

RESEARCH ARTICLE

Open Access

# Genetic population structure of the alpine species *Rhododendron pseudochrysanthum sensu lato* (Ericaceae) inferred from chloroplast and nuclear DNA

Chi-Chun Huang<sup>1†</sup>, Kuo-Hsiang Hung<sup>2†</sup>, Chi-Chuan Hwang<sup>3†</sup>, Jao-Ching Huang<sup>4†</sup>, Hung-Du Lin<sup>1</sup>, Wei-Kuang Wang<sup>1</sup>, Pei-Yin Wu<sup>1</sup>, Tsai-Wen Hsu<sup>4\*</sup> and Tzen-Yuh Chiang<sup>1\*</sup>

## Abstract

**Background:** A complex of incipient species with different degrees of morphological or ecological differentiation provides an ideal model for studying species divergence. We examined the phylogeography and the evolutionary history of the *Rhododendron pseudochrysanthum* s. l.

**Results:** Systematic inconsistency was detected between gene genealogies of the cpDNA and nrDNA. Rooted at *R. hyperythrum* and *R. formosana*, both trees lacked reciprocal monophyly for all members of the complex. For *R. pseudochrysanthum* s.l., the spatial distribution of the cpDNA had a noteworthy pattern showing high genetic differentiation ( $F_{ST} = 0.56-0.72$ ) between populations in the Yushan Mountain Range and populations of the other mountain ranges.

**Conclusion:** Both incomplete lineage sorting and interspecific hybridization/introgression may have contributed to the lack of monophyly among *R. hyperythrum*, *R. formosana* and *R. pseudochrysanthum* s.l. Independent colonizations, plus low capabilities of seed dispersal in current environments, may have resulted in the genetic differentiation between populations of different mountain ranges. At the population level, the populations of Central, and Sheishan Mountains may have undergone postglacial demographic expansion, while populations of the Yushan Mountain Range are likely to have remained stable ever since the colonization. In contrast, the single population of the Alishan Mountain Range with a fixed cpDNA haplotype may have experienced bottleneck/founder's events.

## Background

The analysis of divergence patterns between species has been a major focus in systematics and molecular ecology [1-5]. Species descending from a common ancestor are expected to differentiate from each other and eventually attain reciprocal monophyly. Coalescence theory predicts that sister species are likely to be polyphyletic/paraphyletic in their genetic composition during the early stages after they split from their common ancestor [cf. [6]].

As sufficient time goes by, they become reciprocally monophyletic with the loss of complementary haplotypes. In contrast, lack of monophyly can be a result of several factors, such as incomplete lineage sorting, gene duplication, or hybridization [7-12]. Of them, hybridization leading to speciation has been known as one of the major forces for the diversification in plants [13]. As many hybridizing species remain morphologically discrete [14], hybrids usually share traits with parental species, often resulting in taxonomic difficulties. Interspecific gene flow usually leads to a mixed genetic composition in hybrids, but may have low impacts on parental species' genomes [cf. [15]]. Although molecular tools provide sufficient power in detecting genetic variation, empirical data on species delimitation are often difficult to interpret, since

\* Correspondence: twhsu@tesri.gov.tw; tychiang@mail.ncku.edu.tw

† Contributed equally

<sup>1</sup>Department of Life Sciences, National Cheng-Kung University, Tainan 701, Taiwan

<sup>4</sup>Taiwan Endemic Species Research Institute, Nantou 552, Taiwan

Full list of author information is available at the end of the article

both gene introgression after speciation and shared ancestral polymorphisms can result in species paraphyly [15]. A complex of incipient species with different degrees of morphological or ecological differentiation provides an ideal model for studying species divergence.

Taiwan, an island of the island-arc system along the western edge of the Pacific Ocean, first emerged from the waters via collision between Eurasian and Pacific plates about 9 million years before present (Proto-Taiwan Stage), but attained its modern shape only 5-6 million years ago (MYA) [16,17]. Its rugged topography is characterized by hundreds of steep mountains, which in turn provide diverse habitats along distinct vegetation zones, including tropical, coastal evergreen forests, subalpine shrubs and alpine tundra along the Central Mountain Range [18]. Taiwan was connected to mainland SE Asia by a land bridge during lower sea levels of the last glacial maximum [19,20]; some species may have reached Taiwan by this route and subsequently become endemic [2,21].

Phylogeographical patterns and genetic characteristics of populations/species are shaped by the interaction between historical vicariance and recurrent genetic exchanges. These evolutionary events would leave evolutionary footprints in the genetic polymorphisms within and between populations across the distributional range. Of the geological events, regular historical glacial cycles in the Eurasia Continent had prevalent influences on survival and recolonization of populations/species [11]. Postglacial expansion of alpine species may follow the phalanx model, which describes the effects of slower expansions from refugia due to habitat constraints [2,22,23]. That is, during interglacial periods, alpine species moved upward to peaks and became fragmented. In contrast, down-slope movement of alpine species during glacial periods could connect populations that were previously isolated and allow for genetic exchange among populations [24]. On Taiwan island, only the peaks of high mountains over 3000 m in elevation were covered by ice sheets during the glacial maxima.

*Rhododendron*, one of the largest and most widespread woody plant genera, is distributed from the northern temperate zone, throughout tropical Southeast Asia, to northeastern Australia and consists of over 1000 species [25]. *Hymenanthes*, one of eight subgenera comprising *Rhododendron*, consists of 225 species, most of them distributed in the Himalaya-Southwest China region [25-28]. In Taiwan, 13 taxa occur within *Rhododendron* [29]. The *R. pseudochrysanthum* complex consists of *R. pseudochrysanthum* Hayata sensu stricto (s.s.), *R. morii* Hayata, and *R. rubropunctatum* Hayata, all in subgenus *Hymenanthes* subsection *Maculifera*, and all are phylogenetically related based on molecular analyses [30-32]. The remaining taxa of subsection *Maculifera*

occur in west and central China, disjunctly from those in Taiwan [27]. *R. pseudochrysanthum* s.s. is an alpine species (3000-3900 m a.s.l.), while *R. morii* has a wider range from 1650-3600 m a.s.l., with some overlap in distribution. In contrast, *R. rubropunctatum* is restricted to elevations of 600-1200 m a.s.l. in northern Taiwan. The three species of *R. pseudochrysanthum* complex, all distributed along the central mountain range (CMR) and other mountain ranges west of the CMR, share morphological polymorphisms in leaf shape, petiole and flower shape, thus making taxonomy difficult, despite the differentiation of ecological distributions [33]. Possibly the closest relative of this species complex is *R. hyperythrum* Hayata, which is also endemic to Taiwan, occurring 1400-3700 m a.s.l. [25,28,29]. This species was previously placed in subsection *Pontica* [25,34], but phylogenetic work indicates that it is closely related to *R. pseudochrysanthum* [28]. These two might be sister species, but at present a phylogeny of subsection *Maculifera* is lacking, so this hypothesis cannot yet be tested. *Rhododendron hyperythrum* is also known to hybridise with members of the *R. pseudochrysanthum* complex [33]. Closely related to *R. hyperythrum*, the three species of the *R. pseudochrysanthum* complex with low morphological differentiations may be recognized as a single species (*R. pseudochrysanthum* sensu lato [s.l.]) [25,28,29,35].

The power of molecular markers in elucidating plant phylogeography has been detected in many studies [2,36-41]. Comparative phylogenies between biparental (nuclear) and uniparental (chloroplast and mitochondrial) markers are useful for resolving phylogeographic patterns, origin of hybridization, and assessing the migratory routes of species [40-42]. Because one or both organelle genomes are often maternally inherited in plants [43], they are particularly suitable for investigating processes associated with seed dispersal, such as range expansions [44], contribution of seed movement to total gene flow [44-47], refugial isolation [48,49], and origin of maternal or paternal lineage [50].

Previous studies have investigated the phylogeography of the *R. pseudochrysanthum* in Taiwan, revealing close relationships among these taxa [30,32,51] and a lack of geographical subdivision. However, little attention has been paid to the demographic dynamics of *R. pseudochrysanthum* across the mountain ranges. In contrast to previous studies, sampling in this study covered vaster areas, with the aid of the bureau of the Yushan National Park. The present study analyzed genetic variation of the *R. pseudochrysanthum* s.l. at *atpB-rbcL* intergenic spacer of cpDNA and the internal transcribed spacer regions (ITS) of nrDNA. A phylogeographical approach was used to infer past evolutionary history and processes and to examine species delimitation and population structure of the species complex. The objectives of this study were to investigate the following:

1. Do the lineages of cpDNA correspond with those of nrDNA?

2. Given geographical barriers, does significant genetic differentiation exist among populations of different mountain ranges?

3. Have the populations of different mountain ranges of *R. pseudochrysanthum* s.l. experienced different demographic dynamics?

## Results

### Genetic variability of cpDNA and nrDNA and genetic differentiation in the *Rhododendron pseudochrysanthum* s.l.

The cpDNA *atpB-rbcL* intergenic spacer of *R. formosanum*, *R. hyperythrum*, and *R. pseudochrysanthum* s.l. ranged from 585 to 616 bp. The sequence of alignment of *R. pseudochrysanthum* s.l. had 585 bp, including 41 polymorphic sites, of which 15 were parsimony-informative. Most of these polymorphic sites were single-base and scattered along a sequence. The nrDNA ITS sequences of *R. formosanum*, *R. hyperythrum*, and *R. pseudochrysanthum* s.l. were aligned with a consensus length of 628 bp. No large indels or rearrangements were detected. The sequence alignment of *R. pseudochrysanthum* s.l. had 628 bp, including 133 polymorphic sites, of which 49 were parsimony informative.

With regard to cpDNA, 36 haplotypes (PH01-PH36; accession numbers: HQ850658 - HQ850693) were identified among 70 samples of the *R. pseudochrysanthum* s.l. (Figure 1, additional file 1: Table S1). Nucleotide diversity ( $\pi$ ) detected in *R. pseudochrysanthum* s.l. was 0.0109 with high level of haplotype diversity (0.879) (Table 1). High levels of nucleotide diversity (0.0079) and haplotype diversity (0.900) were detected in the populations of the Yushan Mountain Range, while the single population sampled from the Alishan Mountain Range was fixed at a single haplotype (Table 1). PH01 was the most common haplotype shared among the Alishan, Central, and Sheishan Mountain Ranges. In addition, no haplotypes in the Yushan Mountain Range (PH04-PH07, PH10-PH15, PH23-PH26) were shared by other mountain ranges.

For nrDNA, 35 haplotypes (PR01-PR35; accession numbers: HQ850623 - HQ850657) were identified among 70 samples of the *R. pseudochrysanthum* s.l. (Figure 1, additional file 2: Table S2). Nucleotide diversity detected in *R. pseudochrysanthum* s.l. was 0.0114 with a high level of haplotype diversity (0.881). A high level of nucleotide diversity (0.0174) with a high level of haplotype diversity (0.906) were detected in the populations of the Yushan Mountain Range, while the lowest level of nucleotide diversity (0.0064) was detected in a single population of the Alishan mountains (Table 1).

To investigate the genetic structure of *R. pseudochrysanthum* s.l., a SAMOVA (Table 2) was applied to define groups and to identify any locations of genetic uniqueness

among the 14 populations. For cpDNA, an assumption with two groups ( $K = 2$ ) displayed the greatest value of  $F_{CT}$  and a maximal variance ( $F_{CT} = 0.6167$ ,  $P < 0.05$ ). The groups at  $K = 2$  exactly matched those found in the phylogenetic and network analyses, with those from the Yushan Mountain Range forming one group (YY, YP, YT) and those from the other three mountain ranges forming the other. Genetic structure was significant at three levels based on the spatial apportionment of genetic variation, i.e., among geographical groups (61.67%,  $P < 0.05$ ), among populations within groups (4.93%,  $P < 0.05$ ), and within populations (33.41%,  $P < 0.05$ ).

