani Islands) [15]. Its callus structure shows four teeth, with the inner teeth longer than the outer ones. The four teeth of P. aphrodite, in contrast, are subequal, with the two outer teeth only slightly longer than the inner ones [9] (Fig. 1). The distributions of P. sanderiana and P. amabilis ssp. amabilis are divided by Huxley's Line, and those of P. sanderiana and P. amabilis ssp. moluccana are divided by Wallace's Line [11].

It is known that phylogeographic patterns are determined by the interplay between dispersal and vicariance [16]. Vicariance arises from changes in the earth's surface from changes in sea level or the separation of similar environments by climatic, geological or other changes in intervening regions, resulting in the fragmentation of continuously distributed taxa [17]. Dipersal across isolating barriers, which can disrupt the preexisting vicaricance patterns, might also predominate in the oceanic island biogeography [18]. This species complex is a good tool by which to study biogeography because this P. amabilis complex is distributed throughout the region between Southeast Asia and Australia, where several phylogeographic break lines have been identified. Here, based on the distribution patterns of the taxa in the species complex across these geographical regions, if vicariance exclusively determined the species split, one would expect reciprocal monophyly of P. aphordite, P. amabilis, and P. sanderiana, which were divided by Huxley's and Wallace's Lines. Affinity among P. amabilis ssp. amabilis, ssp. moluccana, and ssp. rosenstromii separated by Wallace's and Weber's Lines remains testable. In other word, Christenson's [9] taxonomy is served as the biogeographical hypothesis.

Table 1 Differentiated characters among six species/subspecies of the Phalaenopsis amabilis species complex

Taxa	Morphological characters				
	Floral color	Lip color	Leaves with anthocyanin pigment	Teeth no. of posterior edge in calli	Midlobe shape of labium
P. amabilis ssp. amabilis	White	Base of lip marked with yellow and red	No	Two	Cruciform
P. amabilis ssp. moluccana	White	Base of lip marked with yellow, white	No	Two	Linear-oblong with a slight dilation toward the base
P. amabilis ssp. rosenstromii	White	Base of lip marked with yellow	No	Two	The shorter, narrowly triangular midlobe of the lip, with inconspicuous teeth at the base
P. aphrodite ssp. aphrodite	White	Base of lip marked with yellow and red	Various	Four[a]	Triangular-hastate midlobe
P. aphrodite ssp. formosana	White	Base of lip marked with yellow	No	Four[a]	Triangular-hastate midlobe
P. sanderiana	White, pink	Base of lip marked with yellow and red	Yes	Four[b]	Triangular-hastate midlobe

[a] outer two teeth longer than the inner two
[b] inner two teeth longer than the outer two

Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 4 of 16

In this study, DNA fragments of the internal transcribed spacer (ITS) of nuclear ribosomal DNA and the *trn*H-*psb*A intergenic spacer of plastid DNA were employed. Several questions were addressed. 1) Is monophyly of the species of the complex supported? 2) Is the geographic disjunction in this species complex due to vicariance or long-distance dispersal? 3) What is the demographic history of *P. amabilis* complex? 4) What is the extent of gene flow between species?

Results

Sequence characteristics, variation and haplotype diversity

For all accessions of the *P. amabilis* species complex (Table 2), the ITS1 sequence lengths ranged from 228 to 233 base pairs (bp); the 5.8S-rDNA sequences were 163 bp, and the ITS2 sequences ranged from 256 to 258 bp (GenBank Nos. AY391515-53; http://dx.doi.org/10.5061/dryad.f8j12). There were no nucleotide substitutions or indels in the 5.8S rDNA of the *P. amabilis* complex. Both variable and potential parsimony-informative sites in the ITS1 region across the species complex are higher than those of ITS2 region. The ITS1 and ITS2 sequences of the 39 accessions of the *P. amabilis* complex were combined and aligned, yielding 494 characters. Eight gap sites, 13 variable sites and 11 potential parsimony-informative sites were found in the aligned matrix (Table 3).

Sequence lengths for the *trn*H-*psb*A spacer ranged from 918 to 946 bp (GenBank Nos. FJ460366-407; http://dx.doi.org/10.5061/dryad.f8j12). The sequences of the 39 accessions of the *P. amabilis* complex were aligned, with a sequence length of 908 characters (http://dx.doi.org/10.5061/dryad.f8j12). Five gap sites, 11 variable sites and eight potential parsimony-informative sites were found in the aligned matrix (Table 3). These sites included three long indels, resulting from differences in the copy number of several types of repeat sequences. Two copies of CAATATCTTGTTCTTAGA (extra copy at position 161–178 of the aligned sequences) were found in the two accessions of *P. amabilis* ssp. *amabilis* from Timor, and two copies of TGAAAT-GAA (extra copy at position 275–283 of the aligned sequences) were found in the four accessions of *P. amabilis* ssp. *rosenstromii*. Most samples from the complex had only one copy of this sequence, but it was absent in *P. amabilis* ssp. *amabilis*-Sabah-kc-342. Two copies of TGAATGAT (extra copy at position 298–306 of the aligned sequences) were found in *P. amabilis* ssp. *amabilis*-Java-kc-97. Most taxa of the complex possessed a single copy of this sequence. Five accessions from *P. aphrodite* ssp. *aphrodite*, including populations from Mindanao and Fuga Island and one accession from the Calayan Island population, did not have this sequence

(Additional file 1: Table S1). Of five indels across the sequence alignment of the *trn*H-*psb*A spacer, four are informative (Table 3).

Estimates of nucleotide diversity, haplotype, and haplotype diversity for different species/subspecies of the *P. amabilis* complex were calculated (Table 4). Total nucleotide diversity (π) in ITS1, 5.8 s rDNA, ITS2, and *trn*H-*psb*A spacer of *P. amabilis* complex was 0.0078, 0, and 0.0039, 0.895, and 0.0035, respectively (Table 3). In the ITS region, nucleotide diversity in the *P. amabilis* (π = 0.0054) was higher than that in *P. aphrodite* and *P. sanderiana*, while similar results were found in haplotype diversity (*P. amabilis*: h = 0.857; *P. aphrodite*: h = 0.222; *P. sanderiana*: h = 0.667) (Table 4). In *trn*H-*psb*A region, nucleotide diversity in the *P. sanderiana* (π = 0.0029) was higher than that in *P. amabilis* and *P. aphrodite*, while haplotype diversity in the *P. amabilis* (h = 0.758) was the hightest among taxa (Table 4). In short, the nucleotide diversity of both ITS1 and ITS 2 regions across the species complex is higher than that of the *trn*H-*psb*A spacer.

Phylogenetic reconstruction

Genealogy was reconstructed based on the ITS1 and ITS2 and plastid DNA haplotypes of the *P. amabilis* complex, rooted at three species of the *P. schilleriana* complex, including *P. schilleriana*, *P. stuartiana*, and *P. philippinensis* (Fig. 2). In the ITS1 and ITS2 sequences, a total of 13 variable sites (nucleotide substitutions) and 7 indels were found among the species/subspecies of the *P. amabilis* complex. Based on the phylogenetic tree, the six taxa of the *P. amabilis* complex were closely related with high bootstrap supporting in both NJ and MP phylogenetic trees (Figs. 2a). *Phalaenopsis aphrodite* ssp. *aphrodite* and *P. aphrodite* ssp. *formosana* clustered together. Subspecies of *P. aphrodite* are not distinguishable from each other, whereas they were separable from other members of the *P. amabilis* complex. Subspecies of *P. amabilis*, except for the Palawan populations, and *P. sanderiana* formed a clade. Within the *P. amabilis*/*P. sanderiana* clade, two subclades were identified, rooted by *P. amabilis* ssp. *amabilis* from Palawan. The first subclades showed distinct geographic subdivision, including *P. sanderiana*, *P. amabilis* ssp. *amabilis* from Java and ssp. *moluccana* from Molucca island. Nevertheless, *P. amabilis* ssp. *amabilis* from Timor and *P. amabilis* ssp. *rosentromii* cannot be distinguished from each other. The other subclade included all *P. amabilis* ssp. *amabilis* from Mentawai Island, Sabah (Borneo) and the one from Palawan (Fig. 2a).

