**Taiwania 67**(3): 366–376, **2022**
*DOI: 10.6165/tai.2022.67.366*

# Incomplete lineage sorting and secondary admixture results in the paraphyly of *Lecanorchis cerina*, *L. suginoana*, and *L. thalassica* (Orchidaceae) in Taiwan

**Shih-Ying HWANG[1,*], Yi-Shao LI[1,†], Tsan-Piao LIN[2,†]**

***1.*** *Department of Life Science, National Taiwan Normal University, 88 Tingchow Road, Section 4, Taipei 11677, Taiwan.* ***2.*** *Institute of Plant Biology, National Taiwan University, 1 Roosevelt Road, Section 4, Taipei 10617, Taiwan.* [†] *These authors contributed equally to this work.* [*]*Corresponding author's email: hsy9347@ntnu.edu.tw; Tel: +886-2-77346250; Fax: +886-2-2931-2904*

(Manuscript received 9 May 2022; Accepted 25 June 2022; Online published 2 July 2022)

ABSTRACT: Inferring the evolutionary history of a group of closely related species can be challenging. Genetic diversity, structure, and multilocus phylogeny were studied to shed light on the evolutionary processes that shapes diversity in three closely related *Lecanorchis* species (*L. cerina*, *L. suginoana*, and *L. thalassica*). In this study, we surveyed nucleotide and genetic variation, respectively, using nuclear internal transcribed spacer 1 (nrITS1) and amplified fragment length polymorphism (AFLP) in individuals of *Lecanorchis* species. Both nrITS1 parsimonious and AFLP neighbor-joining trees showed paraphyly indicated that lineage sorting for these species may be incomplete. ITS1 haplotype sharing was found within *L. cerina*, within *L. suginoana,* and between *L. suginoana* and *L. thalassica*. Higher AFLP diversity was observed in *L. cerina* than that in *L. suginoana* and in *L. thalassica*, indicating larger population size in *L. cerina*. Two contrasting patterns of AFLP genetic differentiation between geographic neighbors relative to geographic distant populations were observed. Two major AFLP genetic components were found in individuals of *Lecanorchis* species examined. Individuals of *L. cerina* and of *L. thalassica* were classified into either component with high probabilities. Most individuals of *L. suginoana* were classified into one of these components, but two individuals showed apparent admixture of these components. The evidence of paraphyly, haplotype sharing, higher *L. cerina* diversity, sharing and admixture of genetic components, and the contrasting patterns of genetic differentiation suggest that both incomplete lineage sorting and secondary gene flow could have occurred in these closely related species, though at different temporal scale.

KEY WORDS: AFLP, incomplete lineage sorting, *Lecanorchis*, nrITS1, secondary gene flow, Vanilleae, Vanilloideae.

## INTRODUCTION

The orchid genus *Lecanorchis* Blume (Orchidaceae) is mycoheterotrophic and comprises 31 accepted names (WCVP, 2022) widely distributed in the Southeast and Eastern Asia. Fourteen *Lecanorchis* species were recorded in Taiwan (Lin, 2019). Mainly growing in moist humus of primary forest, species of *Lecanorchis* in the field were often overlooked due to their small, leafless achlorophyllous habits, and flowers usually have a mixture of white, brown, and purple. The difficulty in precise identification of *Lecanorchis* species is often encountered because of the overall morphological similarity, brief flowering periods, and the lack of important diagnostic characteristics in herbarium specimens (Suetsugu *et al*., 2020). Nonetheless, with the assistance of molecular analysis of nuclear internal transcribed spacer (nrITS) sequences, species-level phylogenetic studies can be more efficient (e.g., Suetsugu *et al*., 2018; Suetsugu *et al*., 2019; Takashima *et al*., 2019; Suetsugu *et al*., 2020; but see Álvarez and Wendel, 2003). Molecular analysis of nrITS clearly separate *L. nigricans* and *L. taiwaniana* even though they are considered synonymous (Suetsugu *et al*., 2018) and did not support the placement of *L. suginoana* (Tuyama) Seriz. as a variety of *L. japonica* as previously interpreted (Takashima *et al*., 2019). Additionally, a new species *L. moritae* (Suetsugu *et al*., 2019) and a new form of *L. japonica*, *L. janopica* f. *lutea* (Suetsugu *et al*., 2020) were identified based on nrITS sequences.

Three *Lecanorchis* species, *L. cerina*, *L. suginoana*, and *L. thalassica*, having overlapping flowering seasons, can be found in the same mountainous area such as Hakanni mountain in northwestern Taiwan (Fig. 1). *L. cerina* (Fig. 2A, B), an endemic species, is the most widely occurred species throughout Taiwan at elevations of 400–1700 m. *L. thalassica* (Fig. 2E, F), an endemic, is restricted to central part of Taiwan and grows at elevations of 1400–2000 m. Whereas *L. suginoana* (Fig. 2C, D), a species also occur to Japan, is restricted to central part of Taiwan and found at elevations of 1200–2000 m. Occasionally, suspected hybrids between these species were observed. For example, the flower of one individual plant from Youluo mountain showed some features similar to *L. cerina* and some to *L. thalassica* (Fig. 3). Hybridization is often inferred from incongruence among gene trees from independent loci (Linder and Rieseberg, 2004) or from trees in which species are not monophyletic (Funk and Omland, 2003). However, incomplete lineage sorting (ILS) can also display these gene tree features. Because of the overall morphological similarity, relationships between *L. cerina*, *L. suginoana*, and *L. thalassica* can be genetically indistinguishable due to hybridization and/or ILS.

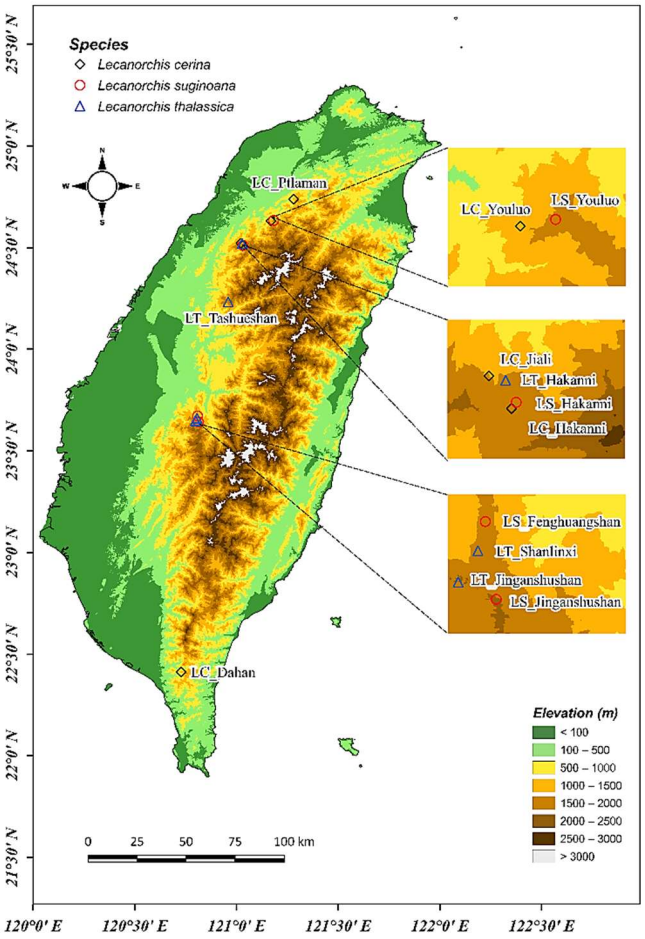


**Fig. 1.** Map of Taiwan with sample locations for ***Lecanorchis*** individuals collected in the present study

Although coalescent-based frameworks have been proposed to distinguish secondary gene flow from ILS (e.g., Nielsen and Wakeley, 2001; Hey, 2010; Csilléry *et al*., 2010; Sunnåker *et al*., 2013), these methods require information regarding to the true phylogenetic relationships among the species examined. Nonetheless, secondary gene flow and ILS can also be disentangled by comparing patterns of genetic differentiation between pairs of neighboring and distantly located populations of different species (Muir and Schlötterer, 2005). Under the secondary gene flow scenario, among geographic neighboring populations of different species, gene flow is expected to be higher resulting in lower levels of interspecific genetic differentiation than between distantly located populations (Petit and Excoffier, 2009). In contrast, the retention of shared ancestral polymorphism due to ILS may result in the low levels of among population divergence across wide geographic distributions of species (Muir and Schlötterer, 2005, Szövényi *et al*., 2008; Stenøien *et al*., 2011; Meleshko *et al*., 2021) and that geographic distant populations are expected to be no more differentiated than neighboring populations. Moreover, it is likely that ILS can be inferred for tangled species relationships if one species harbored significantly higher genetic diversity than that of the others. In this study, we collected floral samples of *L. cerina*, *L. suginoana*, and *L. thalassica* in the same and different mountainous areas in Taiwan. We obtained ITS1 sequences and surveyed the genetic variation using amplified fragment length polymorphism (AFLP) (Vos *et al*., 1995). The aims of this study were to investigate (1) genetic relationships between the three closely related *Lecanorchis* species at individual level, (2) genetic differentiation at species and population level, and (3) sharing and admixture of ITS1 haplotypes and AFLP genetic components between species. Results of this study can be used to gain an insight into whether secondary gene flow and/or ILS played roles in the evolution of the three closely related *Lecanorchis* species.

