Genetic diversity of endangered terrestrial orchids *Spathoglottis plicata* in Peninsular Malaysia based on AFLP markers

Florence C Ginibun*[1,3], Paul Arens[2], Ben Vosman[2], Subha Bhassu[1], Norzulaani Khalid[1] and Rofina Yasmin Othman*[1]

[1]Centre for Research in Biotechnology for Agriculture (CEBAR), Department of Genetics and Molecular Biology, Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia
[2]Wageningen UR Plant Breeding, Wageningen University and Research Center, P.O. Box 386, 6700 AJ Wageningen, The Netherlands
[3]Department of Agriculture, Level 7-17, Wisma Tani, Lot 4G2, No.30, Persiaran Perdana. Presint 4, 62624 Putrajaya, Malaysia

*Corresponding authors: yasmin@um.edu.my; florence@doa.gov.my

Abstract

Spathoglottis plicata is an endangered terrestrial orchid species that have experienced severe threats to its habitat as wild forest sites come under pressure from industrialisation and natural disasters. This orchid species chosen to evaluate their levels of genetic diversity and population genetic structure, which 25-30 accession collected in the different location with different geographical, altitude and habitat. Genomic DNA was extracted from six natural populations (n=172) in Peninsular Malaysia using eleven AFLP markers of *EcoRI*+3 bases/*MseI*+3 base primer combinations. Based on 279 polymorphic bands, a significant degree of genetic population differentiation was found, with a 78.5% variation within populations as measured by AMOVA, indicating a potential restricted gene flow. Two distinct clades generated from a UPGMA dendrogram were further investigated through a Bayesian analysis using STRUCTURE software, producing an estimated population structure at optimal value *K*=4. These results point to the presence of four genetic structures in the *Spathoglottis plicata* population. The Pahang and Terengganu population revealed a higher than average genetic variation (60.25%), indicating that there may be a robust structural division between the population samples and a possible hybridisation between the Northern (Kedah), Southern (Negeri Sembilan and Johor) and Central (Selangor) region populations. In sum, these results suggest that geographical distance is the primary factor contributing to differences among populations and the need for conservation measures to protect the S*pathoglottis plicata* species.

Keywords: AFLP, *Spathoglottis plicata*, terrestrial orchids, genetic diversity.
Abbreviations: AMOVA_analysis of molecular variance; UPGMA_unweighted pair group method with arithmetic mean; SPN9_*Spathoglottis plicata* Negeri Sembilan; SPT_*Spathoglottis plicata* Terengganu; SPK_*Spathoglottis plicata* Kedah; SPJ_*Spathoglottis plicata* Johor; SPP_*Spathoglottis plicata* Pahang; SPS_*Spathoglottis plicata* Selangor

Introduction

Spathoglottis is a genus belonging to a family of Orchidaceae, proposed as a genus by Blume in 1825, citing *Spathoglottis plicata* as the only species. Taxonomists have since identified 40 species, found in the regions of India, South Japan and China, down through the Malay Archipelago, on the islands of the Pacific and in Northern Australia (Teo, 1985; Seidenfaden and Wood, 1992; Comber, 2001). Six species are reported as native and growing wild in Peninsular Malaysia, of which *Spathoglottis plicata* is the most common.
Spathoglottis plicata is one of many wild orchid species in tropical ecosystems. It is a hardy plant usually found growing wild in open fields, poor grassland areas, roadsides, along streams, or on the edge of rock. It is found throughout Malaysia, in both the lowlands and highlands. However, the wild natural population of the *Spathoglottis* species in Malaysia has been exposed to many threats such as erosion, construction and adverse environmental conditions, notably as a result of the development of many wild forest areas into industrial sites or large-scale plantations. These could have caused *Spathoglottis plicata* populations to shrink in size, leading to lower genetic diversity.

Morphologically, the name *Spathoglottis plicata* refers to the plant's pleated or narrow folded leaf blade called a plicate, which is lanceolate, broad and long. Its flowers are about 4 cm broad; the sepal is slightly narrower than the lateral petal; and colours vary from shades of purple, mauve to white (Teo, 1985; Holttum, 1953). Distinguishing genetic variations in the *Spathoglottis plicata* orchid based on morphological characteristics is however unreliable, as the phenotypic expression of these characteristics could have been affected by environmental and developmental conditions. In other words, wild populations of *Spathoglottis plicata* may have undergone adaptations due to the changing environment, which in turn may have led to significant degrees of genetic variation as well as affecting population size.

Because of this possibility, this study uses molecular genetic markers based on genome as an alternative method to characterise and decipher the amount of genetic variations in *Spathoglottis plicata* populations. Using DNA as markers in a structural analysis of the *Spathoglottis* species has a number of advantages over traditional morphological markers. Agarwal et al. (2008) suggested that molecular markers based on DNA polymorphism could replace morphology markers, as the former are more informative and independent of environmental conditions. In practice, several molecular markers such as ISSR, RAPD, DNA barcode and AFLP already been used in a broad range of plant species – including the *Spathoglottis* species (Romeida et al., 2012; Setiawan et al., 2013; Ginibun et al., 2015 and 2010; Pimonrat et al., 2012).

Amplified Fragment Length Polymorphism (AFLP) is one of the most popular molecular marker techniques that has various advantages over other DNA-based markers, such as high reproducibility and a high level of polymorphism widely distributed across the genome, with no prior DNA sequence information needed (Jones et al., 1997). It covers a broad area of the plant genome and produces more independent polymorphic loci in a single analysis than other molecular marker systems (Vos et al., 1995; Sathyanarayana et al., 2011).

