University of Reading

SYSTEMATICS OF PENINSULAR MALAYSIAN SCALY TREE FERNS (CYATHEACEAE): PHYLOGENETICS, COMPUTER-AIDED IDENTIFICATION AND CONSERVATION

A thesis submitted by

Azi Azeyanty Jamaludin

For the Degree of Doctor of Philosophy

Centre of Plant Diversity and Systematics,
School of Biological Sciences, University of Reading

May 2017

DECLARATION

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Azi Azeyanty Jamaludin

بسم الله الرحمن الرحيم

”تعلموا العلم وتعلموا السكينة والوقار“
عمر بن الخطاب

In the name of Allah, the Beneficent, the Merciful

“Acquire knowledge, and learn tranquillity and dignity”
Omar ibn al-Khattab

Untuk suamiku tercinta, Mama dan Abah.

Ayden dan Aaron.

ACKNOWLEDGEMENTS

In the name of God, Most Gracious, Most Merciful.

I am owing so much to the Ministry of Higher Education Malaysia (MOHE) for trusting me with the scholarship and making it possible for me to complete my PhD in the University of Reading here in the UK. I am also forever in debt to my employer, Sultan Idris Education University (SIEU) for giving me a chance to pursue my dreams by giving me a paid study leave, as well as an allowance for my family so that I can come and study here without have to worry about anything back home. Not to forget, thank you for letting me used all of the necessary equipment in the SIEU biological laboratory during my sampling expedition.

Most of the works I did for the research would not be possible without the guidance and advice from my supervisor Dr Alastair Culham whom apart from that, did continuously giving me support and encouragement. I am deeply grateful as he has so much faith to let me be independent and letting me explored the research at my own pace. I had learned to work things out for myself and not to rely so much on others, and that for me are possibly the most important life-lessons from the past three and a half years.

Special thanks to Culham's Research Group (2012-2016), as I learn a lot about research related things from the members. To Dr Maria Christodoulou, thank you for selflessly being my unpaid thesis proof-reader. I am deeply thankful to the Peninsular Malaysia Department of Forestry for giving me permits and access to the States Forestry so that I can collect my Cyatheaceae samples. I would like to thanks, Dr Haja Maideen and Mrs Norazlinda Mohamad from the National University of Malaysia Herbarium and Dr Richard Chung from the Forest Research Institute Malaysia (FRIM) for their willingness to help me with the herbarium specimens.

To my best friend Salwa Shahimi, thank you for being you. I owed you a lot. Only Allah SWT can repay you, if not here maybe in the hereafter. May your PhD journey will be successful.

I owe my very important thanks to my beloved husband, Mohd Hidayat for his unselfishness to quit his well-paid job and accompanied me here. Thank you, Abang for your unfailing love, support and understanding. To my sons Ayden Harith, and Aaron Haydar, you are the reason I did not buy a one-way ticket home.

Last but not least, I would like to express my deepest gratitude to my beloved parents, Rohizan and Jamaludin, and my parents in law Wasa'aran and Mior Abbas for their love, encouragement and infinite prayers.

Thank you all.

ABSTRACT

The work presented in this thesis evaluates the status of Peninsular Malaysian Cyatheaceae and used molecular and morphological identification tools for the local species. 419 Cyatheaceae frond sample were collected from the widest possible range of Peninsular Malaysia to obtain material for morphology and molecular study. 15 *Cyathea* species were identified and the species information for Peninsular Malaysia was updated. The species was incorporated into the existing Cyatheaceae phylogeny by using four plastid regions: *matK, rbcL, trnG-trnR* and *trnL-trnF*. Bayesian MCMC analysis of the concatenated sequence data resulted in a 50% majority rule consensus tree confirm the placement of the four groups: *Cyathea, Alsophila*, *Gymnosphaera* and *Sphaeropteris* in the family. However, the resulting tree representing nested monophyletic groups, proposing Cyatheaceae to be monogeneric, i.e., *Cyathea* with two large groups: *Cyathea* and *Sphaeropteris.* The same plastid regions were then evaluated to develop DNA barcodes. *trnL-trnF* was proposed as a barcode for this family as it almost satisfied the three most important criteria: primer universality, sequence quality and species discrimination. This research also developed a multi-access key for Cyatheaceae field identification based on fifteen taxa identified, by extensive field sampling of the currently recognised species. All of the Cyatheaceae species recognises in this study had also been assessed for the conservation status based on the IUCN Red List criteria. Nine species fall under Least Concern (LC), four species are Near Threatened (NT) and two species are Vulnerable (VU). The thorough knowledge regarding Cyatheaceae in Peninsular Malaysia gained through the work done in this research will benefit in making appropriate conservation strategies for the survival of this family. Overall, the most important outcome of this research was the combination of morphology and molecular data for the purpose of updating taxonomy, identification and conservation of the Cyatheaceae family in Peninsular Malaysia.

TABLE OF CONTENTS

CHAPTER 1 1

GENERAL INTRODUCTION 1

1.1 Malaysian Biodiversity 1

1.2 Study of Pteridophytes in Malaysia 4

1.3 The Scaly Tree Ferns: Cyatheaceae 5

1.4 Generic Delimitation in Cyatheaceae 8

1.5 Importance to Conservation 12

1.6 Thesis Structure 15

1.7 Research Objectives 15

CHAPTER 2 17

FIELD BASED EVALUATION OF CYATHEACEAE IN PENINSULAR MALAYSIA 17

2.1 Introduction 17

2.2 Materials and Methods 18

2.3 Results 34

2.4 Discussion 70

2.5 Conclusion 72

CHAPTER 3 73

THE PHYLOGENETIC RELATIONSHIPS OF PENINSULAR MALAYSIA CYATHEACEAE 73

3.1 Introduction 73

3.2 Materials and Methods 80

3.3 Results 90

3.4 Discussion 99

3.5 Conclusion 102

CHAPTER 4 103

THE DEVELOPMENT OF CHLOROPLAST DNA BARCODING MARKERS FOR PENINSULAR MALAYSIAN CYATHEACEAE SPECIES IDENTIFICATION 103

4.1 Introduction 103

4.2 Evaluation of Some of the Core Chloroplast Coding Regions Used for Ferns 106

4.3 Aims 108

4.4 Materials and Methods 108

4.5 Results 110

4.6 Discussion 133

4.7 Conclusion 135

CHAPTER 5 136

THE DEVELOPMENT OF AN INTERACTIVE MULTI-ACCESS KEY FOR IDENTIFYING PENINSULAR MALAYSIAN CYATHEACEAE 136

5.1 Introduction 136

5.2 Materials and Methods 139

5.3 Results 149

5.4 Discussion 151

5.5 Conclusion 152

CHAPTER 6 153

UPDATING THE PENINSULAR MALAYSIA CYATHEACEAE SPECIES STATUS USING THE IUCN RED LIST CRITERIA 153

6.1 Introduction 153

6.2 Materials and Methods 156

6.3 Results 165

6.4 Discussion 180

6.5 Conclusion 181

CHAPTER 7 182

GENERAL DISCUSSSION 182

REFERENCES 187

Appendix 1 197

Appendix 2 199

Appendix 3 201

Appendix 4 204

CHAPTER 1

GENERAL INTRODUCTION

1.1 Malaysian Biodiversity

Malaysia (Figure 1.1) belongs to the Sundaland biogeographical region which includes the Sunda shelf, a part of the Asian continental shelf that was uncovered during the last glacial period of the Pleistocene (Hall, 1998). It consists of an island part and a part attached to mainland Asia separated by 540 km of the South China Sea. The land area covers approximately 33.27 million hectares (MNRE, 2006; MNRE, 2014; MNRE, 2016). These two parts of Malaysia share a similar landscape that features coastal plains rising to hills and mountains, including Mount Kinabalu at 4095 meters, a UNESCO World Heritage Site and the highest mountain in South East Asia (Metcalfe, 2002; MNRE, 2014; MNRE, 2016). The local climate is equatorial, with temperature ranging from 21°C to 32°C and annual rainfall of 250 cm, along with high humidity and annual southwest (April to October) and northeast (October to February) monsoons (MNRE, 2006; MNRE, 2014; MNRE, 2016; Richmond, 2010).

Figure 1.1 Map showing the location of Malaysia and surrounding countries. The orange line marks the Sundaland shelf today. Source: Google map.

The country is one of the twelve most mega-biodiverse countries in the world (Lee *et al.*, 2010) with more than 170,000 species (Table 1.1), including many endemics with more than 80% endemism occurring in the peninsula alone (Secretariat of the Convention on Biological Diversity, 2010; MNRE, 2014). Much of its diversity survives because two thirds of the land is covered with heavily forested tropical rainforest, parts of which are up to 130 million years old (Lee *et al.*, 2010). There are about 15,000 known species of flowering plants, and more than 1,100 ferns and fern allies occurring in Malaysia (Bidin and Jaman, 1999; MNRE, 2014).

Table 1.1 Summary of Malaysia's overall biodiversity richness (MNRE, 2014).

Group	**Estimated Species**
Mammals	306
Birds	742
Reptiles	567
Amphibians	242
Marine Fishes	1,619
Freshwater Fishes	449
Invertebrates	150,000
Vascular Plants	15,000
Fungi	4,000
Mosses	522
Hard Corals	612

Malaysia is an active party to the Convention on Biological Diversity (CBD) which it ratified in 1994 (Napis *et al.,* 2001). Since then, the National Policy on Biological Diversity had been developed (MSET, 1998) alongside other policies with biodiversity conservation as a focal part of sustainable development (MNRE, 2014). It is also committed to maintain at least 50% of the land area under forest and tree cover in perpetuity and up until 2012, approximately 21.01 million hectares of the country remained forested with 14.5 million hectares designated as permanent forest reserve (MNRE, 2014; MNRE, 2016).

Since its independence in 1957, Malaysia underwent rapid socio-economic growth, which resulted in heavy deforestation (Napis *et al.*, 2001). Activities such as logging and hydroelectricity schemes, led to the endangerment of local biodiversity, raising concerns on the conservation status of species present (Napis *et al.*, 2001; MNRE, 2014). Even though policies for sustainable development are in place, there are few appropriately qualified scientists to monitor progress. The present work is one of the few that focuses on the taxonomic treatment of a plant family susceptible to

development activities. Phylogenetic approaches, such as Bayesian MCMC and DNA Barcoding analysis are used in the evaluation of the species, as well as reviewing the IUCN Red List status for these species and developing a Multi-Access Key for better conservation measurements.

1.2 Study of Pteridophytes in Malaysia

There have been several studies of Malaysian pteridophytes, notably work started by Alfred Russel Wallace in the mid-1800s (Cicuzza, 2014) followed by Ridley (1908, 1912, 1926), and then by Holttum (1963, 1966, 1968) for Flora Malesiana. The fern taxonomy of Malaysia specifically was updated by Bidin (1983, 1985, 1987) in the 1980s. Parris and Latiff (1997) suggested that the overall count of pteridophytes at the time of their study was 1,136 species, 637 of which occurred in Peninsular Malaysia, 718 species in Sabah, and 587 species in Sarawak. There is no current complete key to Malaysian pteridophytes. Efforts in cataloguing plant species in Malaysia, have concentrated on woody plants due to their economic value while pteridophytes have been comparatively neglected. This fact along with a small number of research publications recently show a lack of pteridology expertise in Malaysia even though many species are thought to be threatened.

1.3 The Scaly Tree Ferns: Cyatheaceae

Cyatheaceae, in the order Cyatheales, forms part of the subclass Polypodiidae which includes most of the world's fern diversity (Schuettpelz and Pryer, 2007; Carl J Rothfels *et al.*, 2012; Christenhusz and Chase, 2014). It has trunk-like, erect stems which elevate the fronds above the ground (Figure 1.2) and includes 500 of the estimated 700 species of tree ferns (Conant *et al.*, 1994), along with Metaxyaceae, Dicksoniaceae, and Cibotiaceae (Korall *et al.*, 2006). Regions that are rich in species include the Greater Antilles, Central America, the northern part of Andes including Venezuela, Colombia, Ecuador, Peru, Madagascar, Borneo, Sumatra, the Philippines, and New Guinea (Tryon, 1970; Tryon and Gastony, 1975). Many of the species have confined ranges with few occurring in more than one of these regions (Conant *et al.*, 1995). Even though the geographic ranges of the species are known, the genera ranges are not as there is a lack of agreement on generic boundaries (Conant *et al.*, 1995). This lack of consensus on generic restriction is shown by studies of Tryon and Tryon (1982), Holttum and Edwards (1983) and Lellinger (1987) in which six, one and four genera were recognized, respectively.

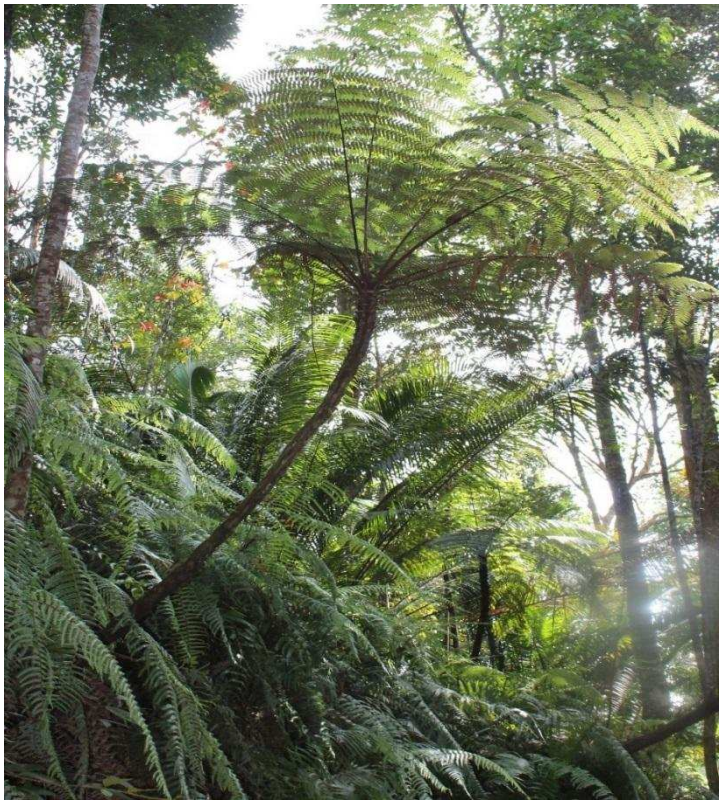


Figure 1.2 Cyathea sp. in Bukit Larut, Perak. Photo© 2013 Azi Jamaludin.

Most of the species are forest plants (Holttum, 1963) and include some of the tallest existing ferns, reaching over 20 m tall (Holttum, 1963; Lehnert, 2009; Korall *et al.*, 2007). The members of this family can be distinguished from the other families by having not just the general pluricellular hairs, but also the presence of different types of scales on their stems and petioles (Figure 1.3) (Kramer, 1990; Korall *et al.*, 2007). However, understanding the relationships between the genera within the family is problematic, since the focus of identification and classification had always been dependent on the scales and indusia morphologies (Korall *et al.*, 2007). These morphological characters have been considered to be frequently subject to homoplasy and of less value in defining major groups of Cyatheaceae (Holttum and Edwards, 1983; Korall *et al.*, 2007).

A. *Cyathea latebrosa* collected in Genting Highlands, Pahang.

B. *Cyathea contaminans* collected in Genting Highlands, Pahang.

Figure 1.3 ***A*** *and* ***B*** *shows two distinctive species of the scaly tree ferns with the presence of scales on their stems and petioles. Photo© 2013 Azi Jamaludin.*

Holttum (1963) counted a total of 36 *Cyathea* species in Malaysia, of which nine species are from Peninsular Malaysia, four species in Sarawak and eight species in Sabah. Eight species can be found throughout Malaysia, five species occurred both in Sarawak and Sabah, while one species can be found in Peninsular Malaysia and Sabah, and Peninsular Malaysia and Sarawak respectively (Holttum, 1963). The Malesian Cyatheaceae was divided by Holttum (1963) into three subfamilies: Cyatheoideae, Cibotioideae and Thyrsopteridoideae and outside Malesia, Metaxyoideae (Latiff, 2015; Holttum, 1963). Holttum (1963) recognised two

subgenera in *Cyathea*: *Cyathea* and *Sphaeropteris* with the latter further divided into two sections: *Sphaeropteris* and *Schizocaena*; and four subsections: *Sphaeropteris, Fourniera, Schizocaena,* and *Sarcopholus.*

1.4 Generic Delimitation in Cyatheaceae

Cyatheaceae have long enthralled scientists and have been the subject of many systematic and taxonomic treatments (Figure 1.4) (Conant *et al.*, 1994; Conant *et al.*, 1995; Conant and Stein, 2001; Korall *et al.*, 2007).

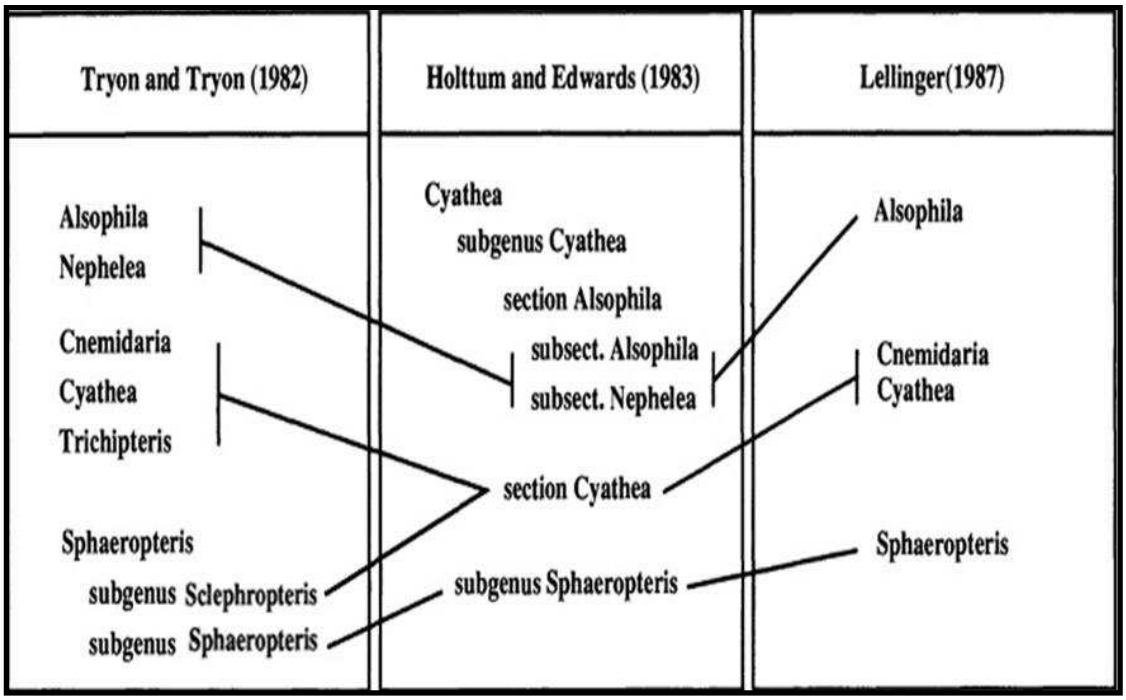


Figure 1.4 Classification systems proposed for the Cyatheaceae (Conant et al., 1994).