## MATERIALS AND METHODS

### Sampling and total DNA extraction

Thirty-nine floral samples of *L. cerina* ($n$ = 17), *L. suginoana* ($n$ = 12), and *L. thalassica* ($n$ = 10) were collected by Tsan-Piao Lin during 2020 and 2021 flowering season. Map of sampling sites (Fig. 1) was created using Quantum Geographic Information System (QGIS version 3.10; http://qgis.osgeo.org) based on the 20 m digital elevation model (DEM) provided by the Taiwan Ministry of Interior (https://data.moi.gov.tw/MoiOD/default/Index.aspx). The color-ramp of DEM was rendered at a 500 m interval for height visualization in the basemap generation. The locations of sampling site were saved as a vector layer assigned to the wgs84 geographic coordinate system. QGIS map composer was used to overlap the vector layer upon the basemap.

Flowers were silica gel dried and ground with liquid nitrogen for genomic DNA extraction based on a modified CTAB (cetyltrimethyl ammonium bromide) procedure (Doyle and Doyle, 1987). DNA was precipitated with ethanol and dissolved in 200 µl TE buffer, pH 8.0 after washing with 70% ethanol. DNA concentration was quantified using a NanoDrop spectrophotometer (NanoDrop Technology, Wilmington, DE, USA) and DNA samples were stored at -20°C.

### DNA sequencing, sequence alignment, and phylogeny assessment of internal transcribed spacer

Double-stranded DNA templates for direct sequencing were amplified by the polymerase chain reaction (PCR). PCR amplification for ITS1 was achieved using ITS1-F (5'-GGAGAAGTCGTAACAAGGTTTCCG-3') and ITS1-R (5'-ATCCTGCAATTCACACCAAGTATCG-3') primers. Reactions of PCR amplification were conducted on a DNA Programmable Thermal Cycler (DNAEngine PTC-0200 Thermal Cycler, Bio-Rad, Foster City, CA, USA). The PCR mixture (20 µl) contained 60 ng of template DNA, 10 µM ITS1-F primer, 10 µM ITS1-R primer, 2.5 mM dNTPs,

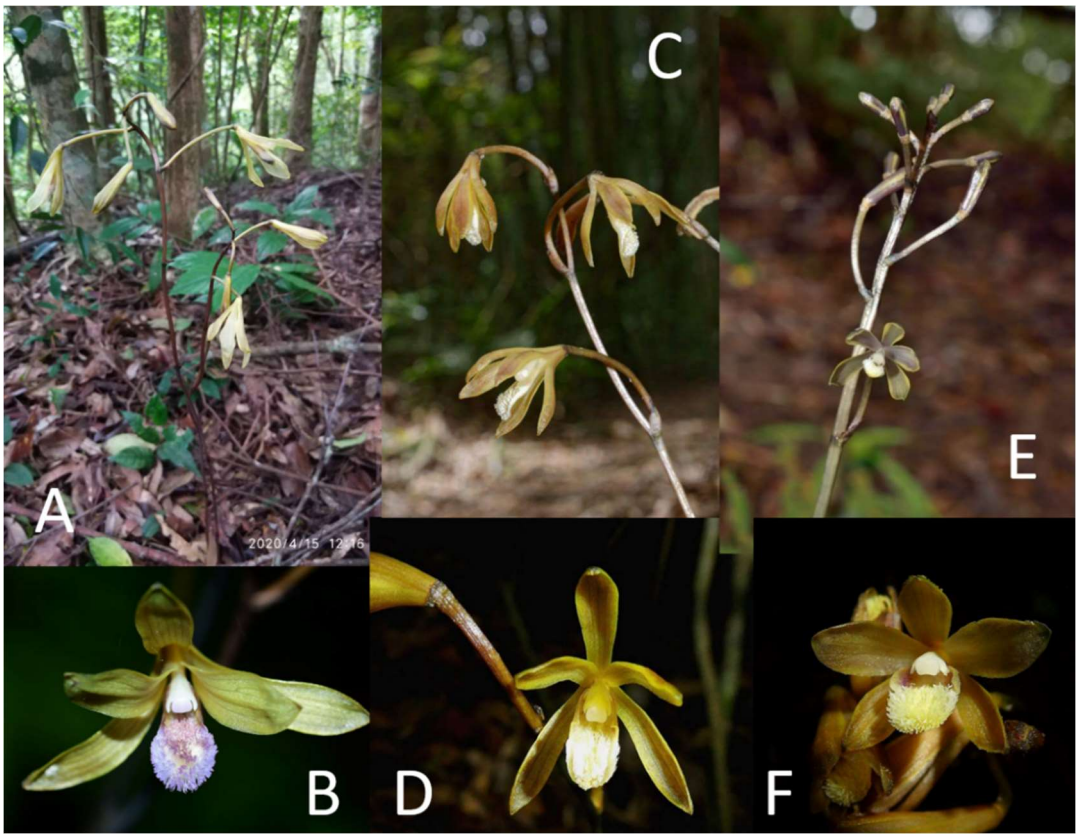


**Fig. 2.** Photos of ***Lecanorchis cerina*** (**A, B**), ***L. suginoana*** (**C, D**) and ***L. thalassica*** (**E, F**). B, taken by Ching-Huang Liu; A and F, taken by Da-Ming Huang; C, D and E, taken by Kuo-Chu Yeuh

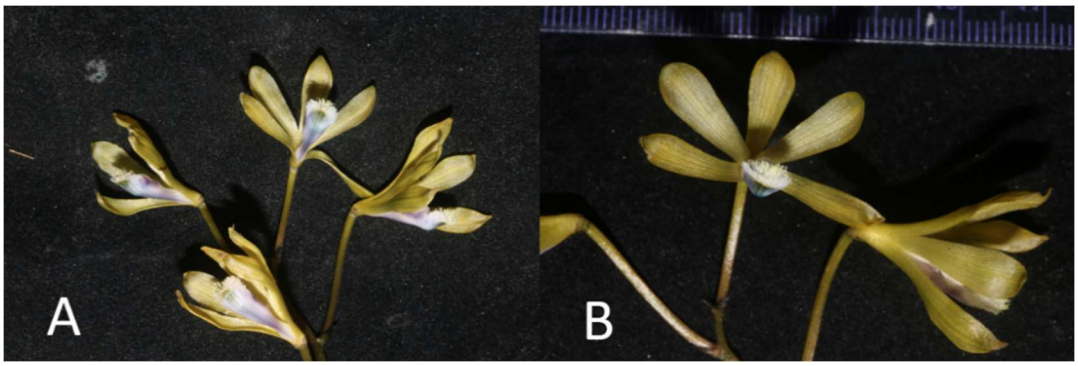


**Fig. 3.** Photos **(A–B)** of one suspected hybrid between ***Lecanorchis*** species investigated. Photos were taken by Da-Ming Huang from Youluo mountain at 1400 m during *Lecanorchis* flowering season in 2019.