The AFLP markers were successfully used in the assessment and determination of genetic diversity and genetic relationships between and within species and wild populations of a wide range of orchid species and hybrids. It is including *Phalaenopsis* (Chang et al., 2009), *Diuris fragrantissima* (Smith et al., 2007), *Liparis loeselli* (Pillon et al., 2007), *Dendrobium* (Xiang et al., 2003), *Orchis mascula* and *Orchis purpurea* (Jacquemyn et al., 2009 and 2007), *Neotinea maculata* (Duffy et al., 2009), *Neotinea ustulata* (Tali et al., 2006), and *Himantoglossum hircinum* (Pfeifer and Jetschlee, 2006). However, there have been only a few reports on molecular markers been used to study the genetic relationships, genetic differentiation or genetic structure in the *Spathoglottis* species. Therefore, the aims of this study are to enhance understanding of the genetic variations that have occurred in the existing wild *Spathoglottis plicata* populations found in different geographical regions of Peninsular Malaysia. In addition, to provide better insights into the factors which contribute to such genetic variation – all with a view to improving the conservation of the wild *Spathoglottis plicata* population in the future.

Results

Genetic diversity analysis of Spathoglottis plicata populations

The eleven AFLP primer combinations screened enabled us to capture the genetic variations within and between populations and to distinguish clearly all the individuals from the six *Spathoglottis plicata* populations of Peninsular Malaysia. These resulted in 279 unambiguous scorable polymorphic bands among the 172 *Spathoglottis plicata* individuals, from 691 bands (Table 3), all of which provided clear amplification profiles. Polymorphic bands both within and among populations were scored between 40 to 300 bp for each primer pair, which produced a variable bands ranging from 16 (IRD 700 E+CAC/M+CCG) to 30 (IRD 700 E+CAG/M+CCA).

An AMOVA conducted on the *Spathoglottis plicata* population at two hierarchical levels showed the proportions of genetic variation between populations to be higher than those within populations. As Table 4 shows, there was 78.51% of molecular variance between populations, but only 21.49% within populations. The average genetic differentiation between the populations amounted to PHIst = 0.785, and was highly significant (P<0.001). All pairwise genetic distances between populations were significant (*P* value), with values ranging from 0.3975 to 0.9953, as shown in Table 5.

A UPGMA dendrogram generated with Nei's (1972) unbiased distance values (in supplementary file). The results indicated three possible subclusters, with the Negeri Sembilan and Johor populations grouped together; Kedah clustered as one subgroup with the Selangor population; and Terengganu and Pahang clustered as a further group with possible genetic admixture.

The AFLP molecular data were further analysed in order to construct a UPGMA dendrogram between individuals within the six *Spathoglottis plicata* populations, based on Jaccard coefficient of similarity for pairwise, using PAST software. The corphenetic correlation based on Jaccard similarity measures was 0.9835. The individuals named according to the State they were located in. The UPGMA dendrogram results revealed two distinct clades (Fig 2). Clade 1 consisted of the Selangor population plus the Kedah, Terengganu and Pahang populations. Clade 2 consisted of the Negeri Sembilan, Johor, Terengganu and Pahang populations. This division could well have arisen due to the geographic separation by the Titiwangsa Range, with the Kedah population representing the Northern population and Terengganu and Pahang representing the Eastern population. One of the clades further subdivided into two distinct clusters – Negeri Sembilan and Johor representing the southern region, and Terengganu and Pahang representing the eastern region. This suggests that these populations could have diverged only recently.

Structure analysis of Spathoglottis plicata populations

It is noteworthy that both eastern region (Pahang and Terengganu) populations were present in both clusters. A further AMOVA analysis therefore conducted to determine the variation between the populations. The detailed results of this AMOVA and the related cluster analysis shown in Table 6.

These populations showed significant differences at P<0.001. The molecular variance between the populations of Terengganu and Pahang was only 39.75%, but within these populations was 60.25%. The genetic differentiation averaged for the overall loci amounted to PHIst = 0.397. All this indicates that the variations were higher within the Terengganu and Pahang populations than between these two populations. It will be important to understand these significant genetic variations between individuals in these two populations and to take account of these in conservation planning.

Further describe the genetic relationships between the populations; a detailed classification of the six *Spathoglottis plicata* populations was produce through a Bayesian analysis conducted with the commonly used STRUCTURE 2.3.3 tool. The number of inferred groups was set at value *K*, ranging from 1 to 6 in line with the number of populations. The STRUCTURE analysis applied using the four-step graphical Evanno method to allow the detection of the true number of group *K*, as shown in Fig 3. The results produced two peaks in the graph representing the structure at *K*=2 and the substructure at *K*=4.

The estimated population structure at *K*=4 is shown in Fig 4. In sum, it proved possible to locate the presence of the *Spathoglottis plicata* population using the 279 AFLP loci in this study. The Negeri Sembilan and Johor populations emerged as belonging to the same genetic group. The Kedah and Selangor populations represented two different genetic groups. Interestingly, Terengganu had a mixture of genetic groups from both the Northern and Southern region, whereas the Pahang populations showed three different mixtures of genotype shared with those from Kedah and Selangor. Overall, the results show four genotypes occurring in the population samples from the different geographical regions, namely: Northern (Kedah), Southern (Negeri Sembilan and Johor), Central (Selangor) and Eastern (Pahang).