Christensen (1905) had separated the family into *Cyathea, Hemitelia* and *Alsophila* based on whether the indusium completely or partially covers the sorus, or is absent altogether. In his study, Christensen (1905) also included *Lophosoria* and *Metaxya* in

Alsophila, but later these two genera (*Lophosoria* and *Metaxya*) were discovered to be remotely associated to the major genera of the family such as *Cyathea, Alsophila* and *Sphaeropteris*.

However, Holttum (1957, 1964) focusing on the Malesian region discovered that the petiole scales provided a useful taxonomic character in classifying the *Cyathea* species. This was because they are associated with other morphological characters and ecological preferences (Holttum, 1957; Holttum, 1964). He also stated that indusium type was not an important general character because it varies broadly within a few species (Holttum, 1957; Holttum, 1964). This led to the family being revised where he used the scales characters to define subgenera and sections, before later proposing *Cyathea* as a single genus (Holttum, 1963; Holttum and Edwards, 1983).

In his study, Tryon (1970) divided the entire Cyatheaceae family into six genera based on the morphological characters of the scales used by Holttum (1957) as well as the presence and absence of indusia and venation characteristics. Tryon (1970) found that indusia ascended from scales on the leaf abaxial surface remote from the margin. Thus he concluded that the indusium should be a derived character in which the absence of the character will be regarded as a primeval state within the family (Tryon, 1970). Tryon's work focused majorly on Neotropical species, contrasting with Holttum who mostly worked on Old World taxa. The six genera that Tryon (1970); Tryon and Tryon (1982) proposed are *Alsophila, Nephelea, Cnemidaria, Cyathea, Trichipteris*, and *Sphaeropteris.*

Even after the studies made by Holttum (1957), Tryon (1970), Tryon and Tryon (1982) and Holttum and Edwards (1983), the classification of Cyatheaceae remains unresolved. This brings Lellinger (1987) to recognise four genera in his study: *Alsophila* (including *Nephelea*), *Cnemidaria*, *Cyathea* (including *Trichipteris*), and *Sphaeropteris*. Lellinger (1987) argued that occasional hybrids occur within *Alsophila* and *Cyathea* as well as between *Cnemidaria* and *Cyathea* but the characteristics of *Alsophila* and *Cnemidaria* were sufficiently different from *Cyathea* to distinguish the genera readily.

There was a clear disagreement between the authors over the relationships and character evolution within this family over the years. With the emergence and advancement of molecular study, investigation using phylogenetic approach led Conant *et al.* (1994, 1995) and Stein, Conant and Valinski (1997) to divide this family into three genera: *Alsophila*, *Cyathea*, and *Sphaeropteris* with *Alsophila* being the most basal group in the family. This classification was used in many Cyatheaceae related studies until 2006 (Conant and Stein, 2001; Smith *et al.*, 2003; Korall *et al.*, 2006).

Smith *et al.* (2006) revised the classification of extant ferns and recognized five genera namely *Alsophila* (including *Nephelea*), *Cyathea* (including *Cnemidaria, Hemitelia*, *Trichipteris*), *Gymnosphaera*, *Hymenophyllopsis* and *Sphaeropteris* (including *Fourniera*). Until then *Hymenophyllopsis* was placed in a monogeneric family (Tryon and Tryon, 1982). Analysis by Wolf *et al.* (1999) suggested a close and well-supported relationship of *Hymenophyllopsis* to Cyatheaceae based on two species: *Hymenophyllopsis hymenophylloides* and *H. dejecta*. Conant and Stein (2001)

and Korall et al. (2007) suggested that *Alsophila* should be divided into two groups: *Alsophila* and *Gymnosphaera* based on broader species and morphology sampling.

Korall et al. (2007) studied the morphology of the scales and indusia based on previous studies by Holttum (1963), Tryon (1970), and Conant et al. (1994), (1995), along with a molecular phylogeny, and separated four groups based on the type of scales (Figure 1.5) and indusia. Korall et al. (2007) then proposed the four groups as genera: *Cyathea, Alsophila, Gymnosphaera* and *Sphaeropteris* with the latter as sister to all others.

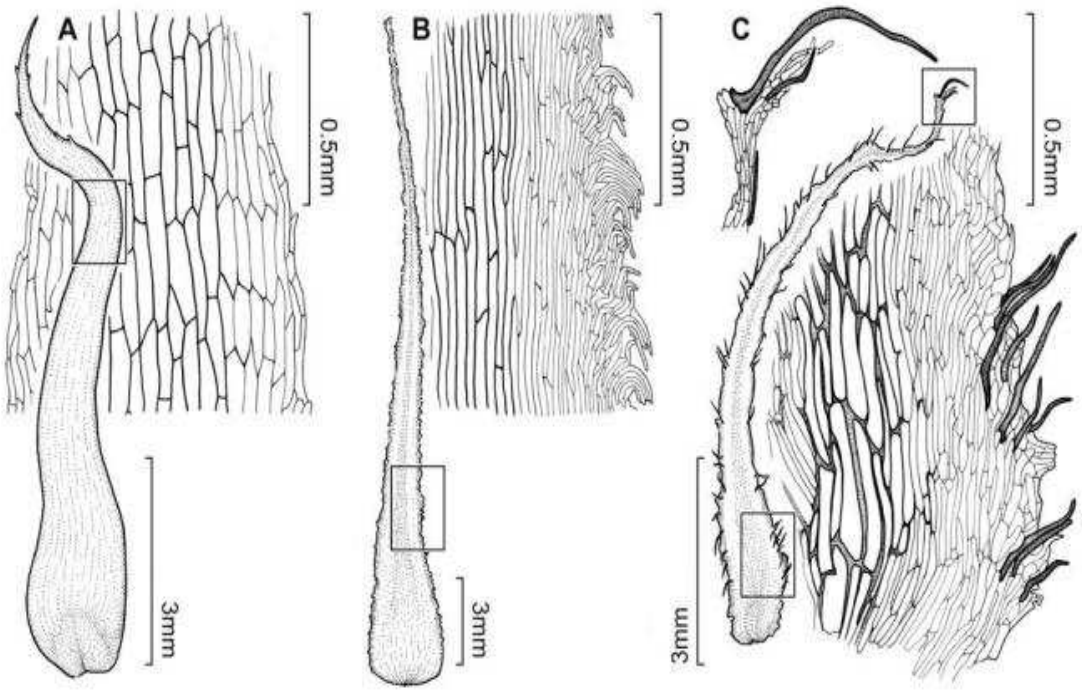


Figure 1.5 (A) Conform scale in Sphaeropteris (B) Marginate scale without apical seta in Cyathea (C) Marginate scale with apical seta in Alsophila (Korall et al., 2007).

1.5 Importance to Conservation

Cyatheaceae and Dicksoniaceae were listed in the Convention on International Trade in Endangered Species (CITES) in 1975 (Oldfield, 1995). Tree ferns have long been used for many socio-economic purposes such as construction, horticulture (Figure 1.6), food, and medicine (Large and Braggins, 2004; Rout *et al.*, 2009) resulting in their heavy exploitation as a source of income (Large and Braggins, 2004). There are many common, non-threatened species used for trading, such as *Cyathea arborea*, *C. biformis*, *C. lepifera*, *Dicksonia antartica*, *D. fibrosa* and *D. sellowiana* (CITES, 2013). However, there are species that may have been threatened locally, mainly because of habitat destruction but there is a need to monitor the species that may be threatened because of the trade (CITES, 2013).

A **B**

Figure 1.6 Examples of the uses of Cyathea species. ***A.*** *Part of the trunk made into ornamental bowl.* ***B****. Roots that have been compressed to be made into orchid mounting medium.*

This group is also ecologically important as a study by Ashton (2000) suggested that the trunks of the tree ferns were favourable sites for the establishment of ground and epiphytic ferns. Another study by Lindenmayer et al. (1994) found that the numbers of mountain short-eared possum increased as the numbers of *C. australis* and *D. antartica* increased. Blows and Schwarz (1991) found that dried fronds of *C. australis* were a favourite living site for *Exoneura bicolor* bees. Fountain-Jones, McQuillan and Grove (2012) observed and sampled 80 individuals of *D. antartica* on which they discovered a total of 108 species of beetles, representing 35 families, living in discrete microhabitats of crown litter, live fronds and trunk. Also, species such as *C. contaminans* can be used as an indicator of forest disturbance in Malaysia highlands as it can be found abundantly inside clearings (personal observation).

Trade-reporting relies on the correct identification of species in the field and correct usage of species names in CITES. The problems arise when different countries tend to report the tree fern trade at different taxonomic levels and use different names. At present, tree fern conservation status has not been updated in Malaysia, specifically none of the species from Malaysian Cyatheaceae have been evaluated for IUCN Red Listing (IUCN, 2015). The lack of effort in updating the conservation status may be due to lack of local expertise in this field. This is where the current work will help re-evaluate the Peninsular Malaysia Cyatheaceae by adapting current taxonomies with modern technologies. It is hoped that this work will contribute towards the better understanding of the overall phylogenetic knowledge which may contribute for better conservation efforts.

DNA barcoding has not only been used as a tool for species identification but also for species discovery as well as clarifying the taxonomic relationships between species (Lahaye *et al.*, 2008). The knowledge acquired will be useful in making appropriate conservation plans for this family in Malaysia (Liao *et al.*, 2011).

While taxonomists work extensively, solving problems affecting trade-reporting and present their findings in journals, keen general users such as local plant collectors and plant nursery traders are sometimes left with insufficient species identification information. Trade-reporting and all of its related fields depend on species identification keys being precise and usable. Most of the dichotomous printed keys are written by taxonomists for similar users in the field, often with very little additional explanation, resulting in difficulties for novice users to access the species information (Lindsay and Middleton, 2009).

Although trade surveys and monitoring rely on experts such as taxonomists, field staff, and wildlife officers (CITES, 2013), general users who are interested in preserving biodiversity can also help by reporting any irregular trade activities to the authorities. The development of a multi-access key for Cyatheaceae in this work aims to facilitate species identification as well as attract the interest of a broader range of people and professions into knowing this family.

1.6 Thesis Structure

This dissertation will be structured based on two aspects: Chapters 2, 5 and 6 focuses on using morphological data to develop electronic key and assessing the species conservation status. Chapters 3, and 4 used molecular data to update the phylogeny, and proposing a DNA barcode markers. As all of the chapters rely heavily on the right identification of *Cyathea* species, Chapter 2 will be the most important as it will determine the research continuation in the succeeding chapters. Finally, the findings of this thesis will be discussed in greater detail in Chapter 7: general discussions.

1.7 Research Objectives

Even though the information on this family has developed over the years, its relationships have not yet been thoroughly understood. This study will aid further in resolving both species and generic identities in the scaly tree ferns. The more specific research objectives include;

- To investigate the phylogenetic relationships of Peninsular Malaysian Cyatheaceae based on DNA sequence data from four plastid DNA regions (*rbcL*, *matK*, *trnG-trnR*, and *trnL-trnF*) to contribute towards resolving and supporting the overall phylogenetic knowledge of the family.
- To develop a barcode based species identification tool.

- To gather the morphological data of identified *Cyathea* species and construct an interactive multi-access key using LucID software to help others identify the species.
- To evaluate and update *Cyathea* species status in the IUCN Red List for better understanding of the conservation status of the family and to help guide conservation measures.

CHAPTER 2

FIELD BASED EVALUATION OF CYATHEACEAE IN PENINSULAR MALAYSIA

2.1 Introduction

Scientific evaluation in the field is important because it allows the observation of the field conditions of the plants under study and the evaluation of any immediate threats to their habitat as well as providing an insight into the natural variation of species. Even though there are 191 *Cyathea* species reported by Holttum (1963) occurring in Malesia, of which 41 species are from Malaysia and 21 species from Peninsular Malaysia, no local scientists specialize in this family. Thus the identification of specimens for the current work relied on the key from Flora Malesiana Series II: Pteridophyta (Holttum, 1963) as well as visual comparison using herbarium specimens from Malaysia National University (UKM) and Kew (K) herbaria .

This fieldwork aims to gather population samples of the widest possible range of Cyatheaceae species from Peninsular Malaysia to provide material for morphological and molecular study. To understand genetic variation within a species and to detect genetic discontinuities between species, a structured sampling strategy is needed that allows investigation of DNA variation within and between populations as well as among species. Herbarium sampling is generally not designed to this purpose as it is focused on species as pre-agreed entities.

2.2 Materials and Methods

Population-level collections were focused on Peninsular Malaysia to allow thorough analysis of the area. Field sampling in the peninsula forests was based on information regarding the species and locality from local floras and herbaria. The locations visited were based on records of previous sightings, but altered following cross-checking with current satellite maps and consultation with local forestry officers. Most of the initial locations had been developed for tourism or agricultural schemes and logging activities, destroying not only the Cyatheaceae populations, but other important pteridophyte families (MNRE, 2006; MNRE, 2014; MNRE, 2016).

The family is recorded under CITES and most of the species are listed in Appendix II, meaning the species are not currently threatened with extinction but trade must be controlled in order to avoid utilization incompatible with their survival (CITES, 2013). Due to this, research and collection permits were obtained from each forestry state department with the condition of collecting a single frond per individual.

2.2.1 Sample Collection

A sampling expedition was undertaken from early September 2013 until late October 2013. The locations are shown in Figure 2.1. The expedition started in Bukit Larut, Perak (1) and continued north to Penang Hill, Penang (2), Mount Perlis, Perlis (3) and Mount Jerai, Kedah (4). The journey east started at Fraser's Hill (5), Mount Berinchang (6) and Genting Highlands (7) all in Pahang, continued with Lake Kenyir, Terengganu (8) and Lojing Highlands, Kelantan (9). The west covered Batang Kali,

Selangor (10) and Mount Angsi, Negeri Sembilan (11). The collection ended in Mount Ledang, Johor (12) in the south. Additional samples were provided courtesy of the National University of Malaysia (UKM) from Fraser's Hill (5) and Bangi Forest, Selangor (HB) (13).

Populations were sampled according to accessibility, ensuring at least 5m gap between samples. A minimum of 10 individuals was collected per population to allow detection of within-population genetic variation. The sampling size was adjusted accordingly, depending on the locations to ensure that sufficient samples were collected without endangering small populations. Parts of the frond, scales and sporangia (where available) were collected for voucher specimens. A machete and secateurs were used to detach the fronds and a long pole with attached secateurs was used for out-of-reach samples. Photographs of each sample were taken in the field, longitude and latitude readings were noted as well as the elevation above sea level.

Figure 2.1 Map showing Peninsular Malaysia and neighbouring countries. The collection sites are marked as white dots and numbered, while state names in white capitals.

The samples were labelled and a small part of the pinna was taken and placed into a re-sealable plastic bag containing silica-gel intended for molecular work (Figure 2.2). The remainder was sprayed with 75% ethanol. Once collection was finished, the samples were sorted and stacked flat between A3-sized corrugated cardboard before being tightly pressed using a wooden press, ready to be dried. Samples were taken to Sultan Idris Educational University (SIEU) laboratory to be placed in an oven at 40°C for seven days (Figure 2.3) before being transferred using air-mail to the School of Biological Sciences, University of Reading. Once arrived, the samples were placed in the freezer at -20°C for 72 hours for decontamination.

Figure 2.2 Silica-dried sample preserved for molecular work. Photo© 2013 Azi Jamaludin.

Figure 2.3 Samples left for drying in the oven. Photo© 2013 Azi Jamaludin.

2.2.2 Sampling Routes (according to the numbers on previous map)

2.2.2.1 Bukit Larut, Perak (BL)/ 1

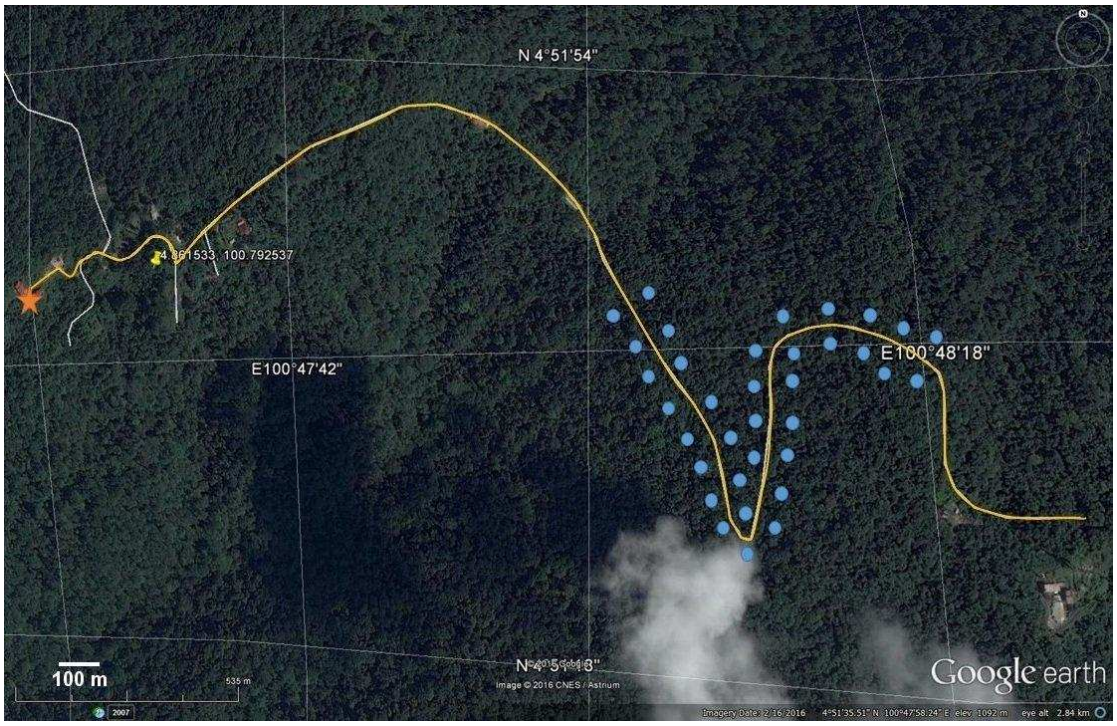


Figure 2.4 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 680m elevation and continued until 1349m. Samples were collected on one side of the trail going up and other side going down with a final sample size of 35. There were no more samples found near the summit as the area was cleared for a communications tower.

