Taiwania 66(1): 20–30, **2021**
DOI: 10.6165/tai.2021.66.20

Genetic diversity, population genetic structure and conservation strategies for *Pleione formosana* (Orchideace)

Wei-Chun CHAO[1], Yea-Chen LIU[2], Ming-Tao JIANG[3], Sha-Sha WU[3], Chun-Li FANG[4], Jia-Fang HO[4], Chun-Lin HUANG[4,*]

1. Department of Forestry and Natural Resources, National Chiayi University, Chiayi.
2. Department of Biological Resources, National Chiayi University, Chiayi.
3. Key Laboratory of National Forestry and Grassland Administration for Orchid Conservation and Utilization at College of Landscape Architecture, Fujian Agriculture and Forestry University, Fuzhou.
4. Laboratory of Molecular Phylogenetics, Department of Biology, National Museum of Natural Science, Taichung.
[*]Corresponding author's email: wagtail@mail.nmns.edu.tw

(Manuscript received 7 October 2020; Accepted 13 December 2020; Online published 3 January 2021)

ABSTRACT: A population genomic approach was employed to investigate the diversity within species to create a more robust, lineage-specific conservation strategy for an endangered orchid. *Pleione formosana* is a species native to southeastern China and Taiwan, where it is distributed at an altitude range of 1,200–2,500 m the foggy mountain area and grows in mosses on half-shaded rocks or tree trunks. To identify whether the level of genetic diversity in the species, we used genotyping by sequencing (GBS) to analysis the sub-populations of the species. Fifty-eight individuals of *P. formosana* were sampled from a total of nine populations within continental island Taiwan and three populations in China as outgroup. Treatment of all samples involved five major steps: sample preparation, library assembly, sequencing, SNP calling and diversity analysis. GBS markers confirmed the China outgroup as distinct, and provided resolution of two clusters of population genetic structure in Taiwan. Outliers provide higher genetic differentiation, and some GBS tags associated with climatic factor were found. Genomic diversity identified among the three clusters suggests that conservation of this species will be best served by considering them as three evolutionary significant units (ESUs). This approach will maximize evolutionary potential among all species during increased isolation and environmental change. According to the genetic consequences, restoration strategies should be carried out in all populations to preserve genetic diversity and evolutionary potential for different environmental factors.

KEY WORDS: Genotyping by sequencing (GBS), next-generation sequencing, Orchidaceae, *Pleione formosana*, population genomics.

INTRODUCTION

The field of conservation genetics is an evolving discipline that applies evolutionary and molecular genetics to the conservation of threatened species and biodiversity (Frankham, 2010). Two principal evolutionary processes influence the distribution of genetic variation among populations. Genetic drift operates randomly and generally has no phenotypic effect. In contrast, natural selection ignores neutral alleles, and favors certain alleles that increase fitness (Frankham *et al.,* 2004). Neutral genetic variation has great potential for investigating gene flow, migration and demographics. Identification of locally adaptive genetic variation can have direct utility for conserving species at risk, especially when management may include action such as translocations for restoration, or genetic rescue (Flanagan *et al.,* 2018).

The use of evolutionary significant units (ESUs) provides an alternative categorization of biological entities (e.g. populations or subspecies) that represent unique genetic diversity for conservation (Moritz 1994). The comprehensive ESUs framework specifically designed to delineate should be based on both neutral and adaptive genetic divergence (Fraser and Bernatchez, 2001, Gebremedhin *et al.,* 2009). Climatic factors such as temperature and precipitation greatly affect the condition of plant communities. Meteorological data may be applied by collecting weather station data or be making use of relevant GIS layers (Hijmans *et al.,* 2005). Next Generation Sequencing (NGS) coupled with advances in computer hardware and software in the field of genomics, have allowed the development of the new genotype–environment association (GEA) methods for comprehensive evaluation of adaptive diversity (Rellstab *et al.,* 2015).

The genus *Pleione* D. Don comprises about 20 species of terrestrial, lithophytic or epiphytic orchids, growing on rocks and mossy substrates of subtropical montane forests (Cribb and Butterfield, 1999). *Pleione formosana* belong to the *P. bulbocodioides* complex, has been the source of persistent confusion and ambiguity due to the close relationships of its taxa as indicated by the small degree of morphological variation, karyotypes and ease of hybridization (Gravendeel *et al.,* 2004). So far, the complete chloroplast genomes of species in the *P. bulbocodioides* complex, including *P. bulbocodioides* (Shi *et al.,* 2018) and *P. formosana* (Jiang *et al.,* 2019), have been annotated. However, population genetics studies are still scarce for *P. formosana*. Reduced-representation

Table 1. Sampling localities chosen for the population study of Pleione formosana accompanied by elevation and sample size.

Acronym	Population	Region	Longitude	Latitude	Elevation (m)	Sample size
AL	Alishan	Taiwan	120.799	23.540	2,346	6
BD	Beidawushan	Taiwan	120.738	22.615	2,120	6
NH	Nanheng	Taiwan	120.918	23.284	2,166	6
CL	Chilanshan	Taiwan	121.448	24.599	1,539	4
DX	Daxueshan	Taiwan	120.986	24.269	2,138	6
HP	Hepinglindao	Taiwan	121.677	24.301	1,482	6
JL	Jialishan	Taiwan	121.013	24.514	2,048	5
RY	Rueiyansi	Taiwan	121.240	24.127	2,288	6
TP	Taipinshan	Taiwan	121.518	24.524	1,511	4
BYS	Baiyunshan	China	119.492	27.164	1,355	3
WYS	Wuyishan	China	117.646	27.736	1,064	3
YSM	Yangshimu	China	114.237	27.559	1,606	3

molecular techniques are particularly useful in that they allow low-cost genome typing of hundreds or thousands of individuals [e.g. genotyping by sequencing (GBS), (Elshire *et al.,* 2011); double digest restriction site-associated DNA sequencing (ddRADseq), (Peterson *et al.,* 2012). Through the use of thousands of markers, researchers and practitioners have the ability to uncover fine-scale structure of a conservation unit regardless of current taxonomic classification.

P. formosana is a species native to southeastern China and Taiwan. It is distributed in mountainous regions at an altitude range of 1,200–2,500 m, where moisture introduced by clouds is more effectively retained, and grows in mosses on half-shaded rocks or tree trunks. Occasionally, some orchids and mosses are accompanied by soil caving, and new colonies can form. This species belongs to the early succession species, so moderate environmental disturbances are beneficial to its survival (Su, 1988). In terms of breeding strategy, *P. formosana* is a self-compatibility species and typically uses a food-deceptive pollination mechanism (Chen *et al.,* 2019). However, its natural seed setting rate is extremely low and most individuals use asexual reproduction to expand the population (Chao *et al.,* 2017). Habitat destruction and limited reproductive abilities of *P. formosana* have endangered this species. For the purpose of protecting *P. formosana*, the government established Taiwan Pleione Nature Reserve, where any alteration of or damage to the natural status is prohibited in order to protect the wild population.