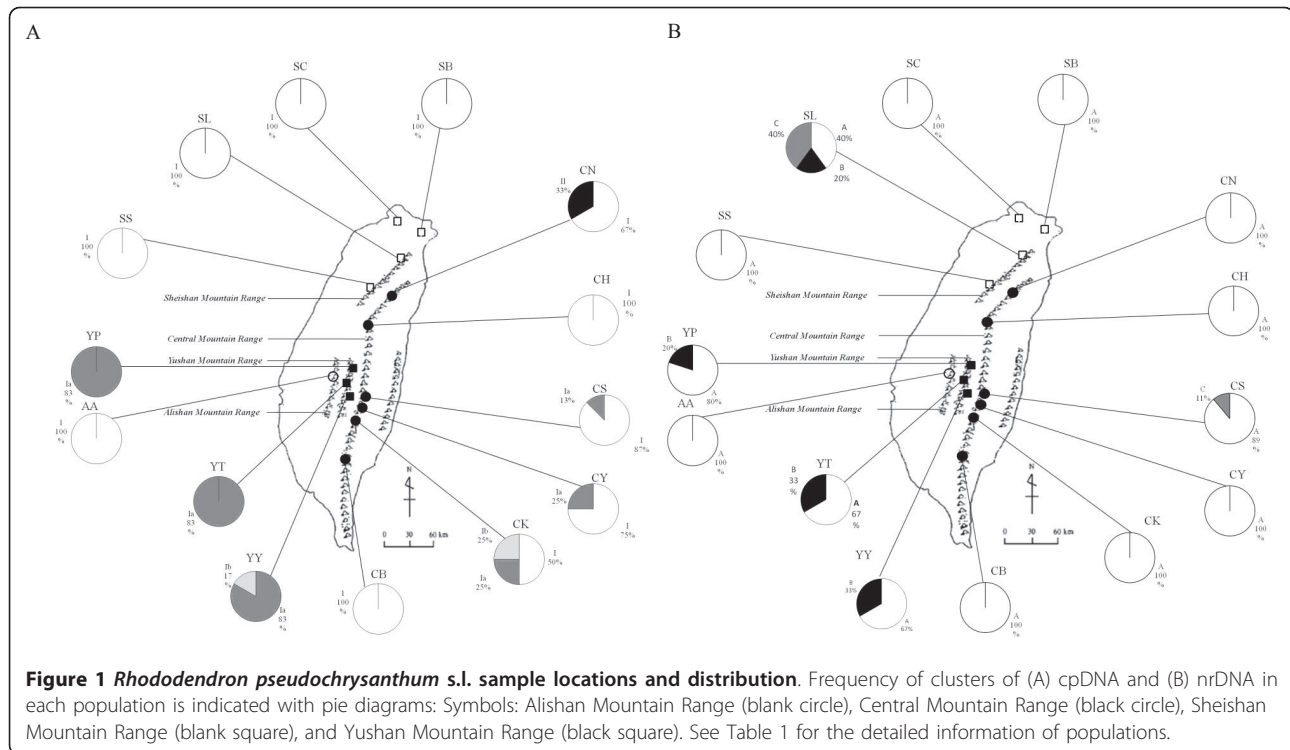
The SAMOVA of nrDNA sequences also revealed the greatest value of  $F_{CT}$  and a statistically maximal variance between two groups ( $F_{CT} = 0.5889$ ,  $P < 0.05$ ) (Table 2). Likewise, a large proportion of the genetic variability was partitioned among groups, although nonsignificantly (57.66%,  $P > 0.05$ ). In contrast, when the three geographical groups of SL vs. SC vs. all others were compared, structuring with significant variabilities distributed among groups (54.78%,  $P < 0.05$ ), among populations within groups (2.81%,  $P < 0.05$ ), and within populations (42.41%,  $P < 0.05$ ) was identified.

Significant genetic differentiation of cpDNA was detected between *R. pseudochrysanthum* s.l. - *R. formosanum* ( $F_{ST} = 0.73$ ) and *R. pseudochrysanthum* s.l. - *R. hyperythrum* ( $F_{ST} = 0.65$ ). Within *R. pseudochrysanthum* s.l., significant genetic differentiation existed between the populations of the Yushan Mountain Range and the populations of the other ranges ( $F_{ST} = 0.56-0.72$ ) (Table 3), whereas there was only little differentiation among the populations of the Alishan, Central and Sheishan Mountain Ranges ( $F_{ST} = 0.00-0.09$ ). In contrast, lower levels of genetic differentiation of nrDNA were detected between *R. pseudochrysanthum* s.l. - *R. formosanum* ( $F_{ST} = 0.02$ ) and *R. pseudochrysanthum* s.l. - *R. hyperythrum* ( $F_{ST} = 0.04$ ), and the populations of different mountain ranges of *R. pseudochrysanthum* s.l. ( $F_{ST} = 0.02-0.08$ ).

### Gene genealogies and phylogeographical analyses of the *Rhododendron pseudochrysanthum* s. l.

The ML tree of the cpDNA *atpB-rbcL* intergenic spacer haplotypes (Figure 2), rooted at *R. formosanum*, identified two major clusters I and II. A minimum-spanning network was also reconstructed based on the mutational steps between haplotypes of cpDNA (Figure 3). *Rhododendron pseudochrysanthum* s.l. was identified with eight mutational steps from *R. formosanum*. Two major clusters I and II, corresponding to the ML tree, were identified with two mutational steps apart.

A ML tree and a minimum-spanning network were reconstructed of the nrDNA haplotypes, respectively (Figures 4, 5). In total, clusters A-C were identified, none of which corresponded to any member of species.



We also utilized the \*BEAST program to analyze the combined data of cp- and nrDNA sequences. A population tree of *R. pseudochrysanthum* s.l., rooted at *R. formosanum* (FH) and *R. hyperythrum* (HH), was reconstructed (Figure 6). In contrast to the nuclear DNA tree, none of the outgroup populations were nested within *R. pseudochrysanthum* s.l.

Mismatch distribution analyses of cpDNA were performed to infer the long-term demographic history of populations. For *R. pseudochrysanthum* s.l., populations of the Central and Sheishan Mountain Ranges had a unimodal distribution (data not shown). Because the data used to produce mismatch distributions are not independent [52], Tajima's *D* was then used to detect departure from population equilibrium. Departures from the null model can be caused by many factors, such as changes in population size (e.g., expansion), which can lead to an excess of low frequency variants and negative *D* values. In contrast, processes such as population subdivision, or recent population bottlenecks can cause an excess of intermediate frequency variants leading to positive *D* values [53,54]. Most values of Tajima's *D* in *R. pseudochrysanthum* s.l. were negative, except in YT (0.325) and YY (1.057) of the Yushan Mountain Range (Table 1).

Using the program IM, we were able to refine six model parameters estimated in MCMC simulations and test for the isolation with migration model in *R. pseudochrysanthum* s.l. All of our parameter estimates had unimodal posterior distributions. The posterior distribution of  $\theta_{\text{other}}$

(scaled effective population size of the Alishan, Central, and Sheishan Mountain Ranges to neutral mutation rate: 95.43) was larger than that of  $\theta_{\text{Yushan}}$  (scaled effective population size of the Yushan Mountain Range: 30.81), whereas the posterior distribution for  $\theta_A$  (scaled effective size of the ancestral population) peaked at 24.80 (Figure 7A). Based on these  $\theta$  values, the populations of the Alishan, Central and Sheishan Mountain Ranges were likely to grow substantially following divergence, whereas the populations of the Yushan Mountain Range remained stable or slightly grew relative to the ancestral population.

When the MLEs for the migration parameters were transformed into population migration rates,  $2N_e m = m^*$  ( $\theta/2$ ) was estimated among the populations of different mountain ranges. Results from the IM revealed that the number of migrants between the Yushan Mountain Range and the other mountain ranges was effectively low. The posterior distribution of *m* (scaled effective migration rate) between the Yushan Mountain Range and the other mountain ranges peaked at 0.005, i.e., both  $m_{\text{Yushan to other mtns}} = 0.005$  (90% HPD = 0.005-1.035) and  $m_{\text{other mtns to Yushan}} = 0.005$  (90% HPD = 0.005-0.195), which were much lower than the *m* values among Alishan, Central, and Sheishan Mountain Ranges ( $m = 0.25-13.45$ ) (Figure 7). All of the non-zero *m* values, nevertheless, implied the occurrence of historical gene flow between populations (Figure 7B).

To elucidate the contribution of hybridization/introgression to the lack of species monophyly of *R. formosanum*, *R. hyperythrum* and *R. pseudochrysanthum* s.l., we

**Table 1 Location sampled within species and populations of *Rhododendron pseudochrysanthum* s.l.**

Species number	Location	Coordinate	Symbol	Nucleotide diversity ( $\pi$ )				Haplotype		Haplotype diversity		Tajima's <i>D</i>	
				cpDNA	nrDNA	cpDNA	nrDNA	cpDNA	nrDNA	cpDNA	nrDNA	cpDNA	nrDNA
<i>R. pseudochrysanthum</i> s.l.				70	70	0.01092 ± 0.00083	0.01135 ± 0.00258	36	35	0.879 ± 0.034	0.881 ± 0.038	-1.223	-2.601**
	Alishan Mountain Range			4	4	0.00000 ± 0.00000	0.00638 ± 0.00178	1	3	0.000 ± 0.000	0.833 ± 0.222	-	0.467
	Alishan	23°31'N 120°49'E	AA	4	4	0.00000 ± 0.00000	0.00743 ± 0.00218	1	3	0.000 ± 0.000	1.000 ± 0.272	-	0.467
	Central Mountain Range			30	30	0.0071 ± 0.00175	0.00721 ± 0.00327	13	17	0.821 ± 0.062	0.874 ± 0.051	-2.016*	-2.616**
	Beidawushan	22°37'N 120°45'E	CB	4	4	0.00170 ± 0.00090	0.00159 ± 0.00084	2	2	0.500 ± 0.265	0.500 ± 0.265	-0.710	-0.710
	Kuanshanlingshan	23°16'N 120°58'E	CK	4	4	0.01138 ± 0.00500	0.00478 ± 0.00207	4	3	1.000 ± 0.711	0.833 ± 0.222	-0.605	-0.809
	Hsiangyangshan	23°17'N 120°59'E	CY	4	4	0.00171 ± 0.00091	0.00637 ± 0.00129	2	5	0.500 ± 0.265	1.000 ± 0.126	-0.710	-1.193
	Sanchashan	23°18'N 121°01'E	CS	8	8	0.00672 ± 0.00240	0.01606 ± 0.00964	5	6	0.786 ± 0.151	0.893 ± 0.111	-1.096	-1.766*
	Nanhutashan	23°23'N 121°26'E	CN	6	6	0.00617 ± 0.00193	0.00350 ± 0.00079	7	6	0.964 ± 0.077	1.000 ± 0.096	-1.11	-0.932
	Hohuanshan	24°08'N 121°16'E	CH	4	4	0.00171 ± 0.00058	0.00399 ± 0.00139	3	3	0.833 ± 0.222	0.833 ± 0.222	-0.710	-0.797
	Sheishan Mountain Range			16	16	0.00193 ± 0.00046	0.01286 ± 0.00465	8	8	0.700 ± 0.127	0.752 ± 0.103	-2.149**	-1.833*
	Sheishan	24°23'N 121°14'E	SS	4	4	0.00170 ± 0.00058	0.00064 ± 0.00380	3	2	0.833 ± 0.222	0.400 ± 0.237	-0.709	-0.817
	Lalashan	24°44'N 121°26'E	SL	5	5	0.00274 ± 0.00082	0.02690 ± 0.00715	4	4	0.900 ± 0.161	0.800 ± 0.172	-1.094	-0.237
	Banpingshan	25°04'N 121°51'E	SB	6	6	0.00171 ± 0.00051	0.00425 ± 0.00109	4	3	0.800 ± 0.172	0.8000 ± 0.122	-1.233	1.219
	Chisingshan	25°10'N 121°32'E	SC	1	1	-	-	-	-	-	-	-	-
	Yushan Mountain Range			20	20	0.00786 ± 0.00117	0.01735 ± 0.00307	14	9	0.900 ± 0.056	0.906 ± 0.038	-0.096	-0.776
	Tataka	23°28'N 120°54'E	YT	9	9	0.00939 ± 0.00112	0.01800 ± 0.00317	7	4	0.917 ± 0.092	0.857 ± 0.082	0.325	1.152
	Paiyuanshanchuang	23°28'N 120°57'E	YY	6	6	0.00944 ± 0.00282	0.01786 ± 0.00487	6	3	1.000 ± 0.096	0.800 ± 0.122	-0.599	1.363
	Patungkuan	23°28'N 120°59'E	YP	5	5	0.00450 ± 0.00101	0.00124 ± 0.00563	5	4	1.000 ± 0.126	0.900 ± 0.161	0.562	-1.068
outgroup													
<i>R. hyperythrum</i>	Nanhutashan	23°23'N 121°26'E	HH	8	8	0.00510 ± 0.00072	0.01549 ± 0.00256	4	2	0.857 ± 0.082	0.571 ± 0.094	1.382	2.484
				8	8	0.00510 ± 0.00072	0.01549 ± 0.00256	4	2	0.857 ± 0.082	0.571 ± 0.094	1.382	2.484
<i>R. formosanum</i>	Heping	24°18'N 121°41'E	FH	2	2	-	0.04306 ± 0.02153	1	2	-	1.000 ± 0.5000	-	-