2.2.2.2 Penang Hill, Penang (PH)/ 2

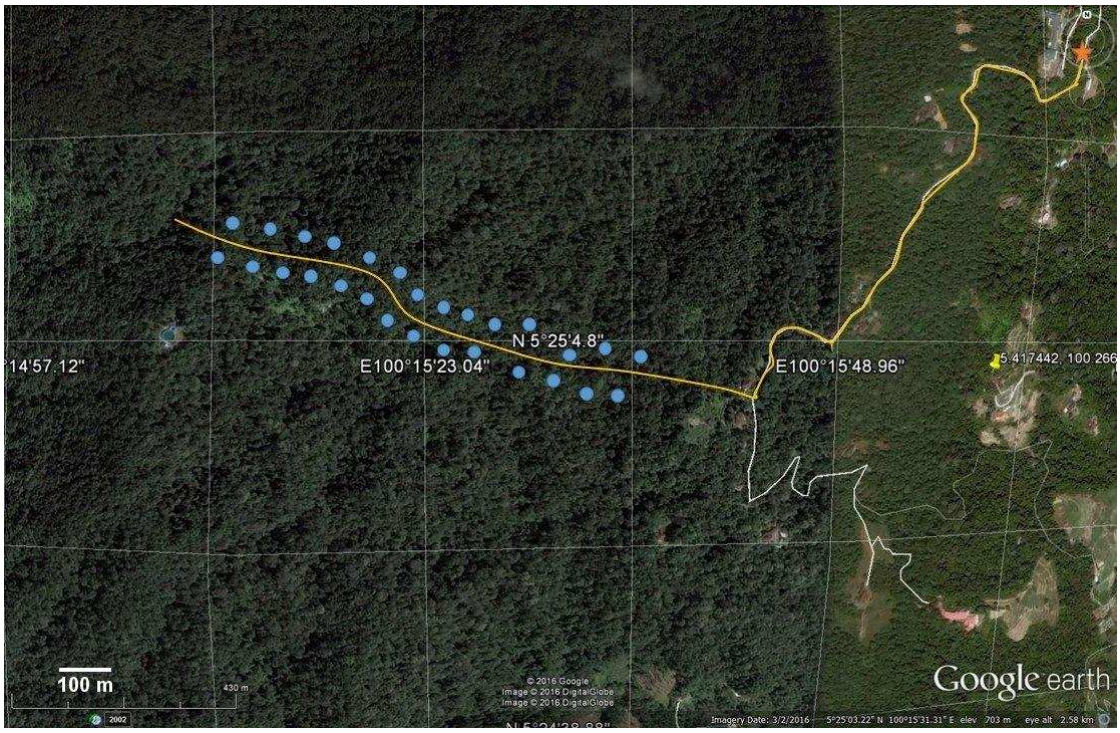


Figure 2.5 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 702m elevation and continued until 755m. Samples were collected on one side of the trail going up and other side going down with a final sample size of 30. There were no more samples found near the summit as it was cleared for a water reservoir. Field identification recorded the majority of samples as *Cyathea borneensis.*

2.2.2.3 Mount Perlis, Perlis (MP)/ 3

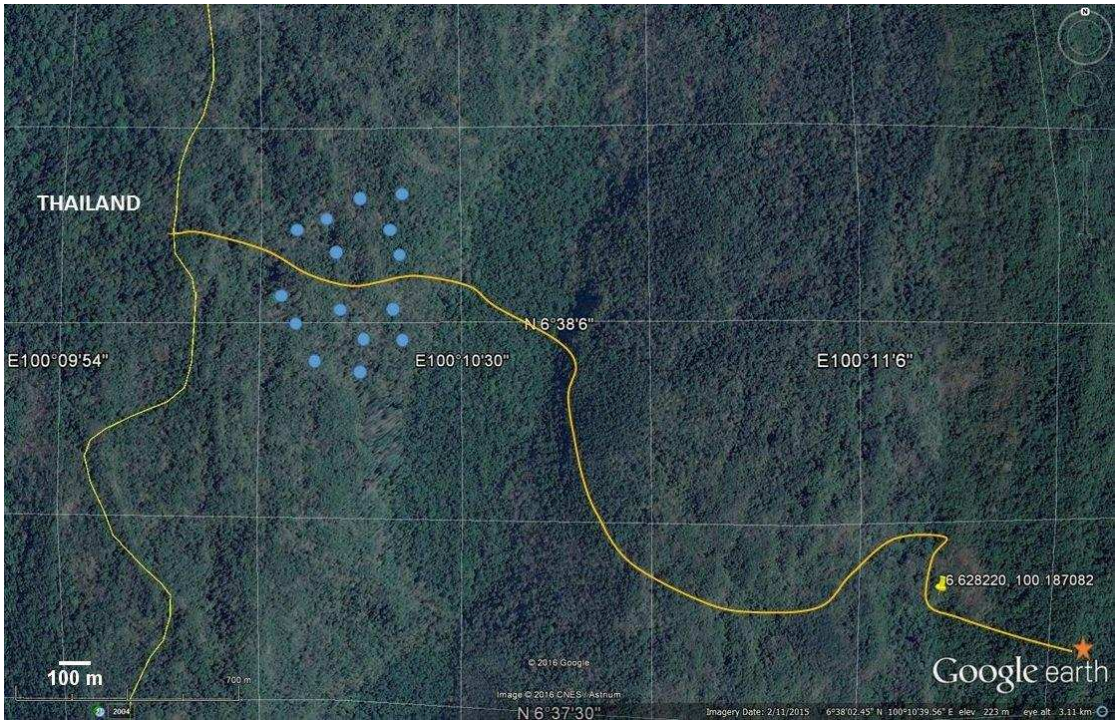


Figure 2.6 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 580m elevation and continued until 590m. Samples were collected on both sides of the trail with a final sample size of 15. There was another population found further up at the summit but collection was not permitted as the area is a border shared with Thailand.

2.2.2.4 Mount Jerai, Kedah (MJ)/ 4

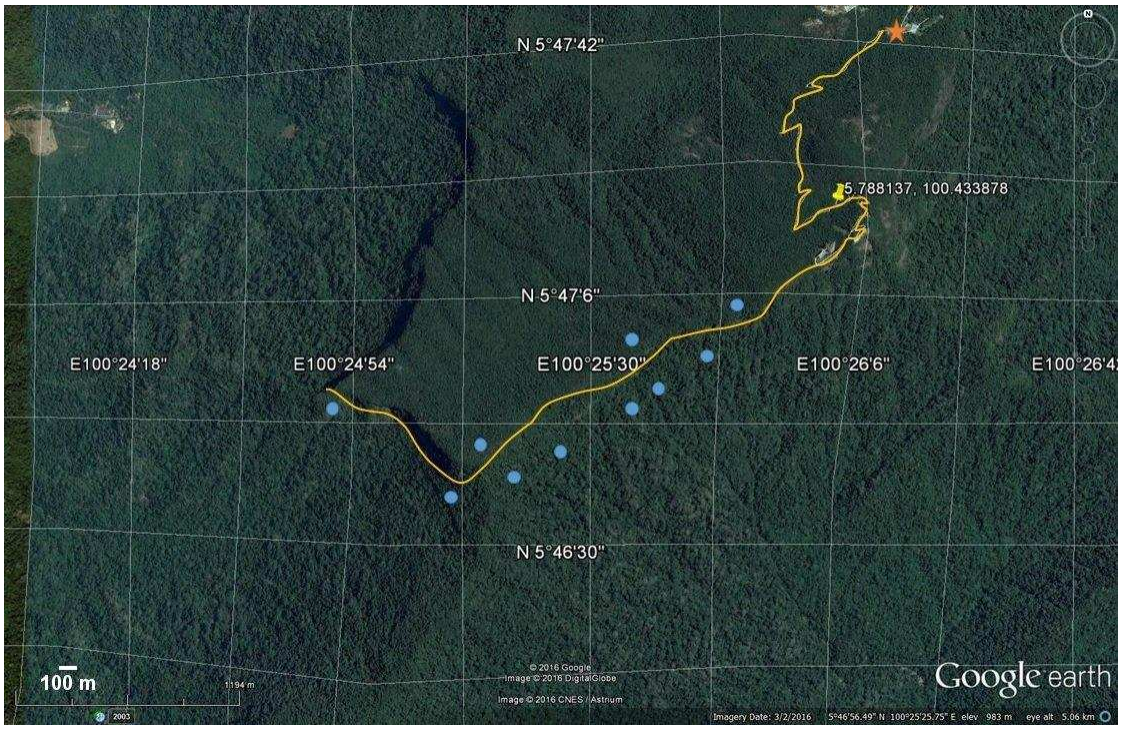


Figure 2.7 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 799m elevation and continued until 1099m. Samples were collected on both sides of the trail with a final sample size of 10. The plants were rare and sparse, with approximate 1000m distance from each sighting. The trail taken was the only one allowed as a military basecamp is stationed at the summit and many of the areas are restricted.

2.2.2.5 Fraser's Hill, Pahang (FH)/ 5

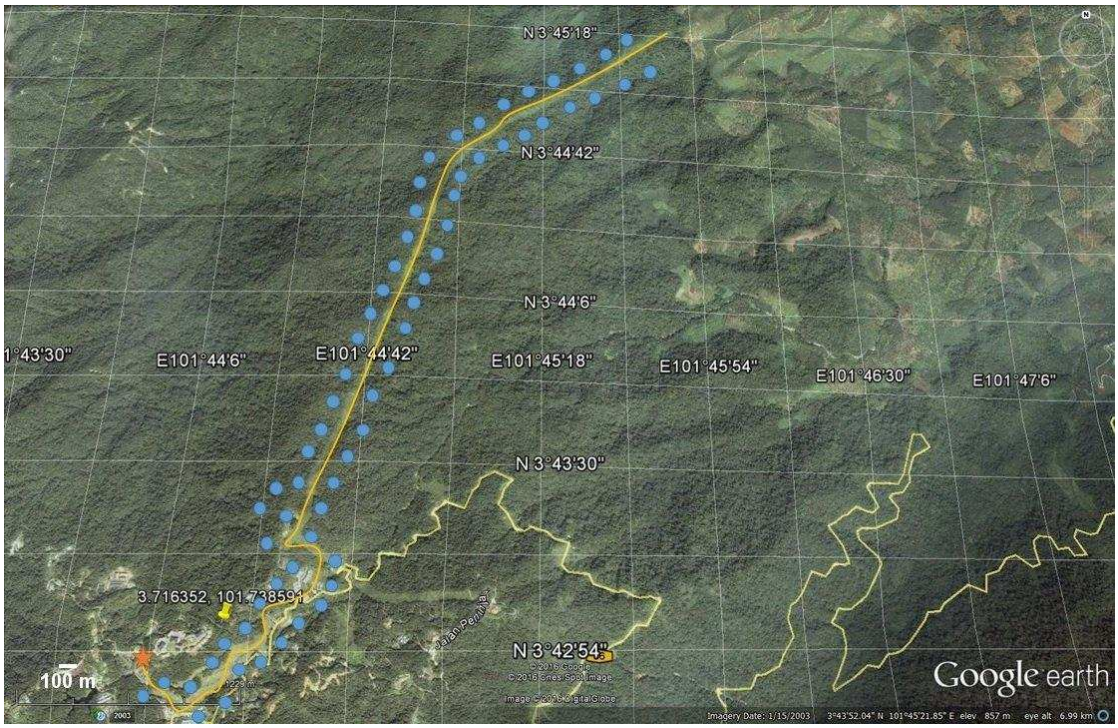


Figure 2.8 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 1242m elevation and continued until 1283m. Samples were collected on one side of the trail going up and other side going down with a final sample size of 65. The population was dense with initial observation found *C. borneensis* and *C. contaminans* dominating the area.

2.2.2.6 Mount Berinchang, Pahang (MB)/ 6

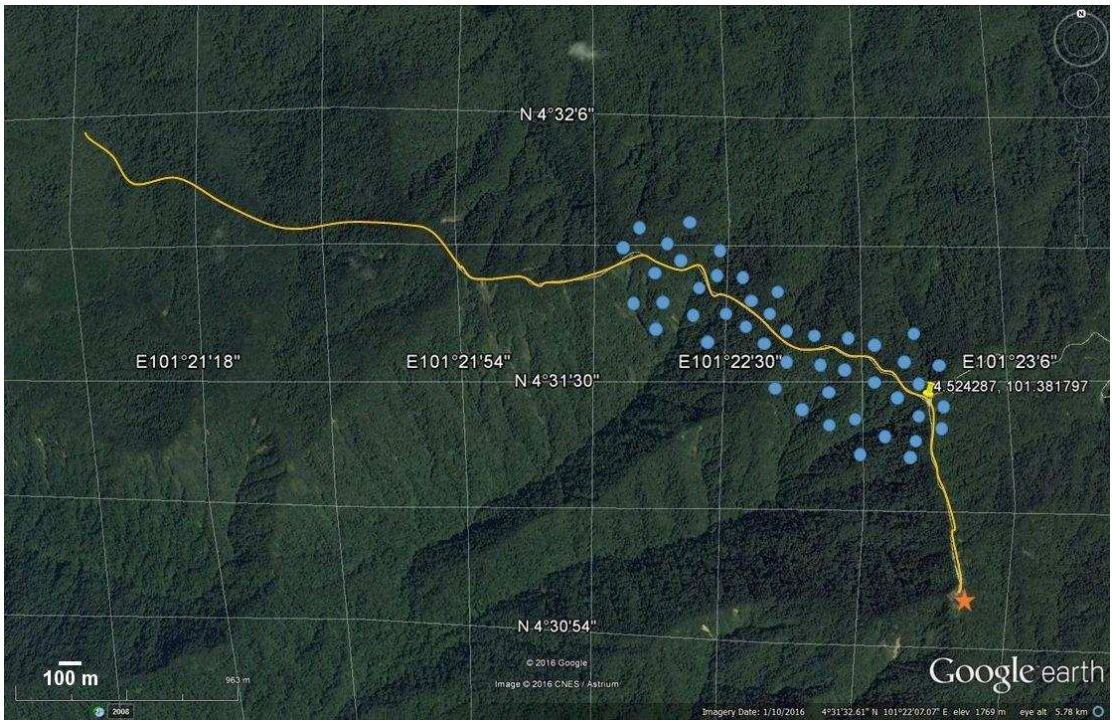


Figure 2.9 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 1867m elevation and continued until 2021m. Samples were collected on one side of the trail going up and the other side going down with a final sample size of 45. There were no more samples found further up until the summit as the area was cleared for a communications tower. The population was dense and the initial observations found *C. borneensis* dominating the area.

2.2.2.7 Genting Highlands, Pahang (GH)/ 7

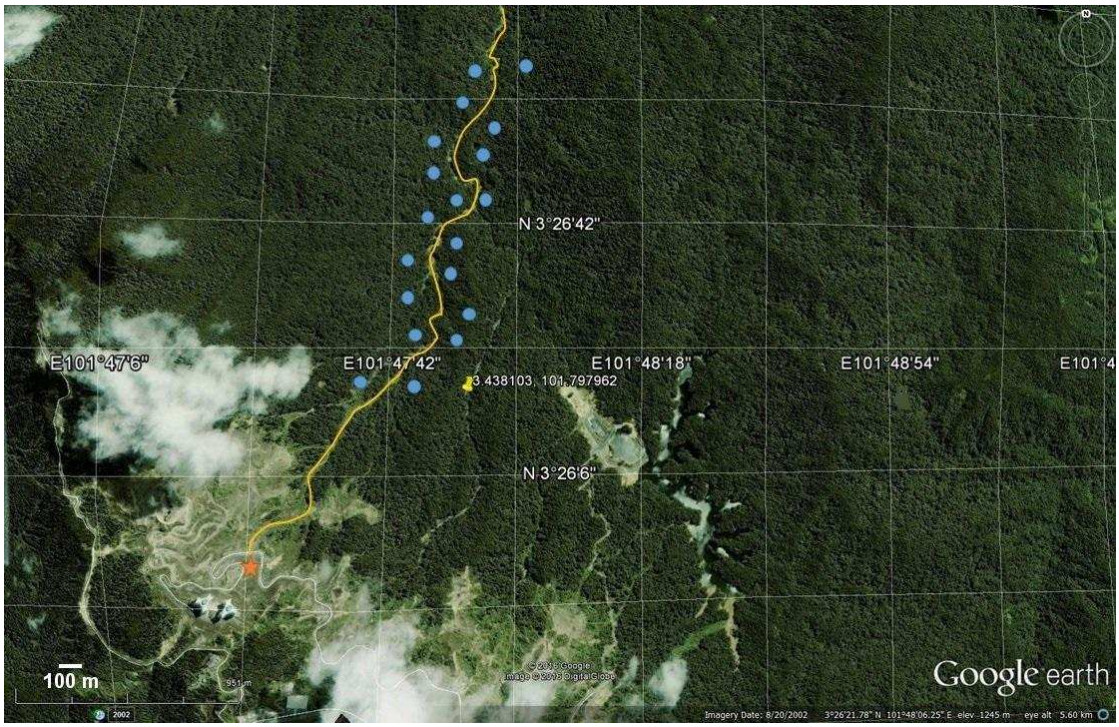


Figure 2.10 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 693m elevation and continued until 1590m. Samples were collected on one side of the trail going up and other side going down with a final sample size of 19. Previous population sighting locations were heavily converted into tourist areas with the current sampling location also impacted. The population was dominated by *C. contaminans.*

2.2.2.8 Lake Kenyir, Terengganu (LK)/ 8

Figure 2.11 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 187m elevation and continued until 203m. Samples were collected on one side of the trail going up and other side going down with a final sample size of 30. Previous population sighting locations were converted into a reservoir for a hydroelectricity scheme, with the current sampling location also impacted. The sampling area was mostly flat as it was near the lake. Initial observation found *C. borneensis* dominating the area.

2.2.2.9 Lojing Highlands, Kelantan (LH)/ 9

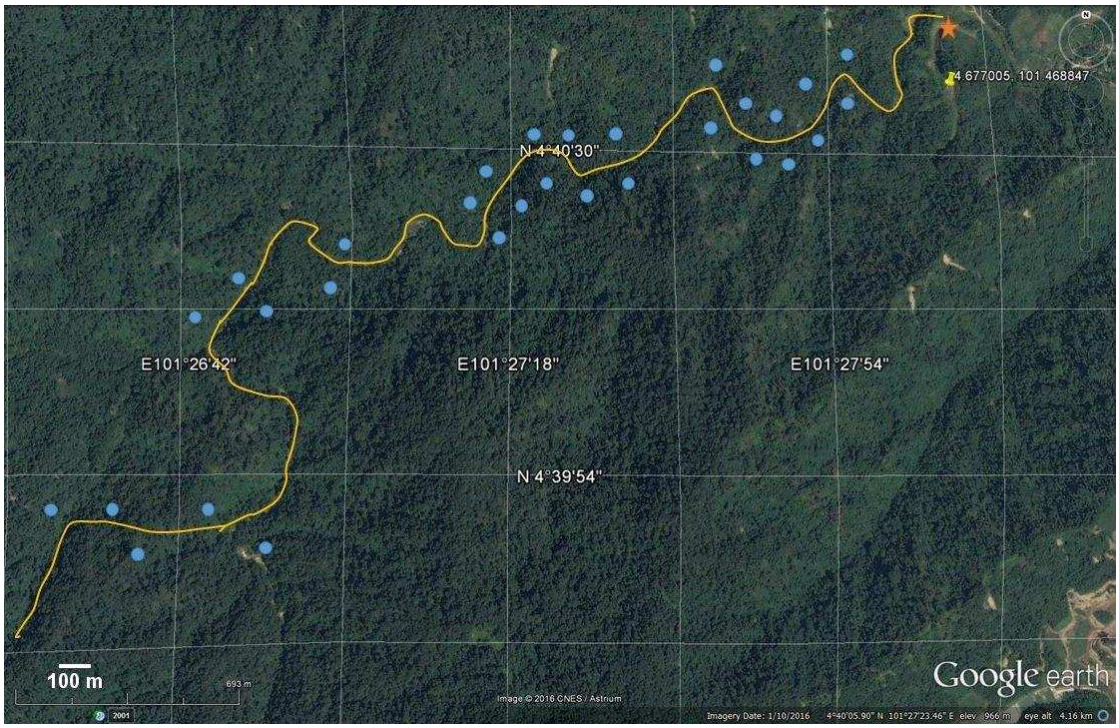


Figure 2.12 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 158m elevation, continued until 675m. Samples were collected on one side of the trail going up and other side going down with a final sample size of 30. The area was heavily deforested by logging and agricultural activities. Initial observation found *C. borneensis* and *C. contaminans* dominating the area.