In the *trn*H-*psb*A spacer, 11 variable sites and 6 indels were found among the species/subspecies of the *P. amabilis* complex. The phylogeny inferred from the plastid DNA was largely consistent with those from the ITS data and provided better resolution in the geographic

Table 2 A list of the 39 accessions of the *Phalaenopsis amabilis* species complex, namely *P. amabilis*, *P. aphrodite*, and *P. sanderiana*, and their different geographical distributions

Taxa and systematic classification[a]	Distribution	Source[b]/Voucher[c]
P. amabilis ssp. *amabilis* 'Java'	Bantam, Java, Indonesia	KDAIS-kc66/Tsai C.C. 1066
P. amabilis ssp. *amabilis* 'Java'	Bantam, Java, Indonesia	KDAIS-kc96/Tsai C.C. 1096
P. amabilis ssp. *amabilis* 'Java'	Bantam, Java, Indonesia	KDAIS-kc97/Tsai C.C. 1097
P. amabilis ssp. *amabilis* 'Mentawai Is.'	Mentawai Is., Sumatra, Indonesia	KDAIS-kc238/Tsai C.C. 1238
P. amabilis ssp. *amabilis* 'Mentawai Is.'	Mentawai Is., Sumatra, Indonesia	KDAIS-kc239/Tsai C.C. 1239
P. amabilis ssp. *amabilis* 'Mentawai Is.'	Mentawai Is., Sumatra, Indonesia	KDAIS-kc240/Tsai C.C. 1240
P. amabilis ssp. *amabilis* 'Palawan'	Brooks Point, Palawan, the Philippines	KDAIS-kc91/Tsai C.C. 1091
P. amabilis ssp. *amabilis* 'Palawan'	Brooks Point, Palawan, the Philippines	KDAIS-kc92/Tsai C.C. 1092
P. amabilis ssp. *amabilis* 'Palawan'	Brooks Point, Palawan, the Philippines	KDAIS-kc93/Tsai C.C. 1093
P. amabilis ssp. *amabilis* 'Sabah'	Sabah, Indonesia	KDAIS-kc327/Tsai C.C. 1327
P. amabilis ssp. *amabilis* 'Sabah'	Sabah, Indonesia	KDAIS-kc342/Tsai C.C. 1342
P. amabilis ssp. *amabilis* 'Sabah'	Sabah, Indonesia	KDAIS-kc444/Tsai C.C. 1444
P. amabilis ssp. *amabilis* 'Timor'	East Timor	KDAIS-kc254/Tsai C.C. 1254
P. amabilis ssp. *amabilis* 'Timor'	East Timor	KDAIS-kc343/Tsai C.C. 1343
P. amabilis ssp. *moluccana*	Celebes, Molucca Is., Indonesia	KDAIS-kc248/Tsai C.C. 1248
P. amabilis ssp. *moluccana*	Celebes, Molucca Is., Indonesia	KDAIS-kc249/Tsai C.C. 1249
P. amabilis ssp. *moluccana*	Celebes, Molucca Is., Indonesia	KDAIS-kc319/Tsai C.C. 1319
P. amabilis ssp. *rosenstromii*	Kaiser Wilhelms, New Guinea	KDAIS-kc94/Tsai C.C. 1094
P. amabilis ssp. *rosenstromii*	Kaiser Wilhelms, New Guinea	KDAIS-kc95/Tsai C.C. 1095
P. amabilis ssp. *rosenstromii*	Kaiser Wilhelms, New Guinea	KDAIS-kc260/Tsai C.C. 1260
P. amabilis ssp. *rosenstromii*	Kaiser Wilhelms, New Guinea	KDAIS-kc329/Tsai C.C. 1329
P. aphrodite ssp. *aphrodite*	Mindanao, the Philippines	KDAIS-kc172/Tsai C.C. 1172
P. aphrodite ssp. *aphrodite*	Mindanao, the Philippines	KDAIS-kc173/Tsai C.C. 1173
P. aphrodite ssp. *aphrodite*	Mindanao, the Philippines	KDAIS-kc174/Tsai C.C. 1174
P. aphrodite ssp. *aphrodite*	Manila, Luzon, the Philippines	KDAIS-kc419/Tsai C.C. 1419
P. aphrodite ssp. *aphrodite*	Manila, Luzon, the Philippines	KDAIS-kc420/Tsai C.C. 1420
P. aphrodite ssp. *aphrodite*	Manila, Luzon, the Philippines	KDAIS-kc421/Tsai C.C. 1421
P. aphrodite ssp. *aphrodite* 'Fuga Is.'	Fuga Is., the Philippines	KDAIS-kc171/Tsai C.C. 1171
P. aphrodite ssp. *aphrodite* 'Calayan Is.'	Calayan Is., the Philippines	KDAIS-kc169/Tsai C.C. 1169
P. aphrodite ssp. *aphrodite* 'Calayan Is.'	Calayan Is., the Philippines	KDAIS-kc181/Tsai C.C. 1181
P. aphrodite ssp. *formosana*	Southern Taiwan	KDAIS-kc179/Tsai C.C. 1179
P. aphrodite ssp. *formosana*	Southern Taiwan	KDAIS-kc180/Tsai C.C. 1180
P. aphrodite ssp. *formosana*	Southern Taiwan	KDAIS-kc198/Tsai C.C. 1198
P. aphrodite ssp. *formosana*	Southern Taiwan	KDAIS-kc199/Tsai C.C. 1199
P. aphrodite ssp. *formosana*	Southern Taiwan	KDAIS-kc202/Tsai C.C. 1202
P. aphrodite ssp. *formosana*	Southern Taiwan	KDAIS-kc253/Tsai C.C. 1253
P. sanderiana	Southern Mindanao, the Philippines	KDAIS-kc35/Tsai C.C. 1035
P. sanderiana	Southern Mindanao, the Philippines	KDAIS-kc175/Tsai C.C. 1175
P. sanderiana	Southern Mindanao, the Philippines	KDAIS-kc176/Tsai C.C. 1176

[a]The classification of *Phalaenopsis* is based on Christenson (2001)
[b]Kaohsiung District Agricultural Improvement Station
[c]Voucher specimens were deposited at the herbarium of National Museum of Natural Science, Taiwan (TNM)

Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 6 of 16

Table 3 Comparisons of sequence divergence and phylogenetic information from variable sites among ITS region of nrDNA and four DNA fragments of chloroplast DNA in *Phalaenopsis amabilis* species complex

DNA region	Length (average)	Average GC content (%)	Nucleotide diversity[a] (π)	Number of variable sites	Number of informative sites[b]	Number of indels	Number of informative indels[c]	Tajima' D[d]
ITS1	228-233	77.1	0.0078	8	7	6	1	−0.2170
5.8S rDNA	163	65	0	0	0	0	0	-
ITS2	255-257	77.1	0.0039	5	4	1	0	−0.4274
*trn*H-*psb*A spacer	918-946	33.3-33.9	0.0035	11	8	5	4	−1.6542

[a]Average of percentage pairwise sequence divergences estimated using the Jukes-Cantor model. The same species were sequenced for these six regions
[b]At a phylogenetically informative sites, a nucleotide substitution is shared by two or more species
[c]At a phylogenetically informative indels, an indel is shared by two or more species
[d]Tajima's test of neutrality for nucleotide substitution. *Significant, $p < 0.05$

subdivision. The *Phalaenopsis amabilis* complex consisted of two clades with high bootstrapping supports in both NJ and MP phylogenetic trees, one for *P. amabilis* with its subspecies and *P. sanderiana* and the other for *P. aphrodite* (Figs. 2b and 3). The clade including *P. amabilis*/*P. sanderiana* consisted of two geographic groups (Fig. 2b), however, the Timor population was not closely related to *P. amabilis* ssp. *rosentromii* but was related to the Mentawei Island population (Fig. 2b)

Divergence time estimates

Based on nuclear and plastid DNAs, tests for the molecular-clock hypothesis using Tajima's *D* revealed no significant violations of the assumption of selective neutrality (Table 3). Based on substitution rates of 1.82×10^{-9} subs/site/yr for the plastid DNA spacer, the coalescence time of the *P. amabilis* complex was estimated to be 0.795 Mya, with 95 % confidence intervals (95 % CI) of 0.484–1.549 Mya (Table 5). Within the *P. amabilis* complex, the coalescence times based on the *trn*H-*psb*A spacer were estimated at 0.589 Mya (95 % CI: 0.310–1.057 Mya), 0.462 Mya (95 % CI: 0.195–0.825 Mya), and 0.138 Mya (95 % CI: 0.018–0.316 Mya) for three taxa within *P. amabilis* complex. In addition, the coalescence time of *P. amabilis* ssp. *moluccana* was tracked back to 0.069 Mya (95 % CI: 0.003–0.309 Mya) (Table 5).

Using nucleotide-substitution rates of $1.11–2.53 \times 10^{-9}$ subs/site/yr for the plastid DNA spacer, the coalescence time of the *P. amabilis* complex was estimated at 0.485 Mya (95 % CI: 0.295–0.945 Mya) and 1.105 Mya (95 % CI: 0.673–2.153 Mya) for the plastid DNA spacer (Table 5). Likewise, the coalescence time based on the substitution rates at the *trn*H-*psb*A spacer were estimated at 0.359–0.819 Mya, 0.282–0.642 Mya, and 0.084–0.192 Mya for three taxa within the *P. amabilis* complex. Furthermore, the coalescence time of *P. amabilis* ssp. *moluccana* was tracked back to 0.042–0.096 Mya (Table 5).

Demographic history and historical biogeography inference

A Bayesian skyline plot was used to estimate population dynamics. Plastid DNA showed historical demographic growth followed by recent steadiness in the *P. amabilis* complex in the late LGM due to extinction of regional populations (Fig. 4). Mismatch analyses of plastid DNA sequences displayed a multimodal distribution pattern with a non-significant sum of squared deviations (SSD) statistic and the raggedness index (HRag) value under both sudden- and spatial-expansion models in the *P. amabilis* complex (Table 6). In addition, negative *Fu's* F_S values suggested demorgraphic expansion under both sudden- and spatial-expansion model in *P. amabilis* (Table 6).