Discussion

An analysis of AFLP data from a combination of eleven primers revealed significant differences in genetic variation between and within the six populations and individuals of *Spathoglottis plicata*. The AMOVA results showed a relatively high differentiation value within the *Spathoglottis plicata* populations (Fst= 0.785) – higher for example than other terrestrial species of *Pseudorchis albida* in Ireland (φpt = 0.306) (Duffy et al., 2011) or *Spiranthes romanzoffiana* elsewhere in Europe (overall mean *G*ST = 0.187) (Forrest et al., 2004). This could be because gene flows among individuals giving rise to higher genetic variations have been limited between geographical ranges. The higher genetic variation shown between and within the *Spathoglottis plicata* populations could imply that this species is predominantly outcrossing. Hamrick and Godt (1990) believed that plants typically display high genetic diversity if they have a wide geographical range of outcrossing. *Spathoglottis plicata* and the other similar species believed by Forbes (1884), Ridley (1930) and Lord (1981) to be both outcrossing and self-pollinating species. A degree of outcrossing is important when

facing an unstable habitat. However, some *Spathoglottis plicata* populations may also reproduce through various forms of self-pollination due to the absence of pollinators. As a result, these populations could become genetically uniform, and their survival may be dependent on the availability of suitable habitat and mass seed production. Our AMOVA and STRUCTURE analyses indicate the genetic isolation of populations located at a distance from each other, with higher variations within populations than between them.

Spathoglottis plicata is reported to be one of the most widespread species of orchid, with many cultivated varieties in Asia Pacific countries, as described by Ridley (1907), Holttum (1953), Teo (1985), Comber (1990, 2001) and Yong (1993). It is widely distributed throughout Peninsular Malaysia, from the North to South and Eastern to Western regions. This is probably a rare long distance seed dispersal event facilitated by wind or human factors. This species has been found with differences in size, colour and in some details of the flower shape, reflecting independent variations in the colour of petals and sepals, side lobes and the callus of the lip. Gene flow constraints could also influence the species. Hamrick and Godt (1989) proposed that reproductive biology is a key factor in determining the genetic structure of plant population. Breeding systems could also be major factor influencing the level of genetic diversity within a species.

STRUCTURE is a simple approach that allows the use of dominant markers such as AFLPs (Falush et al., 2007) to cluster individuals in populations. A STRUCTURE analysis provided a solid basis for identifying the six populations of *Spathoglottis plicata*, with consistent improvements in the probability function for the run with the *K*=4 value. Only one genotype found to be dominant in the *Spathoglottis plicata* populations in the States of Johor and Negeri Sembilan (Southern region). The populations in Kedah (Northern region) and Selangor (Central region) meanwhile showed up two different dominant genotypes. A further dominant genotype found in Pahang (Eastern region). Finally, the Terengganu populations shared a mixture of dominant genotypes from both the Southern and Northern regions.

Geographically, Northern, Eastern and Southern Peninsular Malaysia have different distributions of *Spathoglottis plicata* in terms of range, altitude, climate and habitat, including rainfall and soil taxonomy. The topography of Peninsular Malaysia is characterized by central mountain ranges running from north to south. Literature generally describes *Spathoglottis plicata* as a lowland plant, but it can also found growing in mountain habitats. For our study, the Johor (SPJ) and Negeri Sembilan (SPN9) populations sampled from lowland locations, whereas the Kedah (SPK) one was from a mountainous area. The Selangor (SPS) populations propagated in a nursery. The Terengganu (SPT) population sampled in an open space, but found to be a genetic mixture of lowland and highland populations. The Pahang (SPP) population meanwhile sampled at different latitude.

Another factor that could contribute to differentiation of genotypes is soil taxonomy, which differs between the various regions. The soils of Malaysia can be divide broadly into two groups: (a) sedentary soils formed in the interior on a wide

Table 1. Sampling locations and numbers of accession of *Spathoglottis plicata* populations in Peninsular Malaysia (Link with number to Fig 1/map).

| No.	Location	Geographical Location	Altitude	Habitat	Accession Numbers
1	Simpang Pertang, Kuala Kelawang, Negeri Sembilan	2º56'48.09"N 102º18'07.39"E	356 ft	Roadside	30
2	Sekayu, Terengganu	4º58'07.19"N 102º57' 18.96"E	247 ft	Open fields	29
3	Gunung Jerai, Kedah	5º47'12.23"N 100º25' 59.69" E	3786 ft	Mountain foothills and edges of rocks	29
4	Ulu Sedeli, Johor	1º48'20.04"N 103º56' 59.15"E	127 ft	Open field with loam soil	29
5	Genting Highland, Pahang	3º21'14.32"N 101º46'57.70"E	2553 ft	Mountain foothills	25
6	Old Klang Road, Selangor	3º05'0.9"N 101º40'46.61"E	134 ft	Propagated/cultivated	30
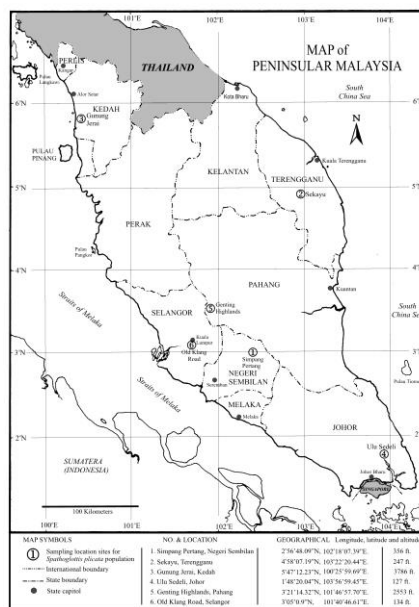
Fig 1. Map of sampling locations for *Spathoglottis plicata* populations in Peninsular Malaysia.

Table 2. List of eleven *Eco*RI + 3 base/*Mse*I + 3 base primer sets generated for *Spathoglottis plicata* population.