2.2.2.10 Batang Kali, Selangor (BK)/ 10

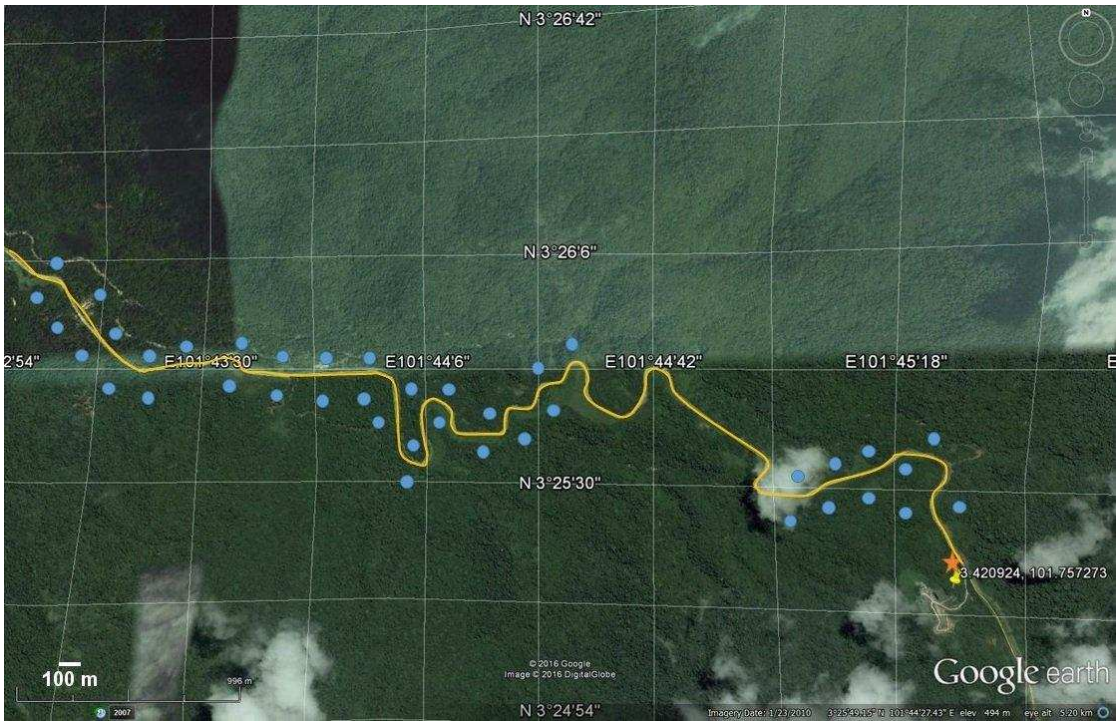


Figure 2.13 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 621m elevation and continued until 912m. Samples were collected on one side of the trail going up and other side going down with a final sample size of 40. There were no more samples found further up as the area was cleared for housing. Initial observation found *C. borneensis* and *C. latebrosa* dominating the area.

2.2.2.11 Mount Angsi, Negeri Sembilan (MA)/ 11

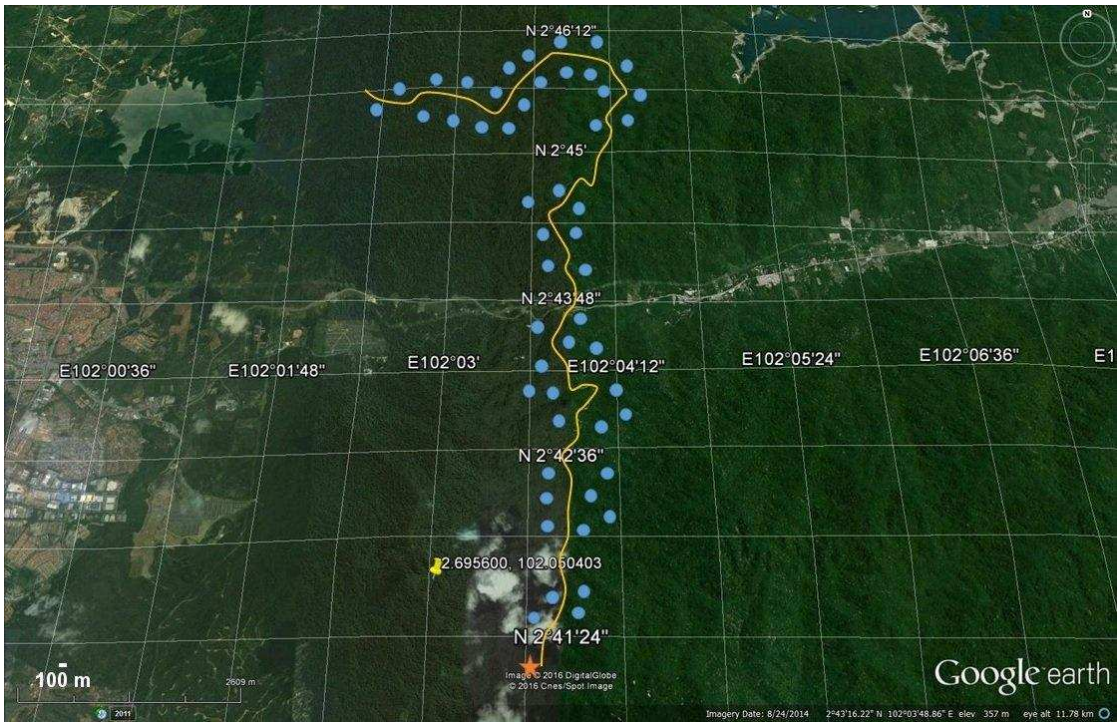


Figure 2.14 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 148m elevation and continued until 221m. Samples were collected on one side of the trail going up and other side going down with a final sample size of 51. The collection was done along the riverbanks at the mountain foot, as the permit to climb the mountain was declined due to bad weather and too dangerous for amateur climbers.

2.2.2.12 Mount Ledang, Johor (MJ)/ 12

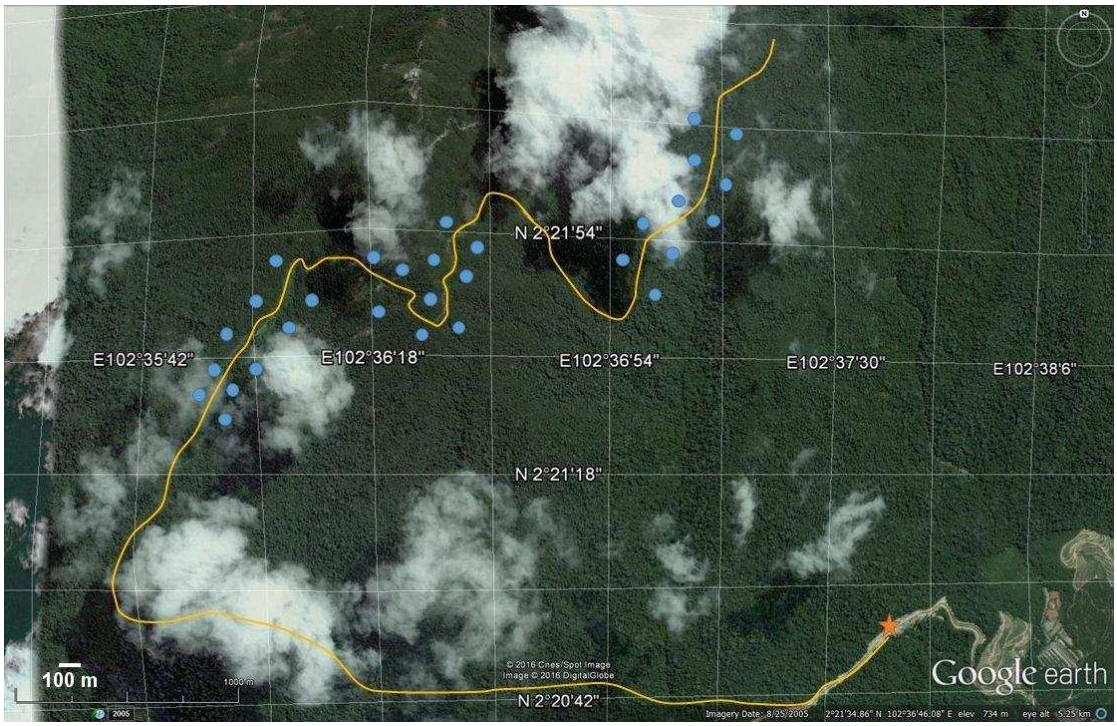


Figure 2.15 Map showing the route taken with a star marking the starting point and the collection sites marked with dots.

The population was first spotted at around 795m elevation and continued until 1122m. Samples were collected on one side of the trail going up and other side going down with a final sample size of 30. There were no more samples collected further up until the summit as the area was gazetted as a national park for wild orchids and *Nepenthes spp.*

2.2.3 Sample Identification

The identification was accomplished with the use of the detailed dichotomous keys from Flora Malesiana Series II: Pteridophyta (Holttum, 1963). Apart from the keys, published descriptions, illustrations and photographs were also used, from books (Large and Braggins, 2004; Piggott, 1988), journals (Lehnert, 2006; Lehnert, 2009; Latiff, 2015), or online sources (FRIM, 2013). The online digitized herbarium https://plants.jstor.org/ and actual herbarium specimens from UKM and Kew herbaria were also referred to for visual comparison. A dissecting microscope (Leica DFC420) with attached digital camera and computer interface was used for work with more detailed characters. A ruler with half millimeter increments was used as a scale for microscopic features. The scale bar was later added to the finished microscope figures using ImageJ software (Abramoff *et al.*, 2004).

2.3 Results

The expedition resulted in collection of 400 samples. Another 19 samples were contributed from UKM herbaria. There were other pteridophytes collected, intended to be used as outgroups in subsequent molecular phylogenetic analysis. The specimens that had been previously identified in the field were separated from unknown specimens. Table 2.1 shows the information on the species and locality extracted from the herbarium samples in Forest Research Institute of Malaysia (FRIM) used for determining the sampling location. Only 94% of the samples could be identified and

the 6% left, remained unidentified based on morphology due to the lamina and pinna being sterile, making sporangium and indusium-based identification problematic.

Table 2.1 Cyathea species list extracted from herbarium samples (FRIM, September 2013).

Collector	Date	State	Area	Lat/Long	Genus	Species
Molesworth-Allen	12 August 1953	Selangor	FRIM	3°14' N, 101°37' E	*Cyathea*	*alternans*
Chee, B.J.	20 October 2002	Terengganu	Rasau Kerteh F.R.	4°35.48' N, 103°17.62' E Alt: 45m	*Cyathea*	*alternans*
Korall, P.	25 August 2006	Pahang	Tanah Rata	4°48.04' N, 101°38.07' E Alt: 1500m	*Cyathea*	*borneensis*
Symington, C.F.	23 April 1931	Perak	Kledang Saiong F.R.	4°47' N, 100°59' E	*Cyathea*	*contaminans*
Molesworth-Allen	30 July 1956	Pahang	Ulu Telom, Sg. Ichat	4°27' N, 101°28' E Alt: 1524m	*Cyathea*	*excavata*
Ng, F.S.P.	17 November 1966	Johor	Panti F.R., Gn. Panti	1°50' N, 103°54' E Alt: 487m	*Cyathea*	*glabra*
Korall, P.	28 August 2006	Kedah	Gn. Jerai F.R., Gn. Jerai	5°80.58' N, 100°43.23' E	*Cyathea*	*glabra*
Imin, K.	20 March 2010	Terengganu	Ulu Brang	4°51.42' N, 102°54.15' E Alt: 1240m	*Cyathea*	*glabra*
Korall, P.	27 August 2006	Pahang	Tanah Rata	4°46.68' N, 101°38.46' E Alt: 1400m	*Cyathea*	*hymenodes*
Holttum, R.E.	6 February 1936	Penang	Penang Hill	5°26' N, 100°16' E Alt: 609m	*Cyathea*	*latebrosa*
Nor Ezzawanis,	8 April 2008	N Sembilan	Berembun F.R., Bkt. Lantai	2°49.90' N, 102°02.27' E Alt: 973m	*Cyathea*	*latebrosa*
Nor Ezzawanis	11 August 2009	Johor	Kluang F.R., Gn. Belumut	2°03.49' N, 103°32.08' E Alt: 580m	*Cyathea*	*latebrosa*
Parris, B.S.	17 March 1985	Pahang	Fraser's Hill	3°43' N, 101°45' E	*Cyathea*	*lurida*
Imin, K.	31 July 2010	Kelantan	Gn. Chamah	5°12.04' N, 101°33.51' E Alt: 1656m	*Cyathea*	*lurida*
Ogata, K.	11 May 1968	Penang	Penang Hill	5°26' N, 100°16' E Alt: 800m	*Cyathea*	*moluccana*
Ogata, K.	27 February 1968	Perak	Changkat Jong F.R.	3°58' N, 101°11' E Alt: 60m	*Cyathea*	*moluccana*
Saw, L.G.	5 March 1989	Terengganu	P. Redang, Bkt. Besar	5°46' N, 103°00' E Alt: 380m	*Cyathea*	*moluccana*
Chee, B.J.	20 October 2002	Terengganu	Rasau Kerteh F.R.	4°35.48' N, 103°17.62' E Alt: 80m	*Cyathea*	*moluccana*
Parris, B.S.	17 March 1985	Pahang	Fraser's Hill	3°43' N, 101°45' E	*Cyathea*	*obscura*
Parris, B.S.	26 August 1986	Johor	Gn. Ledang F.R., Air Panas	2°22' N, 102°37' E Alt: 460m	*Cyathea*	*obscura*
Wyatt-Smith, J.	26 October 1958	Selangor	Ulu Gombak	3°18' N, 101°47' E Alt: 670m	*Cyathea*	*recommutata*
Edwards, P.J.	6 February 1986	Johor	Endau-Rompin S.P.	2°31' N, 103°21' E	*Cyathea*	*squamulata*
Razali, S.	20 February 1984	Selangor	Bangi F.R.	2°55' N, 101°47' E	*Cyathea*	*squamulata*
Kiew, R.	3 September 1985	Johor	Endau-Rompin S.P., Ulu Endau	2°40' N, 103°38' E	*Cyathea*	*trichodesma*

Of the 419 specimens collected, 15 *Cyathea* species were identified (402 specimens), as well as five Cibotiaceae (all *Cibotium barometz*); two Marattiaceae (both *Angiopteris evecta*); one Blechnaceae (*Blechnum fraseri*); seven Tectariaceae (*Pleocnemia olivacea*) and two Athyriaceae (*Diplazium proliferum*) (Table 2.2). Unidentified specimens were sequenced and identification was made against the known specimens as part of the DNA barcoding test. The full list of identified species for each of the populations is presented in Table 2.3. Identified *Cyathea* species are each presented in detail (Figure 2.17 to Figure 2.31). Species descriptions are based on a combination of direct observation with information from Holttum (1963) and Large and Braggins (2004). RHS color chart (color code presented in bracket) was used to describe the color in the description (Grayer, 2009). Specimens will be deposited at University of Reading herbarium (RNG) with specimen code starting with AJ-(accession code)-RNG.

Table 2.2 Cyathea species identified including other pteridophyte species.

Genus	**Species**	**Collecting Sites**
Cyathea	*C. alternans*	MA
	C. assimilis	MB
	C. borneensis	BK, BL, FH, GH, LH, MB, MP, PH
	C. contaminans	BK, BL, FH, GH, LH, LK, MB, MJ, ML, PH
	C. gigantea	ML
	C. glabra	BL, FH
	C. hymenodes	FH, MJ, MP, PH
	C. incisoserrata	FH, LK
	C. latebrosa	BK, BL, FH, LH, LK, MA, MJ, PH, HB
	C. lurida	FH
	C. moluccana	HB
	C. obscura	BL, FH, GH
	C. polypoda	FH, ML
	C. recommutata	ML
	C. trichodesma	BL, MA
Angiopteris	*A. evecta*	BL, FH
Blechnum	*B. fraseri*	MB
Cibotium	*C. barometz*	FH, MJ
Diplazium	*D. proliferum*	BL, FH
Pleocnemia	*P. olivacea*	GH, MA, MJ, MP

Table 2.3 List of species collected throughout Peninsular Malaysia by location.

Location	Species	Accession Code	Collector
Bukit Larut, Perak	*Angiopteris evecta*	BL10	Jamaludin, A.
	C. borneensis	BL02, BL04, BL14, BL20, BL21, BL22, BL24, BL29	Jamaludin, A.
	C. contaminans	BL01, BL08, BL09, BL12, BL13, BL16, BL18, BL31, BL34, BL35	Jamaludin, A.
	C. glabra	BL03, BL05, BL06	Jamaludin, A.
	C. latebrosa	BL17, BL19, BL33	Jamaludin, A.
	C. obscura	BL15, BL23, BL27, B26, BL28, BL30	Jamaludin, A.
	C. trichodesma	BL07, BL32	Jamaludin, A.
	Diplazium proliferum	BL11	Jamaludin, A.
Penang Hill, Penang	*C. borneensis*	PH03, PH04, PH05, PH06, PH09, PH10, PH11, PH12, PH13, PH14, PH15, PH16, PH17, PH18, PH19, PH20, PH21, PH22, PH23, PH24, PH25, PH27, PH28, PH29	Jamaludin, A.
	C. contaminans	PH30	Jamaludin, A.
	C. latebrosa	PH01, PH02, PH07, PH08, PH26	Jamaludin, A.
Mount Perlis, Perlis	*C. borneensis*	MP01, MP02, MP03, MP04, MP05, MP06, MP07, MP08, MP10, MP11, MP13, MP14, MP15	Jamaludin, A.
	C. hymenodes	MP12	Jamaludin, A.
	Pleocnemia olivacea	MP09	Jamaludin, A.
Mount Jerai, Kedah	*C. contaminans*	MJ01	Jamaludin, A.
	C. hymenodes	MJ05, MJ06, MJ07, MJ08	Jamaludin, A.
	C. latebrosa	MJ04, MJ09	Jamaludin, A.
	Cibotium barometz	MJ03, MJ10	Jamaludin, A.
	P. olivacea	MJ02	Jamaludin, A.
Fraser's Hill, Pahang	*A. evecta*	FH53	Jamaludin, A.
	C. borneensis	FH07, FH08, FH17, FH20, FH21, FH23, FH24, FH25, FH27,	Jamaludin, A.