10 U *Taq* DNA polymerase (Zymeset Biotech, Taipei, Taiwan), and 10X PCR buffer (Zymeset). The cycle of the PCR reaction consisted of 3 min at 94°C for pre-denaturation, followed by 10 cycles of 45 s at 94°C for denaturation, 20 s at 62°C for annealing, and 1 min at 72°C for extension. Amplification products were run on 2% agarose gels, purified with a QiaGen purification kit (Thermo Fisher, Taipei, Taiwan), and sequenced using both forward and reverse reactions with a *Taq* BigDye Terminator Cycle Sequencing Kit and a model ABI373XL automated sequencer (Applied Biosystems, Foster City, CA, USA).

We performed multiple sequence alignment using the *msa* function of the R msa package (Bodenhofer, 2015) based on the ClustalW algorithm (Thompson *et al*., 1994) in the R environment (R Development Core team, 2020). Maximum parsimony analysis was performed using the *pratchet* function of the R phangorn package (Schliep, 2011) based on the fitch algorithm and heuristic-search-based subtree pruning and regrafting rearrangement. The amount of support for monophyletic groups was evaluated using 100 bootstrap replicates (Felsenstein, 1985). Gaps were treated as missing data and all characters were accorded equal weight. The consistency

**Table 1.** Sampling sites, species, population codes, population ITS1 haplotypes, and ITS1 nucleotide diversity at species level of three closely related ***Lecanorchis*** species.

| | *N* | Aligned length (bp) | Haplotype | Nucleotide diversity ($\theta\pi$) | Nucleotide diversity ($\theta_S$) |
|---|---|---|---|---|---|
| ***Lecanorchis cerina*** | 17 | 298 | | 0.01387 | 0.01020 |
| Dahan (LCD) | 1 | | F (1) | | |
| Hakanni (LCH) | 1 | | F (1) | | |
| Jiali (LCJ) | 5 | | F (4), G (1) | | |
| Ptlaman (LCP) | 3 | | A (1), D (1), F (1) | | |
| Youluo (LCY) | 7 | | A (1), B (1), F (4), H (1) | | |
| ***Lecanorchis suginoana*** | 12 | 301 | | 0.01563 | 0.01385 |
| Fenghuangshan (LSF) | 2 | | N (1) | | |
| Hakanni (LSH) | 6 | | E (1), K (1), N (2), O (1), U (1) | | |
| Jinganshushan (LSJ) | 2 | | J (1), N (1), T (1) | | |
| Youluo (LSY) | 2 | | K (1), N (1) | | |
| ***Lecanorchis thalassica*** | 10 | 301 | | 0.01307 | 0.01495 |
| Hakanni (LTH) | 3 | | K (1), P (1), Q (1) | | |
| Jinganshushan (LTJ) | 2 | | I (2) | | |
| Shanlinxi (LTS) | 2 | | L (1), S (1) | | |
| Tashueshan (LTT) | 3 | | C (1), M (1), R (1) | | |
| Total | 39 | 301 | | 0.01526 | 0.01006 |

*N*, number of samples

index (CI, Kluge and Farris, 1969) and the retention index (RI, Farris, 1989) were computed using the *CI* and the *RI* function of the R phangorn package, respectively. ITS1 sequences of Japanese *L. japonica* and *L. hokurikuensis* (GenBank number MN221410 and LC504448, respectively) were used as outgroups. ITS1 sequences of Japanese *L. suginoana* and *L. moritae* (GenBank number LC457494 and LC457487, respectively) were also included in the analysis. The ITS1 sequences of *Lecanorchis* examined in this study were deposited in GenBank under the accession numbers ON212619–ON212639 (haplotypes A–U; Table 1). A TCS haplotype network (Clement *et al*., 2002) was depicted using PopART software (Leigh and Bryant, 2015). Nucleotide diversity measures ($\theta_\pi$ and $\theta_S$) were estimated using DnaSP v.6 (Rozas *et al*., 2017). $\theta_\pi$ and $\theta_S$ were calculated based on the average pairwise number of differences between sequences and the number of segregating sites per sequence, respectively.

## AFLP genotyping

Of the 39 samples collected, one sample of *L. suginoana* was excluded for AFLP genotyping due to the scarce of total DNA left after ITS1 amplification experiment. Restriction digestion was performed using 200 ng of total genomic DNA incubated with 10 U *Eco*RI and 10 U *Mse*I in a 10X CutSmart buffer (New England Biolabs, Ipswich, MA, USA) at 37°C for 1.5 h. The reaction was then placed at 65 °C for 15 min to deactivate restriction enzyme activity. The digested DNA products were mixed with 5 μM *Eco*RI adaptor, 50 μM *Mse*I adaptor, 5 U T4 DNA ligase (Thermo Scientific, Vilnius, Lithuania), and 5X ligation buffer (Thermo Scientific) in a 10-μl ligation reaction mixture at 22°C for 1 h.

Pre-selective amplification was performed using 3 μl diluted digested samples (1:9 dilution with $ddH_2O$) as a template in a 20-μl volume containing 12 μM *Eco*RI (E00: 5'GACTGCGTACCAATTC-3') primer, 12 μM *Mse*I (M00: 5'-GATGAGTCCTGAGTAA-3') primer, 2.5 mM dNTPs, 1 U *Taq* DNA polymerase (Zymeset), and 10X PCR buffer (Zymeset). The pre-selective amplification was performed with an initial holding at 72°C for 2 min and pre-denaturation at 94°C for 3 min, followed by 25 cycles of 30 s at 94°C, 30 s at 56°C, and 1 min at 72°C, with a final 5-min holding at 72°C. Eight *Eco*RI-*Mse*I (E00 and M00) primer combinations with additional nucleotide bases added at the ends (Table S1) were used for AFLP selective amplification. *Eco*RI primer was labeled with fluorescent dye (6-carboxyfluorescein or hexachloro-fluorescein) and amplification was performed in a 20-μl volume containing 10 μM *Eco*RI and *Mse*I primers, 2.5 mM dNTPs (2.5 mM), 2 U *Taq* DNA polymerase (Zymeset), 10X PCR buffer (Zymeset), and 1 μl of each diluted pre-selective amplified product (1:19 dilution with $ddH_2O$). PCR conditions were initial holding at 94°C for 3 min, followed by 13 cycles of 30 s at 94°C, 30 s at 65–56°C (decreasing the temperature by 0.7°C each cycle), 1 min at 72°C, then 23 cycles of 30 s at 94°C, 30 s at 56°C, and 1 min at 72°C, with a final 5-min holding at 72°C.

PCR products of selective amplification were electrophoresed on an ABI 3730XL DNA analyzer and scored with Peak Scanner v.1.0 (Applied Biosystem). We scored AFLP fragments using a fluorescent threshold set at 150 units. Amplified fragments peaked by less than one nucleotide in a ± 0.8 base pair window were scored as the same fragment in the range of 100–500 bp. We removed amplified fragments that scored higher than 99% or less than 1% of individuals. Error rate per locus of each AFLP primer combination was calculated based on

amplification replicates obtained from randomly chosen five samples from each species. Loci with an error rate per locus greater than 5% were removed (Bonin *et al*., 2004). The mean error rate for AFLP was 4.51% (Table S1). AFLP data is available from the corresponding author on reasonable request.

**Genetic diversity, differentiation, clustering, and relationships based on AFLP variation**

The proportion of polymorphic loci (%*P*, 95% criterion) and population and species level unbiased expected heterozygosity ($uH_E$) (Nei, 1987) were estimated based on allele frequencies assuming Hardy-Weinberg equilibrium with non-uniform prior distribution (Zhivotovsky, 1999) using AFLP-SURV v.1.0 (Vekemans *et al*., 2002). The number of private alleles for each species was calculated using the *private_alleles* function of the R package poppr (Kamvar *et al*., 2015). $uH_E$ per locus was estimated using ARLEQUIN v.6.0 (Excoffier and Lischer, 2010). Linear mixed effect model (LMM), considering population as a fixed factor and locus as a random factor, was used to test the difference of mean $uH_E$ per locus among species. The *lmer* function of the R package lme4 (Bates *et al*., 2015) was used in the LMM analysis based on reduced maximum likelihood estimation, and significance tested using the *Anova* function of the R package car based on type II Wald $\chi^2$ statistic (Fox and Weisberg, 2011). Tukey's multiple comparison test was applied for pairwise species comparisons using the *lsmeans* function of the R package emmeans (Lenth, 2018). Analysis of molecular variance (AMOVA) was used to estimate the level of genetic differentiation between species ($\Phi_{CT}$) and between populations ($\Phi_{ST}$) of the three *Lecanorchis* species using the *poppr.amova* function of the R package poppr, and significance tested using the *randtest* function of the R package ade4 (Dray and Dufour, 2007) with 9,999 permutations. Pairwise $F_{ST}$ between species and between populations were estimated using ARLEQUIN and significance tested with 10,000 permutations.