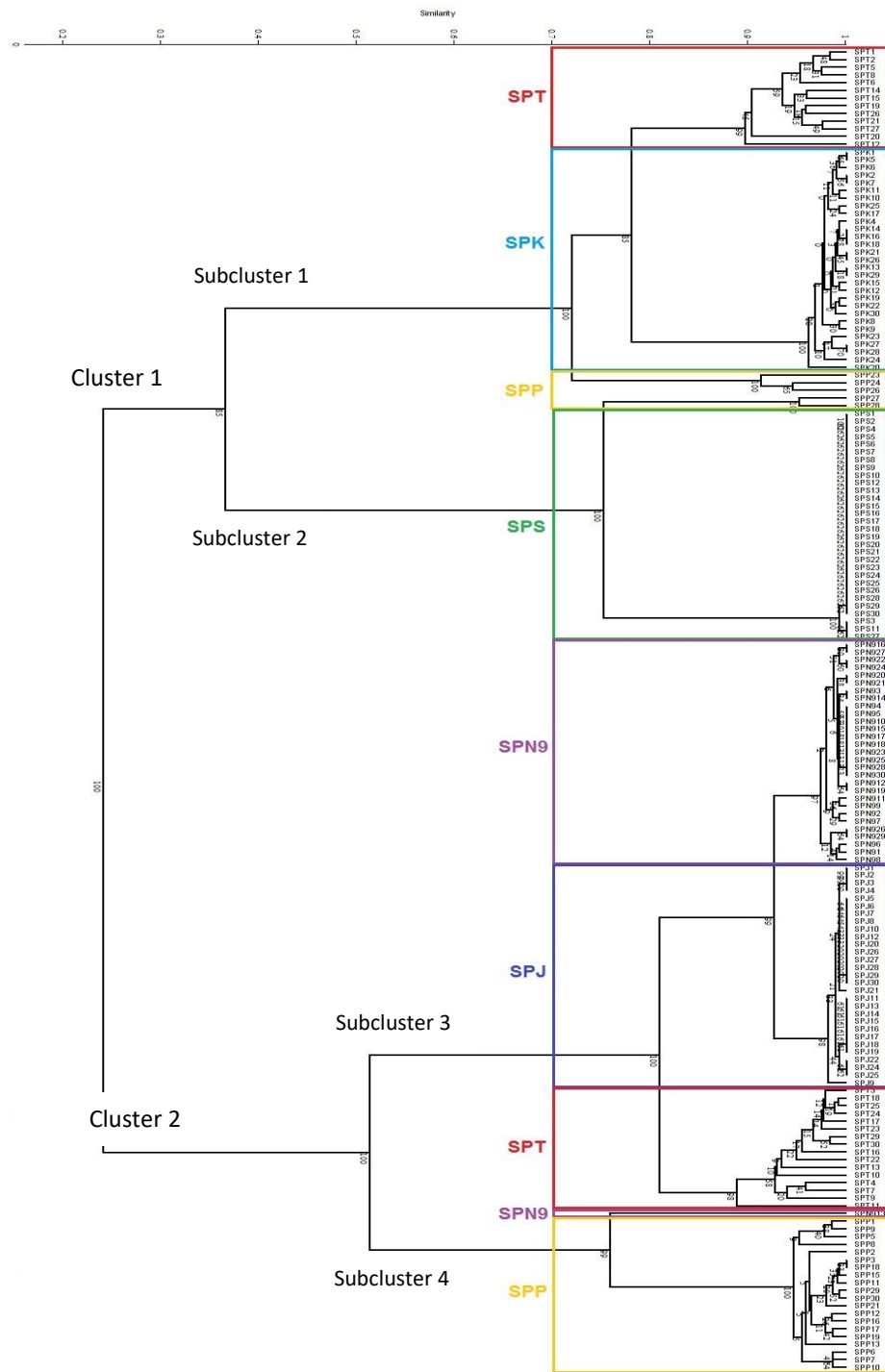| No.	Fluorchrome label of Li-Cor PCR	Selected Primer	
		*Eco*RI	*Mse*I
1	800	AAG	CAA
2	700	AAG	CCG
3	800	ACA	CAG
4	800	ACA	CAT
5	700	ACA	CCT
6	800	ACA	CTG
7	700	ACT	CAA
8	700	ACT	CAC
9	700	ACT	CAG
10	700	CAC	CCG
11	700	CAG	CCA

Fig 2. UPGMA dendrogram based on Jaccard's coeffient of similarity for pairwise of all individuals of the six *Spathoglottis plicata* populations. Individuals are named according to the State they are located in as such SPT_*Spathoglottis plicata* Terengganu (red color); SPK_ *Spathoglottis plicata* Kedah (light blue color); SPP_*Spathoglottis plicata* Pahang (yellow color); SPS_*Spathoglottis plicata* Selangor (green color); SPN9_*Spathoglottis plicata* Negeri Sembilan (purple color) and SPJ_*Spathoglottis plicata* Johor (dark blue color).

Table 3. Primer combination (only showing the three 3'selective nucleotides) used in AFLP, and scored number of polymorphic bands in five wild and one propagated *Spathoglottis plicata* populations

No.	Fluorchrome label of Li-Cor PCR	Selected Primer		Variable bands	Total present bands	% variable bands
		*Eco*RI	*Mse*I			
1	800	AAG	CAA	26	58	44.83
2	700	AAG	CCG	26	40	65
3	800	ACA	CAG	18	50	36
4	800	ACA	CAT	27	80	33.75
5	700	ACA	CCT	26	58	44.83
6	800	ACA	CTG	24	67	35.82
7	700	ACT	CAA	29	94	30.20
8	700	ACT	CAC	29	75	38.67
9	700	ACT	CAG	28	67	41.79
10	700	CAC	CCG	16	34	47.06
11	700	CAG	CCA	30	68	44.11
	Total variable bands			279	691	

Table 4. Analysis of Molecular Variance (AMOVA) based on 279 AFLP loci in five wild and one propagated *Spathoglottis plicata* populations

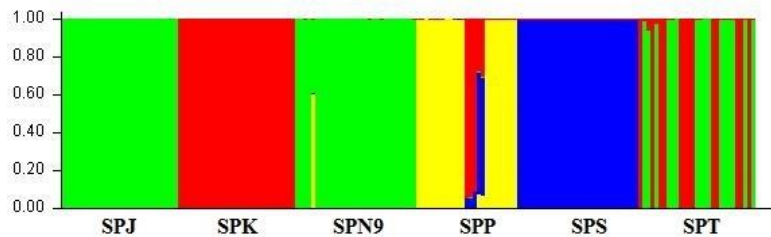| Source	df	SS	MS	Est. Var.	% var	P
Between populations	5	6885.396	1377.079	47.616	78.51%	<0.001
Within populations	166	2163.552	13.033	13.033	21.49%	<0.001
Total	171	9048.948				

Fig 4. STRUCTURE analysis of Malaysian *Spathoglottis plicata* population. Estimated population structure results of admixture analysis at K= 4. K coloured segment (colours) represent different genetic groups and the membership of K cluster.