Samples size, nucleotide diversity, haplotype number, haplotype diversity, and Tajima's *D* are indicated. Significance values are given in asterisks (\*  $P < 0.01$ ; \*\*  $P < 0.05$ ).

**Table 2 Fixation indices corresponding to groups of populations inferred by SAMOVA for *Rhododendron pseudochrysanthum* s.l. populations in Taiwan tested for the cpDNA and nrDNA**

Marker	Group	Population groupings	FCT	P
cpDNA	2	(YY, YP, YT)(AA, CB, CY, CK, CN, CS, CH, SS, SB, SC, SL)	0.61669	0.0039
	3	(YY, YP, YT)(SC)(AA, CB, CY, CK, CN, CS, CH, SS, SB, SL)	0.60668	0.0010
	4	(YY, YT)(YP)(SC)(AA, CB, CY, CK, CN, CS, CH, SS, SB, SL)	0.58933	< 0.0001
	5	(YY)(YT)(YP)(SC)(AA, CB, CY, CK, CN, CS, CH, SS, SB, SL)	0.56974	0.0010
	6	(YY)(YT)(YP)(SC)(CB)(AA, CY, CK, CN, CS, CH, SS, SB, SL)	0.53854	< 0.0001
	nrDNA	2	(SL)(AA, CB, CH, CS, CY, CK, CN, YY, YP, YT, SB, SC, SS)	0.58892
3		(SL)(SC)(AA, CB, CH, CS, CY, CK, CN, YY, YP, YT, SB, SS)	0.54781	0.0088
4		(SL)(SC)(CH)(AA, CB, CS, CY, CK, CN, YY, YP, YT, SB, SS)	0.44184	0.0049
5		(SL)(SC)(CH)(YY)(AA, CB, CS, CY, CK, CN, YP, YT, SB, SS)	0.39814	0.0059
6		(SL)(SC)(CH)(YY)(YT)(AA, CB, CS, CY, CK, CN, YP, SB, SS)	0.37744	0.0059

Alishan Mountain Range: AA; Central Mountain Range: CB, CK, CY, CS, CN, CH; Sheishan Mountain Range: SS, SL, SB, SC; Yushan Mountain Range: YT, YY, YP.

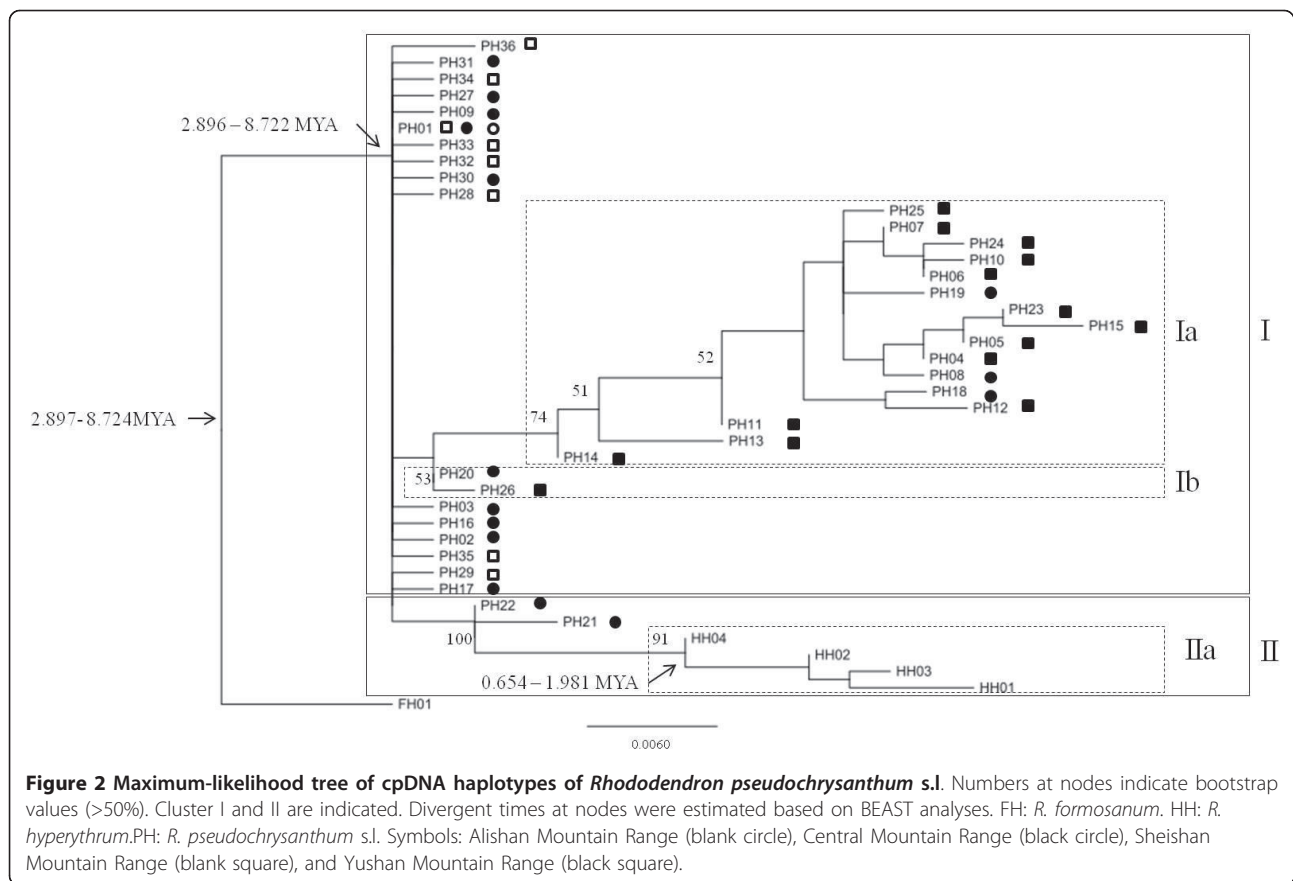
estimated the *m* values among these species. The posterior distribution of *m* among these species peaked at non-zero values ( $m_{\text{pseudochrysanthum s.l. to formosanum}} = 1.35$ , 90% HPD = 0.55-2.65;  $m_{\text{formosanum to pseudochrysanthum s.l.}} = 2.25$ , 90% HPD = 0.75-4.25;  $m_{\text{pseudochrysanthum s.l. to hyperythrum}} = 0.05$ , 90% HPD = 0.05-1.05;  $m_{\text{hyperythrum to pseudochrysanthum s.l.}} = 0.15$ , 90% HPD = 0.05-1.85), suggesting possible interspecific hybridization/introgression.

**Molecular dating of *Rhododendron pseudochrysanthum* s.l.** Bayesian estimates of the divergence time of clades of the sequences of cpDNA and nrDNA were obtained using BEAST v. 1.3. For cpDNA, the sequences

including *R. formosanum*, *R. hyperythrum*, and *R. pseudochrysanthum* s.l. could be traced back to a common ancestor ca. 2.897-8.724 MYA. All sequences of *R. hyperythrum* and *R. pseudochrysanthum* s.l. coalesced at ca. 2.896-8.722 MYA. All cpDNA sequences of *R. pseudochrysanthum* s.l. coalesced at ca. 2.895-8.722 MYA, while those of *R. hyperythrum* coalesced at ca. 0.654-1.981 MYA. Likewise, the nrDNA sequences including *R. formosanum*, *R. hyperythrum*, and *R. pseudochrysanthum* s.l. could be traced back to a common ancestor about 1.905-2.961 MYA. All nrDNA sequences of *R. pseudochrysanthum* s.l. and *R. hyperythrum* coalesced to a common ancestor ca. 1.879-2.912 MYA. All nrDNA

**Table 3 Pairwise  $F_{ST}$  among populations deduced from sequences of cpDNA (below the diagonal) and nrDNA (above the diagonal) for *Rhododendron pseudochrysanthum* s.l.**

cpDNA/nrDNA	Alishan Mountain Range	Central Mountain Range								Sheishan Mountain Range	Yushan Mountain Range						
	AA	CB	CK	CY	CS	CN	CH	SS	SL	SB	SC	YT	YY	YP			
Alishan Mountain Range	-	<b>0.02</b>								<b>0.03</b>	<b>0.08</b>						
AA	-	0	0	0	0.05	0.04	0	0	0.09	0	-	0.12	0.03	0.01			
Central Mountain Range	<b>0.07</b>	-	<b>0.03</b>								<b>0.03</b>	<b>0.06</b>					
CB	0.5	-	0	0	0.06	0.08	0	0	0.11	0	-	0.14	0.04	0.02			
CK	0.05	0.07	-	0	0.01	0.08	0.05	0	0.09	0	-	0.09	0	0			
CY	0	0.29	0.05	-	0.04	0.05	0	0	0.1	0	-	0.11	0.01	0.02			
CS	0	0.19	0.07	0.07	-	0.02	0.04	0.06	0.04	0.06	-	0.04	0.04	0			
CN	0.15	0.3	0.07	0.09	0.08	-	0.05	0.1	0.12	0.07	-	0.1	0.02	0.01			
CH	0	0.3	0.04	0	0	0.1	-	0	0.1	0	-	0.12	0.04	0.02			
Sheishan Mountain Range	<b>0</b>	<b>0.09</b>	<b>0.09</b>								-	<b>0.04</b>					
SS	0	0.33	0.04	0	0	0.11	0	-	0.11	0	-	0.15	0.04	0.02			
SL	0	0.31	0.04	0	0	0.1	0	0	-	0.1	-	0.05	0.02	0.04			
SB	0	0.38	0.04	0	0	0.12	0	0	0	-	-	0.12	0.04	0.02			
SC	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
Yushan Mountain Range	<b>0.72</b>	<b>0.56</b>	<b>0.69</b>								<b>0.69</b>	<b>0.69</b>					
YT	0.71	0.69	0.29	0.61	0.57	0.63	0.66	0.67	0.67	0.68	-	-	0.16	0.01			
YY	0.92	0.88	0.56	0.84	0.57	0.83	0.87	0.88	0.87	0.89	-	-	-	0.05			
YP	0.89	0.86	0.53	0.82	0.71	0.81	0.84	0.86	0.85	0.87	-	-	0.07	0.15			



sequences of *R. pseudochrysanthum* s.l. coalesced to a common ancestor ca. 1.879-2.909 MYA, while those of *R. hyperythrum* coalesced at ca. 1.669-2.620 MYA.