The fragmentation and reduction of the effective population size can have a direct or indirect impact on the evolutionary processes, causing a loss of genetic diversity, changes in intra and inter-population genetic structures and genetic drift. One of the most significant threats to endangered species is the loss of genetic diversity associated with small and fragmented populations. In this research, we assess GBS in the non-model orchid species, *P. formosana*, to develop high quantity of informative single nucleotide polymorphisms (SNP). The specific objectives of this research are to:

- Identify whether *P. formosana* populations have a decreased level of genetic diversity cause by the habitat fragmentation.
- Determine whether there is significant genetic differentiation between populations of *P. formosana.*
- Determine the contribution of natural selection and molecular genetic adaptation to shaping genome-wide variation. We expect higher genetic population differentiation for adaptive SNP than for neutral SNP if natural selection is the principal source of differentiation. Alternatively, if other sources of differentiation (mutation, genetic drift and migration) are relevant, they will equally affect both types of SNP.
- Consider the management implications of the findings, in terms of recommendations for future restoration projects.

MATERIALS AND METHODS

Plant collection & DNA extraction

The materials of *Pleione formosana* were collected during growth season of 2016–2017 from 12 localities in Taiwan and China (Fig. 1, Table 1). The leaf samples collected from the field were preserved in a freezer at -80°C. Total genomic DNA was extracted from leaves using a modified CTAB method (Doyle and Doyle, 1987).

Environmental data collection

To consider the wide distribution of this species, across many different ecological gradients, we obtained 19 climatic variables at 30-s spatial resolution (~1 km) from the WorldClim averaged for the 1950 - 2000 period (Hijmans *et al.,* 2005). Highly correlated environmental variables with $|r| > 0.8$ were identified using a principal component analysis (PCA) (Bothwell *et al.,* 2013). Five variables was selected with the least correlation among them for the studied species range area, including BIO5-Max Temperature of Warmest Month, BIO7-Temperature Annual Range, BIO8-Mean Temperature of Wettest Quarter, BIO15-Precipitation Seasonality (Coefficient of Variation) and BIO19- Precipitation of Coldest Quarter.

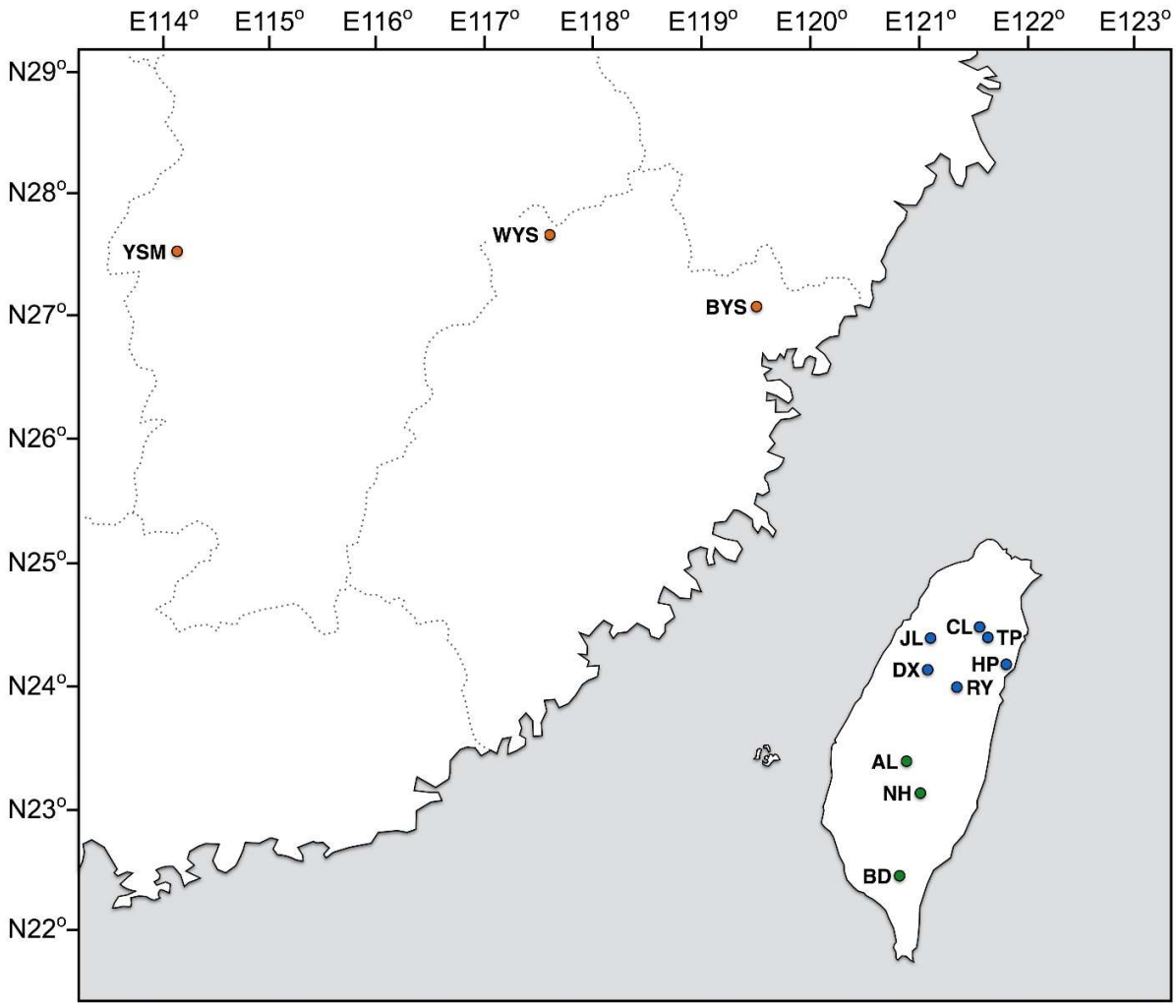

Fig. 1. Sample localities of 12 populations occurring in China and Taiwan. Symbol colors represent the three genetic clusters revealed by the results from STRUCTURE (K = 3)(Fig. 3). See Table 1 for population code acronym.

SNP discovery by GBS

GBS involves five major steps: sample preparation, library assembly, sequencing, SNP calling and diversity analysis. The total amount of extracted DNA was quantified by spectrophotometry using Qbit a Nanodrop 1000 (Thermo Scientific) and integrity verified on Agarose gel (1%). GBS was performed at Beijing Genomics Institute (Beijing, China) as described by (Elshire *et al.,* 2011). Genome complexity was reduced by digesting individual sample genomic DNA with *Ape*KI. The resultant fragments from all samples were directly ligated to a pair of enzyme-specific adapters, and were combined into pools. PCR amplicons between 250 and 600 bp were extracted from an agarose gel and sequenced in an Illumina HiSeq X ten using a 150 bp Paired End protocol (Illumina Inc., USA). Demultiplexing was performed using the barcode sequence, and adapter sequence removal and sequence quality trimming were performed. Candidate single nucleotide polymorphisms (SNPs) were identified in TASSEL using the UNEAK pipeline with default settings, including a minimum coverage per tag of 5, error tolerance of 0.03, maximum frequency of 0.5 and a minimum frequency of 0.05 for minor alleles (Glaubitz *et al.,* 2014; Lu *et al.,* 2013). The unique GBS sequence read that was at least 64-bp long was aligned against GenBank sequence using BLAST tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Detection of Selection Footprints

To identify adaptive SNP (putative loci under selection), we used BayeScan 2.1 (Foll and Gaggiotti 2008) for estimating the posterior probability that a given locus is affected by selection. Briefly, prior odds of 100 were used for identifying the top candidates of the selected loci and a total of 50,000 reversible-jump Markov Chain Monte Carlo chains were run with a thinning interval of 10, following 20 pilot runs of 5,000 iterations each and a burn-in length of 50,000. Loci were considered outliers with an FDR of 0.05.