		FH28, FH29, FH38, FH40, FH41, FH44, FH45, FH46, FH48, FH49, FH50, FH51, FH54, FH58, FH59, FH60	
	C. contaminans	FH01, FH02, FH57, FH61, FH62, FH63, FH64, FH65	Jamaludin, A.
	C. glabra	FH06, FH09, FH26, FH42, FH43, FH47, FH55, FH56	Jamaludin, A.
	C. hymenodes	FH13	Jamaludin, A.
	C. incisoserrata	FH15, FH16	Jamaludin, A.
	C. latebrosa	FH05, FH10, FH12, FH14	Jamaludin, A.
	C. obscura	FH04, FH32, FH33, FH34, FH35, FH39, FH52	Jamaludin, A.
	C. polypoda	FH18, FH31	Jamaludin, A.
	C. barometz	FH03, FH11, FH30, FH36, FH37	Jamaludin, A.
	P. olivacea	FH19, FH22	Jamaludin, A.
Mount Berinchang, Pahang	*Blechnum fraseri*	MB13	Jamaludin, A.
	C. assimilis	MB22, MB42	Jamaludin, A.
	C. borneensis	MB01, MB03, MB04, MB05, MB06, MB08, MB09, MB10, MB14, MB15, MB17, MB18, MB20,MB23, MB24, MB25, MB26, MB27, MB28, MB29, MB30, MB31, MB36, MB37, MB38, MB39, MB41, MB44, MB45	Jamaludin, A.
	C. contaminans	MB02, MB12, MB16, MB19, MB21, MB32, MB33, MB34, MB40	Jamaludin, A.
	C. lurida	MB07, MB11, MB35, MB43	Jamaludin, A.
Genting Highlands, Pahang	*C. borneensis*	GH17	Jamaludin, A.
	C. contaminans	GH04, GH05, GH06, GH07, GH08, GH09, GH10, GH11, GH12, GH13, GH14	Jamaludin, A.
	C. obscura	GH19	Jamaludin, A.

	P. olivacea	GH03, GH18	Jamaludin, A.
	Cyathea cf. *latebrosa*	GH01, GH02, GH15, GH16	Jamaludin, A.
Lake Kenyir, Terengganu	*C. borneensis*	LK16, LK18, LK20, LK21, LK22, LK23, LK24, LK25, LK26, LK27, LK28, LK29, LK30	Jamaludin, A.
	C. contaminans	LK03, LK04, LK05, LK06, LK09, LK12, LK13, LK14, LK15	Jamaludin, A.
	C. incisoserrata	LK02, LK10, LK11, LK17	Jamaludin, A.
	C. latebrosa	LK01, LK07, LK08, LK19	Jamaludin, A.
Lojing Highlands, Kelantan	*C. borneensis*	LH29, LH30	Jamaludin, A.
	C. contaminans	LH01, LH02, LH03, LH04, LH05, LH06, LH07, LH08, LH09, LH10, LH11, LH12, LH13, LH14, LH15, LH16, LH17, LH18, LH20, LH21, LH22, LH23, LH24, LH25, LH26, LH27, LH28	Jamaludin, A.
	C. latebrosa	LH19	Jamaludin, A.
Batang Kali, Selangor	*C. borneensis*	BK03, BK22, BK23, BK24, BK25, BK26, BK27, BK28, BK29, BK30, BK31, BK32, BK33, BK34, BK35, BK36, BK37, BK38, BK39, BK40	Jamaludin, A.
	C. contaminans	BK10	Jamaludin, A.
	C. latebrosa	BK01, BK02, BK04, BK05, BK06, BK07, BK08, BK09, BK11, BK12, BK13, BK14, BK15, BK16, BK17, BK18, BK19, BK20, BK21	Jamaludin, A.
Mount Angsi, Negeri Sembilan	*C. alternans*	MA16	Jamaludin, A.
	C. latebrosa	MA06	Jamaludin, A.
	C. trichodesma	MA03, MA04, MA05, MA07, MA08, MA09, MA10, MA11, MA13, MA14, MA17, MA18, MA19, MA21, MA22,	Jamaludin, A.

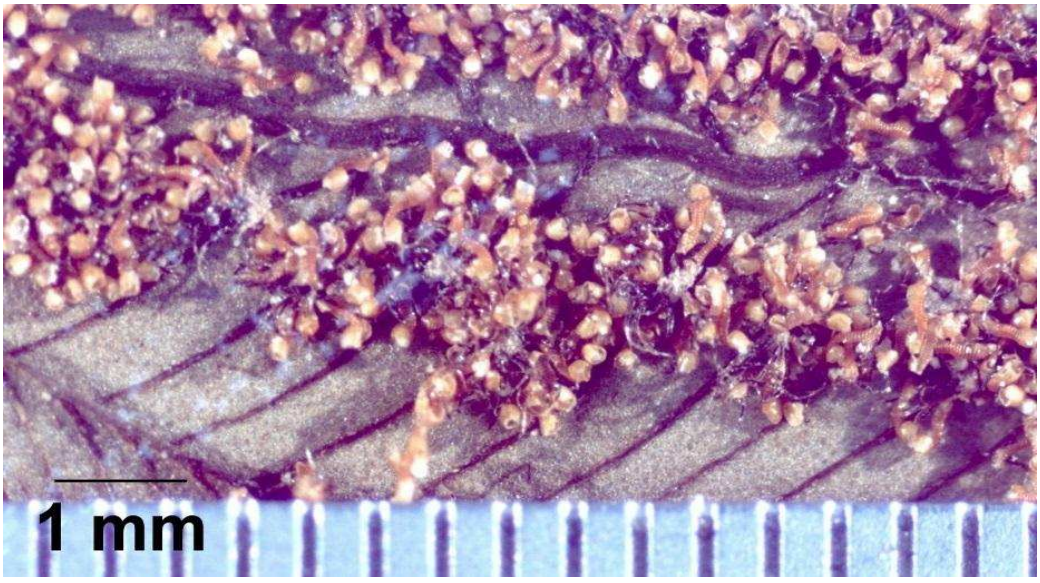
		MA24, MA25, MA26, MA28, MA30, MA31, MA32, MA33, MA34, MA35, MA36, MA41, MA42, MA43, MA44, MA45, MA47, MA48, MA49, MA51	
	P. olivacea	MA12, MA37	Jamaludin, A.
	Cyathea sp.	MA01, MA02, MA15, MA20, MA23, MA27, MA29, MA38, MA39, MA40, MA46, MA50	Jamaludin, A.
Mount Ledang, Johor	*C. contaminans*	ML10, ML11, ML12	Jamaludin, A.
	C. gigantea	ML23, ML24, ML26, ML27, ML28, ML29, ML30	Jamaludin, A.
	C. polypoda	ML06, ML20, ML25	Jamaludin, A.
	C. recommutata	ML07, ML08, ML14, ML15, ML17, ML18, ML19, ML21, ML22	Jamaludin, A.
	Cyathea sp.	ML01, ML02, ML03, ML04, ML05, ML09, ML13, ML16	Jamaludin, A.
Courtesy of UKM herbarium	*C. contaminans*	BF7a, BF7b	Maideen, H
BF (Fraser's Hill)	*C. glabra*	BF8a, BF8b	Maideen, H
HB (Bangi Forest)	*C. latebrosa*	BF1a, BF1b, BF1c, BF5a, BF5b, BF6a, BF6b, HB3	Maideen, H
	C. assimilis	BF2	Maideen, H
	C. moluccana	HB1, HB2	Maideen, H
	C. polypoda	BF3a, BF3b, BF3c, BF4	Maideen, H

2.3.1 Species Identified

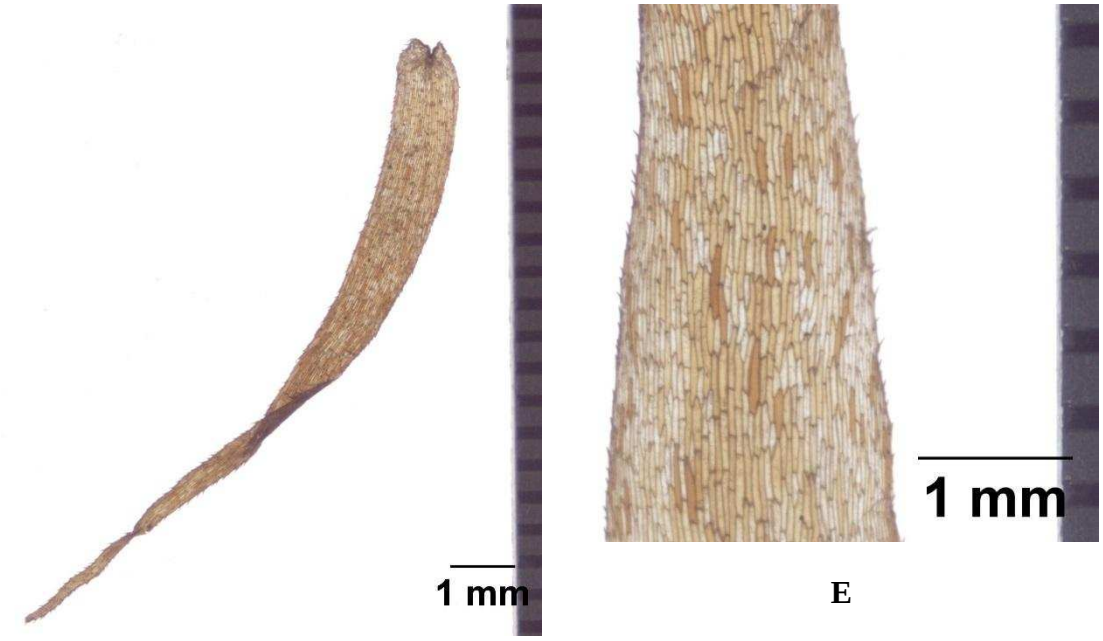
2.3.1.1 *C. alternans* (Wallich ex.W.J. Hooker) C. B. Presl

A

B



C



D **E**

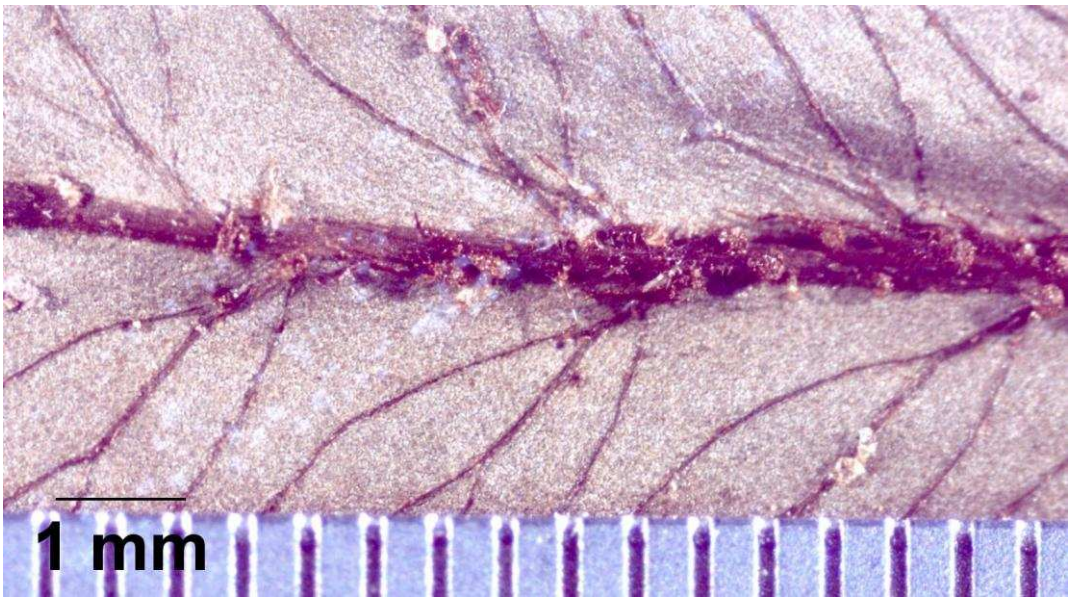
Figure 2.16 ***A****. Plant sample from Mount Angsi.* ***B.*** *Part of pinna and stipe.* ***C****. Sporangia.* ***D****. Scale from stipe.* ***E.*** *Part of the scale. (AJ-MA16-RNG).*

Fronds are pinnate and 1 to 2m long. Lower pinnae a little narrowed at the base, free leaflets as long as lobes, apex not long-acuminate (Holttum, 1963). The stipe is dark (N200A) and generally smooth but has basal scales that are light to medium brown (N199C-N199D). Sori occur in a single row on the either side of the mid-vein. Indusia are present and variable, the form may completely cover the sorus, sometimes in a saucer-like shape (Large and Braggins, 2004).

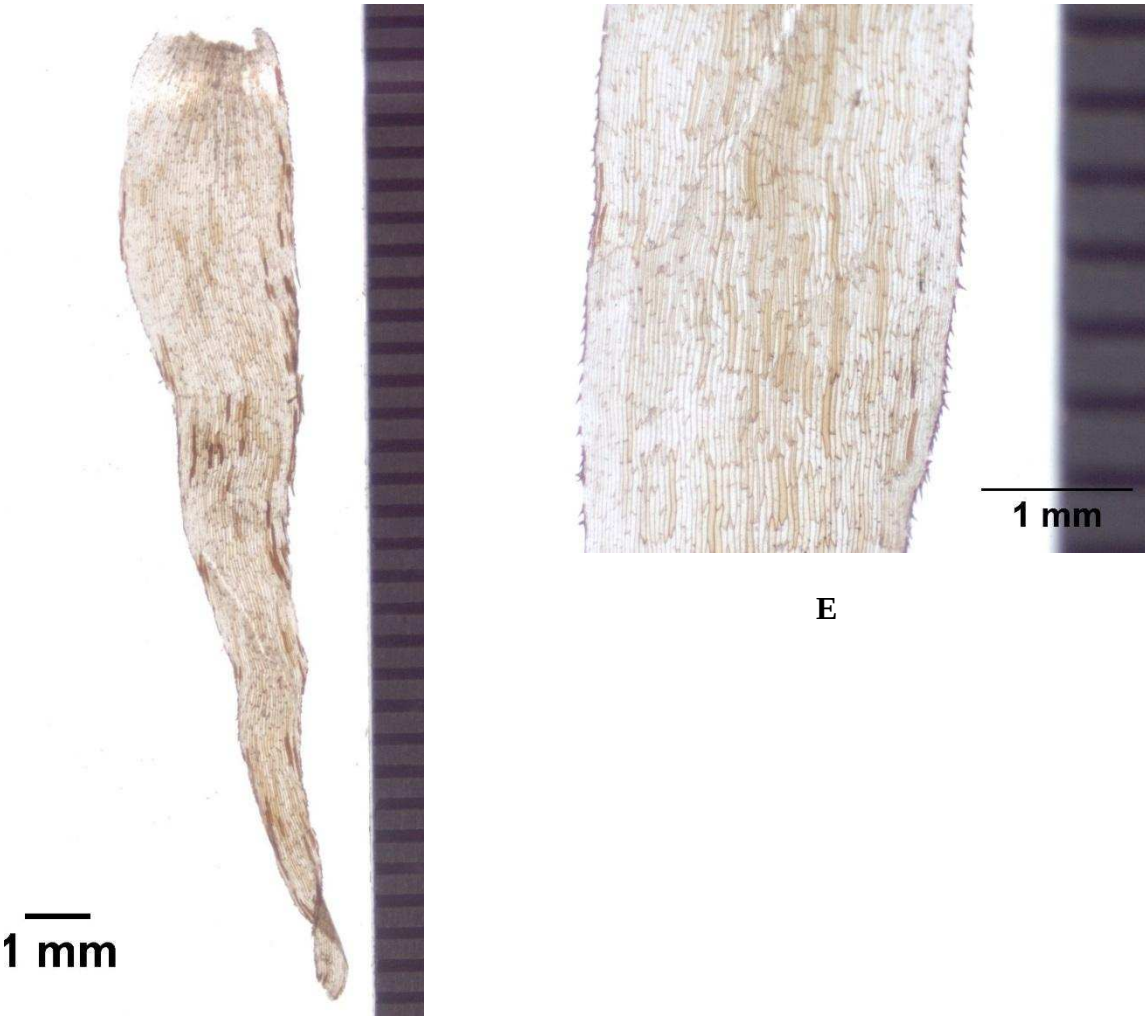
2.3.1.2 *C. glabra* (Blume) Copeland

A

B



C

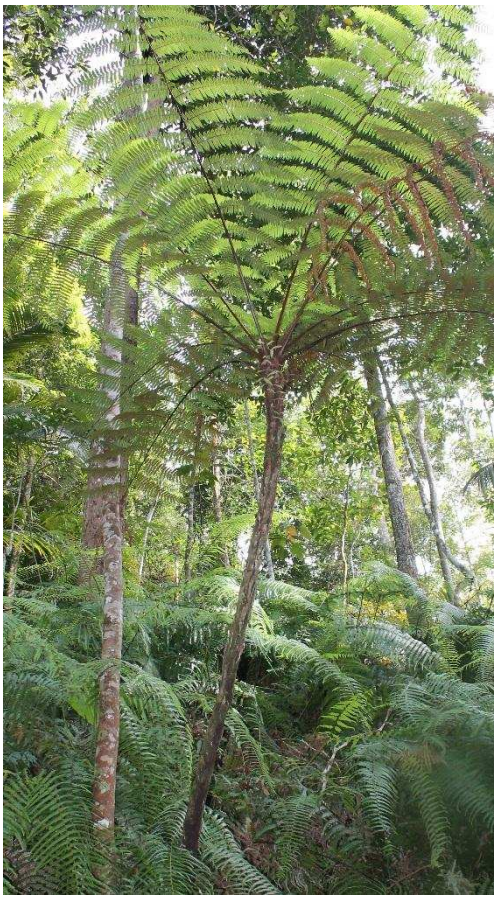


D **E**

Figure 2.17 ***A****. Plant sample from Bukit Larut.* ***B.*** *Part of pinna and stipe.* ***C****. Scales on costa and costules.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-BL05-RNG).*

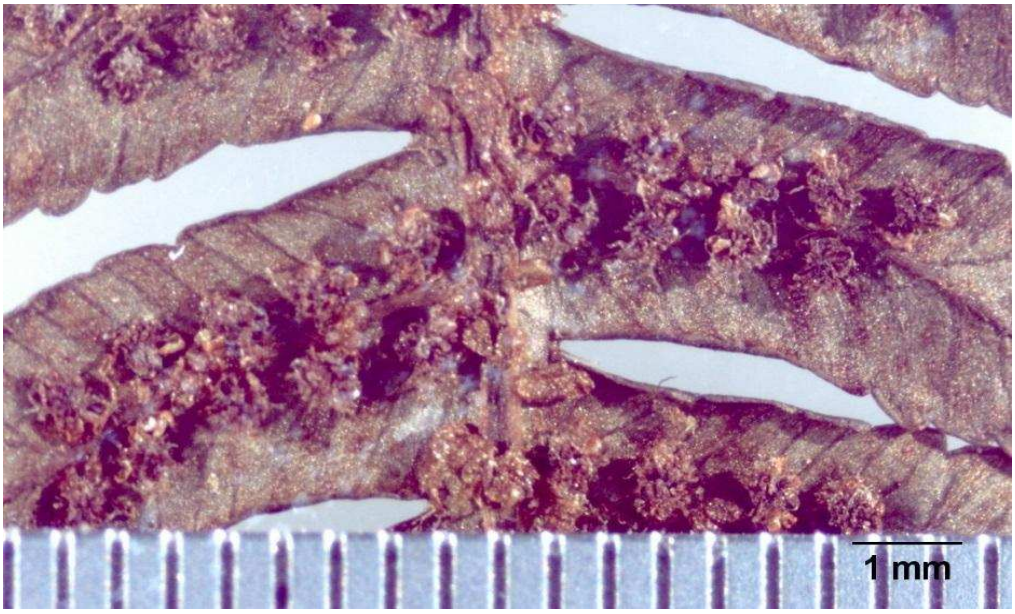
Fronds are bi-pinnate or tri-pinnate and 1 to 2m long. Pinnules almost entire. Lowest pinnae may be reduced. The stipe and rachis are brown to dark brown (200D-200B). Scales are light brown (199A) and glossy. Sori are in groups of one to three and indusia are absent (Large and Braggins, 2004).