Table 5. Pairwise genetic distances (PHIst) between *Spathoglottis plicata* populations.

	SPN9	SPT	SPK	SPJ	SPP	SPS
SPN9	0.0000					
SPT	0.5178	0.0000				
SPK	0.9753	0.5774	0.0000			
SPJ	0.7161	0.5180	0.9884	0.0000		
SPP	0.6991	0.3975	0.7953	0.7197	0.0000	
SPS	0.9779	0.6608	0.9862	0.9953	0.7807	0.0000

Note: SPN9, *Spathoglottis plicata* Negeri Sembilan; SPT, *Spathoglottis plicata* Terengganu; SPK, *Spathoglottis plicata* Kedah; SPJ, *Spathoglottis plicata* Johor; SPP, *Spathoglottis plicata* Pahang; SPS, *Spathoglottis plicata* Selangor.

Table 6. Analysis of molecular variance (AMOVA) based on 53 AFLP loci in two variable *Spathoglottis plicata* populations of Terengganu and Pahang.

Source	df	SS	MS	Est. Var.	% var	P
Between populations	1	725.9446	725.945	25.59057707900	39.75%	<0.0010
Within populations	52	2017.0924	38.790	38.79023872700	60.25%	<0.0010
Total	53	2743.0370				

range of rock types and (b) soils of the coastal alluvial plains (Nieuwolt et al., 1982). The Northern region has various categories of sedentary soil, whereas the Southern region has bris soil. Terengganu and Pahang have alluvium and granite soil at high and low elevations. Sedentary soils develop on igneous, sedimentary and metamorphic rocks, and strongly weathered with mostly kaolinitic clay minerals. Sedentary soils fall into a number of categories, including predominantly fine-textured clay and clay loam soils; peat and organic soils; acid sulphate soils scattered along the west coast plains; and sandy soils (bris soils) spread along the east coast of Peninsular Malaysia.

Genetic diversity within and between populations may be influenced by habitat variability, through a process of differential gene flow and other factors. Gene flow may have increased genetic diversity within the populations of Terengganu and Pahang by gene input from external sources such habitat structure. The genetic diversity of the *Spathoglottis* population could in part be due to different rainfall distribution (Nieuwolt, l982) in the three regions: Kedah rainfall is low compare to the much heavier rainfall in the other two regions, Terengganu and Pahang. In the latter two, the *Spathoglottis plicata* was present with two and three genotypes; this could be due to the higher levels of humidity and moisture in these regions. The climate in the Northern region, on the other hand, has a longer dry period.

One more factor that may have influences genetic diversity, or the relative lack of it, is geographical isolation. Notably, the northern population of *Spathoglottis plicata* is located in Gunung Jerai, at 1000-1500 metres above sea level. This population may have been geographically isolated for a long period and this isolation could be one of the major factors influencing genetic differentiation and limiting the amount of gene flow.

The Southern region population, on the other hand, may be topologically identical where it clustered together genetically. The UPGMA clustered the Pahang and Terengganu populations into two groups, but the STRUCTURE analysis divided the Pahang population into three clusters. This could account for the overlapping of individuals from different populations and the lack of correlation between genetic distance and geographic distance.

Gene differentiation and gene flow are important indices to estimate the population genetic structure of a species. Gene flows between populations and levels of migration could prevent continued divergence between populations (Wright, 1951). The breeding system of the *Spathoglottis* species by cross-pollination and the fact that its seeds can dispersed very long distances by wind could be an important factor, resulting in possibilities for geographical isolation population in Northern and Southern region and a sign of gene flow hybridization within the Pahang and Terengganu population.

Genetic diversity within and between populations is a function of historical events and recent evolutionary processes. Due to our limited knowledge of the evolutionary history and ecology of the *Spathoglottis* species, explanations of the levels and patterns of genetic diversity within and between populations rely primarily on inferences from molecular data. This data could however be further studied to deepen our understanding of the causes of the genetic makeup of the various Spathoglottis *plicata* populations.

Spathoglottis plicata is an endangered orchid species, classified under Appendix II of the CITES because of its indiscriminate extraction. The increased information on its genetic variation gained through this study could help to guide the conservation of this endangered species (Avise and Hamrick, 1996). Changes in the environment could be a critical factor in genetic variation (Frankel and Soule, 1981). In order to preserve the species in future, and enhance its ability to adapt to environmental changes, there should be a sustained and structured programme for genetic conservation, focusing on maintaining the overall genetic diversity of the species (Fischer and Matthies, 1998; Petit et al., 1998 and Fischer et al., 2000). However, a better understanding first needed of the distribution of genetic variation within the gene pool (Qamaruz-zaman et al., 1998).

The authors of this study believe that AFLP provides an efficient method for a first assessment of genetic diversity and variation in *Spathoglottis plicata*. Qamaruz-zaman et al. (1998) used AFLP to decipher the genetic structure of endangered orchid species of *Orchissimia* in the UK, leading to effective conservation management and a higher effective population size. Pfeifer and Jetschke (2006) also used AFLP for genetic diversity and genetic differentiation within *Himantoglossum hircinum*.