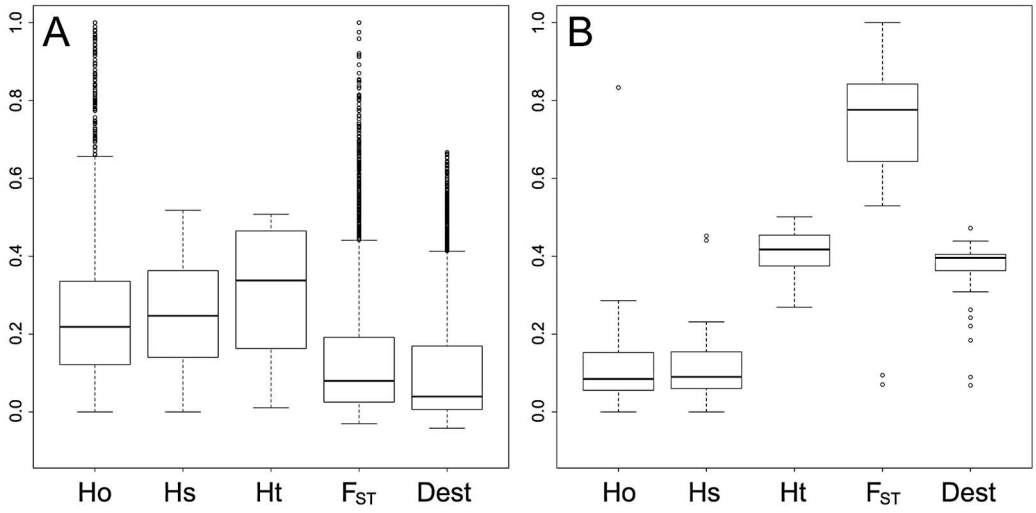

Fig. 2. Summary statistics of genetic variation existing in *Pleione formosana* estimated by a total of 3,737 SNPs (**A**) or 80 adaptive SNPs (**B**). Ho: heterozygosity within population; Hs: genetic diversity within population; Ht: overall gene diversity; F_{ST}: fixation index; Dest: measure of population differentiation.

Because the presence of admixed individuals may impact the power of BayeScan, we also identified candidate SNPs using PCAadpt v2.2.1 (Luu *et al.,* 2017). The number of principal components (K) was established by running the PCA analysis in PCAdapt and then identifies loci that are atypical related to a specific principal component as measured by the latent factors. The Mahalanobis distances were used as the statistic test to look for outlier SNPs, and transformed into p-values to perform multiple hypothesis testing. Finally, the cut-off for detecting outliers was chosen based on the q-value procedure implemented in the qvalue package in R (Storey *et al.,* 2019), using 0.01% as a false discovery rate threshold.

To confirm the adaptive SNP detected by the previous method, the spatial analysis method (SAM) implemented in the program Samβada v0.5.1 (Joost *et al.,* 2007) was used. We conducted the analysis using the 3,737 SNP detected for *P. formosana*. Samβada uses logistic regressions to model the probability of presence of an allelic variant in a polymorphic marker given the environmental conditions of the sampling locations. Five environmental variables previously described were used. Regarding genotypes, each of the states of a given SNP is considered independently as binary presence/absence in each sample. Our biallelic SNPs were recoded as three distinct genotypes (AA, AB, and BB). A maximum likelihood approach is used to fit the models using univariate analyses. Each model for a given genotype was compared to a constant model, where the probability of presence of the genotype was the same at each location. The statistical significance threshold was set to 1% before applying Bonferroni correction. Significance was assessed with log likelihood ratio (G) tests (Dobson and Barnett, 2008) selecting loci/allele that tested higher than the 99th percentile of the G score distribution.

Basic Statistics of Genetic Variation

We made all estimations in parallel with neutral SNP and adaptive SNP. All the results were obtained using the adegenet (Jombart and Ahmed, 2011) and hierFstat (Goudet, 2005) R packages. Basic statistics were estimated including observed heterozygosity (Ho) within population, genetic diversity (Hs) within population and overall gene diversity (Ht). F_{ST} was assessed following Nei 1987. Dest, a measure of population differentiation as defined by Jost 2008, was also calculated. The degree of genetic differentiation among populations is expected to be low for neutral SNP but highly divergent in SNP subject to directional selection. Between population F_{ST} values were calculated with the function "pairwise.WCfst" and significance was tested using 1,000 bootstrap replicates with the function "boot.ppfst". The value of *Nm* can be estimated from F_{ST} by inverting Wright's formula (Wright, 1951), as $Nm \approx (1- F_{ST})/ 4 F_{ST}$, where N is the effective population size of each population and m is the migration rate between populations. This method is effective for estimating gene flow indirectly.

Population Structure

To identify genetically homogeneous groups within *P. formosana*, both parametric and non-parametric approaches were employed. Population structure was determined using the parametric Bayesian model-based clustering method implemented in STRUCTURE v.2.3.4 (Pritchard *et al.,* 2000). For each K (from 1 to 12), ten independent runs were performed applying the admixture model (INFERALPHA = 1), with allele frequencies correlated for SNP markers (FREQSCOR =1), 100,000

Table 2. Adaptive SNP found under selection in both BayeScan and PCAdapt and their significant associations with environmental variables using Samβada.

GBS tag	BayeScan	PCAdapt	Samβada				
		p-value	BIO5	BIO7	BIO8	BIO15	BIO19
TP19155	x	1.08E-04				x	
TP20308	x	1.60E-05					x
TP28006	x	2.96E-04				x	
TP116035	x	3.45E-05					x
TP222607	x	4.39E-04				x	
TP283155	x	1.37E-03					x
TP402610	x	1.48E-06					x
TP407629	x	3.00E-04					x
TP437690	x	6.98E-05					x

BIO5: Max Temperature of Warmest Month; BIO7: Temperature Annual Range; BIO8: Mean Temperature of Wettest Quarter; BIO15: Precipitation Seasonality (Coefficient of Variation); BIO19: Precipitation of Coldest Quarter.

Markov Chain Monte Carlo (MCMC) repetitions, 100,000 burn-in period and RANDOMIZE = 1. The optimal K value was determined by use of the ad-hoc statistic ΔK (Evanno *et al.,* 2005) to evaluate the fit of different clustering scenarios.

To get a simplified picture of the genetic diversity observed among individuals or populations we used Principal Component Analysis (PCA). DAPC function was used to describe the relationships between these clusters. DAPC analysis was performed using the "dapc" function of the adegenet package in R (Jombart and Ahmed, 2011, R Development Core Team, 2015) by retaining principal components explaining >90 % of the variation of data with population information incorporated and depicted in an ordination plot representing the first two linear discriminant axes. Plotting discriminant functions with a minimum-spanning tree connecting the cluster centroids allows a visual representation of affinities between clusters.

Phylogenetic analysis

After the SNP detection, the individual SNPs were used to calculate the distance among individuals. Phylogenetic analysis was performed using the "aboot" function of the poppr package in R (Kamvar *et al.,* 2014, R Development Core Team, 2015), and on this basis, a tree was constructed by neighbor-joining and the bootstrap values were calculated 1,000 times.

RESULTS

Assessment of P. formosana genotypes sampled by GBS using ApeKI

We obtained more than 72 billion bases out of 746,502,750 reads and an average 97.58 % reads keep score $Q \geq 20$ from the samples. The average of 12.87 million sequence reads per sample was obtained (range from 1.16 to 34.85 million). After filtering out alleles with "NN" (unassigned) genotypes, the dataset consisted of 58 individuals with 3,737 binary SNPs. This dataset was subjected to further analysis.