To infer vicariance and dispersal events in the *P. amabilis* complex, ancestral ranges were obtained by

Table 4 Species names, number of haplotypes, haplotype diversity (*h*), nucleotide diversity (π) for each species/subspecies of the *Phalaenopsis amabilis* species complex based on ITS1 + ITS2 region of nrDNA and the *trn*H-*psb*A spacer of chloroplast DNA

Species	No. of accessions	The ITS region/the *trn*H-*psb*A spacer				
		No. of haplotypes	Polymorphic sites	Parsimony informative sites	*h*	π
P. amabilis ssp. *amabilis*	14	6/5	8/4	6/2	0.857/0.758	0.0054/0.0012
P. amabilis ssp. *moluccana*	3	1/1	0/0	0/0	0.000/0.000	0.0000/0.0000
P. amabilis ssp. *rosenstromii*	4	1/2	0/1	0/0	0.000/0.500	0.0000/0.0005
P. aphrodite ssp. *aphrodite*	9	2/1	1/0	0/0	0.222/0.000	0.0005/0.0000
P. aphrodite ssp. *formosana*	6	1/2	0/1	0/1	0.000/0.533	0.0000/0.0006
P. sanderiana	3	2/2	2/4	0/0	0.667/0.667	0.0028/0.0029

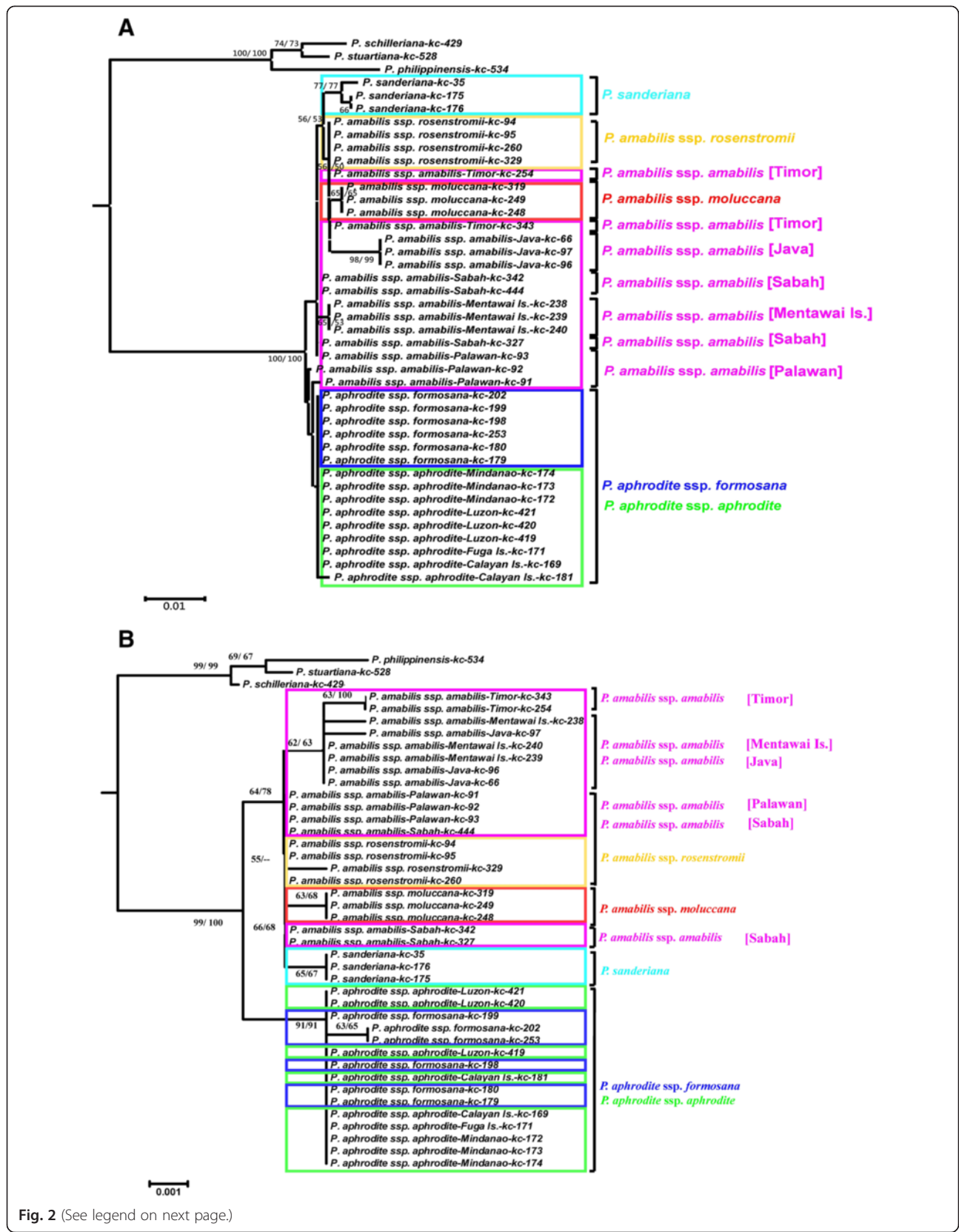Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 7 of 16

Fig. 2 (See legend on next page.)

(See figure on previous page.)
Fig. 2 Phylogenetic relationships using Neighbor joining (NJ) and Maximum Pasimony (MP) methods of the 39 accessions in the *P. amabilis* species complex, plus the three outgroups *P. schilleriana*, *P. stuartiana* and *P. philippinensis*, obtained from sequence comparisons of (**a**) the ITS region of nrDNA and (**b**) the cpDNA *trn*H-*psb*A spacer sequence and generated by MEGA 6.0 and Phylip 3.65. Numbers at nodes represent bootstrap values over 50 % of NJ and MP between major lineages

RASP analysis of plastid DNA (Fig. 3). RASP detected 17 vicariance, 36 dispersal, and 7 extinction events in the whole phylogenetic demography, suggesting a complicatedbiogeographical history in which vicariance, dispersal, and vicariance + dispersal played critical roles in shaping the current distribution in *P. amabilis* species complex. Accordingly, the extant distributions of species and subspecies of the *P. amabilis* complex may have largely been shaped by vicariance events, with relatively rare events of dispersal + vicariance or dispersal. The results supported vicariance events between *P. amabilis* and *P. aphrodite* (Fig. 3), between *P. amabilis* and *P. sanderiana*, and between *P. amabilis* ssp. *rosenstromii* and ssp. *moluccana*/ssp. *amabilis* from Sabah and Palawan/*P. sanderiana* (Fig. 3). In contrast, the analysis of plastid DNA suggested that the extant distribution of *P. aphrodite* was likely shaped by dispersal (Fig. 3). The patterns of the biogeographic reconstruction executed
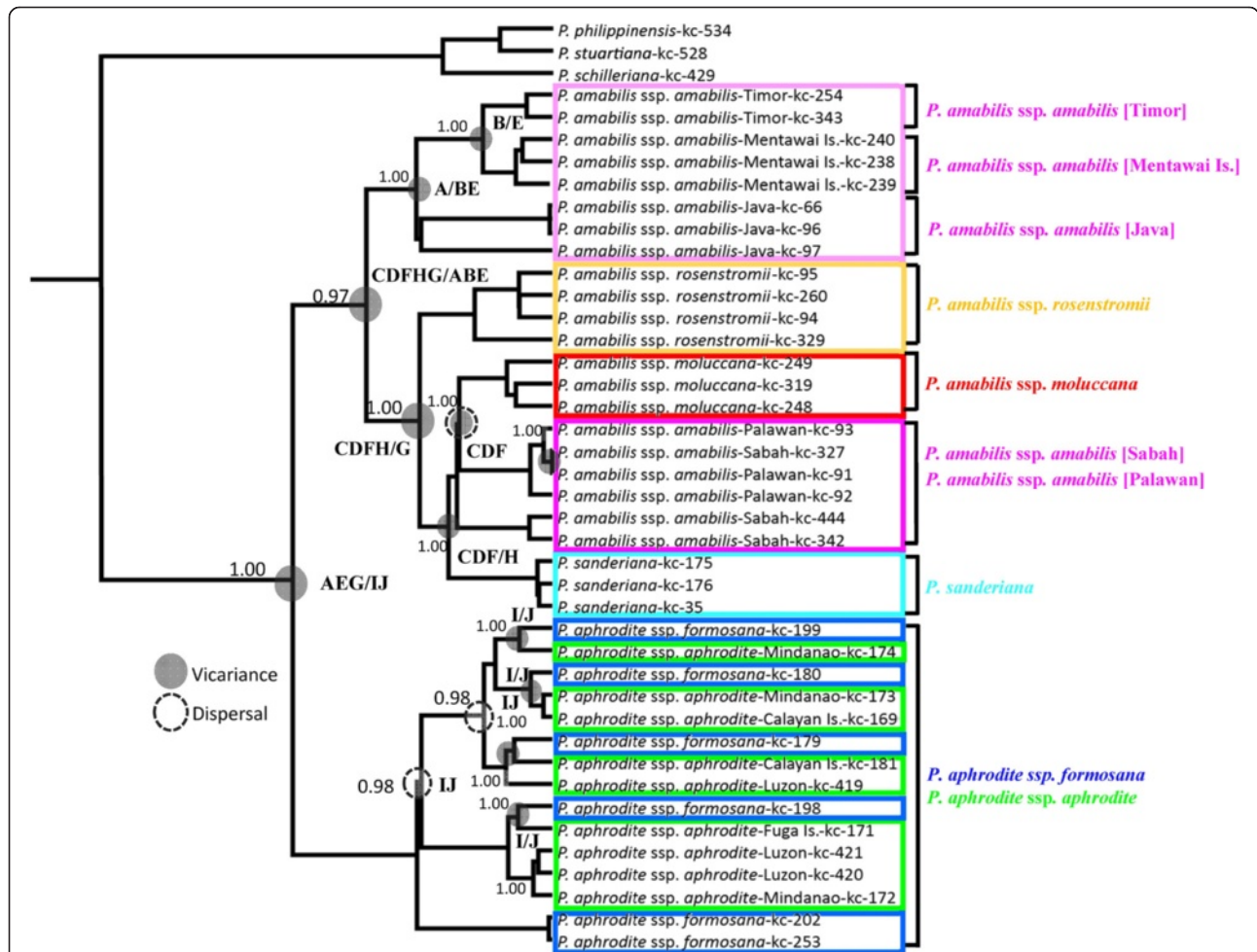