Assignment of *Lecanorchis* individuals into genetic homogeneous groups was assessed using sNMF algorithm of landscape and ecological association (LEA; Frichot and Francois, 2015) and STRUCTURE v.2.3 (Pritchard *et al*., 2000) based on AFLP variation. Individual assignments with $K$ = 1–4 based on least-squares optimization were assessed using the *snmf* function of the R package LEA (Frichot and Francois, 2015). The regularization parameter, iterations, and repetitions in *snmf* were set to 100, 200, and 10, respectively, and other arguments set to defaults. The mean minimal cross-entropy (CE) was used to determine the best $K$ in LEA. Bayesian clustering was also used for individual assignment to different population clusters ($K$ = 1–4) using STRUCTURE v.2.3 (Pritchard *et al*., 2000). STRUCTURE identifies distinct genetic populations and assigns individuals probabilistically to populations. We applied no admixture model with allele frequencies correlated among populations assuming that the genotypes of individuals within a population are derived completely from that population. Prior population information was not incorporated in calculating the posterior probability for individuals belonging to different clusters. For each clustering scenario ($K$ = 1–4), we performed 20 complete analyses each with a 50,000 burn-in period and a sampling period of $5\times10^5$ iterations. Different clustering scenarios were evaluated using the mean log probability, *Ln*P(D) (Pritchard *et al*., 2000) and the change in log probability, $\Delta L$(K) (Evanno *et al*., 2005).

Genetic clustering of *Lecanorchis* individuals was further assessed using discriminant analysis of principal components (DAPC; Jombart *et al*., 2010). The *find.clusters* and *dapc* functions of the R package adegenet (Jombart and Ahmed, 2011) were used in DAPC analysis. Bayesian information criterion (BIC) was used to determine the best $K$ in DAPC.

Genetic relationships among individuals of the three *Lecanorchis* species were assessed using neighbor-joining (NJ) tree based on AFLP variation. The pairwise Nei's genetic distances (Nei, 1978) were calculated using the *nei.dist* function of the R package poppr and used in generating an unrooted NJ tree using the *nj* function of the R package ape (Paradis and Schliep, 2019). The bootstrap support values were calculated based on 1,000 replicates using the *aboot* function of the R package poppr.

## RESULTS AND DISCUSSION

We obtained double stranded ITS1 sequences for all samples studied. The length of ITS1 ranged from 291 to 299 bp before alignment and 301 bp after alignment. Excluding the outgroup, ITS1 had 12 polymorphic sites, 6 of which were parsimony informative and 1 was singleton, yielding 21 haplotypes (Table 1). Similar nucleotide diversity estimated using $\theta_\pi$ and $\theta_S$, based on the ITS1 sequences, were found for the three *Lecanorchis* species examined (Table 1).

Parsimony analysis produced a most parsimonious tree (Fig. 4) with 99 steps. Low CI (0.525) and RI (0.779) values were obtained. Essentially low bootstrap support values (< 50%) were found and the phylogeny can completely collapse into one single lineage. Hence, the short ITS1 length obtained may have low power to resolve phylogenetic relationships of the three *Lecanorchis* species studied, at individual level. Nonetheless, several *L. cerina* individuals were most closely related to the outgroups, *L. japonica* and *L. hokurikuensis. L. suginoana* and *L. thalassica* individuals were nested within *L. cerina* individuals and no reciprocal monophyly was observed for *L. suginoana* and *L. thalassica*. ITS1 sequences of Japanese *L. suginoana* and

**Table 2.** Species and population genetic diversity estimates of ***Lecanorchis cerina***, ***L. suginoana***, and ***L. thalassica*** based on AFLP variation.

| Species | Longitude /latitude | Number of samples | Number of private alleles ($N_P$) | Percent polymorphism (%*P*) | $uH_E$ (SE) |
|---|---|---|---|---|---|
| ***Lecanorchis cerina*** | | 17 | 126 | 81.9 | 0.220 (0.007) |
| Dahan (LCD) | 120.72/22.41 | 1 | | - | - |
| Hakanni (LCH) | 121.03/24.50 | 1 | | - | - |
| Jiali (LCJ) | 121.02/24.52 | 5 | | 61.9 | 0.223 (0.009) |
| Ptlaman (LCP) | 121.28/24.73 | 3 | | 51.4 | 0.256 (0.010) |
| Youluo (LCY) | 121.19/24.64 | 7 | | 78.4 | 0.245 (0.007) |
| ***Lecanorchis suginoana*** | | 11 | 7 | 72.1 | 0.166 (0.007) |
| Fenghuangshan (LSF) | 120.80/23.67 | 1 | | - | - |
| Hakanni (LSH) | 121.03/24.51 | 6 | | 50.5 | 0.154 (0.008) |
| Jinganshushan (LSJ) | 120.81/23.64 | 2 | | 32.7 | 0.202 (0.012) |
| Youluo (LSY) | 121.19/24.65 | 2 | | 27.3 | 0.170 (0.012) |
| ***Lecanorchis thalassica*** | | 10 | 10 | 75.9 | 0.190 (0.007) |
| Hakanni (LTH) | 121.03/24.50 | 3 | | 34.6 | 0.157 (0.009) |
| Jinganshushan (LTJ) | 120.79/23.64 | 2 | | 38.7 | 0.239 (0.012) |
| Shanlinxi (LTS) | 120.80/23.66 | 2 | | 36.8 | 0.226 (0.012) |
| Tashueshan (LTT) | 120.95/24.23 | 3 | | 35.2 | 0.172 (0.010) |

$uH_E$, unbiased expected heterozygosity

*L. moritae* were closely related to Taiwanese *L. suginoana*. Further large scale phylogeographic study may be helpful in resolving the genealogical and phylogenetic relationships of Taiwanese *Lecanorchis* species. Tangled phylogenetic relationships for taxa that have undergone recent, rapid radiation are commonly observed (Rokas and Carroll, 2006). The ITS1 parsimonious and AFLP NJ (Fig. 5) trees reflected paraphyly of individuals of *Lecanorchis* species examined, suggestive of a possibility of rapid radiation of ancestral polymorphism (Pease *et al*., 2016; Ebersbach *et al*., 2017) despite the need for more evidence. Nonetheless, rapid evolutionary transitions in floral characters of orchids are common, including species occurring in Taiwan (e.g., Pérez-Excobar *et al*., 2017; Hu *et al*., 2020; Farminhão *et al*., 2021).

As shown in Fig. S1, most ITS1 haplotypes grouped in cluster 1 contained haplotypes belonged to *L. cerina*, *L. suginoana*, and *L. thalassica*. Cluster 2 contained only the haplotypes belonged to *L. suginoana* and *L. thalassica*. Haplotype A specific to *L. cerina* was found in Youluo and Ptlaman mountains (Table 1, Fig. S1). Haplotype F (the most frequent haplotype in cluster 1) specific to *L. cerina* was found in Youluo, Hakanni, Ptlaman, Jiali, and Dahan mountains. Haplotype N (the most frequent haplotype in cluster 2) specific to *L. suginoana* was found in Hakanni, Jinganshushan, Youluo, and Fenghuangshan mountains. These widespread haplotypes could be parsimoniously explained by ILS (Muir and Schlötterer, 2005) within *L. cerina* and within *L. suginoana*. In cluster 1, mutations of haplotypes found in the Youluo and Ptlaman mountains specific to *L. cerina* may have generated haplotypes of *L. suginoana* and *L. thalassica* occurred in the same and different mountainous areas (Fig. S1). Moreover, haplotype K (cluster 1) was shared between *L. suginoana* (populations LSY and LSH) and *L. thalassica* (population LTH) (Table 1). Although both ILS and secondary gene flow can produce the pattern of haplotype sharing, recent secondary gene flow could be the main cause for the rare, unevenly distributed haplotype K sharing event. Additionally, the finding of only one rare, recent secondary admixture event may due in part to the small sample size collected in every mountainous area.