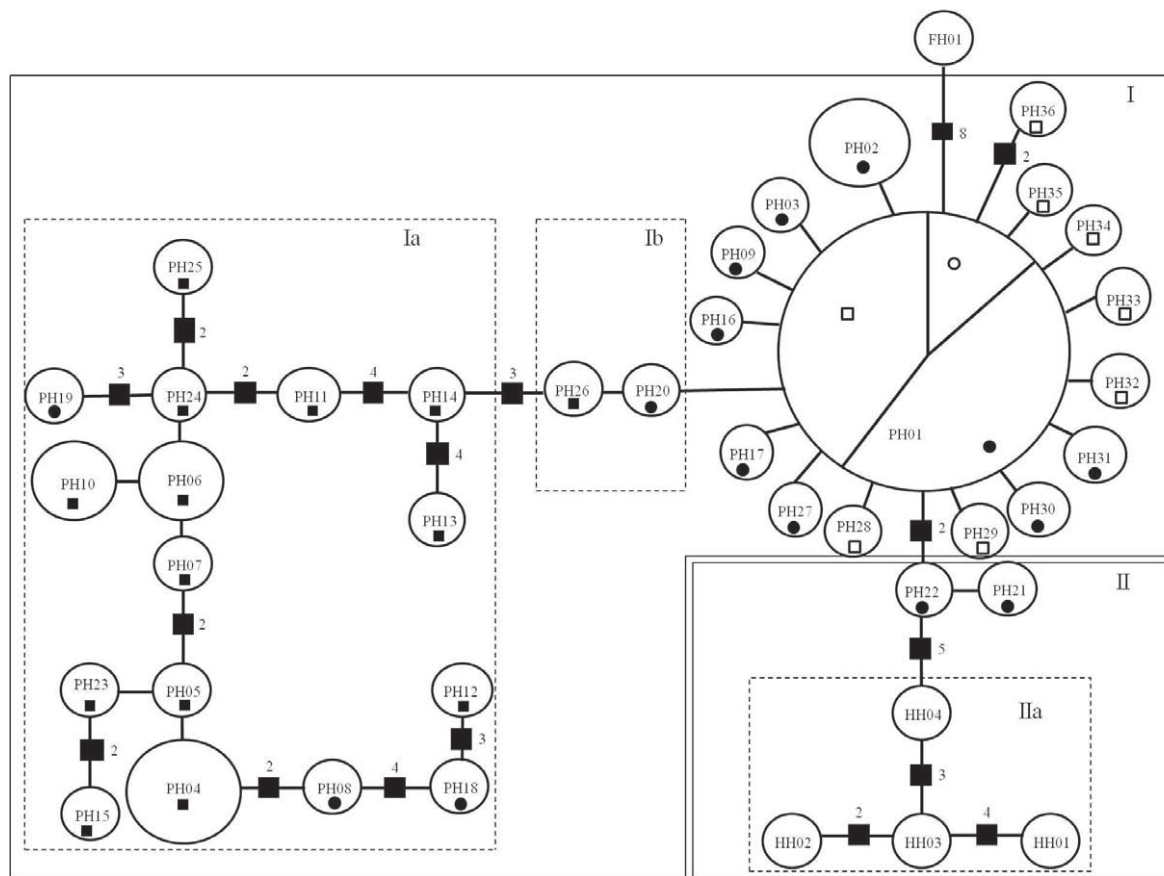
To elucidate the possible evolutionary forces for the lack of species monophyly of *R. formosanum*, *R. hyperythrum* and *R. pseudochrysanthum* s.l., we calculated the expected time to coalescence at cpDNA [ $1.665 \times \theta/4U$ ] and at nrDNA [ $1.665 \times \theta/2U$ ], where  $\theta$  is a population mutation parameter obtained from IM, and  $U$  is the geometric mean of mutation rates, and compared them with the divergence times obtained from the BEAST analyses. According to the IM estimates of  $\theta_{\text{other mtns}} = 95.43$  and  $\theta_{\text{Yushan}} = 30.81$ , the expected time to coalescence of *R. pseudochrysanthum* s.l. ranged from 4.36 to 29.21 MYA based on cpDNA, and from 8.72 to 58.42 MYA based on nrDNA.

## Discussion

### Systematic inconsistency between nuclear and chloroplast DNAs in *Rhododendron pseudochrysanthum* s.l.

Inconsistency in the cpDNA and nrDNA phylogenies (Figure 2 & 3) was detected in *R. formosanum*. The ML tree for the cpDNA agreed with the results of Chamberlain et al. [25] and Goetsch et al. [28] that *R. formosanum* is a monophyletic group relative to *R. pseudochrysanthum* s.l. In contrast, the ML tree of

nrDNA revealed that one accession of *R. formosanum* fell within cluster B with the ingroup species. Lack of species monophyly may result from incomplete lineage sorting, which occurs when populations or species recently diverged, thus retaining ancestral genetic polymorphisms, or from genetic exchanges via hybridization and introgression that introduce foreign alleles into a species or population, resulting in a genetic admixture [10]. For given species, the time for a locus to acquire reciprocal monophyly is about  $1.665 \times \theta/4U$  at cpDNA or  $1.665 \times \theta/2U$  at nrDNA. Willyard et al. [55] suggest that ancestral polymorphisms would likely be shared between species if the coalescence time of the alleles exceeds the divergence time at one locus. In this study, the expected time to coalescence of *R. pseudochrysanthum* s.l. (4.36-29.21 MYA) overlapped the estimates of divergence time (2.895-8.722 MYA) based on cpDNA, in which monophyly was attained. In contrast, at the nrDNA, the expected time for coalescence (8.72-58.42 MYA) predated the divergence time (1.879-2.909 MYA). Accordingly, nuclear loci would less likely be monophyletic in *R. pseudochrysanthum* s.l. Incomplete lineage sorting alone, apparently, cannot completely explain the paraphyly of *R. pseudochrysanthum* s.l. Interspecific hybridization that disturbs the reciprocal



**Figure 3** Minimum-spanning network based on the mutational steps between haplotypes of cpDNA in the populations of *Rhododendron pseudochrysanthum* s.l. Cluster I and II are indicated. FH: *R. formosanum*. HH: *R. hyperythrum*. PH: *R. pseudochrysanthum* s.l.; Symbols: Alishan Mountain Range (blank circle), Central Mountain Range (black circle), Sheishan Mountain Range (blank square), and Yushan Mountain Range (black square). Branches represent single mutational changes, and black squares represent multiple changes (indicated by adjacent numbers). Circle sizes are proportional to the haplotype's frequency.

monophyly needs to be aroused based on the non-zero  $m$  parameters between *R. formosanum*-*R. pseudochrysanthum* s.l. ( $m_{\text{pseudochrysanthum s.l. to formosanum}} = 1.35$ , 90% HPD = 0.55-2.65;  $m_{\text{formosanum to pseudochrysanthum s.l.}} = 2.25$ , 90% HPD = 0.75-4.25) and *R. hyperythrum*-*R. pseudochrysanthum* s.l. ( $m_{\text{pseudochrysanthum s.l. to hyperythrum}} = 0.05$ , 90% HPD = 0.05-1.05;  $m_{\text{hyperythrum to pseudochrysanthum s.l.}} = 0.15$ , 90% HPD = 0.05-1.85). Furthermore, Milne et al. [56,57] revealed that natural hybridization/introgression is not unusual in *Rhododendron*. Thus, interspecific hybridization/introgression would cause the lack of monophyly and prolong the duration of paraphyly, as frequently occurs in most plants, e.g., *Ixeris* [58], *Begonia* [59], and *Ilex* [60].

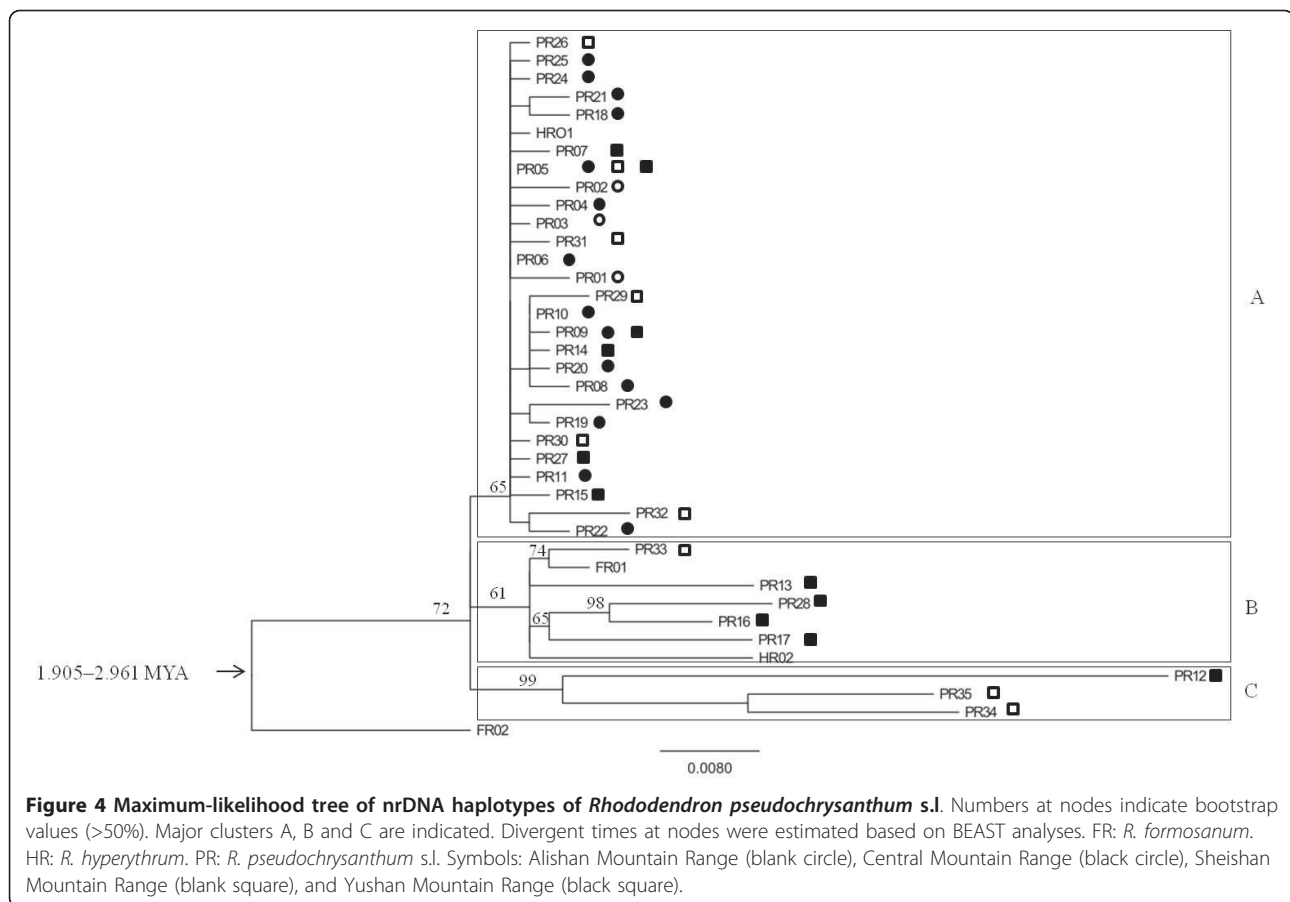
#### Genetic diversity and population structure of *R. pseudochrysanthum* s.l.