Fig. 3 Ancestral distributions reconstructed by RASP. Phylogenetic relationships of the 39 accessions in the *P. amabilis* species complex, plus the three outgroups *P. schilleriana*, *P. stuartiana* and *P. philippinensis*, obtained from sequence comparisons of the cpDNA *trn*H-*psb*A spacer sequence and generated by BEAST. Bayesian credibility values (PP) are indicated above the branch on one of the post-burn Bayesian trees. The distribution areas of extant accessions of *P. amabilis* species are marked in capitals A–J (A: Bantam, Java, Indonesia; B: Mentawai Is., Sumatra, Indonesia; C: Brooks Point, Palawan, the Philippines; D: Sabah, Indonesia; E: East Timor; F: Celebes, Molucca Is., Indonesia; G: Kaiser Wilhelms, New Guinea; H: Mindanao, the Philippines; I: Manila, Luzon, Fuga Is., and Calayan Is. in the Philippines; J: southern Taiwan), respectively. The grey circles indicate the vicariance events and the circles with dashed line indicate the long-distance dispersal (LDD) events obtained from the RASP analysis, respectively

Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 9 of 16

Table 5 Results of calescence time estimations performed with BEAST 1.8.0 for the *Phalaenopsis amabilis* complex based on the *trn*H-*psb*A spacer of chloroplast DNA

| Clade | Coalescence times (Mya) | | | | | |
| | Substitution rate = 1.82×10^{-9} subs/site/yr | | Substitution rate = 1.11×10^{-9} subs/site/yr | | Substitution rate = 2.53×10^{-9} subs/site/yr | |
	Time	95 % CI	Time	95 % CI	Time	95 % CI
P. amabilis complex	0.795	0.484 ~ 1.549	0.485	0.295 ~ 0.945	1.105	0.673 ~ 2.153
P. amabilis	0.589	0.310 ~ 1.057	0.359	0.189 ~ 0.645	0.819	0.430 ~ 1.469
P. aphrodite	0.462	0.195 ~ 0.825	0.282	0.119 ~ 0.503	0.642	0.271 ~ 1.147
P. sanderiana	0.138	0.018 ~ 0.316	0.084	0.011 ~ 0.193	0.192	0.025 ~ 0.439
P. amabilis ssp. *moluccana*	0.069	0.003 ~ 0.309	0.042	0.002 ~ 0.188	0.096	0.004 ~ 0.430

by Lagrange are presented in Fig. 5 and Table 7. Lagrange and RASP methods created similar patterns. However, Lagrange generally estimated in terms of relative probabilities at nodes and a better range of possible ancestral areas at nodes usually inferred a favored reconstruction. The ancestral range of *P. amabilis* complex (node 80) was estimated to have been in the Java and southern Taiwan with 0.174 of relative probability (Fig. 5 and Table 7). Likewise, The ancestral ranges of *P. aphrodite* (node 76), *P. amabilis* (node 65), and *P. sanderiana* (node 60) were likely in southern Taiwan with 0.588 of relative probability, Java / New Guinea with 0.223 of relative probability, and Mindanao with 0.956 of relative probability, respectively.

Discussion

Phylogeny of the *Phalaenopsis amabilis* species complex

The taxonomy of the *Phalaenopsis amabilis* species complex is difficult to resolve [8, 9], given very few diagnosable morphological characters and occurrence of interspecific hybridisation yields numerous intermediate morphs. Nevertheless, unique biogeographical patterns with high levels of endemism on islands exist in the Sunda Shelf and the Philippine Archipelagos [6, 11], reflecting interesting but complicated divergence processes. In this study, two molecular markers provided supplementary information, with patterns in nrITS shaped by concerted evolution [19], whereas plastid DNA genes showed no evidence of genetic recombination [20]. Phylogenetic trees derived from nuclear and plastid DNAs supported the monophyly of the species complex. *P. aphrodite* ssp. *aphrodite* and *P. aphrodite* ssp. *formosana* formed a sister group to other species, but are difficult to distinguish from each other. Morphologically, *P. aphrodite* ssp. *formosana*, which lacks anthocyanin, is only slightly differentiated from *P. aphrodite* ssp. *aphrodite* [9].

Phalaenopsis amabilis is distinguishable from *P. aphrodite*, and both display geographic subdivision based on characteristics of the callus. The reciprocal monophyly
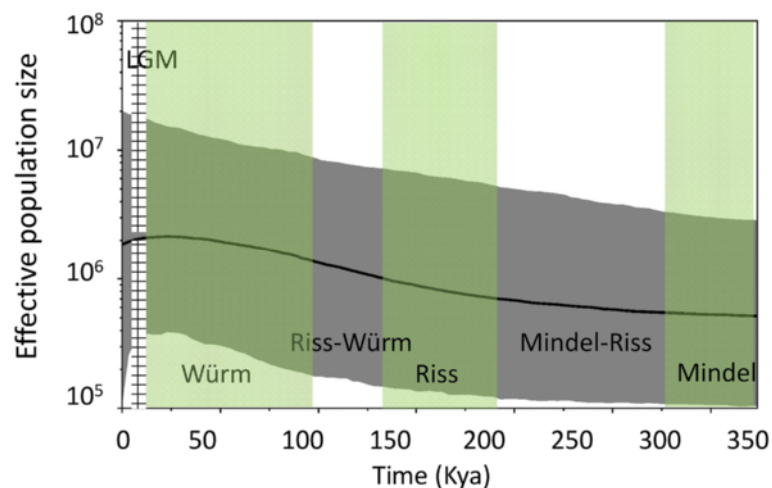

Fig. 4 The effective population size over time for all accessions in the *P. amabilis* species complex based on the cpDNA *trn*H-*psb*A spacer using Bayesian skyline plots analyses. The Last Glacial Maximum (LGM) is green color shaded. Solid lines indicate median estimations; area between gray dash lines indicates 95 % confidence intervals

Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 10 of 16

Table 6 Results of demographic analyses performed with DnaSP and ARLEQUIN for the *Phalaenopsis amabilis* complex based on the *trn*H-*psb*A spacer of chloroplast DNA

Group	Fu's *Fs*	SSD[a]	H_{Rag}^{a}	SSD[b]	H_{Rag}^{b}
*trn*H-*psb*A spacer	−1.645	0.0143	0.0328	0.0153	0.0328
P. amabilis	−2.715*	0.0149	0.1079	0.0149	0.1079
P. aphrodite	−0.133	0.2814	0.3161	0.0008	0.3161
P. sanderiana	———	———	———	———	———

[a]the indices under the sudden expansion model
[b]the indices under the spatial expansion model
*$p < 0.05$, **$p < 0.01$

of three subspecies of *P. amabilis*, with *P. amabilis* ssp. *moluccana* and *P. amabilis* ssp. *rosenstromii* clustering first was supported by the plastid-DNA-based tree (Fig. 2b) but not the nuclear-DNA-based tree (Fig. 2a). The systematic inconsistency between the two trees might derive from differences in the modes of inheritance, with plastid DNA and the ITS region being maternally and biparentally inherited, respectively [21, 22]. Close affinity between sister groups of *P. amabilis* and *P. sanderiana* based on plastid DNA (Fig. 2b) revealed a biogeographcial pattern subdivided by Huxley's Line (Fig. 1) [11]. In addition, the *P. amabilis* population of

Timor and two subspecies of *P. amabilis* distributed in regions east of Wallace's Line (Fig. 1) formed a clade in the nuclear-DNA-based tree, supporting another geographic subdivision by Weber's Line. Nevertheless, some samples from Palawan populations mixed with those of the *P. amabilis* complex in a nuclear-ITS-based tree, indicating possible paternal gene flow (Fig. 1). This result suggests long-distance pollen dispersal in *P. amabilis*. Such dispersal would usually occur with the aid of large pollinators, such as carpenter bees (*Xylocopa*) [9], which are capable of flying [23].

Morphologically, the callus structure of *P. sanderiana* can be distinguished from those of *P. amabilis* and *P. aphrodite* [9]. According to phylogenetic trees derived from nuclear and plastid DNAs, *P. sanderiana* alone formed a clade, nested in *P. amabilis*, making the latter species paraphyletic.