With 38 individuals of *L. cerina, L. suginoana,* and *L. thalassica*, eight selective primer combinations generated a total of 315 AFLP loci (Table S1). The proportion of polymorphic loci ranged from 72.1% (*L. thalassica*) to 81.9% (*L. cerina*) (Table 2). The species level $uH_E$ was ranged from 0.166 (*L. suginoana*) to 0.222 (*L. cerina*). Interestingly, a much larger number of private alleles was found for *L. cerina* ($N_P$ = 126) compared to that of *L. suginoana* ($N_P$ = 7) and of *L. thalassica* ($N_p$ = 6). Analysis with LMM showed overall $uH_E$ significant difference when compared among species (Wald $\chi^2$ = 34.41, $p$ < 0.0001). Pairwise significant differences were also found and *L. cerina* had significantly higher genetic diversity than that of *L. suginoana* and of *L. thalassica* (Table 3). Spearman's rank correlation test found no significant relationship between population sample size and population $uH_E$ (S = 169.18, $\rho$ = -0.025, $p$ = 0.945) and between population sample size and the level of polymorphism (S = 76, $\rho$ = 0.367, $p$ = 0.336), indicating that there is no correlation of the number of samples examined with the level of AFLP variation estimated. Nonetheless, population $uH_E$ estimated may be biased due to the small and unequal sample sizes (Isabel *et al*., 1999; Mariette *et al*., 2002). Owing to shared ancestral variation in homologous regions, closely related *Lecanorchis* species may show similar levels of nucleotide diversity based on the ITS1 sequences. However, genetic diversity is not a constant entity across the genome and hence we found significant differences in the levels of AFLP diversity in these closely related species.

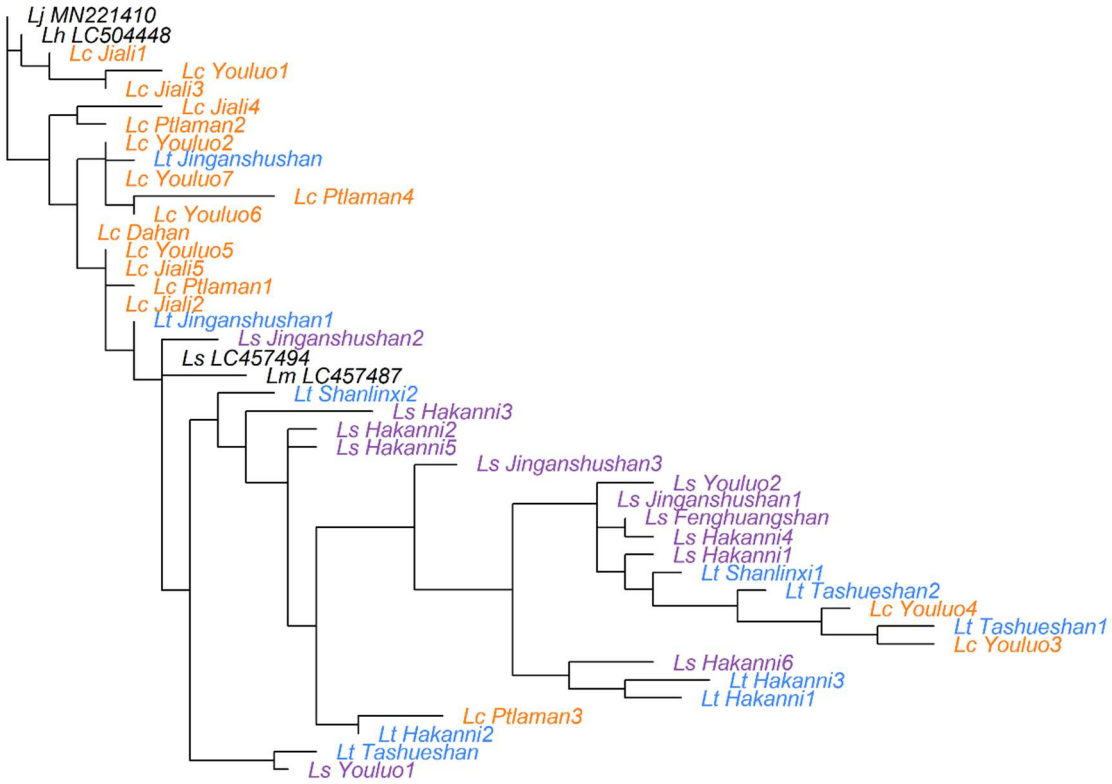


**Fig. 4.** The maximum parsimonious tree of individuals of three *Lecanorchis* species (*L. cerina*, *L. suginoana*, and *L. thalassica*) based on ITS1 DNA sequences. Tip labels for Taiwanese *Lecanorchis* individuals are colored: *L. cerina* (orange), *L. suginoana* (purple), and *L. thalassica* (blue). Japanese *Lecanorchis* species including *L. japonica*, *L. hokurikuensis*, *L. suginoana,* and *L. moritae* included in the analysis are colored in black. Low bootstrap support values (< 50%) were estimated and hence they are not displayed

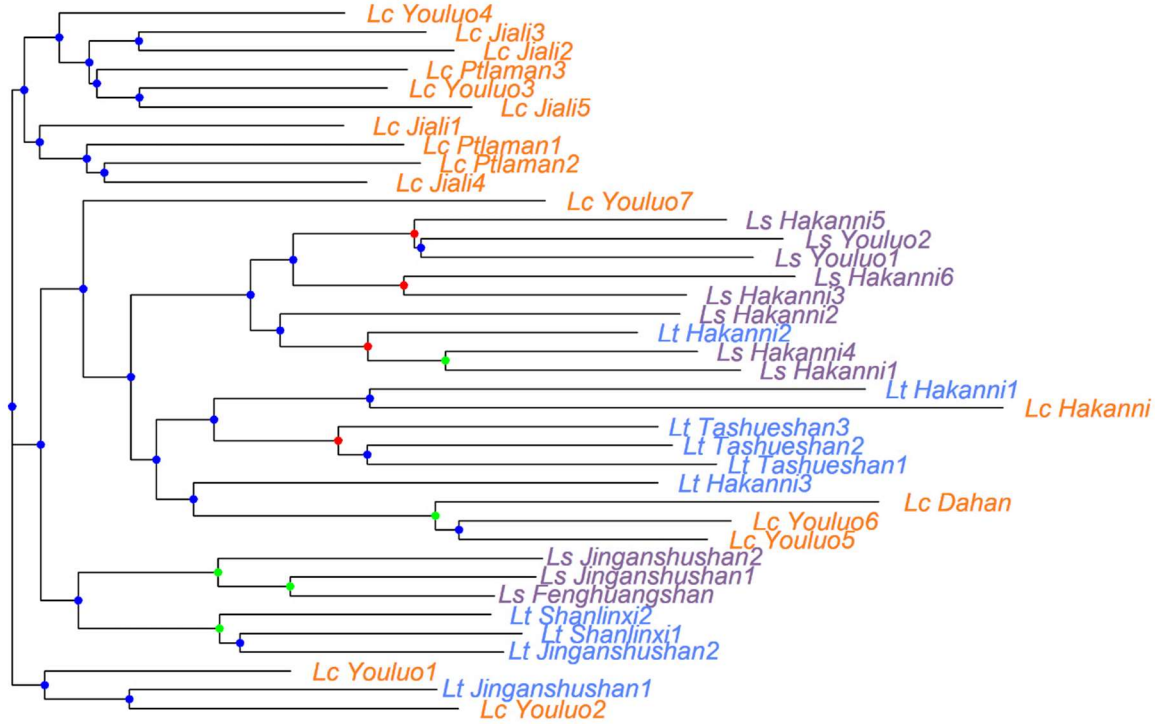


**Fig. 5.** Individual unrooted neighbor-joining tree generated based on AFLP variation. Tip labels for individuals are colored: *L. cerina* (orange), *L. suginoana* (purple), and *L. thalassica* (blue). Bootstrap values < 50% (blue), > 70% (green), and those range between 50% and 70% (red).