Putative under selection SNP (Adaptive SNP)

Using the BayScan approach, 291 outlier SNPs (7.79%) were identified. We detected 307 outlier SNPs (8.22%) in PCAdapt and there were 80 overlap cases between the loci reported by each method. The adaptive SNPs detected by both approaches were limited, possibly reflecting discrepancies in their methodologies. We define the 80 consensus results as the adaptive SNP dataset. Using Samβada we further detected nine significant genotypes associated with precipitation gradients (Table 2). In addition, the twelve BLAST hits in the GenBank for 80 adaptive SNP-bearing tags are in Table 3.

Genome-wide genetic differentiation and variation

The entire population was divided into three genetic clusters (China, North Taiwan and South Taiwan) with a distinct maximum ΔK (ΔK = 435.369 at K = 3, Fig. 3A). We estimated ΔK as the mean of the absolute values of $L''(K)$ averaged over 10 runs divided by the standard deviation of $L(K)$, $\Delta K = m(|L''(K)|)/s[L(K)]$. In this study, the runs at K = 10 to 12 didn't converge, which brought the mean of $L(K)$ down. As ΔK is an absolute value of a second order rate of change, it is sensitive to this type of data. The second-best ΔK = 9 didn't reflect the real population structure. DAPC also provided clear distinction of *P. formosana* populations into three distinctive clusters including cluster China (populations BYS, WYS, YSM), clusterSouth Taiwan (populations AL, BD, NH) and cluster North Taiwan (populations CL, DX, HP, JL, RY, TP)(Fig. 3B). The first linear discriminant axis distinguished the China from Taiwan populations, and the second linear discriminant axis further separated the Taiwan populations into north and south clusters (Fig. 4). The same result of clusters was also obtained using NJ tree based on 3,737 SNP loci (Fig. 5). Fig. 5 also shows that the genetic differentiation was separate for each population in Taiwan.

Using a total 3,737 SNPs (neutral plus outlier), different basic statistics of genetic variation of *P. formosana* shows low to medium genetic diversity level and population differentiation (Fig. 2, Table 4). With 80 adaptive SNPs, the indicators based on heterozygosity and genetic diversity within population decrease, and F statistics show clear differences among populations (Fig. 2). The F_{ST} of a total 3,737 SNPs is 0.245. When only the 80 adaptive SNPs are included, the F_{ST} increase to 0.736.

The pairwise F_{ST} values among populations across different clusters show moderate to strong population structure. The values ranged from 0.168 between populations in the South Taiwan and North Taiwan clusters to 0.407 between populations in the North Taiwan

Table 3. Best BLAST hits in the GenBank for SNP-bearing tags.

GBS tag	GenBank no.	Predicted gene	Sequence
TP20312	XM_020717771.1	*Phalaenopsis equestris* uncharacterized LOC110019902 (LOC110019902)	CAGCAAGGGCAGAAACAAAAT(A/C)GCCACCAGA TGAAGCCCCAAGGGCTACAACAGGGAGCTTCTC
TP51373	XM_020724827.1	*Phalaenopsis equestris* mediator of RNA polymerase II transcription subunit 30 (LOC110024709)	CAGCACGCTCTTCTAATCTCTCAATCTCTGCTTG ATCCACTTTTACTTGCAAGGCAAC(A/G)ATCTC
TP67095	XM_020837626.2	*Dendrobium catenatum* probable methyltransferase PMT23 (LOC110107384), transcript variant X1	CAGCAGATCATATGTTCT(A/G)GGATATGTATTAA ATGATTCACACCAGTCATGATATATTCCAATT
TP87242	XM_020848676.2	*Dendrobium catenatum* BTB/POZ domain and ankyrin repeat-containing protein NPR2 (LOC110115447)	CAGCATACTGTGATTC(C/T)AAAATAGTTGCAGA AGTTCTGGATATGGGATTGGCCAATGTTAATCT
TP116035	XM_020841691.2	*Dendrobium catenatum* flowering time control protein FCA-like (LOC110110293), transcript variant X1	CAGCCAAGACAGCAAGCACAATTGCAGATCCTT CAGCCTTCTTTAGCTGT(A/C)AGTAAAATTTGAC
TP175631	XM_020830683.2	*Dendrobium catenatum* ACT domain-containing protein ACR6-like (LOC110102363), transcript variant X1	CAGCCGCCGCTCGGTGTGAGTCATTCCAATAGA AATCGCAGTTTTGGC(A/G)CTTTGCTTGTCCTTG
TP198849	XM_020733003.1	*Phalaenopsis equestris* zinc-finger homeodomain protein 6-like (LOC110030334)	CAGCCTCCACCCAACTCTCTCCGCGAAACC(A/C) AGCATCTTCTCCTTCTGCTCCGCCGTAAAAGTT
TP340206	XM_020841659.2	*Dendrobium catenatum* V-type proton ATPase subunit B 2 (LOC110110267), transcript variant X1	CAGCTATATCTTGAGTTCCTGGACAAATTTGAGA GGAAGTTTGTTTCTCAAGGGGCTTA(C/T)GATA
TP462208	XM_021870395.1	*Chenopodium quinoa* WD repeat-containing protein 44-like (LOC110693240), transcript variant X1	CTGCC(G/T)ACCAGTCCCAACCTCTCTCAGCTTA TTCCACATCCCATCTTCCCTAAACTCCTTCACC
TP472563	XM_018962128.1	*Juglans regia* 40S ribosomal protein S9-2-like (LOC108988775)	CTGCCTGATTAGGACCCTGGCATGGTGAATGGA CTTGGCCATTCCCGACTTAAAGACCTGAGT(C/T)
TP490415	XM_024039389.1	*Quercus suber* NAC domain-containing protein 35 (LOC112007079)	CTGCGGGAGCCTATACTC(A/G)TTCATGATCCAG CTGGTTCGGATGCCCTTCGGAGCCTTTCCTGAA
TP522599	XM_020728072.1	*Phalaenopsis equestris* ABC transporter G family member 39-like (LOC110026881)	CTGCTTCCATAGGCAAGC(A/G)TTGCACTGAGTC AAGAATGATTGAGAGAATTGGGTTGGAAAGAAA