In this study, we investigated genetic diversity at the cpDNA and nrDNA loci of *R. pseudochrysanthum* s.l.

High levels of nucleotide diversity were detected at both cpDNA and nrDNA loci, agreeing with a pattern found in many East Asian plants, that is considered to have been shaped by a common glaciation history [cf. [11]]. During the Quaternary glaciations, Taiwan, as a refuge, provided shelters for species from a range of biomes and thus harbored much plant biodiversity. Examples include *Cunninghamia konishii* [21], *Pinus luchuensis* ssp. *taiwanensis* [2], *Cycas taitungensis* [41], and *Cycas revoluta* [1]. A potential refugium for *R. pseudochrysanthum* s.l. in Taiwan was inferred to harbor the genetic variations. Furthermore, highly diverse environments may also have created a wide spectrum of habitats to accommodate mutations [38].

The populations of the Yushan Mountain Range had the highest levels of genetic diversity (0.0079) in the cpDNA. A similar pattern was found in Taiwan fir (*Abies kawakamii*), a dominant species of subalpine forests at 2500-3700 m a.s.l. [61]. This common result is





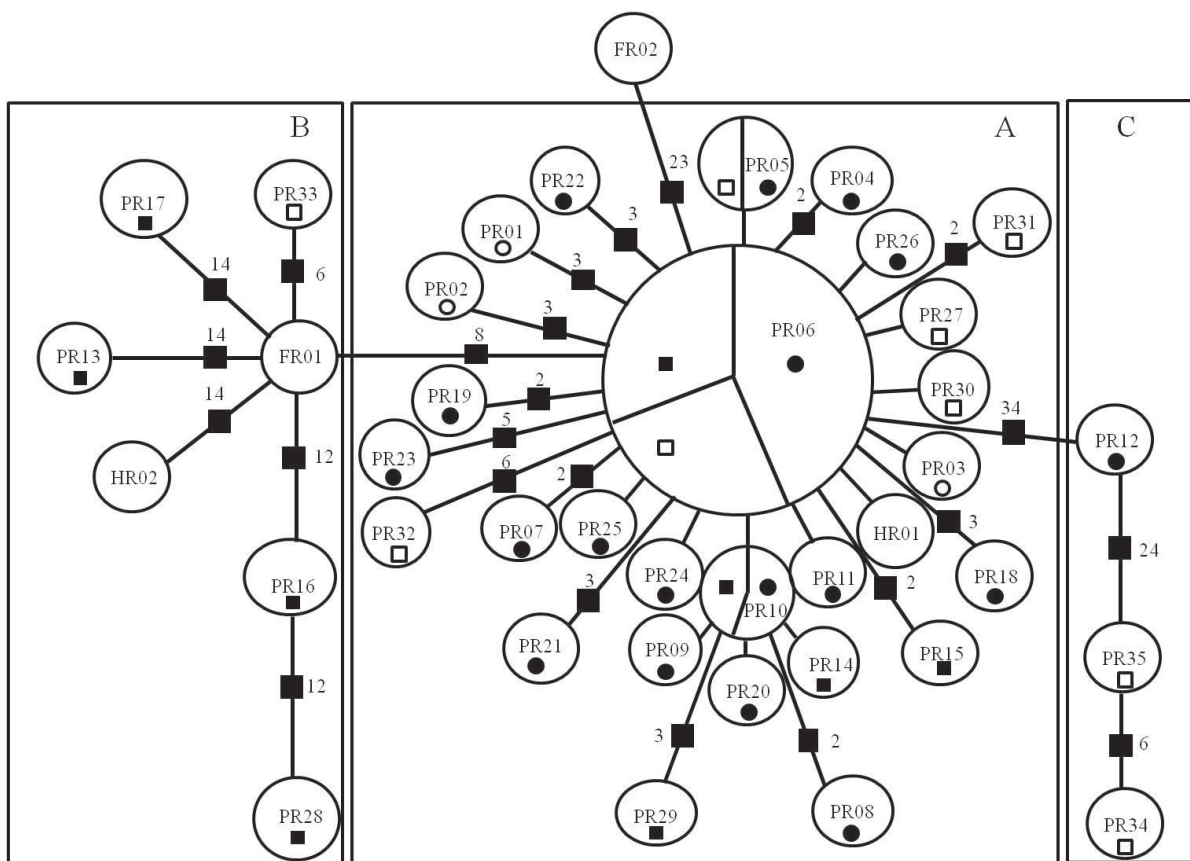
likely to be attributable to a stable population size in the Yushan Mountain Range. The habitat ranges of the Yushan Mountain Range, with the highest peak at 3952 m, are wider than those of the other mountain ranges (3886 m at most), providing diverse habitats for *R. pseudochrysanthum* s.l. Given a stable, effective population size under diverse habitats, the populations of the Yushan Mountain Range fostered a great number of genetic polymorphisms. In contrast, genetic diversity can be lost by chance in fragmented populations, as occurred in the population of the Alishan Mountain Range. PH01, the interior dominant haplotype, is shared among the Alishan, Central, and Sheishan mountain ranges. A single population of the Alishan Mountain Range was fixed for the dominant haplotype PH01, indicating a possible genetic bottleneck due to a founder effect and/or genetic drift during a period when population sizes were small [62].

#### Genetic structure of *R. pseudochrysanthum* s.l.

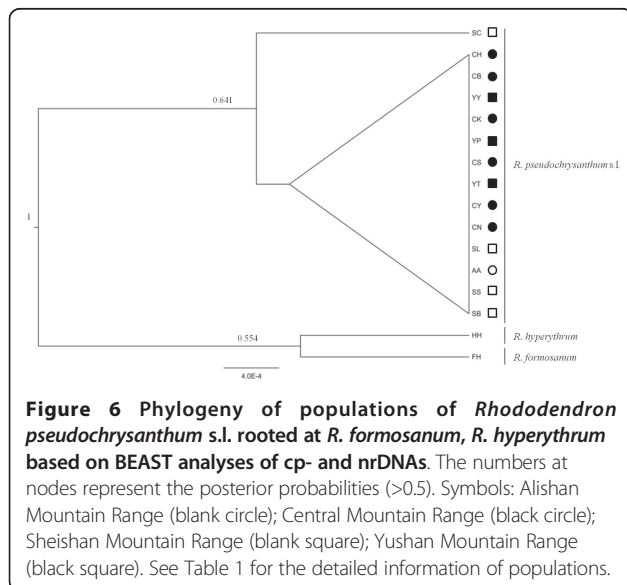
*R. pseudochrysanthum* s.l. inhabits mountaintops isolated by surrounding low-elevation habitats, forming small and fragmented patches. Genetic differentiation would be detected between these isolated populations of

different mountain ranges. Nevertheless, genetic analysis of the cpDNA variation revealed very weak genetic differentiation among the populations within a mountain range. The short time of isolation since the last glacial retreat is seemingly not long enough for differentiating these populations within a mountain range and thus resulted in sharing ancestral genetic polymorphisms.

Specifically, spatial distribution of cpDNA revealed a noteworthy pattern of high differentiation between populations of the Yushan Mountain Range and the other mountain ranges ( $F_{ST} = 0.56-0.72$ ) (Table 3). The divergence time for *R. pseudochrysanthum* s.l. is about 2.895-8.722 MYA, and given that within *R. pseudochrysanthum* s.l., the populations from different mountain ranges are more differentiated than the species, then it follows that this divergence time likely reflects an estimated divergence time between mountain populations. Molecular dating revealed that populations among different mountain ranges split some 2.895-8.722 MYA, a time range approximating or predating the formation of the Central Mountain Range of Taiwan [16]. Independent colonizations via individuals with two different subclusters, likely from adjacent mainland over Pleistocene, may result in high genetic differentiation within



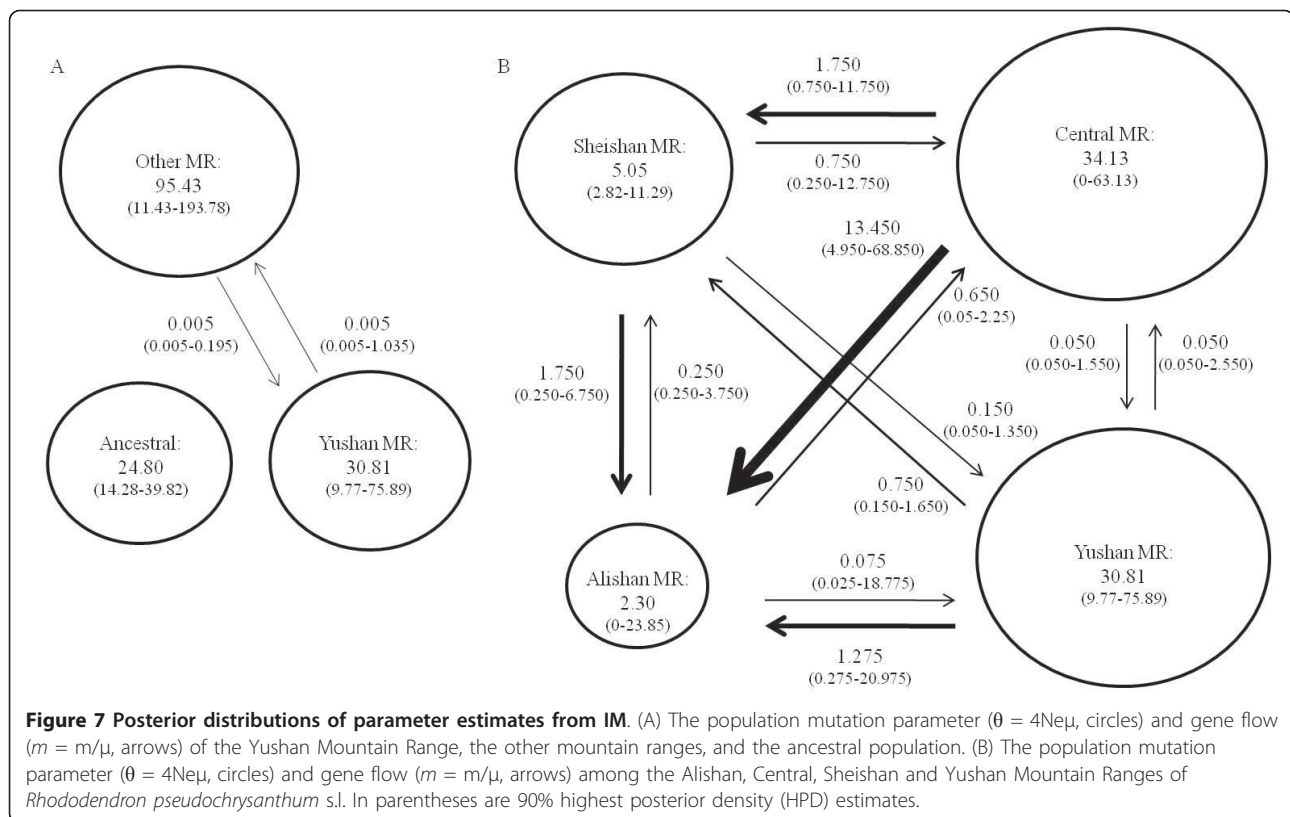
**Figure 5** Minimum-spanning network based on the mutational steps between haplotypes of nrDNA in the populations of *Rhododendron pseudochrysanthum* s.l. Major clusters A, B and C are indicated. FR: *R. formosanum*. HR: *R. hyperythrum*. PR: *R. pseudochrysanthum* s.l.; Symbols: Alishan Mountain Range (blank circle), Central Mountain Range (black circle), Sheishan Mountain Range (blank square), and Yushan Mountain Range (black square). Branches represent single mutational changes, and black squares represent multiple changes (indicated by adjacent numbers). Circle sizes are proportional to the haplotype's frequency.