2.3.1.3 *C. borneensis* Copeland



A

B



C

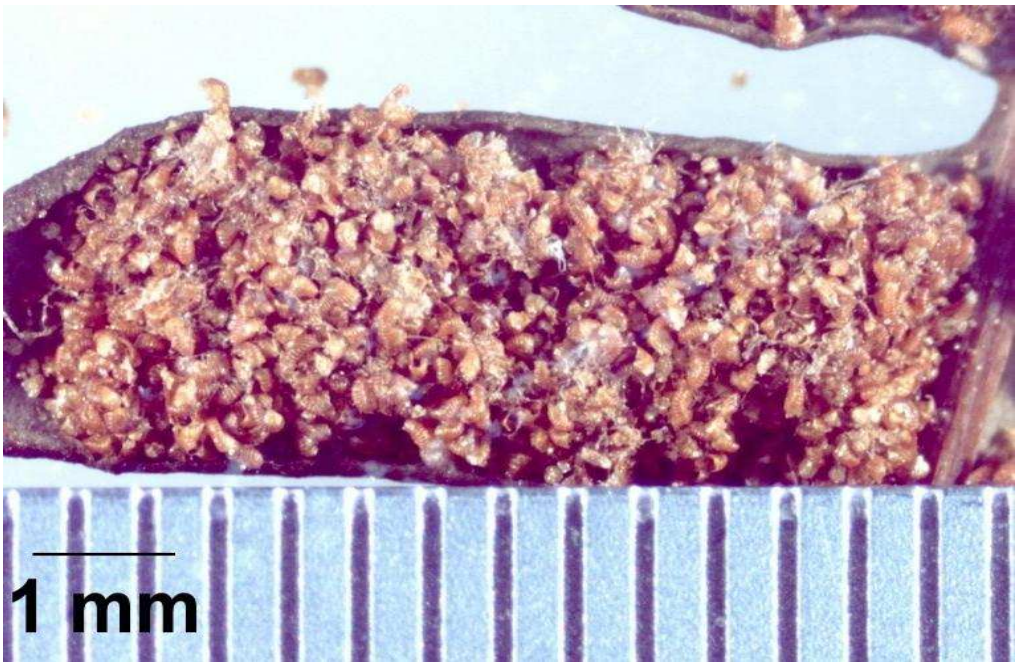
Figure 2.18 ***A.*** *Plant sample from Penang Hill.* ***B.*** *Part of pinna and stipe.* ***C.*** *Sporangia.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-PH04-RNG).*

Fronds are bi-pinnate or tri-pinnate and may reach 2 to 3m long. Pinnules not articulate (Holttum, 1963). The stipe is medium brown (N199C), spiny and warty and has scales that are dark brown (N200A) and glossy. Sori are close to the mid-vein and covered with thin indusia (Large and Braggins, 2004).

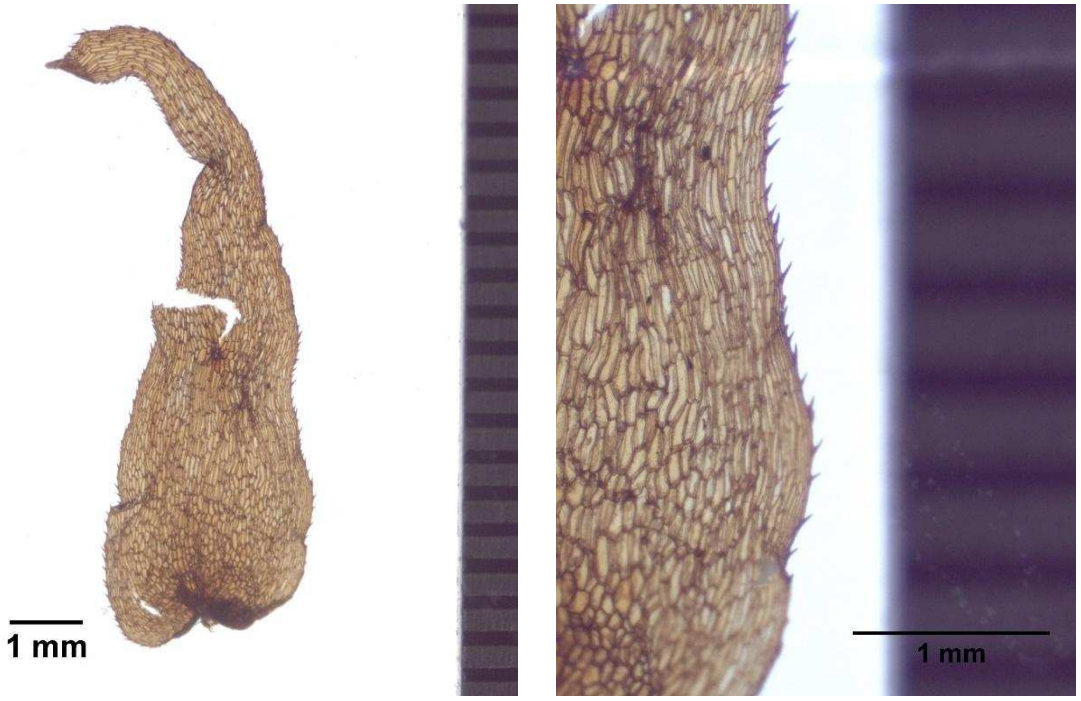
2.3.1.4 *C. contaminans* (Wallich ex W. J. Hooker) Copeland

A

B



C



D **E**

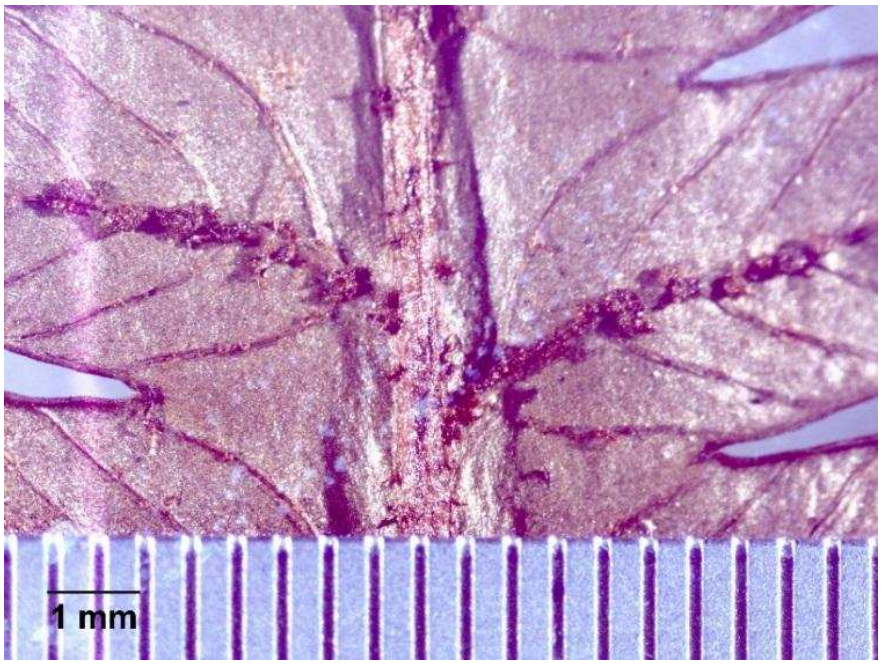
Figure 2.19 ***A****. Plant sample from Bukit Larut.* ***B.*** *Part of pinna and stipe.* ***C****. Sporangia.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-BL35-RNG).*

Fronds are bi-pinnate or tri-pinnate and may reach 3 to 4m long or more. Pinnules commonly cut almost to costa. The stipe is purplish (N187A) toward the base and has spines and scales that are light brown to brown (199A-N199B). The rachis is pale and spiny. Sori occur in rows close to pinnule mid-vein and lack indusia (Large and Braggins, 2004).

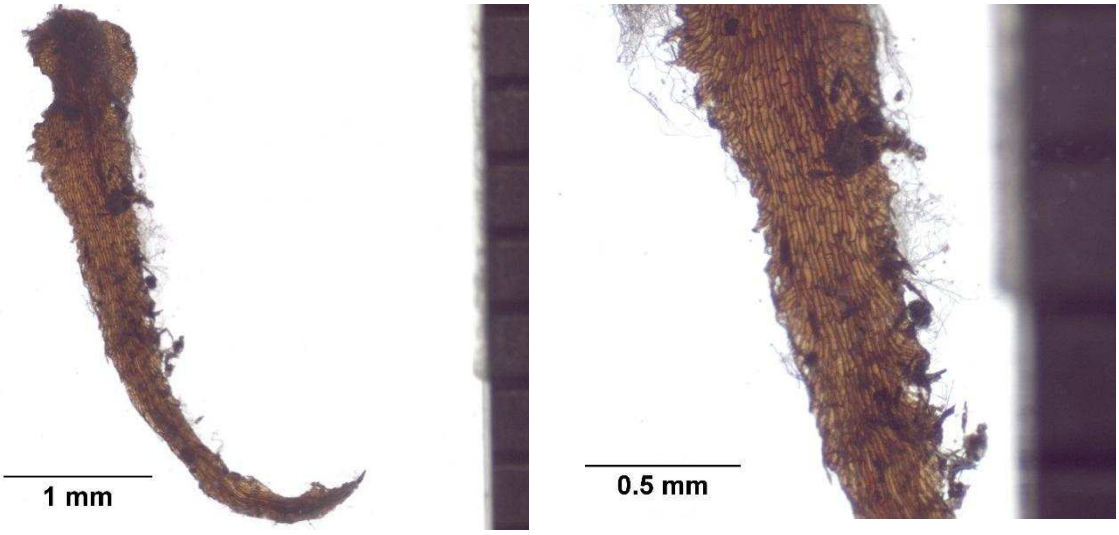
2.3.1.5 *C. gigantea* (Wallich ex W. J. Hooker) Holttum

A

B



C

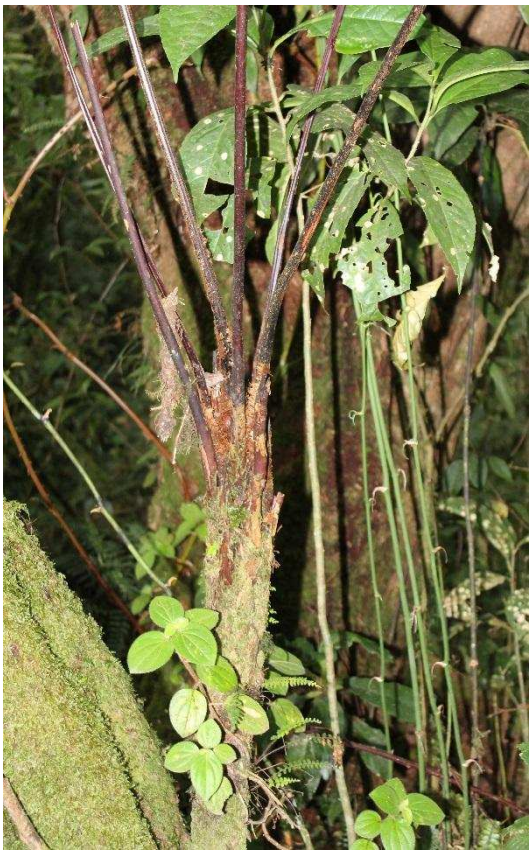


D **E**

Figure 2.20 ***A****. Plant sample from Mount Ledang.* ***B.*** *Part of pinna and stipe.* ***C****. Scales on costa and costules.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-ML24-RNG).*

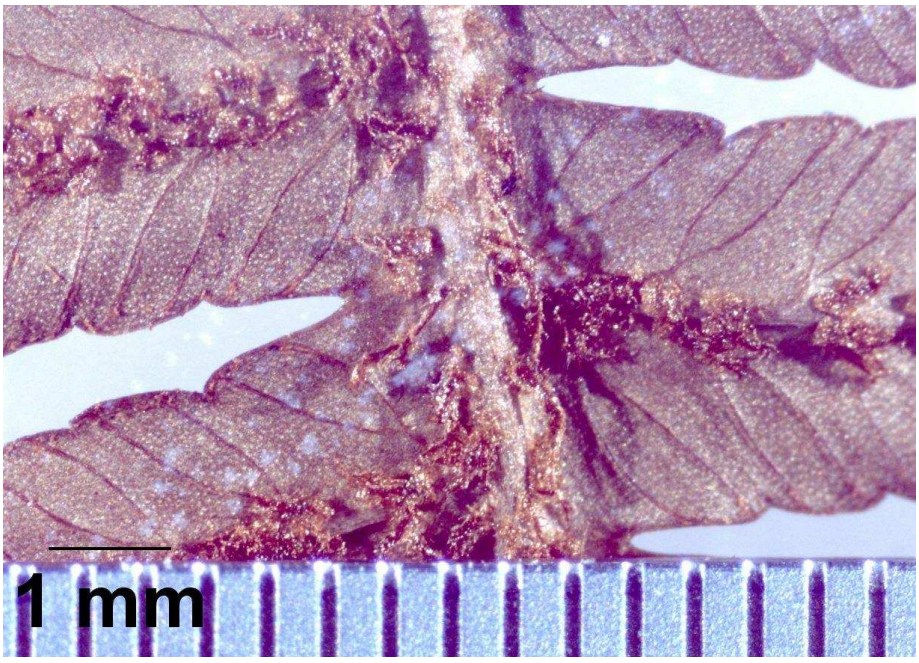
Fronds are bi-pinnate or tri-pinnate and 2 to 3m long. Pinnules not lobed more than 2/3 towards costa. The stipe and rachis are dark or black (202A). Scales are dark brown (N200A) and glossy. Sori are round and lack indusia (Large and Braggins, 2004).

2.3.1.6 *C. lurida* (Blume) Copeland



A

B



C

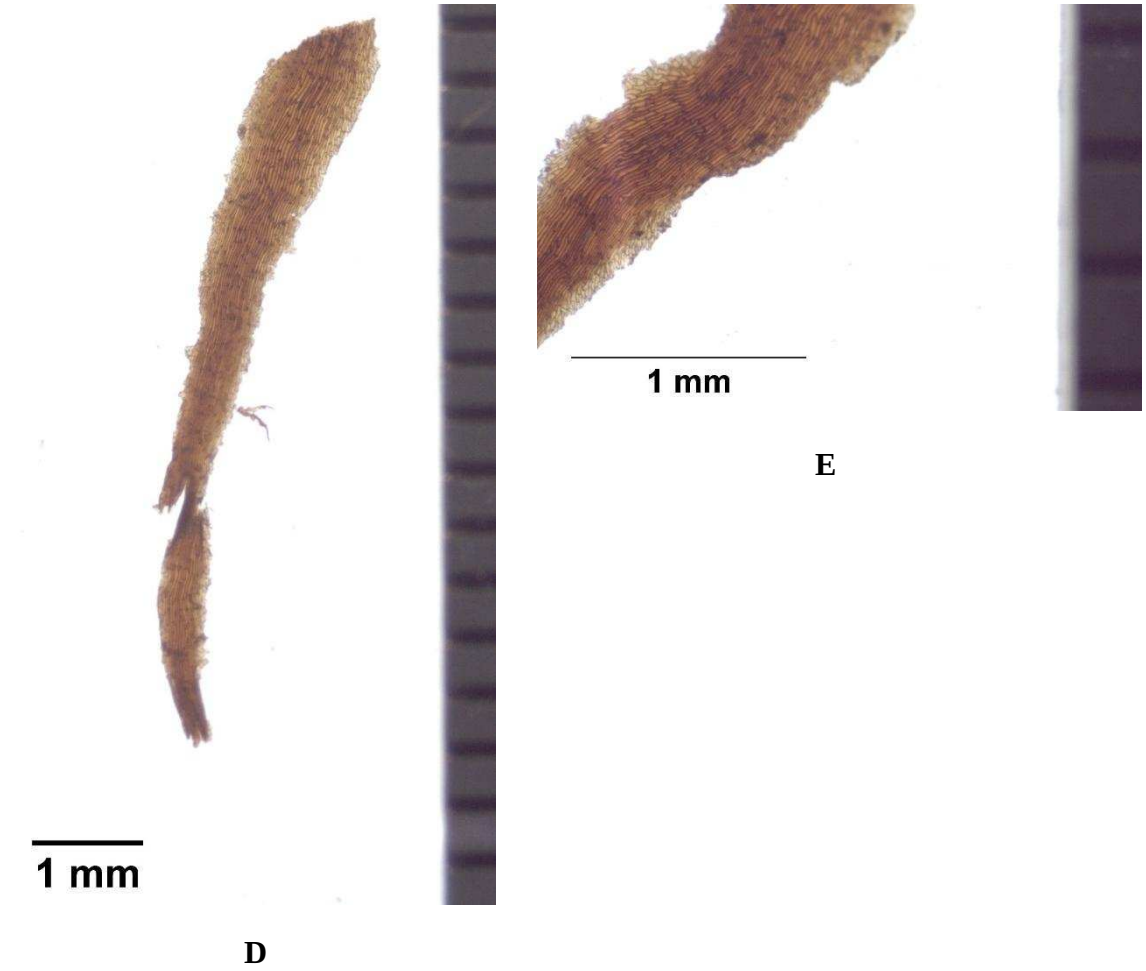


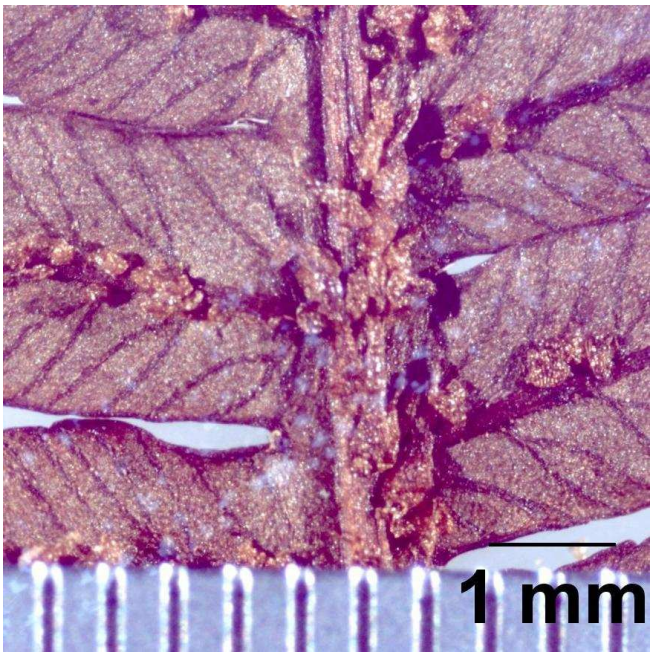
Figure 2.21 ***A****. Plant sample from Mount Berinchang.* ***B.*** *Part of pinna and stipe.* ***C****. Scales on costa and costules.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-MB11-RNG).*

Fronds are bi-pinnate and 1 to 2m long. Pinnules distinctly lobed. The stipe is long and dark (202A). Scales are medium brown to dark brown (N199B-200C). Sori almost cover the lower surface of pinnule and lack indusia (Large and Braggins, 2004).