Biogeography and evolutionary trends
Gene genealogies were used to extract data on historical gene flow [24], hybridisation [25] and divergence between lineages [26–29]. According to the phylogenies based on nuclear (Fig. 2) and plastid DNA (Fig. 3), *P. aphrodite* ssp. *aphrodite* and ssp. *formosana* distributed
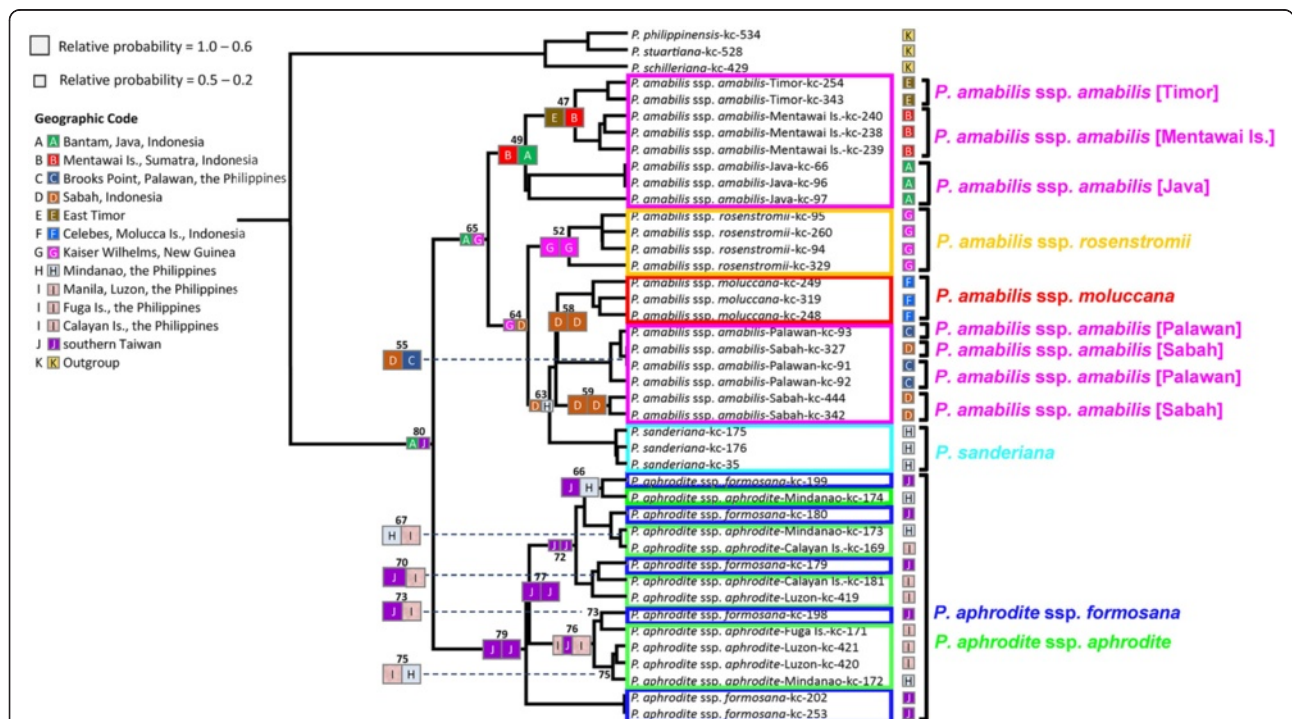

Fig. 5 Historical biogeographical reconstruction using Lagrange on the *P. amabilis* species complex topology. Coloured squares indicate reconstructed ancestral ranges and the square size is proportional to the probability of the reconstructions (see Table 7). The geographic ranges of species are displayed at right. [left | right]: 'left' and 'right' are the ranges inherited by each descendant branch (in the printed tree, 'left' is the upper branch, and 'right' the lower branch). The distribution areas of extant accessions of *P. amabilis* species are marked in capitals A–J (A: Bantam, Java, Indonesia; B: Mentawai Is., Sumatra, Indonesia; C: Brooks Point, Palawan, the Philippines; D: Sabah, Indonesia; E: East Timor; F: Celebes, Molucca Is., Indonesia; G: Kaiser Wilhelms, New Guinea; H: Mindanao, the Philippines; I: Manila, Luzon, Fuga Is., and Calayan Is. in the Philippines; J: southern Taiwan), respectively

Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 11 of 16

Table 7 The ancestral areas inferred through Lagrange. Relative probability is the fraction of the global likelihood of a split

Node	Ancestral range	-ln(L)	Relative probability
80	[A|J]	−100.1	0.174
79	[J|J]	−98.84	0.588
77	[J|J]	−98.68	0.689
76	[I|I]	−98.43	0.886
75	[I|H]	−98.41	0.913
73	[J|I]	−98.44	0.875
72	[J|J]	−99.22	0.401
70	[J|I]	−98.45	0.862
67	[H|I]	−98.36	0.942
66	[J|H]	−98.53	0.796
65	[A|G]	−99.81	0.223
64	[G|D]	−99.61	0.271
63	[D|H]	−99.32	0.363
59	[D|D]	−98.32	0.981
58	[D|D]	−98.82	0.596
55	[D|C]	−113.5	1.000
52	[G|G]	−98.50	0.825
49	[B|A]	−98.76	0.635
47	[E|B]	−98.76	0.638

The distribution areas of extant accessions of *P. amabilis* species are marked in capitals A–J (A: Bantam, Java, Indonesia; B: Mentawai Is., Sumatra, Indonesia; C: Brooks Point, Palawan, the Philippines; D: Sabah, Indonesia; E: East Timor; F: Celebes, Molucca Is., Indonesia; G: Kaiser Wilhelms, New Guinea; H: Mindanao, the Philippines; I: Manila, Luzon, Fuga Is., and Calayan Is. in the Philippines; J: southern Taiwan), respectively

in the Philippines and Taiwan is a sister species to the rest of the *P. amabilis* complex. *P. aphrodite* ssp. *formosana* is geographically restricted to Taiwan. This subspecies represented the northernmost limitation in distribution for *Phalaenopsis*, while *P. aphrodite* is distributed throughout the Philippines [8]. In addition, the Philippines are the centre of diversity for species of this group of *Phalaenopsis* spp. (*P. amabilis* complex and *P. schilleriana* complex) [30], as all members of the section *Phalaenopsis*, except for *P. amabilis*, were found in the Philippines [8, 9]. In addition, several old islands of the Philippines, including Palawan, Mindoro and Zamboanga, were separated from the Eurasian Plate when the crust of the plate hit Borneo approximately 5–10 Mya [3–5], providing a refuge for the diverging species.

Phalaenopsis aphrodite is absent in Palawan and southern Mindanao [8, 9]. These two islands were interconnected during the glacial periods [6] (Fig. 1), providing opportunities for population expansion. Furthermore, *P. aphrodite* probably dispersed into Taiwan during the glacial maxima, as the Taiwan-Luzon-Mindoro belt was formed in the late Eocene [4]. Phylogenetically, *P. aphrodite* ssp. *formosana* and *P. aphrodite* are very similar based on the nuclear ITS and the chloroplast *trn*H-*psb*A spacer. The coalescence of *P. aphrodite* can be traced to approximately 0.535 Mya (95 % CI: 0.257–0.953 Mya), ranging from 0.312 to 0.758 Mya (plastid *trn*H-*psb*A) during the Pleistocene glaciations. The Philippine Archipelagos, except for Palawan, formed the super island [31] and were periodically linked and unlinked to the Sunda Shelf by the land bridge between Mindoro, Palawan, and Borneo. The most recent link likely formed approximately 0.01–1.8 Mya as a result of lowered Pleistocene sea levels [6, 32] and *P. aphrodite* probably dispersed among Mindoro, Palawan, Borneo, and Taiwan. This geographic disjunction likely contributed to the vicariance between the Philippine Archipelagos and continental Asia.

Geographically, *P. amabilis* is separated from *P. amabilis* ssp. *moluccana* and *P. amabilis* ssp. *rosenstromii* by Wallace's Line. There is divergence between floral and faunal species in the regions west and east of Wallace's Line [6, 33, 34]. Biogeographical subdivision also occurred between *P. amabilis* ssp. *moluccana* (Sulawesi) and *P. amabilis* ssp. *rosenstromii* (New Guinea/Australia). The coalescence of *P. amabilis* can be dated to 0.589 Mya (95 % CI: 0.310–1.057 Mya) by the plastid *trn*H-*psb*A, likely in association with the Würm glaciation. This finding is in agreement with the phylogeographical break described by Weber [11]. The current biota in this geographic area therefore show evidence of vicariace from the Sunda Shelf west of New Guinea/Australia [33, 34].

Inconsistent phylogenetic demographies between nuclear and plastid markers indicate historical introgression/hybridisation [35, 36] between *P. amabilis* and *P. aphrodite* (Fig. 2) may due to recombination by paternal gene flow . The *P. amabilis* distributed in Palawan is closely related to the *P. amabilis* distributed in Borneo based on maternal molecular data and historical connection by land bridges [37, 38] between the Sunda Shelf and the Philippine Archipelagos, however, *P. amabilis* was paraphyletic according to the nuclear ITS, with one major clade distributed in Palawan and one haplotype clustering with *P. aphrodite*. The coalescence time based on plastid data was less than 0.10 Mya between the Palawan and Borneo populations, indicating a very recent split. That result is consistent with the topophysiographic data [39].

Phalaenopsis sanderiana forms a clade (Figs. 2, 3 and 5) linked to *P. amabilis* of Borneo/Palawan. Geographically, *P. sanderiana* is exclusively found in southern Mindanao [8, 9]. Several islands between southern Mindanao and Borneo may have formed a land bridge during the glacial maxima, providing opportunities for dispersal from Borneo to southern Mindanao [6, 33, 34].

RASP analyses also suggested that the current disjunct distribution of the *P. amabilis* complex might reflect vicariant relics of an ancestral population (Fig. 3). This finding concurs with findings for other disjunctions at

Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 12 of 16

the Sunda Shelf and Philippine Archipelagos or Sunda Shelf and New Guinea/Australia [33, 34]. In addition, several vicariance, dispersal, and vicariance + dispersal events detected using RASP (Fig. 3) revealed geographic isolation among populations/taxa during the Pleistocene, likely as a result of the landbridge submergence [38, 40, 41].