Materials and Methods

Sample collection

Our sampling consisted of five wild populations of *Spathoglottis plicata* located in Peninsular Malaysia (Fig 1), and one vegetatively propagated population obtained from the Selangor Orchid Nursery as control samples. A total of 25 to 30 individuals were sampled in each location by collecting leaf material. In total, 172 accessions from the six populations were sampled (Table 1).

DNA extraction

Genomic DNA of the *Spathoglottis* extracted from fresh leaf material using a Qiagen DNeasy plant mini kit (Qiagen, Germany) according to the manufacturer's instructions. DNA concentration measured using a NanoDrop ND-1000 Spectrophotometer (ISOGEN Lifescience). The quality checked by electrophoresis on a 0.8% agarose gel in tris-borate-EDTA (TBE) buffer.

Amplified Fragment Length Polymorphism (AFLP) protocol

The AFLP method conducted according to Vos et al. (1995) with some modifications. Genomic DNA (300ng) digested with the restriction enzymes *Eco*RI and *Mse*I (Biolegio B, Nijmegen, The Netherlands). *Eco*RI and *Mse*I adapters ligated in a final volume of 50 µl. The reaction contained 0.05 RL-buffer, 10 mM ATP, 50 pMol MseI adapter, 5 pMol EcoR1 adapter, 5 U *Mse*I, 5 U *Eco*R1 and 1 U T4 DNA Ligase (Fermentas, USA). The reactions gently mixed and incubated at 37ºC overnight. After

restriction-ligation, the reaction mixture diluted 10 times with TE buffer (10 mM Tris-HCl; 0.1 mM EDTA pH 8.0).

Pre-amplification was carried out in a final volume of 20 µl, consisting of 30 ng unlabelled Eco-primer (*Eco*RI+A and *Eco*RI+C), 30 ng unlabelled Mse-primer (*Mse*I+C), 5 mM dNTPs (Invitrogen Life Technologies, Breda, The Netherlands), 1x Dream Taq Buffer, 0.4 U Dream Taq (Fermentas, USA) and 5 µl diluted restriction-ligation product. A Polymerase Chain Reaction (PCR) was carried out in a Gene Amp PCR System 2700 (AB Applied Biosystems, USA) programmed at 94ºC for 5 min, followed by 24 cycles denaturation at 94ºC for 30 seconds, 56ºC annealing for 30 seconds, 72ºC extension for 60 seconds and final incubation at 72ºC for 7 min. The pre-amplification products diluted 20 times in TE buffer and used as a template for selective amplification.

Selective amplification was carried out in 10 µl reaction volume containing 50 ng unlabelled Mse-primer, 1 pmol/µl IRD (700 or 800)-labelled Eco-primer, 5 mM dNTPs, 1 X Dream Taq buffer, 0.2 U/µl Dream Taq and 5 µl of diluted pre-amplification product.

The PCR profile was a touchdown PCR carried out in a Gene Amp PCR System 2700 (AB Applied Biosystems, USA) starting at 94ºC for 5 min, followed by 12 cycles denaturation at 94ºC for 30 seconds, 65-56ºC annealing for 30 seconds, with a decrease of 0.7ºC for each cycle; a 72ºC extension for 60 seconds; then continuing with 24 cycles at 94ºC for 30 seconds, 56ºC annealing for 30 seconds, 72ºC extension for 60 seconds and a final incubation at 72ºC for 7 min. Meanwhile, the amplified fragments were separated on a 6.0% denaturing polyacrylamide gel (Long Ranger) mix with Tetramethylethylenediamine (TEMED) and 10% APS, following which they were analysed on a Li-Cor 4300 DNA analyzer (Li-Cor Biosciences, Lincoln, NE, USA).

A total of 56 AFLP primers were initially screened. However, eleven *Eco*RI + 3 base/*Mse*I + 3 base primer sets which showed clearly scorable and highly polymorphic fragments (Table 2) were selected for the fluorescent AFLP reactions in all the 172 samples of six *Spathoglottis plicata* populations.

Data analysis

The AFLP profiles (polymorphic bands) manually scored from Li-Cor Tiff images as present (1) and absent (0) as band (binary) data ranging from 40 to 300 bp. The analysis of the AFLP markers based on the assumption that the AFLP is a dominant marker, with dominant alleles coding at the present band at the given locus and a recessive null-allele coding for the absence of a band.

A statistical analysis between and within the six investigated *Spathoglottis plicata* populations was carried out using analysis of molecular variance (AMOVA-PREP 1.01) (Miller, 1998), (AMOVA 1.55) (Excoffier et al., 1992), Paleontological Statistics (PAST) version 2.03 (Hammer et al., 2001) and STRUCTURE 2.3.3 (Pritchard et al., 2010). Pairwise genetic distances within the six investigated populations and their level of significance were estimated using an AMOVA with 1,000 permutations. A cluster analysis was performed using the unweighted pair group method with arithmetic mean analysis (UPGMA), based on Jaccard's coefficient of similarity for pairwise of all

individuals of the six populations. A dendrogram was drawn of the six *Spathoglottis plicata* populations using Paleontological Statistics (PAST) version 2.03. Bootstrap values (based on 1000 resamplings) were used to estimate the reliability of the clustering pattern.

STRUCTURE 2.3.3 was also used to test which accessions might contain one or more genetic units, via the Bayesian population clustering approach. Markov Chain Monte Carlo (MCMC) simulations were used to numerically approximate the posterior distributions. A Structure computation was carried out in 5 replicates for each user-defined *K* value from 1 to 6, with a burn-in period of 60,000 iterations and MCMC sampling run of 300,000 iterations. The admixture model estimates the proportion of each accession genome that descended from each of the *K* inferred groups. The model choice criterion implemented in STRUCTURE aimed to detect the true *K* by estimating the posterior probability of the data for a given *K*, *Pr (X|K)* (Pritchard et al., 2000).