2.3.1.7 *C. hymenodes* Mattenius

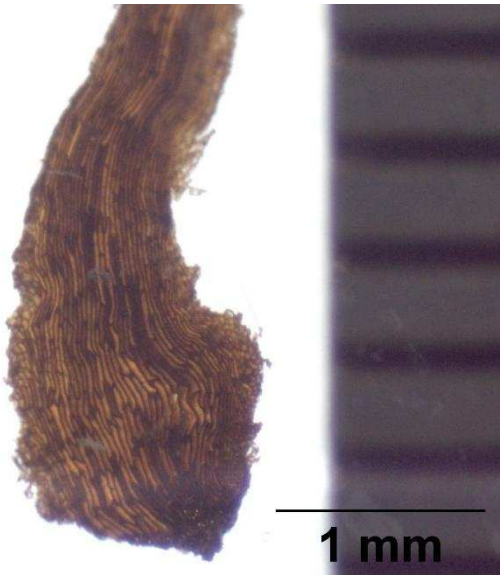
A

B



C

D



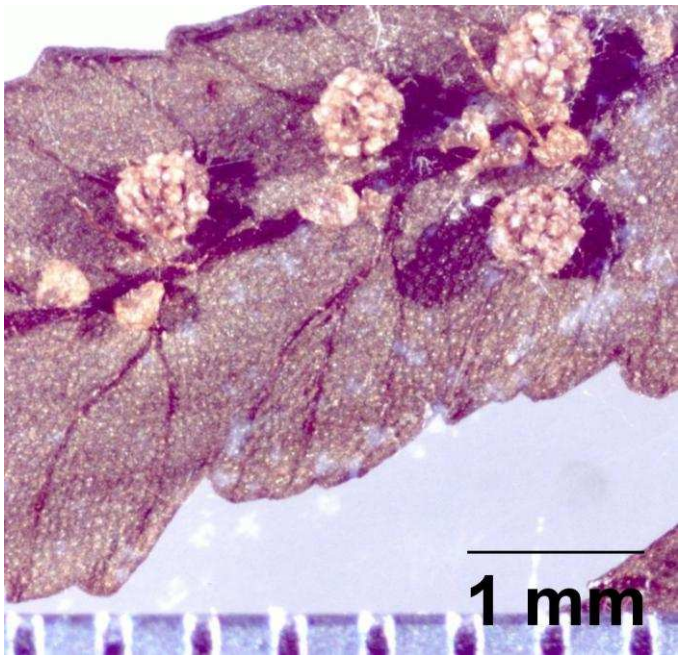
E

Figure 2.22 ***A****. Plant sample from Fraser's Hill.* ***B.*** *Part of pinna and stipe.* ***C****. Scales on costa and costules.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-FH13-RNG).*

Fronds are bi-pinnate or tri-pinnate and 1 to 2m long; lowest pinnae may be reduced. Pinnules without free basal segment. The stipe is medium to dark brown (N199B-200D) and covered with dark brown (N200A) scales. Sori occur near the mid-vein and covered by brown, saucer-like indusia (Large and Braggins, 2004).

2.3.1.8 *C. incisoserrata* Copeland

A **B**



C

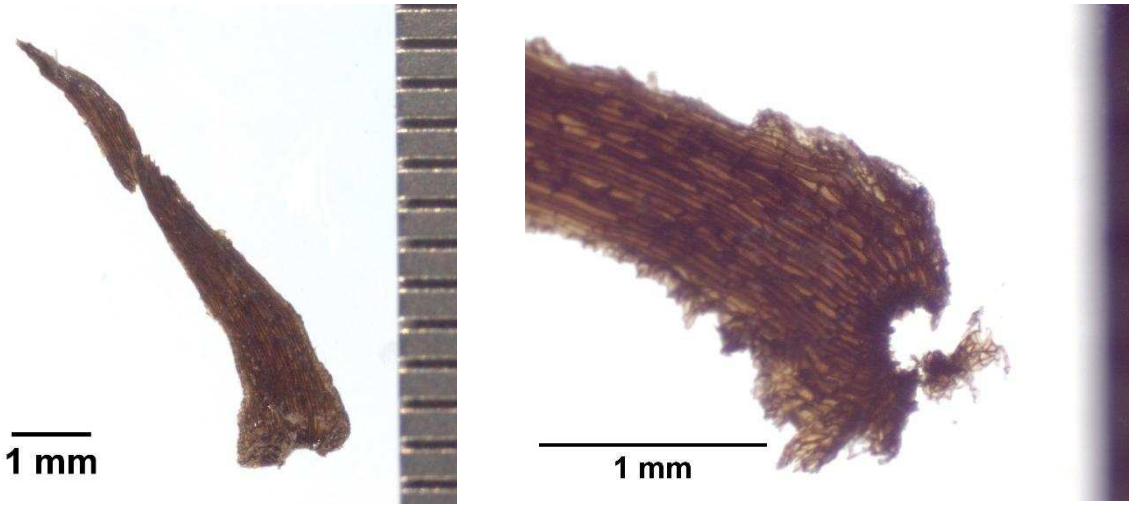


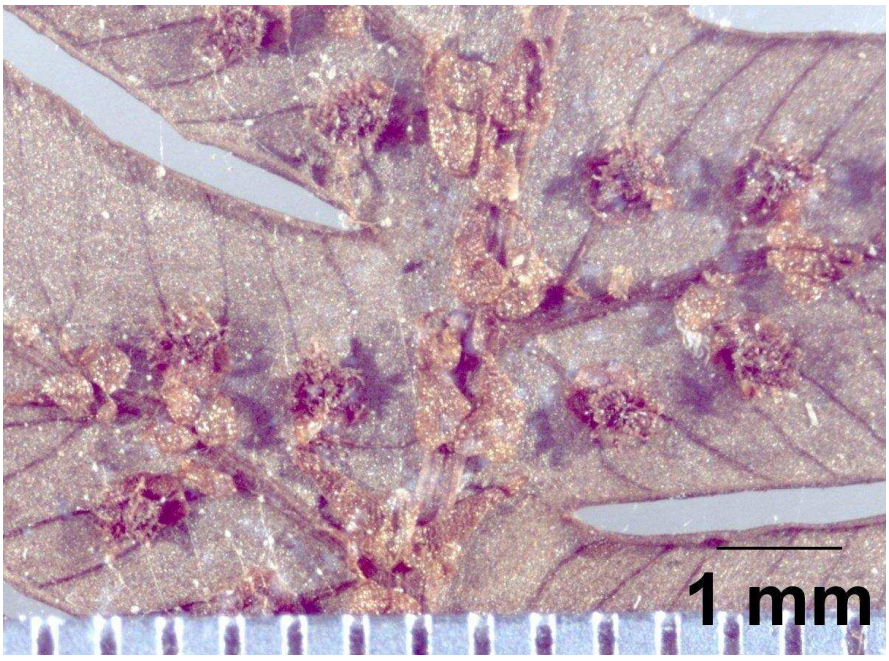
Figure 2.23 ***A****. Plant sample from Lake Kenyir.* ***B.*** *Part of pinna and stipe.* ***C****. Sporangia.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-LK11-RNG).*

Fronds are bi-pinnate or tri-pinnate and 1 to 2m long. Pinnules commonly cut almost to costa throughout and lower pinnules sessile or nearly so (Holttum, 1963). The stipe is light brown (N199C), has warts and spines, and lightly covered with dark brown scales (N200A) that are small and fringed. Sori occur near the mid-vein and are covered by small, bilobed indusia (Large and Braggins, 2004).

2.3.1.9 *C. latebrosa* (Wallich ex W. J. Hooker) Copeland

A

B



C

D **E**

Figure 2.24 ***A****. Plant sample from Penang Hill.* ***B.*** *Part of pinna and stipe.* ***C****. Scales on costa and costules with sporangia.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-PH08-RNG).*

Fronds are bi-pinnate or tri-pinnate and about 2m long. Pinnules not articulate (Holttum, 1963). The stipe is light medium brown (N199B), has spines and scales near the base. The scales are dark brown (N200A) and glossy. Sori occur near the mid-vein and are covered by small, bilobed, scale-like indusia (Large and Braggins, 2004).

2.3.1.10 *C. assimilis* W.J. Hooker (Blume) Copeland

A

B

Figure 2.25 ***A****. Part of pinna.* ***B****. Stipe. Photo courtesy of UKM herbaria.*

Fronds are pinnate or bi-pinnate, and 1 to 2m long. Pinnules distinctly lobed. The stipe is medium to dark brown (N199B-200D), has fine warts and scales near the base; the scales are brown (200C) and glossy. Sori occur near the mid-vein and are covered with thin indusia (Large and Braggins, 2004).

2.3.1.11 *C. moluccana* R. Brown in Desvaux

A **B**

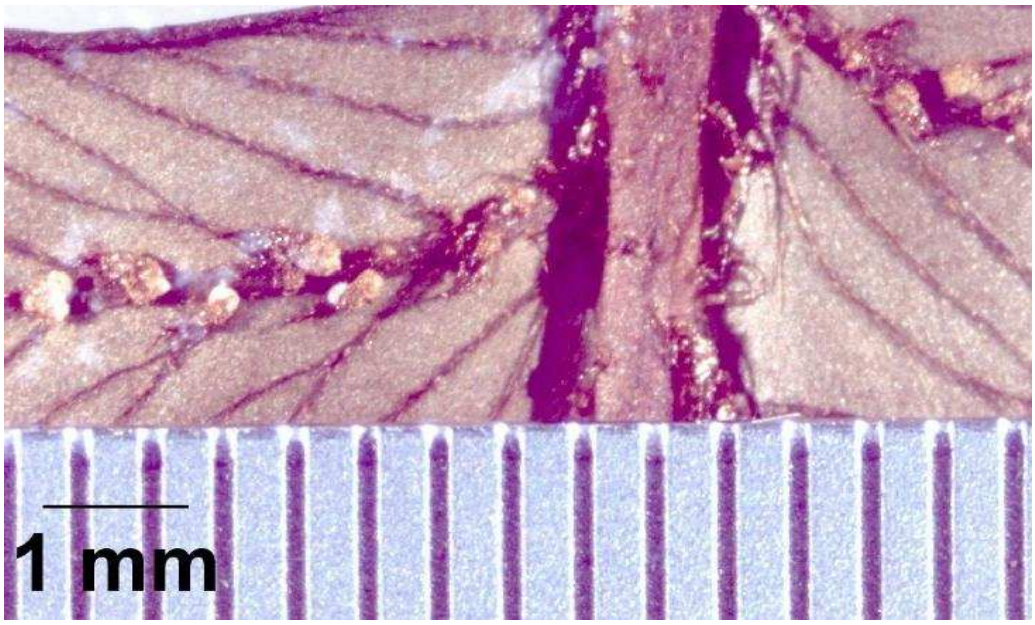
Figure 2.26 ***A****. and* ***B****. Part of pinna. Photo courtesy of UKM herbaria.*

Fronds are pinnate and 1.75 to 3m long. Pinnae not long-acuminate, upper usually sessile (Holttum, 1963). The stipe is light brown (N199D) and has scales. Scales are medium brown (N199C). Sori occur in one to three rows on either side of the mid-vein and are covered by translucent indusia (Large and Braggins, 2004).

2.3.1.12 *C. obscura* (Scort) Copeland

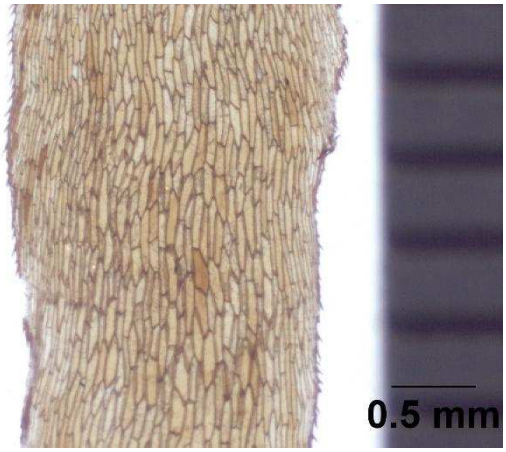
A

B



C

D



E

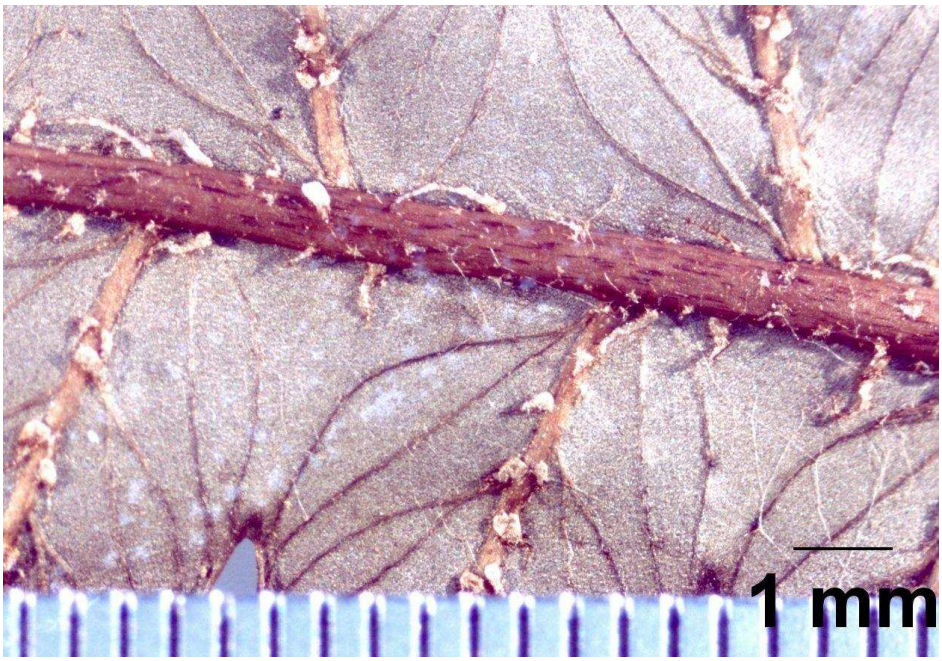
Figure 2.27 ***A****. Plant sample from Bukit Larut.* ***B.*** *Part of pinna and stipe.* ***C****. Scales on costa and costules.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-BL26-RNG).*

Fronds are pinnate to bi-pinnate and 1 to 2m long. Pinnules cut 2/3 to costa. The stipe is medium brown to dark brown (N199B-200C) and densely scaly toward the base. Scales are medium brown (N199C) and glossy. Sori occur on three pairs of veins about halfway between the mid-vein, and indusia are absent (Large and Braggins, 2004).

2.3.1.13 *C. polypoda* Baker

A

B



C

D **E**

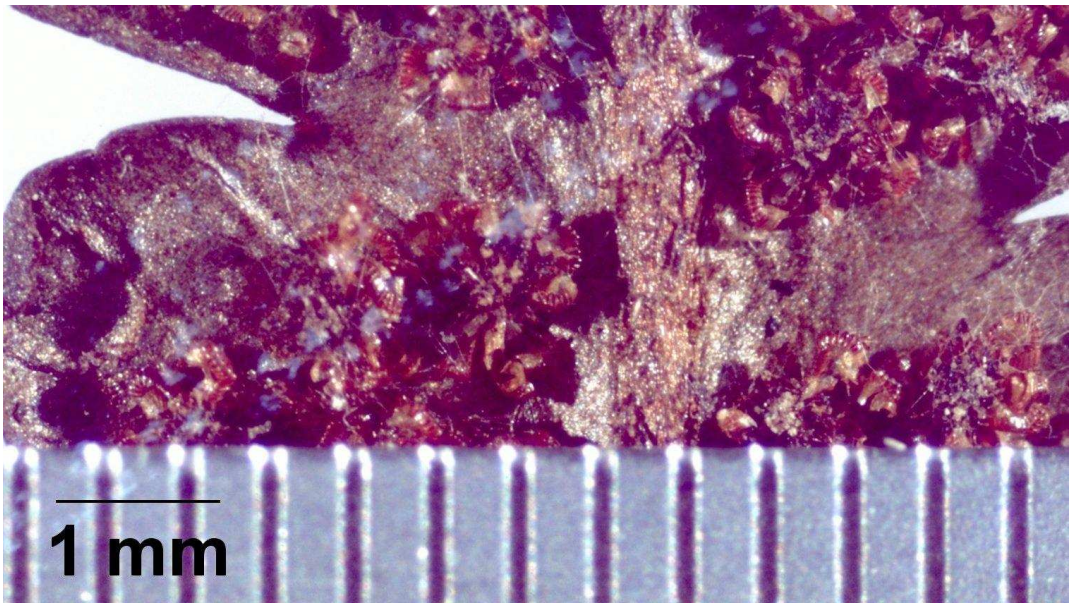
Figure 2.28 ***A****. Plant sample from Bukit Larut.* ***B.*** *Part of pinna and stipe.* ***C****. Scales on costa and costules.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-FH31-RNG).*

Fronds are pinnate or bi-pinnate and 1 to 2m long. Pinnules on stalks and distinctly lobed. The stipe is long, dark brown (N200A), and densely scaly toward the base. Scales are light brown (199A) and glossy. Sori occur near the mid-vein, and indusia are absent (Large and Braggins, 2004).