Conservation genetics of *P. amabilis* complex

Phalaenopsis amabilis complex represents a group of epiphytes of typical evergreen forests throughout tropical Asia and the larger islands of the Pacific Ocean; the whole region was the ancient Sunderland and Great Sunda Island during the Pleistocene glaciation [6]. As epiphytic species widespread in tropical Asia, the *P. amabilis* complex is expected to have had a large ancestral population. In this study, population growth was detected with sudden- and spatial-mismatch distribution models for three taxa and the species complex based on the Bayesian skyline plots using BEAST program for species. At plastid marker, a continuous demographic expansion from Mindel to Würm pleistocene period of the *P. amabilis* complex was detected, followed by subsequent reversed population slowdown after LGM (Fig. 4). Climate changes in the Quaternary that triggered isolation and local extinction based on 7 extinction events detected by RASP (thus promoting speciation) may explain the increased differentiation rates during that period [42, 43]. During the Pleistocene, geographic fluctuations in the Sunda Shelf, Philippine Archipelagos, and New Guinea/Australia resulted in population fragmentation. At the same, colonization via long distance dispersal also occurred. The ancestral population of the *P. amabilis* complex was likely to split into *P. aphrodite*, *P. amabilis*, and *P. sanderiana* by geological vicariance associated with Huxley's Line [44] and Wallace's Line [11].

At the species level, the nucleotide diversity of nuclear DNA (*P. amabilis*: 0.0054, *P. aphrodite*: 0.0005, *P. sanderiana*: 0.0028) and cpDNA (*P. amabilis*: 0.0012, *P. aphrodite*: 0.0000, *P. sanderiana*: 0.0029) was low (Table 4) compared to other flowering plants from the same geographic region, such as *Kandelia candel* (nucleotide diversity $\theta = 0.02652$, Chiang et al. [45]), *Ceriops tagal* (nucleotide diversity $\pi = 0.01703$, Liao et al. [46]), *Bruguiera gymnorrhiza* ($\pi = 0.08$, Urashi et al. [47]). As an epiphyte growing in vulnerable habitats across a wide distribution, all local populations tend to be small, strengthening the effects of genetic drift [48]. Therefore, taxa of the *P. amabilis* complex tend to have low genetic diversity at the intraspecific level in both nuclear and plastid DNAs and populations are fixed at a single haplotype (Table 4).

Moreover, the nucleotide diversity of other sympatrically distributed species and the *P. amabilis* complex reflects the effects of a common geological history, likely including the Quaternary glacial circle [38, 40, 49, 50].

At the population and species levels, the combination of low genetic diversity and high levels of genetic differentiation in the *P. amabilis* complex indicate that populations of this species were isolated from each other and experienced either extreme bottlenecks [29, 51–53] or founder effects [54, 55]. Furthermore, the unique interactions between orchids and pollinators [56] and the habitat preferences [57] strengthened the genetic differentiation among populations/species. Given such genetic background, most species, apparently, face threats of human exploitation of natural forests that inevitably led to shrinkage of populations of the species complex. Conservation of these evolving species/subspecies is therefore urgent and required.

Conclusions

In summary, variations in nuclear and plastid DNA revealed that the disjunct distribution of the *P. amabilis* complex likely resulted from vicariance or vicariance + dispersal in the middle–late Pleistocene. Due to the submergence of the Sunda Shelf, Philippine Archipelagos, and New Guinea/Australia, *P. aphrodite* was separated from other members of the *P. amabilis* complex. Within *P. aphrodite*, the admixture of populations made subspecies indistinguishable, suggesting a history of long-distance dispersal. In contrast, the species and subspecies of *P. amabilis* with disjunct distributions were differentiated due to vicariance. After divergence, the species complex experienced isolation and *in situ* range expansion as well as population growth during subsequent climatic oscillations in the Würm glacial period. Population slowdown following the end of the LGM may have occurred by an extinction of local or island populations owing to a reduction in suitable habitats and formed the contemporary pattern of geographic distribution.

Methods

Plant materials

Thirty-nine accessions of the *P. amabilis* complex were obtained from 13 different populations, and three species of the *P. schilleriana* complex were used as outgroups (Table 2) [37, 58]. Four individuals were sampled from each accession. Leaf materials were taken from living plants. All samples examined in this study were collected from wild populations and cultivated at the green house of the Kaohsiung District Agricultural Improvement Station (KDAIS) in Taiwan, and the voucher specimens were deposited at the herbarium of the National Museum of Natural Science, Taiwan (TNM). The geographic distributions and flower photos of species of this complex are shown in Fig. 1.

DNA extraction, PCR amplification and sequencing

Using the cetyltrimethylammonium bromide (CTAB) method [59], total DNA was extracted from fresh etiolated

Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 13 of 16

leaves. Ethanol-precipitated DNA was dissolved in TE (Tris-EDTA) buffer and stored at −20 °C. Qiagen (Valencia, CA, USA) columns were used to clean the DNA samples that were difficult by amplify by PCR. Approximate DNA yields were determined using a spectrophotometer (model U-2001, Hitachi).

Amplification protocols were as follows: each 50-µL reaction contained 40 mM Tricine-KOH (pH 8.7), 15 mM KOAc, 3.5 mM Mg(OAc)$_2$, 3.75 µg ml^{-1} BSA, 0.005 % Tween 20, 0.005 % Nonidet-P40, four dNTPs (0.2 mM each), primers (0.5 µM each), 2.5 units of Advantage 2 DNA polymerase (Clontech, USA), 10 ng of genomic DNA, and 50 µL of mineral oil. Amplification reactions were completed in a dry-block with two-step thermal cycles (Biometra, Germany). Two primers designed to amplify *Phalaenopsis* ITS and the PCR conditions were described in Tsai et al. [58], and the primers and conditions for the *trn*H-*psb*A spacer were described in Taberlet et al. [60]. The PCR conditions for the *trn*H-*psb*A spacer were as follows: incubation at 94 °C for 3 min, 10 cycles of denaturation at 94 °C for 45 s, annealing at 54 °C for 30 s, and extension at 72 °C for one min, 30 cycles of denaturation at 94 °C for 45 s, annealing at 52 °C for 30 s, extension at 72 °C for one min, and a final extension for five min at 72 °C. PCR products were stained with 0.5 µg ml^{-1} ethidium bromide, detected by agarose-gel electrophoresis (1.0 % w v^{-1}, in TBE), and photographed under UV light.

All PCR products of different DNA fragments from the plant material were recovered by glassmilk (BIO 101, California) and sequenced directly by the dideoxy chain-termination method, using an ABI3730 automated sequencer with a Bigdye^{TM} Terminator Cycle Sequencing Ready Reaction Kit (PE Biosystems, California). The sequencing primers were the same as those used for PCR. These reactions were performed according to the recommendations of the manufacturers.

Sequence alignment and genetic-diversity analysis
The sequences of the nrITS region and *trn*H-*psb*A spacer were deposited in GenBank under accession numbers of AY391515-AY391553 and FJ460366-FJ460407, respectively. The sequences were aligned using the MUSCLE multiple-alignment program in MEGA 6.0 [61]. Genetic diversity was evaluated based on indices of haplotype diversity [62] and nucleotide diversity [62] using DnaSP 5.10.1 [63].

Phylogenetic reconstruction
Insertion/deletions (indels) were coded as missing data. For phylogenetic analyses, the three outgroups were included to determine whether all recovered haplotypes formed a monophyletic lineage. Phylogenetic relationships were inferred using NJ in MEGA 6.0 [61] and MP

in PHYLIP 3.65 [64], with the best-fitting Tamura 3-parameter model (T92) using a discrete Gamma distribution (+G) for the nrITS region (Additional file 2: Table S2) and Tamura 3-paramenter model (T92) [65] for the *trn*H-*psb*A spacer (Additional file 3: Table S3) selected by a model test using MEGA 6.0 [61], with indels treated as missing data. Bootstrapping (1000 replicates) was carried out to estimate the support for NJ and MP topology [66, 67]. The NJ phylogenetic tree was drawn with bootstrapping values using the MEGA 6.0 and combined the bootstrapping information from MP phylogenetic tree.

Divergence time estimation
To avoid the noise of introgression and recombination due to paternal gene flow, we estimated the divergence time and historical demographic evaluation exclusively based on non-coding cpDNA regions. For estimating divergence time, a Bayesian estimate of group delimitation and the ages of TMRCA of the moth orchid clades on each node were inferred using Yule model methods in BEAST 1.8.0 [68], with the best-fitting HKY model selected by a model test using MEGA 6.0 [61], and indels treated as missing data. For Bayesian analysis, four parallel runs and four Markov Chain Monte Carlo (MCMC) chains, each with a different starting seed, were run for 100,000,000 generations. A tree was sampled every 10,000 generations after a burn-in period of 10,000,000 generations, after which the standard deviation of the split frequencies was below 0.01, per the suggestion in the manual. BEAST 1.8.0 was used to reconstruct the gene trees and to simultaneously estimate divergence between and within clades. As the taxon-specific substitution rate had not previously been calibrated for non-coding cpDNA regions of the *P. amabilis* complex, we used the synonymous (*Ks*) substitution rates of 0.36 ± 0.14 per site for cpDNA coding regions between *P. aphrodite* and grasses [69] as the reference, based on the coalescence time of 99 Mya, which is the age of the oldest known Asparagales [70]. We estimated a substitution rate of 1.82×10^{-9} subs/site/yr and set 1.11×10^{-9} subs/site/yr as the lower limit and 2.53×10^{-9} subs/site/yr as the upper limit for the nucleotide-substitution rate of the cpDNA spacer in our analyses. All summary statistics of the output values were summarised using Tracer 1.5 [71], and both log and tree files of the last four runs were combined using LogCombiner 1.6.1 [72]. TreeAnnotator 1.6.1 [72] and FigTree 1.3.1 [73] were used to summarise and display the sampled trees, respectively.