**Table 3.** Summary of Tukey's post-hoc pairwise population comparisons of the mean unbiased expected heterozygosity ($uH_E$) per locus using a linear mixed effect model. In linear mixed effect model, species was treated as a fixed factor and locus as a random factor based on AFLP variation of the three ***Lecanorchis*** species investigated.

| Contrast | Estimate | SE | *t* | *p* |
|---|---|---|---|---|
| LC - LS | 0.0758 | 0.0129 | 5.866 | < 0.0001 |
| LC - LT | 0.0376 | 0.0129 | 2.912 | 0.0104 |
| LS - LT | -0.0382 | 0.0129 | -2.954 | 0.0091 |

LC, *Lecanorchis cerina*; LS, *Lecanorchis suginoana*; LT, *Lecanorchis thalassica*; SE, standard error; *t*, *t* statistic

**Table 4.** Genetic differentiation between three ***Lecanorchis*** species (***Lecanorchis cerina***, ***L. suginoana***, *and* ***L. thalassica***) and between populations of these three species based on AFLP variation using analysis of molecular variance (AMOVA).

| Source of variation | df | Sum of square | Percent variation | *Φ* Statistics (*p*) |
|---|---|---|---|---|
| Between species | 2 | 182.33 | 8.10 | $\Phi_{CT}$ = 0.081 (0.0001) |
| Within species | 35 | 1531.93 | 91.90 | |
| Total | 37 | 1714.26 | 100 | |
| Between populations | 9 | 568.79 | 13.96 | $\Phi_{ST}$ = 0.140 (0.0001) |
| Within populations | 25 | 1017.96 | 86.04 | |
| Total | 34 | 1586.74 | 100 | |

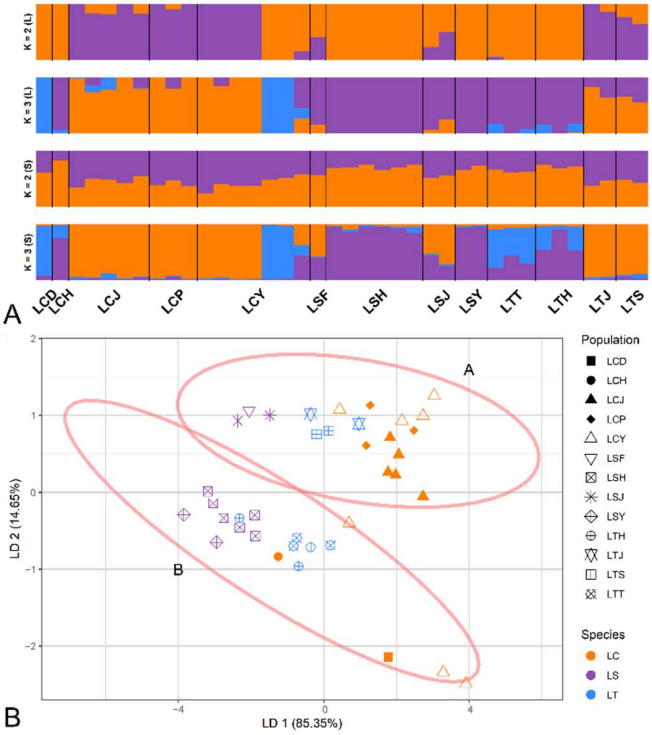


**Fig. 6.** Analysis of genetic homogeneous groups of 38 individuals of three closely related ***Lecanorchis*** species (***L. cerina***, ***L. suginoana***, and ***L. thalassica***) based on AFLP variation. **A.** Bar plots of LEA (L) and STRUCTURE (S) analyses. The clustering scenarios for *K* = 2–3 were displayed. **B.** Scatterplot of the first two linear discriminants (LDs) from discriminant analysis of principal components (DAPC). See Table 1 for population codes

Neutral theory predicts that genetic diversity will increase with a larger effective population size (Kimura, 1983) and positive relationship between population size and genetic diversity have been found in many wild plant species, including orchid species (e.g., Leimu *et al*., 2006; George *et al*., 2009; Jacquemyn *et al*., 2009; Szczecińska *et al*., 2016). A significantly larger AFLP diversity found in *L. cerina* suggests a larger historical population size in *L. cerina* than in *L. suginoana* and in *L. thalassica*.

AMOVA based on AFLP revealed low, but significant species differentiation ($\Phi_{CT}$ = 0.081, $p < 0.001$; Table 4). Significantly moderate level of differentiation was found between populations ($\Phi_{ST}$ = 0.140, $p < 0.001$). Pairwise species comparisons showed low levels of $F_{ST}$ between species (average $F_{ST}$ = 0.079; Table 5) and moderate level of population differentiation ($F_{ST}$ = 0.166). Although sample size may influence the estimation of genetic diversity, large number of loci compensate the biased estimation of population differentiation from small sample size (Isabel *et al*., 1999; Mariette *et al*., 2002). Nevertheless, we found contrasting pair $F_{ST}$ within and between species. $F_{ST}$ for neighboring population pair LCY and LSY was 0.122, higher than the $F_{ST}$ between population LCY and geographic distant populations of different species including populations LSJ ($F_{ST}$ = 0.070), LTJ ($F_{ST}$ = 0.068), and LTS ($F_{ST}$ = 0.106). However, comparison between population LSH and population LTH located in Hakanni mountain had an $F_{ST}$ = 0.043, which is smaller than the comparisons of the population LSH with geographic distant populations such as LSJ ($F_{ST}$ = 0.227), LTJ ($F_{ST}$ = 0.287), LTS ($F_{ST}$ = 0.301), and LTT ($F_{ST}$ = 0.208). Gene f low is expected to be higher resulting in lower levels of interspecific genetic differentiation between geographic neighbor populations than between distantly located populations (Petit and Excoffier, 2009). However, geographically distant populations may show lower genetic differentiation than geographically neighboring populations because of ILS (Muir and Schlötterer, 2005, Szövényi *et al*. 2008; Stenøien *et al*. 2011; Meleshko *et al*., 2021). The finding of two contrasting patterns of AFLP genetic differentiation between geographic neighbors relative to geographic distant populations suggest that both ILS and secondary gene flow may have occurred in these closely related species.

Both CE and BIC were minimized at *K* = 3 (Fig. S2) in the LEA and DAPC analyses based on AFLP variation (Fig. 6). However, DAPC revealed only two genetic clusters (Fig. 6B). Therefore, we interpreted the results based on *K* = 2. Both DAPC clusters comprised individuals of all three *Lecanorchis* species under investigation (Fig. 6B). Essentially, all individuals of *L. cerina, L. suginoana,* and *L. thalassica* shared genetic variation of orange and purple components based on the result of STRUCTURE (Fig. 6A). *L. cerina* comprised two genetic components (orange and purple components) and the population LCY contained individuals belonged to either component based on the

**Table 5.** Pairwise $F_{ST}$ (below diagonal) and *P* values (above diagonal) between three ***Lecanorchis*** species (***L. cerina***, ***L. suginoana***, and ***L. thalassica***) and between populations of these three ***Lecanorchis*** species based on AFLP variation using ARLEQUIN with 10,000 permutations. See Table 1 for population code. Populations with only one individual were excluded from between population comparisons.

**Between species**

| | LC | LS | LT |
|---|---|---|---|
| | | 0.004 | 0.000 |
| LS | 0.107 | | 0.008 |
| LT | 0.057 | 0.073 | |

**Between populations**

| | LCJ | LCP | LCY | LSH | LSJ | LSY | LTH | LTJ | LTS | LTT |
|---|---|---|---|---|---|---|---|---|---|---|
| LCJ | | 0.384 | 0.296 | 0.002 | 0.046 | 0.048 | 0.018 | 0.050 | 0.045 | 0.018 |
| LCP | 0.007 | | 0.183 | 0.011 | 0.096 | 0.102 | 0.106 | 0.094 | 0.098 | 0.102 |
| LCY | 0.013 | 0.025 | | 0.001 | 0.137 | 0.029 | 0.163 | 0.111 | 0.052 | 0.049 |
| LSH | 0.186 | 0.213 | 0.144 | | 0.041 | 0.180 | 0.055 | 0.036 | 0.026 | 0.012 |
| LSJ | 0.174 | 0.122 | 0.070 | 0.227 | | 0.335 | 0.102 | 0.334 | 0.332 | 0.097 |
| LSY | 0.207 | 0.235 | 0.122 | 0.040 | 0.222 | | 0.096 | 0.322 | 0.329 | 0.099 |
| LTH | 0.115 | 0.156 | 0.036 | 0.043 | 0.165 | 0.089 | | 0.111 | 0.099 | 0.097 |
| LTJ | 0.141 | 0.123 | 0.068 | 0.287 | 0.188 | 0.270 | 0.220 | | 0.340 | 0.096 |
| LTS | 0.168 | 0.169 | 0.106 | 0.301 | 0.250 | 0.303 | 0.230 | 0.086 | | 0.100 |
| LTT | 0.211 | 0.234 | 0.103 | 0.208 | 0.247 | 0.256 | 0.113 | 0.250 | 0.307 | |

result of LEA (Fig. 6A). The LEA result showed that *L. suginoana* individuals were mostly belonged to orange component. However, admixture of purple and orange components was observed in the only individual of the population LSF and one individual of the population LSJ. However, *L. thalassica* individuals can be distinguished into those belonging to orange (populations LTT and LTH) and those with high probability estimates belonging to purple component (populations LTJ and LTS) based on the LEA analysis.