The true number of populations (*K*) is often identified using the maximum value of *L(K)* returned by the structure (Zeisset and Beebee, 2001; Ciofi et al., 2002; Vernesi et al., 2003; Hampton et al., 2004). However, in most simulations, once the real *K* is reached, *L(K)* at larger *K*s plateaus or continues increasing slightly (a phenomenon mentioned in the STRUCTURE manual, Pritchard and Wen, 2003) and the variance between runs increases.

Estimating the *K* number of the ancestral genetic populations and the ancestry membership proportions of each individual in these clusters were done based on the Evanno method (Evanno et al., 2005). Different runs can produce different likelihood values and different calculations on the data sets carried out to quantify the amount of variation of the likelihood for each *K*. The true number of *K* groups was determined by plotting the mean likelihood of *K* [$L(K)(\pm SD)$]; the rate of change of the likelihood distribution (mean ±SD) calculated as $L'(K) = L(K) - L(K-1)$; absolute values of the second order rate change of the likelihood distribution (mean ±SD): calculated according to the formula $|L''(K)| = |L'(K+1) - L'(K)|$ and Delta *K* calculated as $\Delta K = m|L''(K)|/s|L(K)|$.

Conclusion

Spathoglottis plicata are wild orchids about which little is known on their distribution, adaptation or genetics in Malaysia. The populations studied showed significant variation, indicating possible gene flows, but there were also signs of isolation through distance between these populations in line with their geographical ranges. The data from this study should provide a good basis for further research on conservation planning for this endangered species, taking into account the significant variations discovered between populations, which points to a lack of gene flow between them. This could further tested to implement the traceability of the species necessary for a worldwide management guide and to stimulate further research on the species. Specifically, our discovery of the existence of three sub clusters within Peninsular Malaysia should help support further efforts in conservation management by the forest authorities in order to protect this species. This is vital, as industrialization will

continue to affect forest area coverage, and this type of data could be pivotal in ensuring the survival of this species in the future.

References

Agarwal M, Shrivasta N, Padh H (2008) Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep. 27:617-631.

Avise JC, Hamrick JL (1996) Conservation genetics: case histories from nature. Chapman and Hall, New York.

Chang YK, Veilleux RE, Iqbal MJ (2009) Analysis of genetic variability among *Phalaenopsis* species and hybrids using amplified fragment length polymorphism. J Am Soc Hortic Sci. 134(1):58-66.

Ciofi C, Milinkovitch MC, Gibbs JP, Caccone, A, Powell, JR (2002) Microsatellite analysis of genetic divergence among populations of giant Galápagos tortoises. Mol Ecol. 11:2265–2283.

Comber JB (1990) Orchids of Java. Royal Botanic Gardens. Kew: Bentham-Moxon Trust.

Comber JB (2001) Orchids of Sumatra. Kew Publishing.

Duffy KJ, Scopece G, Cozzolino S, Fay MF, Smith RJ, Stout JC (2009) Ecology and genetic diversity of the dense-flowered orchid, *Neotinea maculata*, at the centre and edge of its range. Ann Bot. 104:507–16.

Duffy KJ, Fay MF, Smith RJ, Stout JC (2011) Population genetics and conservation of the small white orchid, *Pseudorchis albida*, in Ireland. Biology and Environment: Proceedings of the Royal Irish Academy. 111B:1-9.

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

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Applications to human mitochondrial DNA restriction data. Genetics. 131:479–491.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data:dominant markers and null alleles. Mol Ecol Notes. 7(4):574-578.

Fischer M, Matthies D (1998) RAPD variation in relation to population size and plant fitness in the rare *Gentianella germanica* (Gentianaceae). Am J Bot. 85:811–819.

Fischer M, Husi R, Prati D (2000) RAPD variation among and within small and large populations of the rare clonal plant *Ranunculus reptans* (Ranunculaceae). Am J Bot. 87:1128-1137.

Forbes HO (1884) On the contrivances for ensuring self-fertilization in some tropical orchids. In Journal of Linnean Society. Botany. 21:538-549.

Forrest AD, Hollingsworth ML, Hollingworth PM, Sydes C, Bateman RM (2004) Population genetic structure in European populations of *Spiranthes romanzoffiana* set in the context of the genetic studies on orchids. Heredity. 92:218-227.

Frankel OH, Soule ME (1981) Conservation and evolution. Cambridge University Press.

Ginibun FC, Bhassu S, Khalid N, Arens P, Vosman B, Othman RY (2015) A preliminary amplified fragment length polymorphism (AFLP) primer selection for *Spathoglottis* species. Acta Hortic. 1078:167-174.

Ginibun FC, Saad MRM, Hong TL, Othman RY, Khalid N, Bhassu S (2010) Chloroplast DNA barcoding of *Spathoglottis* species for genetic conservation. Acta Hortic. 878:453-459.

Hammer O, Harper DAT, Ryan PD (2001) PAST: Paleontological statistics software package for education and data analysis. Palaeontol Electron. 4(1):9.

Hampton JO, Spencer PBS, Alpers DL, Twigg LE, Woolnough AP, Doust J, Higgs T, Pluske J (2004) Molecular techniques, wildlife management and the importance of genetic population structure and dispersal: a case study with feral pigs. J Appl Ecol. 41:735–743.

Hamrick Jl, Godt MJW (1989) Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding and genetic resourses, Sunderland, MA: Sinauer Associates, p 43–63.

Hamrick JL, Godt MJW (1990) Allozyme diversity in the grasses. In: Cheplick G (ed) Population biology of grasses, Cambridge University Press, Cambridge, p 11-29.