Historical demographic evaluation
To examine the demographic history of *P. amabilis* complex, nucleotide sequences were tested using Tajima's D test of neutrality [74] and Fu's *Fs* statistics [75], which are powerful for detecting population growth [76], using

Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 14 of 16

DnaSP 5.10.1. Mismatch analysis was performed to evaluate the range expansions of the *P. amabilis* complex under the sudden-expansion model using Arlequin var. 3.5.1.2 [77] and was compared to Poisson distributions. The SSD between the expected and observed mismatch distributions and *P*-values were calculated. The HRag and its significance for observed distributions were used to compute the smoothness of the mismatch distributions [78]. Low raggedness is typical of recently non-stationary, expanding populations.

To investigate the historical demographics, a coalescence approach with a Bayesian skyline plot was used to evaluate the dynamic history of the *P. amabilis* complex using BEAST 1.8.0. The best-fitting substitution model, substitution rate, and general parameter setting followed the phylogenetic reconstruction using BEAST described above. All summary statistics of the output values and the Bayesian skyline plot were generated using Tracer 1.5 [71], and the log and tree files of the last four runs were combined using LogCombiner 1.6.1 [72].

Biogeographic inference using RASP and Lagrange

To infer the biogeographic history of the *P. amabilis* complex and thereby distinguish the effects of vicariance from dispersal, we condcuted two different approaches. First, we used the Statistical dispersal–vicariance analysis (S-DIVA) and Bayesian binary MCMC (BBM) analysis, performed in RASP 2.0 [79]. RASP is one of the most widely used methods for inferring biogeographic histories, and distingushing vicariance from dispersal. Second, we used Lagrange v. 20130526 [80] to test the complex models of directional biogeographic expansion and inferred the geographic origin of the clade based on a dispersal-extinction-cladogenesis (DEC) model employing a maximum likelihood framework for testing models of geographic range evolution. To clarify the vicariance and dispersal events, the geographic distributions of the *P. amabilis* complex were mapped onto 10 areas across islands: (A) Bantam, Java, Indonesia; (B) Mentawai Is., Sumatra, Indonesia; (C) Brooks Point, Palawan, the Philippines; (D) Sabah, Boeneo, Indonesia; (E) East Timor; (F) Celebes, Molucca Is., Indonesia; (G) Kaiser Wilhelms, New Guinea; (H) Mindanao, the Philippines; (I) Manila, Luzon, Fuga Is., and Calayan Is. in the Philippines; and (J) southern Taiwan. The tree topologies constructed with BEAST and 10,000 trees were used in S-DIVA and Lagrange analysis. At each node, the number of maximal areas endorsed in ancestral distributions was set as 3 and other parameters were automatically implemplemented in S-DIVA. The range matrix defining the region from each specimen of moth orchids and a phylogenetic tree was uploaded to the Lagrange configurator [80], biogeographic models were defined as allowing for two potential ancestral regions per node in the Lagrange analyses.

Availability of supporting data

The data used to perform the phylogenetic analyses is available at Dryad [http://dx.doi.org/10.5061/dryad.f8j12].

Additional files

> **Additional file 1: Table S1.** One part of the cpDNA *trn*H-*psb*A spacer sequence alignment. Three different types of repeat sequences were found and are marked by gray regions.
>
> **Additional file 2: Table S2.** The model test for *Phalaenopsis amabilis* species complex based on the nrITS region using MEGA 6.0.
>
> **Additional file 3: Table S3.** The model test for *Phalaenopsis amabilis* species complex based on the plastid noncoding region using MEGA 6.0.

Abbreviations

NJ: Neighbor joining; MP: Maximum parsimony; BEAST: Bayesian evolutionary analysis sampling trees; RASP: Reconstruct Ancestral State in Phylogenies; LGM: Last glacial maximum; Mya: Million years ago; KDAIS: Kaohsiung District Agricultural Improvement Station; TNM: National Museum of Natural Science, Taiwan; CTAB: Cetyltrimethylammonium bromide; T92: Tamura 3-paramenter model; MCMC: Markov Chain Monte Carlo; SSD: Sum of squared deviations; HRag: The raggedness index; S-DIVA: Statistical dispersal–vicariance analysis; BBM: Bayesian binary MCMC; DEC: dispersal-extinction-cladogenesis.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CHC, TYC and YCC supervised the project. CCT collected plant sample in the field. CCT conducted the experiment. YCC, YZK and TYC analyzed the data. CCT, CHC, HVW TYC, and YCC wrote the manuscript. All authors read and approved the final manuscript.

Author details

[1]Crop Improvement Division, Kaohsiung District Agricultural Improvement Station, Pingtung 900, Taiwan. [2]Graduate Institute of Biotechnology, National Pingtung University of Science and Technology, Pingtung 912, Taiwan. [3]Department of Biological Sciences, National Sun Yat-sen University, Kaohsiung 804, Taiwan. [4]Research Center for Biodiversity, China Medical University, Taichung 404, Taiwan. [5]Department of Life Science, National Cheng Kung University, Tainan 701, Taiwan. [6]Department of Biomedical Science and Environment Biology, Kaohsiung Medical University, Kaohsiung 807, Taiwan.

References

1. Aurelio MA, Barrier E, Rangin C, Muller C. The Philippine Fault in the late Cenozoic tectonic evolution of the Bondoc-Masbate-N, Leyte area, central Philippines. J SE Asian Earth Sci. 1991;6:221–38.
2. Quebral RD, Pubellier M, Rangin C. The Mindanao: a transition from collision to strike-slip environment. Tectonics. 1994;15:713–26.
3. Karig DE, Sarewitz DR, liaeck GD. Role of strike-slip faulting in the evolution of allochthonous terraces in the Philippines. Geology. 1986;14:852–5.
4. Stephan JF, Blanchet R, Rangin C, Pelletir B, Letouzey J, Muller C. Geodynamic evolution of the Taiwan-Luzon-Mindoro belt since the Late Eocene. Tectonophysics. 1986;125:245–68.
5. Hall R. Reconstructing Cenozoic SE Asia. In: Hall R, Blundell DJ, editors. Tectonic Evolution of SE Asia, The Geological Society, vol. 106. 1996. p. 153–84.
6. van Oosterzee P. Where Worlds Collide: The Wallace Line. Ithaca, New York: Cornell University Press; 1997.

Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 15 of 16

7. Moss SJ, Wilson MEJ. Biogeographic implications of the Tertiary palaeogeographic evolution of Sulawesi and Borneo. In: Hall R, Holloway JD, editors. Biogeography and Geological Evolution of SE Asia. Leiden: Backhuys; 1998.

8. Sweet HR. The genus Phalaenopsis. Pomona, California: The Orchid Digest; 1980.

9. Christenson EA. *Phalaenopsis*. Portland, Oregon: Timber Press; 2001.

10. Turner H, Hovenkamp P, van Welzen PC. Biogeography of Southeast Asia and the West Pacific. J Biogeogr. 2001;28:217–30.

11. Pianka ER. Biogeography and Historical Constraints. In: Evolutionary Ecology. 5th ed. New York: HarperCollins College Publisher; 1994. p. 15–40.

12. Lydekker R. A Geographical History of Mammals. United Kingdom: Cambridge University Press; 1896.

13. Sweet HR. A revision of the genus *Phalaenopsis*-VII. Amer Orchid Soc Bull. 1969;38:681–94.

14. Tsai CC, Huang SC, Huang PL, Chou CH. Phylogeny of the genus *Phalaenopsis* (Orchidaceae) with emphasis on the subgenus *Phalaenopsis* based on the sequences of the internal transcribed spacers 1 and 2 of rDNA. J Hort Sci Biotech. 2003;78:879–87.

15. Sweet HR. A revision of the genus *Phalaenopsis*-VIII. Amer Orchid Soc Bull. 1969;38:888–901.

16. Briggs JC. Historical biogeography: The pedagogical problem. J Biogeogr. 1991;18:3–6.

17. Brundin LZ. Phylogenetic biogeography. In: Myers AA, Giller PS, editors. Analytical biogeography. London: Chapman and Hall; 1988. p. 343–69.

18. Tryon Jr RM. Development and evolution of fern floras of oceanic islands. Biotropicn. 1970;2:76–84.

19. Hillis DM, Moritz C, Porter CA, Baker RJ. Evidence for biased gene conversion in concerted evolution of ribosomal DNA. Science. 1991;251:308–10.

20. Chiu WL, Sears BB. Recombination between chloroplast DNAs does not occur in sexual crosses of Oenothera. Mol Gen Genet. 1985;198:525–8.

21. Chang SB, Chen WH, Chen HH, Fu YM, Lin YS. RFLP and inheritance patterns of chloroplast DNA in intergeneric hybrids of *Phalaenopsis* and *Doritis*. Bot Bull Acad Sin. 2000;41:219–23.

22. Tsai CC, Huang SC, Huang PL, Chen FY, Su YT, Chou CH. Molecular evidences of a natural hybrid origin of *Phalaenopsis × intermedia* Lindl. J Hort Sci Biotech. 2006;81:691–9.

23. Pasquet RS, Peltier A, Hufford MB, Oudin E, Saulnier J, Paul L, et al. Long-distance pollen flow assessment through evaluation of pollinator foraging range suggests transgene escape distances. Proc Natl Acad Sci U S A. 2008;105:13456–61.