Considering the widespread distribution of *L. cerina*, the higher AFLP diversity in *L. cerina*, and individuals of *L. suginoana* and *L. thalassica* nested within the *L. cerina* individuals in both the ITS1 parsimonious and the AFLP NJ trees (Figs. 4 and 5), *L. suginoana* and *L. thalassica* may have been derived from *L. cerina* in the same and different mountainous regions. Moreover, we observed ITS1 haplotype sharing within *L. cerina*, within *L. suginoana*, and between *L. suginoana* and *L. thalassica* (Table 1 and Fig. S1) as well as the contrasting patterns of higher or lower population divergence in geographic neighbors than in distantly located populations based on AFLP variation (Table 5). In addition to these results, the sharing and admixture of two AFLP genetic components observed in individuals of the three closely related *Lecanorchis* species (Fig. 6) also suggest that both ILS and secondary gene flow did occur in the *Lecanorchis* species examined. Disentangling of ILS from secondary gene flow in the interpretation of phylogenetic relationships between species is problematic if these species are diverged recently evolutionarily (Qu *et al*., 2011). Complex evolutionary history of plant evolution caused by secondary gene flow, radiation, and ILS are commonly seen (e.g., Muir and Schlötterer, 2005; Koopman *et al*., 2008; Szövényi *et al*., 2008; Goetze *et al*., 2017; Zhou *et al*., 2017). Descendant lineages are expected to share polymorphic alleles with the ancestral population for some time (i.e. incomplete lineage sorting). In contrast, lineages that have diverged completely may also share genetic polymorphism in parts of the genome due to secondary admixture (Gow *et al*., 2006). Incomplete sorting and secondary admixture produce similar genetic patterns, and the sharing of alleles in two lineages may thus be the result of either process, or both. Our results suggest that both ILS and secondary gene flow may have occurred in the *Lecanorchis* species examined. However, these evolutionary events may have occurred at different temporal scale with the ILS process occurred due to a possibility of rapid radiation and followed by secondary admixture at a later time. Nonetheless, further study using larger sample size, low-copy nuclear genes, and even next generation sequencing such as restriction site-associated sequencing is required in order to gain more insights into the evolution of the three closely related species investigated. Although we observed suspected morphological hybrid between the three *Lecanorchis* species examined, future investigation of the association of genotype with morphology can be helpful for a better understanding of the evolution of these species.

## ACKNOWLEDGMENTS

We are most grateful for the assistance of Da-Ming Huang in collecting samples from Ptlaman and Youluo mountains and of Kuo-Chu Yeuh in collecting samples from Fenghuangshan, Jinganshushan, and Tashueshan mountains. This work was partly supported by the Taiwan Ministry of Science and Technology under grant number of MOST109-2313-B-003-001-MY3.

## LITERATURE CITED


**Álvarez, I., J.F. Wendel.** 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. **29(3):** 417–434.

**Bates, D., M. Maechler, B. Bolker and S. Walker.** 2015. Fitting linear mixed effects models using lme4. J. Stat. Soft. **67(1):** 1–48.

**Bodenhofer, U., E. Bonatesta, C. Horejš-Kainrath and S. Hochreiter.** 2015. msa: an R package for multiple sequence alignment. Bioinformatics **31(24):** 3997–3999.

**Bonin, A., E. Bellemain, E.P. Bronken, F. Pompanon, C. Brochmann and P. Taberlet.** 2004. How to track and assess genotyping errors in population genetics studies. Mol. Ecol. **13(11)**: 3261–3273.

**Clement, M.J., Q. Snell, P. Walke, D. Posada and K.A. Crandall.** 2002. TCS: estimating gene genealogies. Proc. 16th Int. Parallel Distrib. Process Symp. **2:** 184.

**Csilléry, K., M.G.B. Blum, O.E. Gaggiotti and O. François.** 2010. Approximate Bayesian computation (ABC) in practice. Trends Ecol. Evol. **25(7):** 410–418.

**Doyle, J.J. and J.L. Doyle.** 1987. A rapid DNA isolation procedure from small quantities of fresh leaf material. Phytochem. Bull. **19(1):** 11–15.

**Dray, S. and A.-B. Dufour.** 2007. The ade4 package: implementing the duality diagram for ecologists. J. Stat. Soft. **22(4):** 1–20.

**Ebersbach, J., J. Schnitzler, A. Favre and A.N. Muellner-Riehl.** 2017. Evolutionary radiations in the species-rich mountain genus *Saxifraga* L. BMC Evol. Biol. **17(1):** 119.

**Evanno, G., S. Regnaut and J. Goudet.** 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. **14(8):** 2611–2620.

**Excoffier, L. and H.E. Lischer.** 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. **10(3):** 564–567.

**Farminhão, J.N.M., S. Verlynde, E. Kaymak, V. Droissart, M. Simo-Droissart, G. Collobert, F. Martos and T. Stévart.** 2021. Rapid radiation of angraecoids (Orchidaceae, Angraecinae) in tropical Africa characterized by multiple karyotypic shifts under major environmental instability. Mol. Phylogenet. Evol. **159:** 107105.

**Farris, J.S.** 1989. The retention index and the rescaled consistency index. Cladistics **5(4):** 417–419.

**Felsenstein, J.** 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution **39(4):** 783–791.

**Fox, J. and S. Weisberg.** 2011. An R Companion to Applied Regression, 2nd ed. Sage, Thousand Oaks, USA.

**Frichot, E. and O. Francois.** 2015. LEA: an R package for landscape and ecological association studies. Methods Ecol. Evol. **6(8):** 925–929.

**Funk, D.J. and K.E. Omland.** 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Ann. Rev. Ecol. Evol. and Syst. **34(1):** 397–423.

**George, S., J. Sharma and V.L. Yadon.** 2009. Genetic diversity of the endangered and narrow endemic *Piperia yadonii* (Orchidaceae) assessed with ISSR polymorphisms. Amer. J. Bot. **96(11):** 2022–2030.

**Goetze, M., C.M. Zanella, C. Palma-Silva, M.V. Büttow and F. Bered.** 2017. Incomplete lineage sorting and hybridization in the evolutionary history of closely related, endemic yellow-flowered *Aechmea* species of subgenus *Ortgiesia* (Bromeliaceae). Amer. J. Bot. **104(7):** 1073–1087.

**Gow, J.L., C.L. Peichel and E.B. Taylor.** 2006. Contrasting hybridization rates between sympatric three spine sticklebacks highlight the fragility of reproductive barriers between evolutionarily young species. Mol. Ecol. **15(3):** 739–752.

**Hey, J.** 2010. Isolation with migration models for more than two populations. Mol. Biol. Evol. **27(4):** 905–920.

**Hu, A.-Q., S.W. Gale, Z.-J. Liu, S. Suddee, T.-C. Hsu, G.A. Fischer and R.M.K. Saunders.** 2020. Molecular phylogenetics and floral evolution of the *Cirrhopetalum* alliance (*Bulbophyllum*, Orchidaceae): Evolutionary transitions and phylogenetic signal variation. Mol. Phylogen. Evol. **143:** 106689.

**Isabel, N., J. Beaulieu, P. Thériault and J. Bousquet.** 1999. Direct evidence for biased gene diversity estimates from dominant random amplified polymorphic DNA (RAPD) fingerprints. Mol. Ecol. **8(3):** 477–483.

**Jacquemyn, H., R. Brys, D. Adriaens, O. Honnay and I. Roldán-Ruiz.** 2009. Effects of population size and forest management on genetic diversity and structure of the tuberous orchid *Orchis mascula*. Conserv. Genet. **10(1):** 161–168.