**Figure 6** Phylogeny of populations of *Rhododendron pseudochrysanthum* s.l. rooted at *R. formosanum*, *R. hyperythrum* based on BEAST analyses of cp- and nrDNAs. The numbers at nodes represent the posterior probabilities (>0.5). Symbols: Alishan Mountain Range (blank circle); Central Mountain Range (black circle); Sheishan Mountain Range (blank square); Yushan Mountain Range (black square). See Table 1 for the detailed information of populations.

*R. pseudochrysanthum* s.l. Unexpected genetic structuring was detected with populations of the Yushan Mountain Range highly differentiated from other mountain ranges, despite this range lying between the Alishan and the Central Mountain Ranges. The single population of the Alishan Mountain Range was fixed at the dominant haplotype PH01, which was shared by the populations of the Central and Sheishan Mountain Ranges. Some other haplotypes may have been randomly lost from the Alishan population due to possible bottleneck/founder's events. Despite the fact that the Alishan Mountain Range is geographically distant to Central and Sheishan Mountain Ranges, sharing the dominant PH01 made these populations clustered together and genetically undifferentiated. Furthermore, subsequent glacial retreat and mountain uplifting triggered the geographical isolations and may have thereby enhanced the divergence of populations among different mountain ranges.

Today, over an interglacial period, *Rhododendron* populations of different mountain ranges represent



some isolated patches. Given such geographical barriers, seeds are less likely to disperse across mountain ranges. Low capabilities of seed dispersal in current environments, plus possible independent colonization events followed by limited historical gene flow (Figure 7), may have together contributed to the high levels of genetic differentiation between populations of the Yushan Mountain Range and the other mountain ranges ( $F_{ST} = 0.56-0.72$ ). In contrast, low levels of genetic differentiation at cpDNA detected among the populations of Alishan, Central and Sheishan Mountain Ranges ( $F_{ST} = 0.00-0.09$ ) implied that there may exist historical gene flow during the glacial maxima, as the alpine populations migrated downwards and became connected to each other (Table 3).

Compared to the high genetic differentiation among populations between Yushan Mountain Range and the other mountain ranges based on cpDNA, nrDNA revealed much less differentiation among populations of different mountain ranges ( $F_{ST} = 0.02-0.08$ ) (Tables 2, 3). Here, biparentally inherited nrDNA was used to estimate the extent of pollen flow. In contrast to the limited seed dispersal, pollen of *Rhododendron* species seemingly can move long distances [63,64]. Long-distance pollen dispersal of *R. pseudochrysanthum* s.l. across the discontinuity of mountain ranges reduced the genetic differentiation.

#### Ancestral population size, gene flow, and population dynamics of *R. pseudochrysanthum* s.l

Given their different inheritances, nrDNA and cpDNA markers provide extensive aspects to the understanding of the evolutionary forces shaping fates of *R. pseudochrysanthum* s.l. Combining cpDNA and nrDNA data, IM analysis was used to evaluate the ancestral population size and gene flow. Populations of the Alishan, Central, and Sheishan Mountain Ranges were characterized by a recent population expansion as  $\theta_{\text{other mtns}}$  was approximately four times larger than ancestral  $\theta_A$  [65]. Star-like genealogies of cpDNA and negative Tajima's *D* of cpDNA and nrDNA (Table 1) also supported a scenario of recent demographic expansion in populations of the Central and Sheishan Mountain Ranges, a common pattern also occurring in *Castanopsis carlesii* [66], *Cunninghamia konishii* [67], and *Trema dielsiana* [68]. In contrast, the value of  $\theta_{\text{Yushan}}$  approximated ancestral  $\theta_A$ , indicating that the effective population size of the Yushan Mountain Range may have remained stable relative to the ancestral population.

Given low seed dispersal capabilities of *Rhododendron* in today's environments, non-zero migration rates ( $2Nem$ ) detected among populations ( $m = 0.05-13.45$ ) likely represented historical gene flow across mountain ranges. Here, compared to the gene flow among populations of Alishan, Central, and Sheishan Mountain

Ranges, with  $m = 0.25$ -13.45, much lower levels of migration rates were detected between populations of the Yushan Mountain Range and other mountain ranges ( $m_{\text{Yushan to other mtns}} = 0.005$ ,  $m_{\text{other mtns to Yushan}} = 0.005$ ). Even so, higher levels of migration rates detected from populations of the Yushan Mountain Range to those of the Alishan ( $m = 1.275$ ) and Sheishan ( $m = 0.750$ ) Mountain Range further indicated asymmetric gene flow among them. In contrast, historical dispersal may have occurred freely among other populations. That is, during glacial maxima, there may exist corridors among the mountain ranges. These corridors allowed frequent gene flow among populations of Alishan, Central, and Sheishan Mountain Ranges, and occasional gene flow between populations of the Yushan Mountain Range and the other mountain ranges. Altogether, populations of these mountain ranges may have been mediated with different levels of historical gene flow during the glacial maxima.

## Conclusions

In this study, inconsistent gene genealogies of cpDNA and nrDNA were detected. In the cpDNA tree, agreeing with Chamberlain et al. [25], *R. pseudochrysanthum* s.l. is phylogenetically more closely related to *R. hyperythrum* than to *R. formosanum*. In contrast, nrDNA failed to identify reciprocal monophyly of ingroup vs. outgroup, indicating possible incomplete lineage sorting and interspecific hybridization/introgression. For *R. pseudochrysanthum* s.l., the spatial distribution of cpDNA revealed a noteworthy pattern of high differentiation between the populations of the Yushan Mountain Range and the populations of the other mountain ranges. Molecular dating revealed that the split of the populations at different mountain ranges approximated or predated the formation of the Central Mountain Range of Taiwan island, likely suggesting independent colonizations via individuals from previously differentiated populations on the mainland. At the population level, the populations of Central, and Sheishan Mountain Ranges may have undergone postglacial demographic expansion, while populations of the Yushan Mountain Range are likely to have remained stable ever since the colonization. In contrast, the single population of the Alishan Mountain Range with a fixed cpDNA haplotype may have experienced bottleneck/founder's events.

## Methods

### Sample collection, DNA extraction, PCR, and DNA sequencing

We assessed the levels of genetic variation of cpDNA (*atpB-rbcL* intergenic spacer) and nrDNA ITS sequences among populations of *Rhododendron pseudochrysanthum* s.l. (Table 1). In addition, *R. formosanum* Hemsl. (subgenus *Hymenanthes* subsection *Argyrophylla* [25]), and

*R. hyperythrum* (subgenus *Pontica* subsection *Maculifera*) were selected to serve as outgroups. In the field, *R. hyperythrum* is sympatric with *R. pseudochrysanthum* s.l., while *R. formosanum* is allopatric from two other species. Morphologically, *R. hyperythrum* and *R. formosanum* are distinguishable from *R. pseudochrysanthum* s.l. According to Goetsch et al. [28], *R. pseudochrysanthum* s.l. is phylogenetically more related to *R. hyperythrum* than to *R. formosanum*. A total of 70 *R. pseudochrysanthum* s.l. individuals were collected from 14 populations throughout their natural range, mostly from four major mountain ranges in Taiwan, i.e., Alishan, Central, Sheishan and Yushan Mountain Ranges (Figure 1 and Table 1). One to 12 individuals about 100 m apart were sampled based on the natural size of each population. Young, healthy leaves were collected in the field, rinsed with tap water, and dried in silica gel. All samples were stored at  $-70^{\circ}\text{C}$  until processing.

Leaf tissue was ground to powder in liquid nitrogen and stored at  $-70^{\circ}\text{C}$ . Total genomic DNA was extracted from the powdered tissue using a cetyltrimethyl ammonium bromide (CTAB) procedure [69]. PCR amplification was carried out in a 100- $\mu\text{L}$  reaction. The reaction was optimized and programmed on a MJ Thermal Cycler (PTC 100) (MJ, Alameda, California, USA) as one cycle of denaturation at  $95^{\circ}\text{C}$  for 4 min; 30 cycles of 45 s denaturation at  $92^{\circ}\text{C}$ , 1 min 15 s annealing at  $52^{\circ}\text{C}$ , and 1 min 30 s extension at  $72^{\circ}\text{C}$ ; followed by 10 min extension at  $72^{\circ}\text{C}$ . Template DNA was denatured with reaction buffer,  $\text{MgCl}_2$ , NP-40 and ddH<sub>2</sub>O for 4 min (first cycle), and cooled on ice immediately. Primers, for the ITS (ITS5: 5'-GGAAGTAAAAGTCGTAACAAGG-3' and ITS4: 5'-TCCTCCGCTATATGATATGC-3') [70] or cpDNA *atpB-rbcL* intergenic spacer (*atpB*-1: 5'-ACATCKAR-TACKGGACC AATAA-3' and *rbcL*-1: 5'-AACACCA GCTTTRAATCCAA-3') [71], dNTP and *Taq* polymerase (Promega, Madison, Wisconsin, USA) were added to the above ice-cold mix. The reaction was restarted at the first annealing at  $52^{\circ}\text{C}$ .

PCR products were purified by electrophoresis in 1.0% agarose gel using 1X TAE buffer. The gel was stained with ethidium bromide, and the desired DNA band was cut and eluted using agarose gel purification (Qiagen, Valencia, California, USA). Both direct and subcloning sequencing was conducted to check PCR error. Chloroplast DNA sequences were determined by direct sequencing, while nrDNA was sequenced via subcloning. For subcloning, the desired DNA fragments were cloned into the pGEM-T vector system (Promega, Madison, Wisconsin, USA). Seven clones of the plasmid DNA were selected randomly and purified using a plasmid mini kit (Qiagen). DNA sequencing in both directions was conducted with an Applied Biosystems Model 377A automated sequencer (Applied Biosystems, Foster,

California, USA). No within-individual variation was detected for either DNA marker.

#### Data analysis

Nucleotide sequences were aligned with the program MAFFT version 6 [72] and later adjusted visually. After the alignment, maximum-likelihood (ML) analyses (HKY model) were performed with PhyML v3.0 [73], and bootstrap consensus values were calculated with 1000 bootstrap pseudoreplicates. Nucleotide substitution models were determined with the Akaike Information Criterion by ModelTest 3.7 analysis [74,75]. The HKY + I + G model with 6 substitution categories was determined to be the most suitable model by Modeltest and was used for all subsequent nucleotide analyses. The number of mutations between DNA sequence haplotypes in pairwise comparisons was used to construct a minimum-spanning network with the aid of the MINSPNET [76] in a hierarchical manner [77].