Hedren M, Fay MF, Chase MW (2001) Amplified fragment length polymorphisms (AFLP) reveal details of polyploid evolution in *Dactylorhiza* (Orchidaceae). Am J Bot. 88(10):1868–1880.

Holttum RE (1953) A revised flora of Malaya: An illustrated systematic account of the Malayan flora, including commonly cultivated plants, volume I Orchids of Malaya, Government printing office, Singapore.

Jacquemyn H, Vandepitte K, Brys R, Honnay O, Rolda N-Ruiz I (2007) Fitness variation and genetic diversity of small, remnant populations of the food deceptive orchid *Orchis purpurea*. Biol Conserv. 139:203–210.

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

Jones CJ, Edwards K J, Castiglione S, Winfield MO, Sala F, Van De Wiel C, Bredemeijer G, Vosman B, Matthes M, Daly A, Brettschneider R, Bettini P, Buiatti M, Maestri E, Malcevschi A, Marmiroli N, Aert R, Volckaert G, Rueda J, Linacero R, Vazquez A, Karp A (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a Network of European laboratories. Mol Breeding 3:381-390.

Lord EM (1981) Cleistogamy: A tool for the study on flora morphogenesis, function and evaluation. Bot Rev. 47:421-449.

Miller MP (1998) AMOVA-PREP 1.01: A program for the preparation of AMOVA input files from dominant-markers raw data. Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ.

Nei M (1972) Genetic distance between populations. Am Nat. 106:283–292.

Nieuwolt S, Zaki MG, Gopinathan B (1982) Agro-ecological regions in Peninsular Malaysia. MARDI, Serdang, Selangor, Malaysia, p 22.

Nieuwolt S (1982) Climate and agricultural planning in Peninsular Malaysia. MARDI Special Report No. ASAS-01-83, p 141.

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conserv Biol. 12:844–855.

Pfeifer M, Jetschke G (2006) Influence of geographical isolation on genetic diversity of *Himantoglossum hircinum* (Orchidaceae). Folia Geobot. 41:3-20.

Pillon Y, Qamaruz-Zaman F, Fay MF, Hendoux F, Piquo Y (2007) Genetic diversity and ecological differentiation in the endangered fen orchid (*Liparis loeselii*). Conserv Genet. 8:177-184.

Pimonrat P, Suraninpong P, Wuthisuthimethavee S (2012) Acute effect of gamma radiation on stable characteristic of *Spathoglottis plicata* Blume. Acta Hortic. 953:173-180.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotypes. Genetics. 155:945–959.

Pritchard JK, Wen W (2003) Documentation for STRUCTURE software: Version 2. http://pritch.bsd.uchicago.edu.

Pritchard JK, Wen X, Falush D (2010) Documentation for the structure: Version 2.3.3. http://pritch.bsd.uchicago.edu/structure.html.

Qamaruz-Zaman, F, Fay MF, Parker JS, Chase MW (1998) The use of AFLP fingerprinting in conservation genetics: a case study of *Orchis simia* (Orchidaceae). Lindleyana. 13:125–133.

Ridley HN (1907) Material for a flora of the Malay Peninsula. Methodist Publishing House Singapore.

Ridley HN (1930) The dispersal of plant throughout the world. L. Reeve and Co.

Romeida A, Sutjahjo SH, Purwito A, Sukma D, Rustikawati (2012). Genetic variation of *Spathoglottis plicata* Blume.

orchid mutants based on ISSR markers. J Agron Indonesia. 40(3):218-224.

Sathyanarayana N, Leelambia M, Mahesh S, Jaheer M (2011) AFLP assessment of genetic diversity among Indian Mucuna accessions. Physiol Mol Biol Pla. 17:171-180.

Seidenfaden G, Wood JJ (1992) The Orchids of Peninsular Malaysia and Singapore, Olsen and Olsen, Fredensburg, Denmark.

Setiawan A, Shochicha AN, Pramana AAC, Restiyanti, Daryono BD (2013) Pengembangan marka molekular untuk karakterisasi varietas anggrek tanah unggul (*Spathoglottis*) hasil poliploidisasi dengan kolkisin. artikel.dikti.go.id/index.php/PKM-P/article/view/11.

Smith ZF, James EA, McLean CB (2007) Investigation of the phylogenetic relationships between *Diuris fragantissima* and its closest relatives using AFLPs. Environmentalist. 27:217-226.

Tali K, Fay MF, Bateman RM (2006) Little genetic differentiation across Europe between early-flowering and late-flowering populations of the rapidly declining orchid *Neotinea ustulata*. Biol J Linn Soc. 87:13-25.

Teo CKH (1985) Native orchids of Peninsular Malaysia. Times books International Singapore

Vernesi C, Crestanello B, Pecchioli E, Tartari D, Caramelli D, Hauffe H, Bertorelle G (2003) The genetic impact of demographic decline and reintroduction in the wild boar (Sus scrofa): a microsatellite analysis. Mol Ecol. 12:585–595.

Vos P, Hogers R, Bleeker M, Reijans M, Van De Lee T, Hones M, Fritjers A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP-a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-4414.

Wright S. (1951) The genetical structure of populations. Ann Eugenic. 15(4):323-354.

Xiang N, Hong Y, Lam-Chan LT (2003) Genetic analysis of tropical orchids hybrid (*Dendrobium*) with fluorescence amplified fragment length polymorphism (AFLP). J Am Soc Hortic Sci. 128(5):731-735.

Yong HS (1993) Orchid flowers with male fruit-fly attractants. Malayan Orchids Bulletin. 6:11-13.

Zeisset I, Beebee TJC (2001) Determination of biogeographical range: an application of molecular phylogeography to the European pool frog *Rana lessonae*. Proceedings of the Royal Society B, Biological Sciences. 268:933–938.