**Jombart, T. and I. Ahmed.** 2011. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics **27(21):** 3070–3071.

**Jombart, T., S. Devillard and F. Balloux.** 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. **11(1):** 94.

**Kamvar, Z.N., J.C. Brooks and N.J. Grünwald.** 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Front. Genet. **6:** 208.

**Kimura, M. 1983.** The Neutral Theory of Molecular Evolution. Cambridge Univ. Press, Cambridge, UK.

**Kluge, A.G. and J.S. Farris.** 1969. Quantitative phyletics and the evolution of Anurans. Syst. Zool. **18(1):** 1–32.

**Koopman, W.J.M., V. Wissemann, K. de Cock, J. van Huylenbroeck, J. de Riek, G.J.H Sabatino, D. Visser, B. Vosman, C.M. Ritz, B. Maes, G. Werlemark, H. Nybom, T. Debener, M. Linde and J.M. Smulders.** 2008. AFLP markers as a tool to reconstruct complex relationships: A case study in *Rosa* (Rosaceae). Amer. J. Bot. **95(3):** 353–366.

**Leigh, J.W. and D. Bryant.** 2015. PopART: full-feature software for haplotype network construction. Methods Ecol. Evol. **6(9):** 1110–1116.

**Leimu, R., P. Mutikainen, J. Koricheva and M. Fischer.** 2006. How general are positive relationships between plant population size, fitness and genetic variation? J. Ecol. **94(5):** 942–952.

**Lenth, R.** 2018. emmeans: Estimated marginal means, aka least-squares means. R package version 1.3.1. https://CRAN.R-project.org/package=emmeans. Accessed Apr 7. 2018.

**Lin, T.P.** 2019. The Orchid Flora of Taiwan, a collection of line drawings. NTU Press, Taipei, Taiwan. 1012 pages.

**Linder, C.R. and L.H. Rieseberg.** 2004. Reconstructing patterns of reticulate evolution in plants. Amer. J. Bot. **91(10):** 1700–1708.

**Mariette, S., V. Le Corre, F. Austerlitz and Kremer, A.** 2002. Sampling within the genome for measuring within-population diversity: trade-offs between markers. Mol. Ecol. **11(7):** 1145–1156.

**Meleshko, O., M.D. Martin, T.S. Korneliussen, C. Schröck, P. Lamkowski, J. Schmutz, A. Healey, B.T. Piakowski, A.J. Shaw, D.J. Weston, K.I. Flatberg, P. Szövényi, K. Hassel and H.K. Stenøien.** 2021. Extensive genome-wide phylogenetic discordance is due to incomplete lineage sorting and not ongoing introgression in a rapidly radiated bryophyte genus. Mol. Biol. Evol. **38(7):** 2750–2766.

**Muir, G. and C. Schlötterer.** 2005. Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.) Mol. Ecol. **14(2):** 549–561.

**Nei, M.** 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89(3):** 583–590.

**Nei, M.** 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, USA. 512 pp.

**Nielsen, R. and J. Wakeley.** 2001. Distinguishing migration from isolation. A Markov chain Monte Carlo approach. Genetics **158(2):** 885–896.

**Paradis, E. and K. Schliep.** 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics **35(3):** 526–528.

**Pease, J.B., D.C. Haak, M.W. Hahn and L.C. Moyle.** 2016. Phylogenomics reveals three sources of adaptive variation during a rapid radiation. PLoS Biol. **14(2):** e1002379.

**Pérez-Excobar, O.A., G. Chomicki, F.L. Condamine, A.P. Karremans, D. Bogarin, N.J. Matxke, D. Silvestro and A. Antonelli.** 2017. Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. New Phytol. **215(2):** 891–905.

**Petit, R.J. and L. Excoffier.** 2009. Gene flow and species delimitation. Trends Ecol. Evol. **24(7):** 386–393.

**Pritchard, J.K., M. Stephens and P. Donnelly.** 2000. Inference of population structure using multilocus genotype data. Genetics **155(2):** 945–959.

**Qu, Y.H., X. Luo, R.Y. Zhang, G. Song, F.S. Zou and F.M. Lei.** 2011. Lineage diversification and historical demography of a montane bird *Garrulax elliotii*—implications for the Pleistocene evolutionary history of the eastern Himalayas. BMC Evol. Biol. **11(1):** 174.

**R Development Core Team.** 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

**Rokas, A. and S.B. Carroll.** 2006. Bushes in the tree of life. PLoS Biol. **4(11):** e352.

**Rozas, J., A. Ferrer-Mata, J.C. Sánchez-DelBarrio, S. Guirao-Rico, P. Librado, S.E. Ramos-Onsins and A. Sánchez-Gracia.** 2017. DnaSP 6: DNA sequence polymorphism analysis of large datasets. Mol. Biol. Evol. **34(12):** 3299–3302.

**Schliep, K.P.** 2011. phangorn: phylogenetic analysis in R. Bioinformatics **27(4):** 592–593.

**Stenøien, H.K., A.J. Shaw, B. Shaw, K. Hassel and U. Gunnarsson.** 2011. North American origin and recent European establishments of the amphi-Atlantic peat moss *Sphagnum angermanicum*. Evolution **65(4):**1181–1194.

**Suetsugu, K., S. Kaida, H. Fukunaga and S. Sawa.** 2020. A new form of the mycoheterotrophic plant *Lecanorchis japonica* (Orchidaceae) from Japan. Acta Phytotax. Geobot. **71(3):** 243–248.

**Suetsugu, K., S. Kaida, T.-C. Hsu and S. Sawa.** 2019. *Lecanorchis moritae* (Orchidaceae, Vanilloideae), a new mycoheterotrophic species from Amamai-Oshima Island, Japan based on morphological and molecular data. Phytotaxa **404(4):** 137–145.

**Suetsugu, K., C. Shimaoka, H. Fukunaga and S. Sawa**. 2018. The taxonomic identity of three varieties of *Lecanorchis nigricans* (Vanilleae, Vanilloideae, Orchidaceae) in Japan. PhytoKeys **92:** 17–35.

**Sunnåker, M., A.G. Busetto, E. Numminen, J. Corander, M. Foll and C. Dessimoz.** 2013. Approximate Bayesian computation. PLoS Comput. Biol. **9(1):** e1002803.

**Szczecińska, M., G. Sramkó, K. Wolosz and J. Sawwicki.** 2016. Genetic diversity and population structure of the rare and endangered plant species *Pulsatilla patens* (L.) Mill in East Central Europe. PLoS One **11(3):** e0151730.

**Szövényi, P., S. Terracciano, M. Ricca, S. Giordano and A.J. Shaw.** 2008. Recent divergence, intercontinental dispersal and shared polymorphism are shaping the genetic structure of amphi-Atlantic peatmoss populations. Mol Ecol. **17(24):** 5364–5377.

**Takashima, M., Y. Yamashita, A. Kinoshita, M. Ohta and T. Yukawa.** 2019. Taxonomic status of *Lecanorchis suginoana* (Tuyama) Seriz. (Orchidaceae). Bull. Natl. Mus. Nat. Sci. Ser. B. **45(3):** 119–129.

**Thompson, J.D., D.G. Higgins and T.J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. **22(22):** 4673–4680.

**Vekemans, X., T. Beauwens, M. Lemaire and I. Roldán-Ruiz.** 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. Mol. Ecol. **11(1):** 139–151.

**Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van der Lee, M. Hornes, A. Friters, J. Pot, J. Paleman, M. Kuiper and M. Zabeau.** 1995. AFLP: a new technique for DNA fingerprinting. Nucl. Acids Res. **23(21):** 4407–4414.

**WCVP.** 2022, World Checklist of Vascular Plants, version 2.0. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; http://wcvp.science.kew.org. Accessed 11 Mar. 2022.

**Zhivotovsky, L.A.** 1999. Estimating population structure in diploids with multilocus dominant DNA markers. Mol. Ecol. **8(6):** 907–913.

**Zhou, Y., L. Duvaux, G. Ren, L. Zhang, O. Savolainen and J. Liu.** 2017. Importance of incomplete sorting and introgression in the origin of shared genetic variation between two closely related pines with overlapping distributions. Heredity **118(3):** 211–220.

**Supplementary materials are available from Journal Website.**