To reconstruct the phylogeny of populations, the combined data of both cp- and nrDNA sequences were analyzed by the program \*BEAST v1.5.4 (Bayesian Evolutionary Analysis Sampling Trees)[78]. This method uses coalescent theory to provide joint inferences of a species tree topology, divergence times, population sizes and gene trees from multiple genes sampled from multiple individuals across a set of closely related species.

Levels of genetic diversity within populations and species were quantified with measures of nucleotide diversity ( $\pi$ ) [79] using DnaSP Version 5.1 [80]. Patterns of geographical subdivision and levels of genetic differentiation among populations were estimated hierarchically with the aid of DnaSP. Mismatch distribution analyses and Tajima's  $D$  were also conducted [81].

SAMOVA (spatial analysis of molecular variance) was applied to identify groups of geographically adjacent populations of *R. pseudochrysanthum* s.l. that were maximally differentiated based on sequence data [82]. We performed the analyses based on 100 simulated annealing steps and examined maximum indicators of differentiation ( $F_{CT}$  values) when the program was assigned to identify  $K = 2-6$  partitions of populations.

#### Isolation with migration and estimating of ancestral population size

We used the simulation program IM [83] to investigate the scenario of isolation with migration between populations. By applying coalescence simulations and Bayesian computation procedures, IM yielded six model parameters, including the population-split time ( $t = t_{\mu}$ ), scaled migration rates ( $m = m/\mu$ ), and scaled effective population size ( $\theta = 4Ne\mu$ ) for the ancestral and two current populations. The posterior probability densities of these parameters were generated by Markov chain Monte Carlo

(MCMC) simulations, and three of these simulations were run with individual simulations being updated 50 million times. Within each simulation, we used a procedure to swap among 10 heated chains (Metropolis coupling). Each simulation yielded a marginal density histogram for the six parameters of interest, and the peaks of the resulting distributions were considered as the maximum likelihood estimates (MLE) of the parameter with credibility intervals equaling the 90% highest posterior density (HPD) intervals. Here we compared populations of the Yushan Mountain Range with other populations of three different mountain ranges based on the results of  $F_{ST}$  analyses. Parameters in the IM model are scaled by the mutation rates per generation for the loci, i.e.,  $\mu = UG$ , where  $U$  is the geometric mean of mutation rates of all the loci and  $G$  is the generation time.

#### Estimation of divergence times of cpDNA and nrDNA lineages

In the present study, we estimated the divergence time of the sister lineages of *R. pseudochrysanthum* s.l. split for both cp- and nrDNAs using a Bayesian relaxed clock method with BEAST. We excluded sequences of 18S, 5.8S and 26S RNA regions of nrDNA from the calculation of the divergent time. To estimate the divergence between lineages a well-supported mutation rate is required. In this study, applying evolutionary rates of  $1.0-3.0 \times 10^{-9}$  [84] and  $5.0-7.8 \times 10^{-9}$  [85] substitutions per site per year for synonymous sites of cpDNA and nrDNA, respectively, we estimated the divergence times of cpDNA and nrDNA lineages for *R. hyperythrum* and *R. pseudochrysanthum* s.l. Additionally, Bayesian estimates of the divergence time of the MRCA of the *R. pseudochrysanthum* s.l. clusters were obtained using BEAST v. 1.3; available from [http://beast.bio.ed.ac.uk/Main\\_Page](http://beast.bio.ed.ac.uk/Main_Page)[86]. Following the result of ModelTest analysis [75], the HKY model of nucleotide substitution with estimated base frequencies, gamma shape distribution (with six categories), proportion of invariant sites, and a relaxed clock with uncorrelated log normal distribution of branch lengths were chosen. Posterior estimates of the mutation rate and divergence time were obtained by Markov chain Monte Carlo analysis, with samples drawn every 500 steps over a total of 1 million steps. All operators were automatically optimized. Convergence of parameters and mixing of chains were followed by visual inspection of parameter trend lines and checking effective sampling size (ESS) values with three pre-runs. The ESS parameter was found to exceed 100, which suggests acceptable mixing and sufficient sampling. Adequate sampling and convergence to a stationary distribution were checked using TRACER v. 1.3 [87].

There are two primary processes that may cause a locus to show non-monophyly for a group of closely

related species, i.e., incomplete lineage sorting and hybridization/introgression. Distinguishing between these two processes is difficult, nevertheless, estimates of divergence time can provide a means of assessing the probability that incomplete lineage sorting alone could account for the observed non-monophyly. The expected time for a locus to become monophyletic following divergence from a sister lineage is  $1.665 \times NeG$  and  $1.665 \times 2NeG$  for cp- and nuclear DNAs, respectively [55,88]. According to IM [83], the effective population size can be estimated as  $Ne = \theta/(4UG)$ . Therefore, the former is then equal to  $1.665 \times \theta/4U$ , while the latter is  $1.665 \times \theta/2U$ . Here,  $\theta$  values were obtained from IM analyses, and  $U$  is the geometric mean of mutation rates of cp- and nrDNAs.

## Additional material

**Additional file 1: Table S1. Absolute frequency of chloroplast DNA haplotypes of *Rhododendron pseudochrysanthum* s.l.** Numbers of chloroplast haplotypes distributions from 14 populations of *Rhododendron pseudochrysanthum* s.l. Cluster I and II are indicated. N: number of samples. Alishan Mountain Range: AA; Central Mountain Range: CB, CK, CY, CS, CN, CH; Sheishan Mountain Range: SS, SL, SB, SC; Yushan Mountain Range: YT, YY, YP. See Table 1 for the detailed information of populations.

**Additional file 2: Table S2. Absolute frequency of nuclear DNA haplotypes of *Rhododendron pseudochrysanthum* s.l.** Numbers of nuclear haplotypes distributions from 14 populations of *Rhododendron pseudochrysanthum* s.l. Cluster A, B and C are indicated. N: number of samples. Alishan Mountain Range: AA; Central Mountain Range: CB, CK, CY, CS, CN, CH; Sheishan Mountain Range: SS, SL, SB, SC; Yushan Mountain Range: YT, YY, YP. See Table 1 for the detailed information of populations.

## Acknowledgements

The authors thank the financial support provided by a grant of the Yushan National Park, Taiwan and reviewers for providing comments that improved the manuscript.

## Author details

<sup>1</sup>Department of Life Sciences, National Cheng-Kung University, Tainan 701, Taiwan. <sup>2</sup>Graduate Institute of Bioresources, National Pingtung University of Science and Technology, Pingtung 912, Taiwan. <sup>3</sup>Department of Engineering Science, National Cheng-Kung University, Tainan 701, Taiwan. <sup>4</sup>Taiwan Endemic Species Research Institute, Nantou 552, Taiwan.

## Authors' contributions

CCH, KHH, JCH, HDL and WKW carried out the molecular genetic studies and the statistical analysis, and participated in the sequence alignment. CCH, TWH and PYW took part in material sampling. CCH, TWH and TYC designed the research and drafted the manuscript. All authors read and approved the final manuscript.

Received: 1 July 2010 Accepted: 19 April 2011 Published: 19 April 2011

## References

- Chiang YC, Hung KH, Moore SJ, Ge XJ, Huang S, Hsu TW, Schaal BA, Chiang TY: Paraphyly of organelle DNAs in *Cycas* Sect. *Asiorientales* due to ancient ancestral polymorphisms. *BMC Evol Biol* 2009, **9**:161-179.
- Chiang YC, Hung KH, Schaal BA, Gest XJ, Hsu TW, Chiang TY: Contrasting phylogeographical patterns between mainland and island taxa of the *Pinus luchuensis* complex. *Mol Ecol* 2006, **15**:765-779.
- Avise JC, Ball RM: Principles of genealogical concordance in species concepts and biological taxonomy. *Oxf Surv Evol Biol* 1990, **7**:45-67.
- Avise JC: *Phylogeography: the history and formation of species* Cambridge, Massachusetts: Harvard University Press; 2000.
- Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ: *Plant systematics: A phylogenetic approach* Sunderland, Massachusetts: Sinauer Associates, Inc; 2002.
- Ting CT, Tsaur SC, Wu CI: The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. *Proc Natl Acad Sci USA* 2000, **97**:5313-5316.
- Maddison WP: Gene trees in species trees. *Syst Biol* 1997, **46**:523-536.
- Eckert AJ, Carstens BC: Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. *Mol Phylogenet Evol* 2008, **49**:832-842.
- Meng C, Kubatko LS: Detecting hybrid speciation in the presence of incomplete lineage sorting using gene tree incongruence: A model. *Theor Popul Biol* 2009, **75**:35-45.
- Chiang TY: Lineage sorting accounting for the disassociation between chloroplast and mitochondrial lineages in oaks of southern France. *Genome* 2000, **43**:1090-1094.
- Chiang TY, Schaal BA: Phylogeography of plants in Taiwan and the Ryukyu archipelago. *Taxon* 2006, **55**:31-41.
- Rieseberg LH, Brouillet L: Are many plant species paraphyletic? *Taxon* 1994, **43**:21-32.
- Baack EJ, Rieseberg LH: A genomic view of introgression and hybrid speciation. *Curr Opin Genet Dev* 2007, **17**:513-518.
- Yatabe Y, Kane NC, Scotti-Saintagne C, Rieseberg LH: Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. *Genetics* 2007, **175**:1883-1893.
- Wang WK, Ho CW, Hung KH, Wang KH, Huang CC, Araki H, Hwang CC, Hsu TW, Osada N, Chiang TY: Multilocus analysis of genetic divergence between outcrossing *Arabidopsis* species: evidence of genome-wide admixture. *New Phytol* 2010, **188**:488-500.
- Sibuet JC, Hsu SK: Geodynamics of the Taiwan arc-arc collision. *Tectonophysics* 1997, **274**:221-251.
- Sibuet JC, Hsu SK: How was Taiwan created? *Tectonophysics* 2004, **379**:159-181.
- Hsieh CF: Composition, endemism and phytogeographical affinities of the Taiwan Flora. *Taiwania* 2002, **47**:298-310.
- Ota H: Geographic patterns of endemism and speciation in amphibians and reptiles of the Ryukyu archipelago, Japan, with special reference to their paleogeographical implications. *Res Popul Ecol* 1998, **40**:189-204.
- Wang PX, Wang LJ, Bian YH, Jian ZM: Late Quaternary paleoceanography of the South China Sea: surface circulation and carbonate cycles. *Mar Geol* 1995, **127**:145-165.
- Lu SY, Peng CI, Cheng YP, Hong KH, Chiang TY: Chloroplast DNA phylogeography of *Cunninghamia konishii* (Cupressaceae), an endemic conifer of Taiwan. *Genome* 2001, **44**:797-807.
- DeChaine EG, Martin AP: Historic cycles of fragmentation and expansion in *Parnassius smintheus* (Papilionidae) inferred using mitochondrial DNA. *Evolution* 2004, **58**:113-127.
- Hewitt G: The genetic legacy of the Quaternary ice ages. *Nature* 2000, **405**:907-913.
- Schmitt T, Hewitt GM, Muller P: Disjunct distributions during glacial and interglacial periods in mountain butterflies: *Erebia epiphron* as an example. *J Evol Biol* 2006, **19**:108-113.
- Chamberlain D, Hyam R, Argent G, Fairweather G, Walter KS: *The genus Rhododendron, its classification and synonymy* Oxford, U. K.: Royal Botanic Garden Edinburgh; 1996.
- Irving E, Hebda R: Concerning the origin and distribution of *Rhododendrons*. *J Am Rhododendron Soc* 1993, **47**:139-162.
- Fang R-C, Ming T-L: The floristic study on the genus *Rhododendron*. *Acta Bot Yunnanica* 1995, **17**:359-379. (in Chinese with English abstract).
- Goetsch L, Eckert AJ, Hall BD: The molecular systematics of *Rhododendron* (Ericaceae): A phylogeny based upon RPB2 gene sequences. *Syst Bot* 2005, **30**:616-626.
- Li HL, Lu SY, Yang YP, Tseng YH: Ericaceae. In *Flora of Taiwan. Volume 4.* 2 edition. Editorial Committee of the Flora of Taiwan. Taipei; 1998:17-39.
- Chung JD, Lin TP, Chen YL, Cheng YP, Hwang SY: Phylogeographic study reveals the origin and evolutionary history of a *Rhododendron* species complex in Taiwan. *Mol Phylogenet Evol* 2007, **42**:14-24.