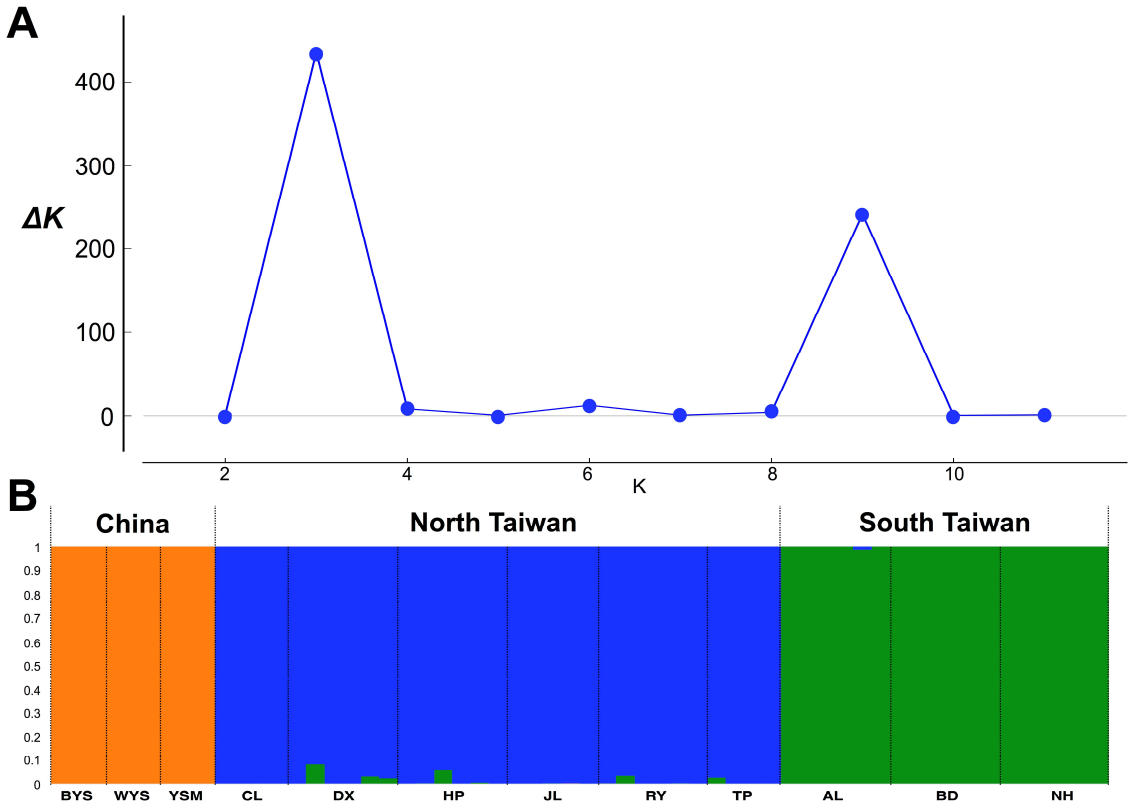

Fig. 3. STRUCTURE clustering results deduced from 3,737 SNPs within populations of *Pleione formosana*. **A**. *ΔK* values as a function of the K values according to 10 run outputs; and **B**. STRUCTURE results at K = 3, with different colors representing different clusters.

Table 4. Summary statistics of genetic variation existing in *Pleione formosana* estimated by total (3,737 SNPs) and outlier (80 SNPs) datasets

Dataset	Ho	Hs	Ht	F_{ST}	Dest
total (3,737 SNPs)	0.267	0.237	0.314	0.245	0.110
outliers (80 SNPs)	0.122	0.109	0.413	0.736	0.373

Ho: heterozygosity within population; **Hs**: genetic diversity within population; **Ht**: overall gene diversity; **F_{ST}**: fixation index; **Dest**: measure of population differentiation.

and China clusters (Table 5). The North Taiwan cluster maintains the highest genetic variation and differentiation. The minimum spanning tree indicates that the North Taiwan cluster is closer to China cluster, suggesting that the South Taiwan cluster stems from there (Fig. 4).

Nm is a combination of parameters that indicates the relative strengths of gene flow and genetic drift. Genetic drift will result in substantial local differentiation if $Nm < 1$ but not if $Nm > 1$ (Wright, 1951). The gene flow among the China populations was generally at a high level (mean value for *Nm* is 3.244), especially between BYS and WYS, where *Nm* peaked 6 (Table 6). In addition, gene flow exhibited higher levels among the populations distributed close to each other, such as AL with BD (*Nm* = 2.002) in the South Taiwan cluster, CL with TP (*Nm* = 2.875) and DX with JL (*Nm* = 2.691) in the North Taiwan cluster (Table 6).

DISCUSSION

The Orchidaceae are typically characterized by relatively low genetic differentiation among populations (F_{ST} = 0.146). In many cases, genetic differentiation (F_{ST}) within rare orchid species remains very low (0.092), compared with more common species (0.279) (Phillips *et al.,* 2012). Although the distribution of *P. formosana* was fragmented and isolated recently, it still maintained moderate genetic diversity (Ht = 0.314) and genetic differentiation (F_{ST} = 0.245)(Fig. 2, Table 4).

When examining patterns of *P. formosana* population differentiation at neutral versus adaptive SNP, we have detected several distinct differences. Primarily stochastic processes drive differentiation of neutral SNP, whereas both stochastic and selective processes drive that of adaptive SNP (Galindo *et al.,* 2009, Stronen *et al.,* 2015, Guo *et al.,* 2016). Directional selection may lead to an increase in F_{ST} of an adaptive SNP. When only the 80 adaptive SNPs were included, the genetic differentiation was much higher (F_{ST} = 0.726) (Fig. 2). This result supports the hypothesis that putative under selection of SNPs are being affected by adaptation.

Irrespective of the mechanism underlying these changes (drift or selection), this illustrates that, in the natural distribution of *P. formosana*, specific nucleotides can undergo drastic changes within only a hundred kilometers of distance. Our results exhibited spatial patterns differentiating environmental variations in loci which could occur in genes for important functions of *P. formosana* across their habitat. For example, the GBS tags of adaptive SNP TP116035 (Table 3) show a high identity with *Dendrobium catenatum* flowering time control protein FCA-like (LOC110110293), transcript variant X1. It plays a major role in the promotion of the transition of the vegetative meristem to reproductive development by decreasing FLOWERING LOCUS C mRNA levels.

Trochodendron aralioides has been generally assigned as an indicator of the cloud forest in Taiwan (Su, 1984). It has higher genetic diversity in the north to central mountainous area that might be the major refugium during the last glaciations (Huang *et al.,* 2004). The divergence of *P. formosana* populations, prevalent under cloudy forest, is consistent with a north-south phylogeographical divergence found in *T. aralioides*. The genetic diversity and population structure of *P. formosana* might reflect a similar post-glacial colonization history with cloudy forest species, and also can be paralleled by its unique local adaptation.

There is no immediate risk to the survival of *P. formosana* in Taiwan, because most of the populations are located in places that cannot be easily reached by humans, but the remote locations it also pose challenges for conducting research. The species is also benefitted by the current artificial reproduction and sale in Mei-Feng Farm, which is affiliated with National Taiwan University, and Nan-zhuang Farm. These practices help decrease illegal *P. formosana* poaching and trade. In addition, since *P. formosana* uses the mechanism of food-deceptive pollination, there will be no specific pollinators. We recommend experimenting with artificial pollination or finding possible pollinators for the species to increase its pollination rate and increase the amount of capsular fruits and seeds. Although topographical factors render the individuals difficult to reach, identification of possible pollinators and further pollination research are critical for the development of a successful conservation strategy.

The critical status of *P. formosana* may be exacerbated by the following aspects of climate change: (1) If the temperature continues to increase in summer, it may cause the plant death, and high temperatures in winter may affect the plant vernalization and cause the failure of plants to flower in the following spring; (2) Uneven rainfall can cause mortality in moss and epiphyte orchid species during long-term drought, and even if there is rain following drought, mortality of moss and *P. formosana* can be high; (3) Using a hierarchical model to predict the migration of several forest types and epiphytes under climate change, results show the distinct decline in the montane cloud forest in Taiwan and *P. formosana* is one of the epiphyte species having restricted distribution in this particular regime (Hsu *et al.,* 2012). The decrease in montane cloud forest may cause the decreasing habitat of *P. formosana*. We need more

Table 5. Pairwise F_{ST} among *Pleione formosana* populations based on 3,737 SNPs. F_{ST} below diagonal, 95% confidence interval above diagonal, based on 1,000 bootstrap replicates. F_{ST} estimates with lower 95% CI above zero highlighted in bold font.