24. Hsu TW, Shih HC, Kuo CC, Chiang TY, Chiang YC. Characterization of 42 microsatellite markers from poison ivy, *Toxicodendron radicans* (Anacardiaceae). Int J Mol Sci. 2013;14:20414–26.

25. Chiang YC, Huang BH, Chang CW, Wan YT, Lai SJ, Huang S, et al. Asymmetric introgression in the horticultural living fossil Cycas Sect, Asiorientales using a genome-wide scanning approach. Int J Mol Sci. 2013;14:8228–51.

26. Donnelly P, Tavare S. Coalescent and genealogical structure under neutrality. Annu Rev Genet. 1995;29:401–21.

27. Hudson RR. Island models and the coalescent process. Mol Ecol. 1998;7:413–8.

28. Templeton AR. Gene tree: A powerful tool for exploring the evolutionary biology species and speciation. Plant Spec Biol. 2000;15:211–22.

29. Ge XJ, Hung KH, Ko YZ, Hsu TW, Gong X, Chiang TY, et al. Genetic divergence and biogeographical patterns in *Amentotaxus argotaenia* species complex. Plant Mol Biol Rep. 2014;33(2):264–80. doi:10.1007/s11105-014-0742-0.

30. Vavilov NI. Centers of origin of cultivated plants. Trends Pract Bot Genet Sel. 1926;16:3–248.

31. Concepcion R, Dimalanta C, Yumul G, Faustino-Eslava D, Queaño K, Tamayo R, et al. Petrography, geochemistry, and tectonics of a rifted fragment of Mainland Asia: evidence from the Lasala Formation, Mindoro Island, Philippines. Int J Earth Sci. 2012;101:273–90.

32. Brown RM, Diesmos AC. Application of lineage-based species concepts to oceanic island frog populations: the effects of differing taxonomic philosophies on the estimation of Philippine biodiversity. Silliman J. 2002;42:133–62.

33. Tougard C. Biogeography and migration routes of large mammal faunas in South-East Asia during the Late Middle Pleistocene: focus on the fossil and extant faunas from Thailand. Palaeogeogr Palaeoclimatol Palaeoecol. 2001;168:337–58.

34. Lohman DJ, de Bruyn M, Page T, von Rintelen K, Hall R, Ng PKL, et al. Biogeography of the Indo-Australian Archipelago. Annu Rev Ecol Evol Syst. 2011;42:205–26.

35. Liao PC, Shih HC, Yen TB, Lu SY, Cheng YP, Chiang YC. Molecular evaluation of interspecific hybrids between *Acer albopurpurascens* and *A. buergerianum* var. *formosanum*. Bot Stud. 2010;51:413–20.

36. Liao PC, Tsai CC, Chou CH, Chiang YC. Introgression between cultivars and wild populations of *Momordica charantia* L. (Cucurbitaceae) in Taiwan. Int J Mol Sci. 2012;13:6469–91.

37. Tsai CC, Chiang YC, Huang SC, Chen CH, Chou CH. Molecular phylogeny of *Phalaenopsis* Blume (Orchidaceae) based on the plastid and nuclear DNA. Plant Syst Evol. 2010;288:77–98.

38. Ge XJ, Hsu TW, Hung KH, Lin CJ, Huang CC, Huang CC, et al. Inferring mutiple refugia and phylogeographical patterns in *Pinus massoniana* based on nucleotide sequence variation and fingerprinting. PLoS One. 2012;7:e43717.

39. Fowlie JA. Speciation amongst the Orchidaceae as a function of climate change and topophysiography. Orch Dig. 1981;45:44–9.

40. Chiang TY, Schaal BA. Phylogeography of plants in Taiwan and the Ryukyu Archipelago. Taxon. 2006;55:31–41.

41. Chiang YC, Hung KH, Moore SJ, Ge XJ, Huang S, Hsu TW, et al. Paraphyly of organelle DNAs in *Cycas* Sect. Asiorientales due to ancient ancestral polymorphisms. BMC Evol Biol. 2009;9:161.

42. Graham J, Marshall B, Squire GR. Genetic differentiation over a spatial environmental gradient in wild Rubus ideaus populations. New Phytol. 2003;157:667–75.

43. Gavrilets S, Losos JB. Adaptive radiation: Contrasting theory with data. Science. 2009;323:732–7.

44. Huxley TH. On the classification and distribution of the Alectoromorphae and Heteromorphae. Proc Zool Soc. 1868;14:294–319.

45. Chiang TY, Chiang YC, Chou CH, Havanond S, Hong TN, Huang S. Phylogeography of *Kandelia candel* in East Asiatic mangroves based on nucleotide variation of chloroplast and mitochondria DNAs. Mol Ecol. 2001;10:2697–710.

46. Liao PC, Chiang YC, Huang S, Wang JC. Gene flow of *Ceriops tagal* (Rhizophoraceae) across the Kra Isthmus in the Thai Malay Peninsula. Bot Stud. 2009;50:193–204.

47. Urashi C, Teshima KM, Minobe S, Koizumi O, Inomata N. Inferences of evolutionary history of a widely distributed mangrove species, *Bruguiera gymnorrhiza*, in the Indo-West Pacific region. Ecol Evol. 2013;3:2251–61.

48. Hamrick JL, Godt MJ. Effects of life history traits on genetic diversity in plant species. Phil Trans Roy Soc Lond B Biol Sci. 1996;351:1291–8.

49. Juan II, Emerson BC, Orom II, Hewitt GM. Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. Trends Ecol Evol. 2000;15:104–9.

50. Huang S, Chiang YC, Schaal BA, Chou CH, Chiang TY. Organelle DNA phylogeography of *Cycas taitungensis*, a relict species in Taiwan. Mol Ecol. 2001;10:2669–81.

51. Birky CW, Fuerst P, Maruyama T. Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics. 1989;121:613–27.

52. Wang JF, Gong X, Chiang YC, Kuroda C. Phylogenetic patterns and disjunct distribution in *Ligularia hodgsonii* Hook. (Asteraceae). J Biogeogr. 2013;40:1741–54.

53. Ho CS, Shih HC, Liu HY, Chiu ST, Chen MH, Ju LP, et al. Development and characterization of 16 polymorphic microsatellite markers from Taiwan cow-tail fir, *Keteleeria davidiana* var. *formosana* (Pinaceae) and cross-species amplification in other Keteleeria taxa. BMC Res Notes. 2014;7:255.

54. Schwaegerle KE, Schaal BA. Genetic variability and founder effect in the pitcher plant *Sarracenia purpurea* L. Evolution. 1979;33:1210–8.

55. Chiang YC, Huang BH, Liao PC. Diversification, biogeographic pattern, and demographic history of Taiwanese *Scutellaria* species inferred from nuclear and chloroplast DNA. PLoS One. 2012;7:e50844.

56. Boberg E, Alexandersson R, Jonsson M, Maad J, Agren J, Nilsson LA. Pollinator shifts and the evolution of spur length in the moth-pollinated orchid *Platanthera bifolia*. Ann Bot. 2014;113:267–75.

57. Mallet B, Martos F, Blambert L, Pailler T, Humeau L. Evidence for isolation-by-habitat among populations of an epiphytic orchid species on a small oceanic island. PLoS One. 2014;9:e87469.

58. Tsai CC, Huang SC, Chou CH. Molecular phylogeny of *Phalaenopsis* Blume (Orchidaceae) based on the internal transcribed spacers 1 and 2 of nuclear ribosomal DNA. Plant Syst Evol. 2006;256:1–16.

59. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11–5.

Tsai *et al. BMC Plant Biology* (2015) 15:202

Page 16 of 16

60. Taberlet P, Gielly L, Pautou G, Bouvet J. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol. 1991;17:1105–9.

61. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol. 2013;30:2725–9.

62. Nei M. Molecular Evolutionary Genetics. New York: Columbia University Press; 1987.

63. Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25:1451–2.

64. Felsenstein J. PHYLIP (Phylogeny Inference Package) Version 3.6. Seattle: Department of Genome Sciences, University of Washington; 2004.

65. Tamura K. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C content biases. Mol Biol Evol. 1992;9:678–87.

66. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 1985;39:783–91.

67. Hillis DM, Bull JJ. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst Biol. 1993;42:182–92.

68. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol. 2012;29:1969–73.

69. Chang CC, Lin HC, Lin IP, Chow TY, Chen HH, Chen WH, et al. The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae): comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. Mol Biol Evol. 2006;23:279–91.

70. Ramírez SR, Gravendeel B, Singer RB, Marshall CR, Pierce NE. Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. Nature. 2007;448:1042–5.

71. Rambaut A, Drummond AJ. Tracer vs. 1.5. MCMC trace analysis tool. 2009. [http://tree.bio.ed.ac.uk/software/tracer/].

72. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007;7:214–21.

73. Rambaut A. FigTree version 1.3.1. Institute of Evolutionary Biology, University of Edinburgh. 2009. [http://tree.bio.ed.ac.uk/software/FigTree/].

74. Tajima F. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics. 1989;123:585–95.

75. Fu YX. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 1997;147:915–25.

76. Ramos-Onsins SE, Rozas J. Statistical properties of new neutrality tests against population growth. Mol Biol Evol. 2002;19:2092–100.

77. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010;10:564–7.

78. Harpending HC. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum Biol. 1994;66:591–600.

79. Yu Y, Harris AJ, He XJ. RASP (Reconstruct Ancestral State in Phylogenies) 2.1 beta. 2013 [http://mnh.scu.edu.cn/soft/blog/RASP].

80. Ree RH, Smith SA. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst Biol. 2008;57:4–14.