31. Hwang SY, Hsu KK: Molecular phylogeny of eight Taiwanese *Rhododendron* species based on chloroplast *trnF-trnL* DNA sequences. *Taiwan J For Sci* 2001, **16**:153-160.
32. Tsai CC, Huang SC, Chen CH, Tseng YH, Huang PL, Tsai SH, Chou CH: Genetic relationships of *Rhododendron* (Ericaceae) in Taiwan based on the sequence-of the internal transcribed spacer of ribosomal DNA. *J Hortic Sci Biotechnol* 2003, **78**:234-240.
33. Lu SY, Yang YP: A revision of *Rhododendron* (Ericaceae) of Taiwan. *Bull Taiwan For Res Ins* 1989, **4**:155-166.
34. Milne RI: Phylogeny and biogeography of *Rhododendron* subsection *Pontica*, a group with a tertiary relict distribution. *Mol Phylogenet Evol* 2004, **33**:389-401.
35. Pan FJ: Altitudinal variation of *Rhododendron pseudochrysanthum* and *R. morii*. *Quart J Chin For* 1983, **21**:99-102.
36. Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA: Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc Natl Acad Sci USA* 2006, **103**:9578-9583.
37. Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA: Phylogeographic studies in plants: problems and prospects. *Mol Ecol* 1998, **7**:465-474.
38. Yuan QJ, Zhang ZY, Peng H, Ge S: Chloroplast phylogeography of *Dipentodon* (Dipentodontaceae) in southwest China and northern Vietnam. *Mol Ecol* 2008, **17**:1054-1065.
39. Holderegger R, Abbott RJ: Phylogeography of the arctic alpine *Saxifraga oppositifolia* (Saxifragaceae) and some related taxa based on cpDNA and its sequence variation. *Am J Bot* 2003, **90**:931-936.
40. Huang SC, Wang WK, Peng CI, Chiang TY: Phylogeography and conservation genetics of *Hygrophila pogonocalyx* (Acanthaceae) based on *atpB-rbcL* noncoding spacer cpDNA. *J Plant Res* 2005, **118**:1-11.
41. 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**:2669-2681.
42. Ouborg NJ, Piquot Y, Van Groenendael JM: Population genetics, molecular markers and the study of dispersal in plants. *J Ecol* 1999, **87**:551-568.
43. Birky CW: Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc Natl Acad Sci USA* 1995, **92**:11331-11338.
44. Cruzan MB, Templeton AR: Paleocology and coalescence: phylogeographic analysis of hypotheses from the fossil record. *Trends Ecol Evol* 2000, **15**:491-496.
45. McCauley DE: Contrasting the distribution of chloroplast DNA and allozyme polymorphism among local populations of *Silene alba*: implications for studies of gene flow in plants. *Proc Natl Acad Sci USA* 1994, **91**:8127-8131.
46. McCauley DE: The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends Ecol Evol* 1995, **10**:198-202.
47. Orive ME, Asmussen MA: The effects of pollen and seed migration on nuclear dicytoplasmic systems. II. A new method for estimating plant gene flow from joint nuclear cytoplasmic data. *Genetics* 2000, **155**:833-854.
48. Stehlik I: Glacial history of the alpine herb *Rumex nivalis* (Polygonaceae): A comparison of common phylogeographic methods with nested clade analysis. *Am J Bot* 2002, **89**:2007-2016.
49. Stehlik I, Blattner FR, Holderegger R, Bachmann K: Nunatak survival of the high alpine plant *Eritrichium nanum* (L.) Gaudin in the central Alps during the ice ages. *Mol Ecol* 2002, **11**:2027-2036.
50. Song BH, Wang XQ, Wang XR, Sun LJ, Hong DY, Peng PH: Maternal lineages of *Pinus densata*, a diploid, hybrid. *Mol Ecol* 2002, **11**:1057-1063.
51. Hwang SY, Lin HW, Kuo YS, Lin TP: RAPD variation in relation to population differentiation of *Chamaecyparis formosensis* and *Chamaecyparis taiwanensis*. *Bot Bull Acad Sinica* 2001, **42**:173-179.
52. Groth JG, Barrowclough GF: Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. *Mol Phylogenet Evol* 1999, **12**:115-123.
53. Tajima F: Statistical-Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics* 1989, **123**:585-595.
54. Fay JC, Wu CI: A human population bottleneck can account for the discordance between patterns of mitochondrial versus nuclear DNA variation. *Mol Biol Evol* 1999, **16**:1003-1005.
55. Willyard A, Cronn R, Liston A: Reticulate evolution and incomplete lineage sorting among the ponderosa pines. *Mol Phylogenet Evol* 2009, **52**:498-511.
56. Milne RI, Abbott RJ, Wolff K, Charberlain DF: Hybridization among sympatric species of *Rhododendron* (Ericaceae) in Turkey: Morphological and molecular evidence. *Am J Bot* 1999, **86**:1776-1785.
57. Milne RI, Terzioğlu S, Abbott RJ: A hybrid zone dominated by fertile F(1)s: maintenance of species barriers in *Rhododendron*. *Mol Ecol* 2003, **12**:2719-2729.
58. Denda T, Yokota M: Hybrid origins of *Ixeris nakazonei* (Asteraceae, Lactuceae) in the Ryukyu Archipelago, Japan: evidence from molecular data. *Bot J Linn Soc* 2003, **141**:379-387.
59. Peng CI, Chiang TY: Molecular confirmation of unidirectional hybridization in *Begonia* × *taipeiensis* Peng (Begoniaceae) from Taiwan. *Ann Mo Bot Gard* 2000, **87**:273-285.
60. Setoguchi H, Watanabe I: Intersectional gene flow between insular endemics of *Ilex* (Aquifoliaceae) on the Bonin Islands and the Ryukyu Islands. *Am J Bot* 2000, **87**:793-810.
61. Shih FL, Hwang SY, Cheng YP, Lee PF, Lin TP: Uniform genetic diversity, low differentiation, and neutral evolution characterize contemporary refuge populations of Taiwan fir (*Abies kawakamii*, Pinaceae). *Am J Bot* 2007, **94**:194-202.
62. Ellstrand NC, Elam DR: Population genetic consequences of small population size: implications for plant conservation. *Annu Rev Ecol Syst* 1993, **24**:217-242.
63. Kameyama Y, Isagi Y, Naito K, Nakagoshi N: Microsatellite analysis of pollen flow in *Rhododendron metternichii* var. *hondoense*. *Ecol Res* 2000, **15**:263-269.
64. Kameyama Y, Isagi Y, Nakagoshi N: Patterns and levels of gene flow in *Rhododendron metternichii* var. *hondoense* revealed by microsatellite analysis. *Mol Ecol* 2001, **10**:205-216.
65. Peters JL, Gretes W, Omland KE: Late Pleistocene divergence between eastern and western populations of wood ducks (*Aix sponsa*) inferred by the 'isolation with migration' coalescent method. *Mol Ecol* 2005, **14**:3407-3418.
66. Cheng YP, Hwang SY, Lin TP: Potential refugia in Taiwan revealed by the phylogeographical study of *Castanopsis carlesii* Hayata (Fagaceae). *Mol Ecol* 2005, **14**:2075-2085.
67. Chung JD, Lin TP, Tan YC, Lin MY, Hwang SY: Genetic diversity and biogeography of *Cunninghamia konishii* (Cupressaceae), an island species in Taiwan: a comparison with *Cunninghamia lanceolata*, a mainland species in China. *Mol Phylogenet Evol* 2004, **33**:791-801.
68. Chiang YC, Schaal BA, Ge XJ, Chiang TY: Range expansion leading to departures from neutrality in the nonsymbiotic hemoglobin gene and the cpDNA *trnL-trnF* intergenic spacer in *Trema dielsiana* (Ulmaceae). *Mol Phylogenet Evol* 2004, **31**:929-942.
69. Doyle JJ, Doyle JL: A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 1987, **19**:11-15.
70. Chiang TY, Hong KH, Peng CI: Experimental hybridization reveals biased inheritance of the internal transcribed spacer of the nuclear ribosomal DNA in *Begonia* × *taipeiensis*. *J Plant Res* 2001, **114**:343-351.
71. Chiang TY, Schaal BA, Peng CI: Universal primers for amplification and sequencing a noncoding spacer between the *atpB* and *rbcL* genes of chloroplast DNA. *Bot Bull Acad Sinica* 1998, **39**:245-250.
72. Katoh K, Toh H: Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC Bioinformatics* 2008, **9**:212-224.
73. Guindon S, Gascuel O: A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003, **52**:696-704.
74. Posada D, Buckley TR: Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Syst Biol* 2004, **53**:793-808.
75. Posada D, Crandall KA: MODELTEST: testing the model of DNA substitution. *Bioinformatics* 1998, **14**:817-818.
76. Excoffier L, Smouse PE: Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. *Genetics* 1994, **136**:343-359.
77. Chiang TY, Schaal BA: Phylogeography of North American populations of the moss species *Hylocomium sylvendens* based on the nucleotide sequence of internal transcribed spacer 2 of nuclear ribosomal DNA. *Mol Ecol* 1999, **8**:1037-1042.

78. Heled J, Drummond AJ: **Bayesian Inference of Species Trees from Multilocus Data.** *Mol Biol Evol* 2010, **27**:570-580.
79. Watterson GA: **On the number of segregating sites in genetical models without recombination.** *Theor Popul Biol* 1975, **7**:256-275.
80. Librado P, Rozas J: **DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.** *Bioinformatics* 2009, **25**:1451-1452.
81. Tajima F: **Statistical method for testing the neutral mutation hypothesis by DNA polymorphism.** *Genetics* 1989, **123**:585-595.
82. Dupanloup I, Schneider S, Excoffier L: **A simulated annealing approach to define the genetic structure of populations.** *Mol Ecol* 2002, **11**:2571-2581.
83. Hey J, Nielsen R: **Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D-persimilis*.** *Genetics* 2004, **167**:747-760.
84. Wolfe KH, Li WH, Sharp PM: **Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs.** *Proc Natl Acad Sci USA* 1987, **84**:9054-9058.
85. Kay KM, Whittall JB, Hodges SA: **A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects.** *BMC Evol Biol* 2006, **6**:36-44.
86. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A: **Relaxed phylogenetics and dating with confidence.** *PLoS Biol* 2006, **4**:699-710.
87. Rambaut A, Drummond AJ: *Tracer - MCMC trace analysis tool* Oxford, UK: University of Oxford; 2004.
88. Rosenberg NA: **The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly, and polyphyly in a coalescent model.** *Evolution* 2003, **57**:1465-1477.

doi:10.1186/1471-2148-11-108

**Cite this article as:** Huang *et al.*: Genetic population structure of the alpine species *Rhododendron pseudochrysanthum* sensu lato (Ericaceae) inferred from chloroplast and nuclear DNA. *BMC Evolutionary Biology* 2011 **11**:108.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