	South Taiwan			North Taiwan						China		
	AL	BD	NH	CL	DX	HP	JL	RY	TP	BYS	WYS	YSM
AL	-	0.102, 0.121	0.108, 0.126	0.197, 0.221	0.158, 0.178	0.222, 0.247	0.176, 0.200	0.216, 0.242	0.186, 0.209	0.360, 0.392	0.356, 0.387	0.333, 0.362
BD	0.111	-	0.123, 0.143	0.208, 0.233	0.176, 0.197	0.223, 0.249	0.197, 0.220	0.236, 0.262	0.190, 0.214	0.373, 0.405	0.365, 0.398	0.341, 0.371
NH	0.116	0.133	-	0.219, 0.243	0.185, 0.206	0.235, 0.260	0.205, 0.228	0.246, 0.270	0.205, 0.229	0.365, 0.398	0.360, 0.393	0.338, 0.368
CL	0.209	0.220	0.231	-	0.120, 0.140	0.160, 0.184	0.120, 0.140	0.195, 0.220	0.072, 0.090	0.332, 0.363	0.328, 0.359	0.309, 0.339
DX	0.168	0.186	0.195	0.129	-	0.170, 0.193	0.077, 0.093	0.128, 0.149	0.103, 0.123	0.317, 0.346	0.308, 0.339	0.294, 0.321
HP	0.235	0.236	0.248	0.172	0.182	-	0.180, 0.205	0.243, 0.270	0.140, 0.164	0.367, 0.399	0.362, 0.395	0.350, 0.381
JL	0.189	0.209	0.216	0.131	0.085	0.192	-	0.163, 0.185	0.100, 0.119	0.329, 0.359	0.327, 0.357	0.308, 0.339
RY	0.229	0.249	0.257	0.207	0.138	0.257	0.174	-	0.180, 0.204	0.390, 0.423	0.386, 0.417	0.371, 0.402
TP	0.197	0.202	0.216	0.080	0.113	0.152	0.110	0.191	-	0.317, 0.348	0.310, 0.343	0.295, 0.326
BYS	0.375	0.388	0.382	0.347	0.331	0.382	0.344	0.407	0.333	-	0.025, 0.055	0.110, 0.140
WYS	0.371	0.381	0.377	0.344	0.324	0.379	0.343	0.402	0.327	0.040	-	0.099, 0.127
YSM	0.348	0.356	0.354	0.325	0.307	0.366	0.324	0.387	0.311	0.125	0.112	-

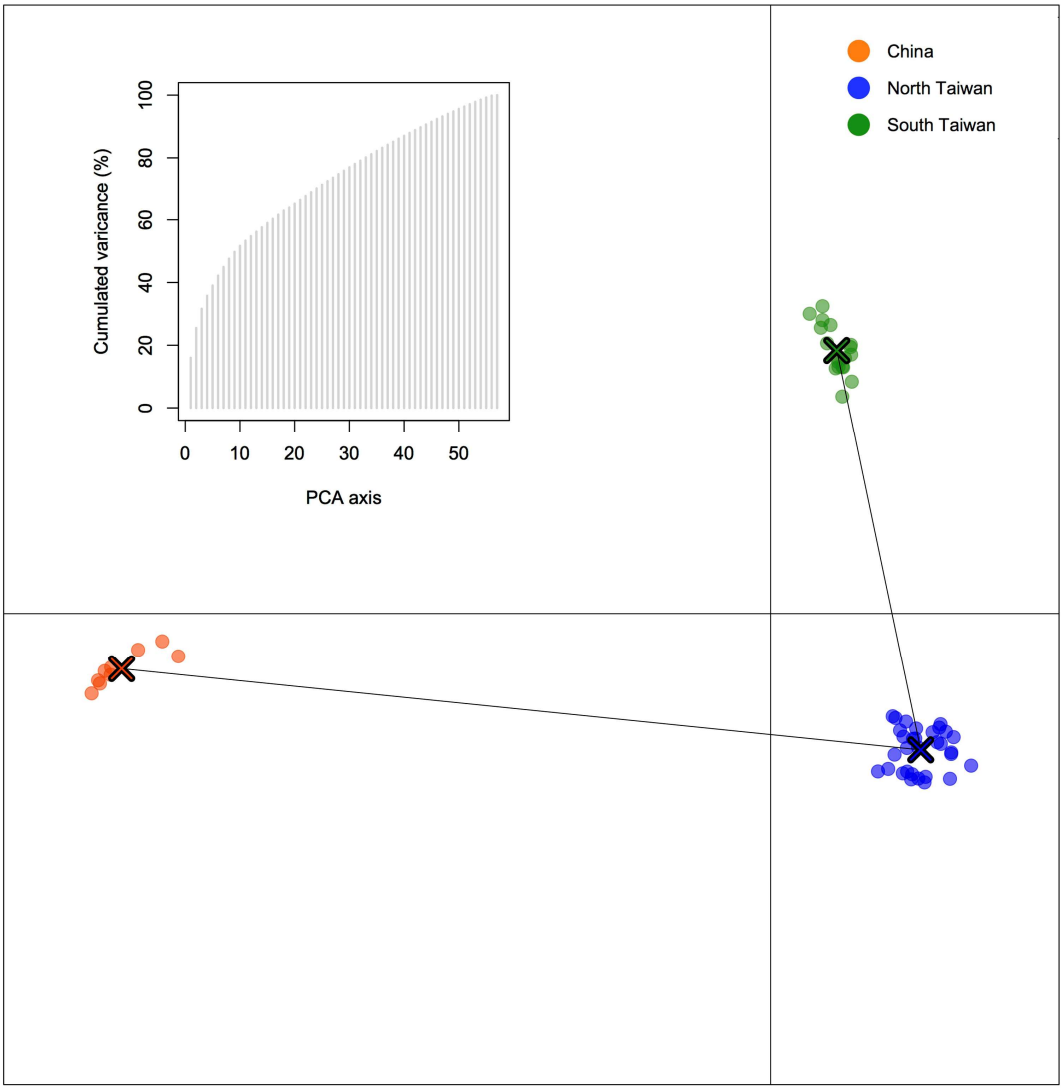

Fig. 4. Discriminant Analysis of Principal Components (DAPC) scatterplot drawn using 3,737 SNPs across 58 *Pleione formosana* individuals. Dots represent individuals, with colors denoting cluster allocation. Percentages of cumulated variance explained by Principal Components are shown in the top left corner. Minimum spanning tree based on the (squared) distances between clusters within the entire space is shown.

Table 6. Gene flow (*Nm*) among 12 *Pleione formosana* population.

	South Taiwan			North Taiwan						China		
	AL	BD	NH	CL	DX	HP	JL	RY	TP	BYS	WYS	YSM
AL												
BD	2.002											
NH	1.905	1.630										
CL	0.946	0.886	0.832									
DX	1.238	1.094	1.032	1.688								
HP	0.814	0.809	0.758	1.203	1.124							
JL	1.073	0.946	0.907	1.658	2.691	1.052						
RY	0.842	0.754	0.723	0.958	1.562	0.723	1.187					
TP	1.019	0.988	0.907	2.875	1.962	1.395	2.023	1.059				
BYS	0.417	0.394	0.404	0.470	0.505	0.404	0.477	0.364	0.501			
WYS	0.424	0.406	0.413	0.477	0.522	0.410	0.479	0.372	0.515	6.000		
YSM	0.468	0.452	0.456	0.519	0.564	0.433	0.522	0.396	0.554	1.750	1.982	

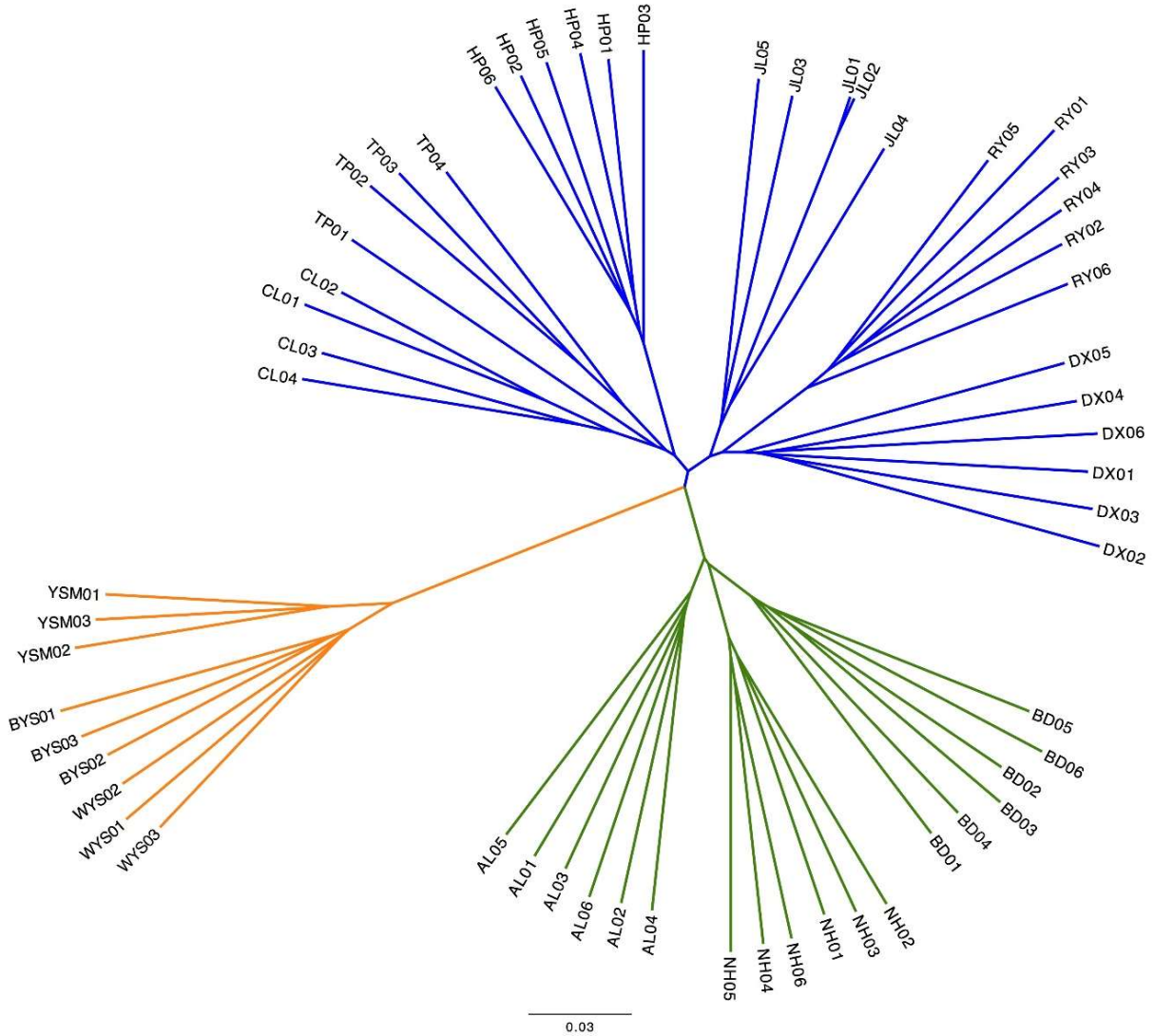

Fig. 5. Neighbor-join (NJ) tree (unrooted) based on 3,737 SNPs. Colored branches represent the corresponding group, colors are as indicated in Fig. 3 and Fig. 4. All major nodes have 100% bootstrap support.

research linking climate change and its effect on the distribution dynamic of *P. fomosana*.

The maintenance of a maximal amount of genetic diversity is a fundamental objective for conserving endangered species, because diversity is associated with a species' evolutionary potential and viability, as well as with its ability to adapt to ever-changing fitness thanenvironments (Reed and Frankham 2003). However, hybridization between regions cannot be recommended because a mixture of individuals from genetically distinct populations may result in outbreeding depression (Edmands and Timmerman 2003). Moreover, the introduction of genotypes with lower the local genotypes can have negative consequences for the successful establishment of restored populations (Hufford and Mazer 2003). Possible restoration actions for *P. formosana* should consider a number of viable populations as propagule donors within the same evolutionary significant units (ESUs) to create reinforced populations to increase the geographical range and reduce the risk of extinction. Given the genetic differentiation of *P. formosana* among China and Taiwan populations, we suggested treating them as different management units in utilization of resources. The larger population in the north cluster are distributed in the Daxueshan Forest Recreation Area (DX) and Rueiyansi Major Wildlife Habitat (RY). The main population in the south cluster is distributed in the Alishan Timber Management Area (AL). Beside the conservation action in each native habitat to avoid the decline in germplasm, the population genetic structure should be taken into consideration for the exhibition of *P. formosana* to avoid the artificial genetic admixture.

P. formosana is protected by the Cultural Heritage Preservation Act in Taiwan because it is a rare species seriously impacted by stealing and difficulties with breeding. Meanwhile, we established Taiwan Pleione Nature Reserve which prohibits any alteration of or damage to the natural state in order to protect the wild population. This Act is suitable for cultural heritage, but not for the early-succession plant species, like *P. formosana*. The early-succession plant species will disappear when succession is ongoing, so the *P. formosana* in Nature Reserve will be replaced by other late-succession plant species. *P. formosana* might continue to live in habitats receiving some disturbance rendering and prolonging an early-succession environment. In addition to natural disturbances (earthquake, typhoon, etc), human interference, such as selectively cutting the canopy or weeding herbs, can renew the environment. Otherwise, many populations of *P. formosana* are living outside the natural reserves, and the distribution will change as new suitable habitat occurs. Therefore, according to our research, the Taiwanese Cultural Heritage Preservation Act is no longer suitable for managing the distribution of *P. formosana*. We suggest the development of a new Act similar to the Taiwanese Wildlife Conservation Act to protect *P. formosana*, as well as other early-succession plant species.

ACKNOWLEDGMENTS

The authors thank the Chiayi Forest District Office for funding the project through the grant No. 104CRC01.

LITERATURE CITED

Bothwell, H., S. Bisbing, N.O. Therkildsen, L. Crawford, N. Alvarez, R. Holderegger and S. Manel. 2013. Identifying genetic signatures of selection in a non-model species, alpine gentian (*Gentiana nivalis* L.), using a landscape genetic approach. Conserv. Genet. **14(2)**: 467–481.

Chao, W.-C., C.-L. Huang, Y.-C. Liu. 2017. The genetic diversity analysis and population dynamic changes project of *Pleione formosana* in Taiwan Pleione Nature Reserve, Chiayi Forest District Management Office, Forestry Bureau, Concil of Agriculture, Executive Yuan, Taiwan.

Chen, J.-L., Y.-Z. Zhou, S.-S. Wu, S.-R. Lan and D.-H. Peng. 2019. Pollination mechanism and breeding system of *Pleione formosana*. Journal of Forest and Environment **39**: 460–466.

Cribb, P., and I. Butterfield. 1999. The Genus Pleione, 2nd ed. Richmond: Royal Botanic Gardens, Kew.

Dobson, A.J. and A.G. Barnett. 2008. An introduction to generalized linear models, 3rd ed. Chapman and Hall/CRC, London, UK.

Doyle, J.J. and J.L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin **19**: 11–15.

Edmands, S. and C.C. Timmerman. 2003. Modeling factors affecting the severity of outbreeding depression. Conserv. Biol. **17(3)**: 883–892.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler and S.E. Mitchell. 2011. A robust, simple genotyping by sequencing (GBS) approach for high diversity species. PLoS One **6(5)**: e19379.

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.

Flanagan, S.P., B.R. Forester, E.K. Latch, S.N. Aitken and S. Hoban. 2018. Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. Evol. Appl. **11(7)**: 1035–1052.

Foll, M. and O. Gaggiotti. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics **180(2)**: 977–993.

Frankham, R. 2010. Where are we in conservation genetics and where do we need to go? Conserv. Genet. **11(2)**: 661–663.

Frankham, R., J.D. Ballou and D.A. Briscoe. 2004. A primer of conservation genetics. Cambridge University Press, Cambridge, UK.

Fraser, D.J. and L. Bernatchez. 2001. Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Mol. Eco. **10(12**: 2741–2752.

Galindo, J., P. Morán and E. Rolán-Alvarez. 2009. Comparing geographical genetic differentiation between

candidate and noncandidate loci for adaptation strengthens support for parallel ecological divergence in the marine snail *Littorina saxatilis*. Mol. Eco. **18(5)**: 919–930.

Gebremedhin, B., G. Ficetola, S. Naderi, H. R. Rezaei, C. Maudet, D. Rioux, G. Luikart, Ø. Flagstad, W. Thuiller and P. Taberlet. 2009. Frontiers in identifying conservation units: from neutral markers to adaptive genetic variation. Anim. Conserv. **12(2)**: 107–109.

Glaubitz, J. C., T. M. Casstevens, F. Lu, J. Harriman, R. J. Elshire, Q. Sun and E. S. Buckler. 2014. TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. PLoS One **9(2)**: e90346.

Goudet, J. 2005. HIERFSTAT, a package for R to compute and test hierarchical *F*-tatistics. Mol. Eco. Notes **5(1)**: 184–186.

Gravendeel, B., M. C.M. Eurlings, C. van den Berg and P.J. Cribb. 2004. Phylogeny of *Pleione* (Orchidaceae) and parentage analysis of its wild hybrids based on plastid and nuclear ribosomal ITS sequences and morphological data. Syst. Bot. **29(1)**: 50–63.

Guo, B., Z. Li and J. Merilä. 2016. Population genomic evidence for adaptive differentiation in the Baltic Sea herring. Mol. Eco. **25(12)**: 2833–2852.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. **25(15)**: 1965–1978.

Hsu, R. C.-C., W. L.M. Tamis, N. Raes, G.R. de Snoo, J.H.D. Wolf, G. Oostermeijer and S.-H. Lin. 2012. Simulating climate change impacts on forests and associated vascular epiphytes in a subtropical island of East Asia. Divers. Distrib. **18(4)**: 334–347.

Huang, S.-F., S.-Y. Hwang, J.-C. Wang and T.-P. Lin. 2004. Phylogeography of *Trochodendron aralioides* (Trochodendraceae) in Taiwan and its adjacent areas. J. Biogeogr. **31(8)**: 1251–1259.

Hufford, K.M. and S.J. Mazer. 2003. Plant ecotypes: genetic differentiation in the age of ecological restoration. Trends Ecol. Evol. **18(3)**: 147–155.

Jiang, M.-T., W.-C. Chao, C.-L. Huang, S.-R. Lan, Z.-J. Liu and S.-S. Wu. 2019. The complete chloroplast genome of *Pleione formosana* (Orchidaceae). Mitochondrial DNA Part B **4(1)**: 1044–1046.

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.

Joost, S., A. Bonin, M.W. Bruford, L. Després, C. Conord, G. Erhardt and P. Taberlet. 2007. A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. Mol. Eco. **16(18)**: 3955–3969.

Jost, L. 2008. GST and its relatives do not measure differentiation. Mol. Eco. **17(18)**:4015–4026.

Kamvar, Z.N., J.F. Tabima and N.J. Grünwald. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ **2**: e281.

Lu, F., A.E. Lipka, J. Glaubitz, R. Elshire, J.H. Cherney, M.D. Casler, E.S. Buckler and D.E. Costich. 2013. Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. PLoS Genetics **9(1)**: e1003215.

Luu, K., E. Bazin and M.G.B. Blum. 2017. Pcadapt: an R package to perform genome scans for selection based on principal component analysis. Mol. Ecol. Resour. **17(1)**:67–77.

Moritz, C. 1994. Defining 'evolutionarily significant units' for conservation. Trends Ecol. Evol. **9(10)**: 373–375.

Nei, M. 1987. Molecular evolutionary genetics. Columbia university press, New York, USA.

Peterson, B.K., J.N. Weber, E.H. Kay, H.S. Fisher and H.E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One **7(5)**: e37135.

Phillips, R.D., K.W. Dixon and R. Peakall. 2012. Low population genetic differentiation in the Orchidaceae: implications for the diversification of the family. Mol. Eco. **21(21)**: 5208–5220.

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

R Development Core Team. 2015. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at http://www.R-project.org.

Reed, D.H. and R. Frankham. 2003. Correlation between fitness and genetic diversity. Conserv. Biol. **17(1)**: 230–237.

Rellstab, C., F. Gugerli, A.J. Eckert, A.M. Hancock and R. Holderegger. 2015. A practical guide to environmental association analysis in landscape genomics. Mol. Eco. **24(17)**: 4348–4370.

Shi, Y., L. Yang, Z. Yang and Y. Ji. 2018. The complete chloroplast genome of *Pleione bulbocodioides* (Orchidaceae). Conserv. Genet. Resour. **10(1)**: 21–25.

Storey, J.D., A.J. Bass, A. Dabney and D. Robinson. 2019. Qvalue: Q-value estimation for false discovery rate control. R package version 2.16.0, https://github.com/jdstorey/qvalue.

Stronen, A.V., B. Jędrzejewska, C. Pertoldi, D. Demontis, E. Randi, M. Niedziałkowska, T. Borowik, V.E. Sidorovich, J. Kusak and I. Kojola. 2015. Genome-wide analyses suggest parallel selection for universal traits may eclipse local environmental selection in a highly mobile carnivore. Ecol. Evol. **5(19)**: 4410–4425.

Su, H.-J. 1984. Studies on the climate and vegetation types of the natural forest in Taiwan (II). Altitudinal vegetation zones in relation to temperature gradient. Quarterly Journal of Chinese Forestry **17**: 57–73.

Su, H.-J. 1988. Research Study of Vegetation Ecosystems in Alishan Taiwan Pleione Nature Reserve. Forestry Bureau, Council of Agriculture, Executive Yuan. Taipei. (press in Chinese)

Wright, S. 1951. The genetical structure of populations. Ann. Hum. Biol. **15(1)**: 323–354.