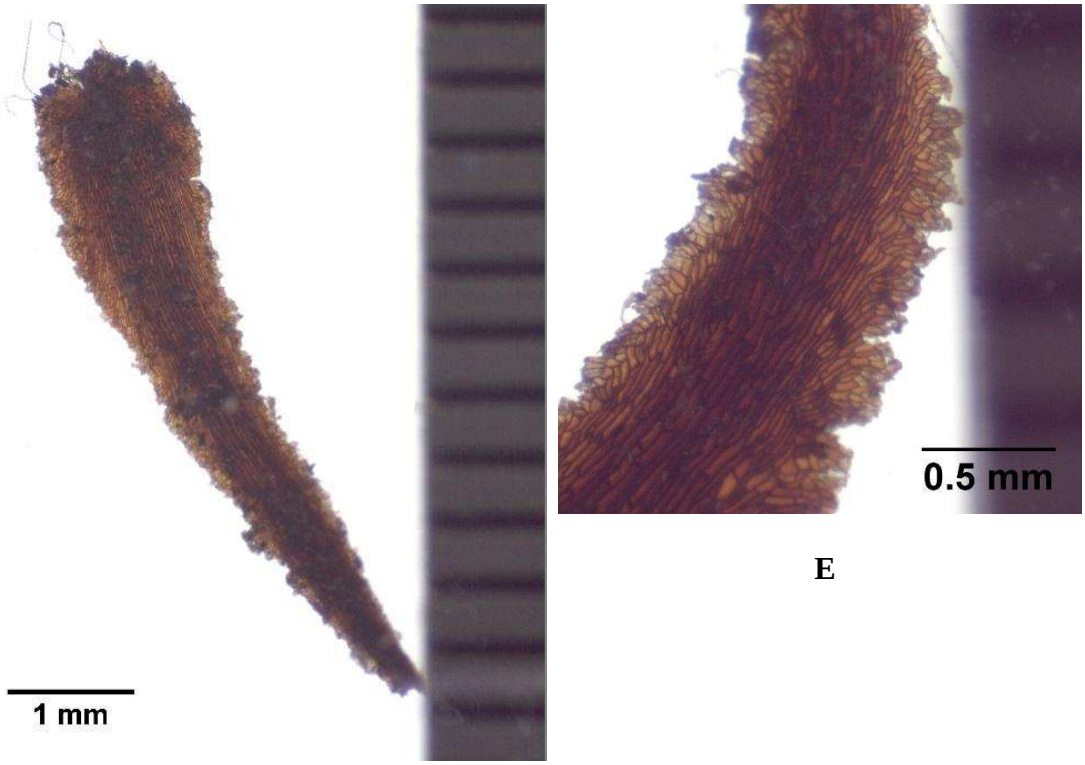
2.3.1.14 *C. recommutata* Copeland

A

B



C



D **E**

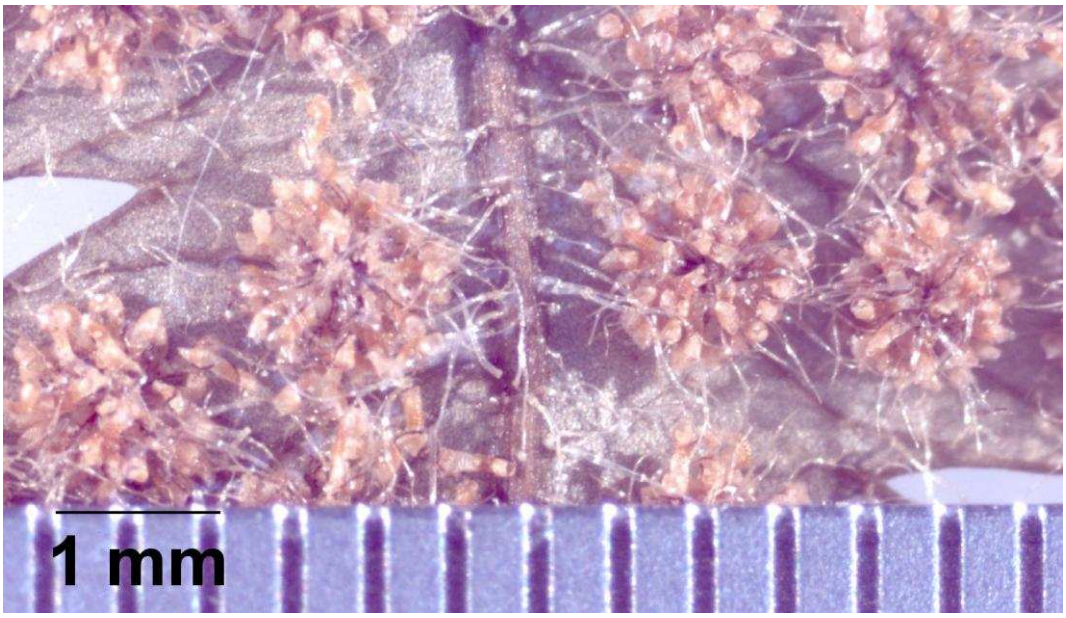
Figure 2.29 ***A****. Plant sample from Mount Ledang.* ***B.*** *Part of pinna and stipe.* ***C****. Sporangia.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-ML14-RNG).*

Fronds are bi-pinnate and 1 to 2m long. Reduced pinnae present at base of stipe, separated from normal pinnae. Fertile pinnules are smaller than sterile. The stipe is dark (202A) and has scales toward the base. Scales are dark brown (N200A) and glossy. Sori occur near the mid-vein, and indusia are absent (Large and Braggins, 2004).

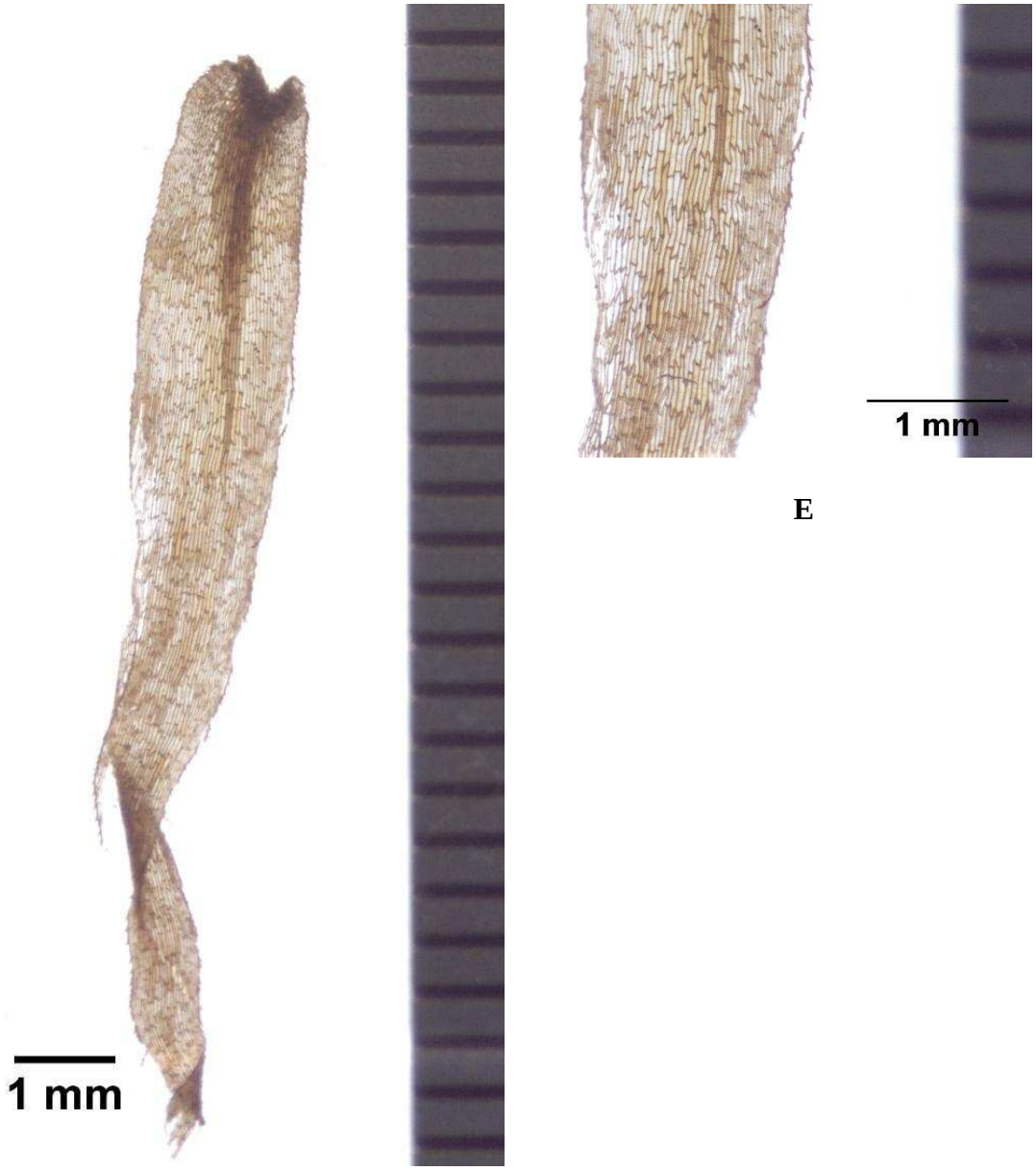
2.3.1.15 *C. trichodesma* (Scort) Copeland

A

B



C



D **E**

Figure 2.30 ***A****. Plant sample from Mount Angsi.* ***B.*** *Part of pinna and stipe.* ***C****. Sporangia.* ***D****. Scale from stipe.* ***E.*** *Part of the scale (AJ-MA47-RNG).*

Fronds are pinnate or bi-pinnate and 1 to 2m long. Pinnules cut 2/3 to costa. The stipe is long, light to medium brown (N199C-N199D), and warty at the base. Scales are light to medium brown (N199C-N199D) and glossy. Sori occur near the mid-vein and indusia are absent (Large and Braggins, 2004).

2.4 Discussion

Referring to the keys from Holttum (1963), the specimens were identified using several morphological characters such as scales, indusia, pinnules shape, veins, hairs and stipe. As the common name suggests (i.e. scaly tree ferns), Cyatheaceae identification relies heavily on scale morphology and often requires microscopy. However scale morphology is more valuable for generic delimitation than for species as demonstrated by Conant *et al.*, (1994, 1995) and Stein, Conant and Valinski (1997). Their works show that three genera were separated based on analyses of restriction site data as well as scale morphology in a maximum parsimony framework. Specifically, Conant *et al.,* (1994, 1995) and Stein *et al.,* (1997) confirmed that samples with conform scales belonged in *Sphaeropteris,* marginate scales without apical seta in *Cyathea*, and marginate scales with apical seta in *Alsophila* with the latter as a sister to all of the other groups. However, the relationship between *Alsophila* was found to be weakly supported, thus Conant *et al.,* (1994, 1995) and Stein *et al.*, (1997) suggest that marginate scales are plesiomorphic within the family, with a transition to conform scales in *Sphaeropteris.*

Apart from the scales, species identification relies on sporangium and indusium characters. For approximately 200 of the 419 samples collected, the laminae and pinnae were sterile, making sporangium and indusium-based identification difficult. During field collection, many of the plant individuals were too tall (more than 3m) to observe the fertile frond without cutting them.

Based on the key, pinnule shape was observed and this character was used to assist species delimitation by examining the size and depth of the lobes. Apart from that, other morphological characters such as hairs and scales were sometimes present on pinnules. These characters were used to compare the sterile with fertile laminae (which had been previously identified) thus assisting the identification of the unknown samples (Figure 2.31).

C. obscura *C. borneensis*

Figure 2.31 Pinnules shape differences from the identified species. Photo© 2013 Azi Jamaludin.

The vein morphology on the pinnules was also noted and used as one of the characters to compare with other samples. Even though the scale, indusium, and pinnule characters were mainly used for identification, the stipe and rachis were also observed in detail. Using the key, each character such as the color of the stipe and rachis, presence or absence of hairs, scales, or warts, and in some species, the length and color of the spines was taken into account. The work done was to ensure that each sample was identified as far as possible with the morphological features available.

2.5 Conclusion

The Cyatheaceae sampling expedition took place in most of the mountains and highlands in Peninsular Malaysia. The fieldwork allowed observation that major habitat conversion caused by anthropogenic effects has happened since Holttum's time. This fact has not only affected the distribution and abundance of Cyatheaceae but most of the seed plants and pteridophyte in Peninsular Malaysia. The primary motivation for this study was to gather population samples of the widest possible range of *Cyathea* species from Peninsular Malaysia to provide material for morphological and molecular study. Although there were 15 *Cyathea* species successfully identified using Holttum (1963) key from Flora Malesiana Series II: Pteridophyta, the information provided had not been revised and needed to be updated. In terms of methodology, the outcome of this work found a wider range of characters essential for Cyatheaceae field identification, especially when identifying species based on sterile individuals. Through the use of powerful molecular techniques such as phylogenetic inference and DNA barcoding methods presented in the succeeding chapters, current knowledge regarding species information and relationships on this Cyatheaceae family can be greatly understood. This work can also be added to the collection of studies done towards this family, offering updated information regarding the species in Peninsular Malaysia. Achieving species identification of nearly all of the collected Cyatheaceae sample is an indication of how important having an identification key is, particularly the updated and revised version.

CHAPTER 3

THE PHYLOGENETIC RELATIONSHIPS OF PENINSULAR MALAYSIA CYATHEACEAE

3.1 Introduction

Until 1970, classification of ferns was unstable and problematic as there were different ideas on the interpretation of the accessible evidence (Tryon, 1952; Christenhusz *et al.*, 2011). Work on fern families, genera and species classification before the 1970s were summarised in detail by Pichi-Sermolli (1973) as cited by Fraser-Jenkins (2009) in which he continued to improve and add more details to fern classification later on. However, the understanding of fern relationships encountered a major change in the mid-1990s, with the emergence of plastid DNA studies (Gastony and Yatskievych, 1992). The arrival of molecular phylogenetics further added to the knowledge regarding fern classification by redefining many of the genera and families (Christenhusz and Chase, 2014). Numerous molecular phylogenetic, as well as morphological studies have been executed since the advent of DNA sequence analysis (Hasebe *et al.*, 1994; Hasebe *et al.*, 1995; Conant *et al.*, 1994; Conant *et al.*, 1995; Pryer *et al.*, 1995; Pryer *et al.*, 2001; Pryer *et al.*, 2004; Wolf *et al.*, 1999). In year 2006, fern classification appeared to be more stable after Smith *et al.*, (2006) published their findings. Their work sums up molecular results to that date and provided synapomorphies for the accepted families (Christenhusz and Chase, 2014).

Since then, a number of fern families were further studied in greater detail, resolving many taxonomic problems. These were then incorporated in the updated classifications of ferns, for example Schuettpelz and Pryer (2007); Smith *et al.*, (2008); Christenhusz, Zhang and Schneider (2011); Lehtonen (2011); Rothfels *et al.*, (2012) (Figure 3.1). The classifications of Smith *et al.* (2006, 2008) and Christenhusz, Zhang and Schneider (2011) essentially decreased the number of genera, causing an expansion of several others, such as *Asplenium, Blechnum, Hymenophyllum* and *Cyathea,* but also resulted in the acceptance of narrower generic concepts in other groups such as Hymenophyllaceae, Polypodiaceae and Pteridaceae (Christenhusz and Chase, 2014).

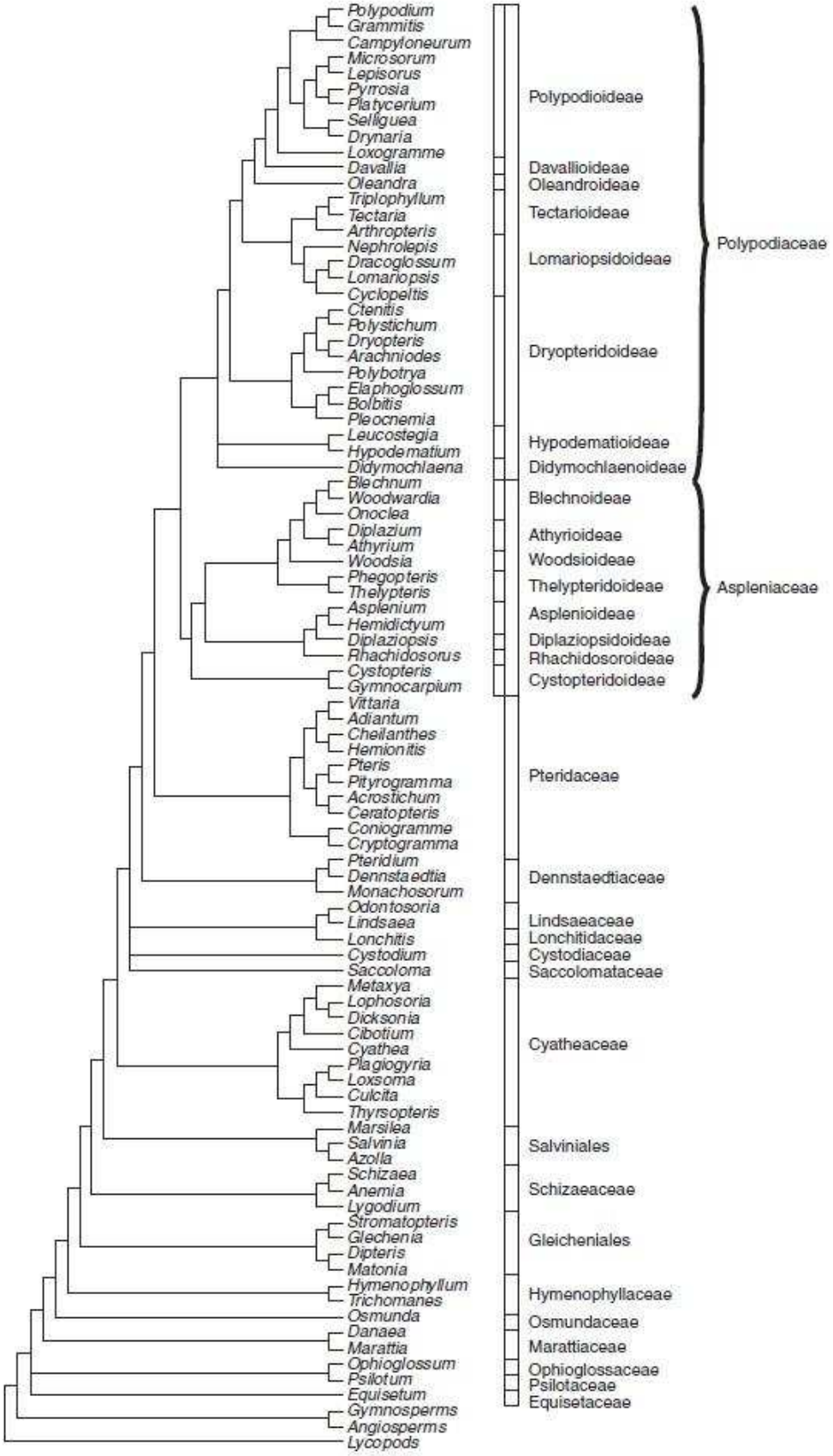


Figure 3.1 Summary phylogenetic tree showing relationships of a representative selection of fern genera based on molecular data, modified from (Schuettpelz and Pryer, 2007; Lehtonen, 2011; Rothfels et al., 2012; Schneider et al., 2013; Christenhusz and Chase, 2014).

Further studies in phylogenetics were made towards the tree fern clade Cyatheales. The clade is usually divided into eight families with Cyatheaceae representing a large proportion of the total tree ferns (Korall *et al.*, 2007). Tree ferns are all minimally genetically divergent, which is perhaps an outcome of the much longer generation time of these plants (Christenhusz and Chase, 2014). Tree ferns are highly divided at the family level and the lineages should still be updated taxonomically on the basis of synapomorphies and monophyly (Christenhusz and Chase, 2014).

One family of the tree fern clade, Cyatheaceae, attracted many taxonomists even before the arrival of DNA sequence analysis in phylogenetics (Korall *et al.*, 2007). The family has been the focus of many systematic and taxonomic treatments (Holttum, 1963; Tryon, 1970; Tryon and Gastony, 1975; Tryon and Tryon, 1982; Conant *et al.*, 1994; Conant *et al.*, 1995; Conant and Stein, 2001) but despite the attention, there remain many unanswered questions regarding relationships and character evolution within this group (Korall *et al.*, 2006; Korall *et al.*, 2007). The advent of DNA sequence analysis in phylogenetics led Conant *et al.* (1994, 1995) to recognise three major Cyatheaceae lineages, the genera *Alsophila*, *Cyathea* and *Sphaeropteris* (Figure 3.2). *Alsophila* was sister to the other two genera. Based from the study by Conant *et al.* (1994, 1995), Korall *et al.* (2007) made an investigation using five plastid regions: *rbcL*, *rbcL-accD, rbcL-atpB, trnG-trnR* and *trnL-trnF* that resulted in demonstration of a basal dichotomy within the family phylogeny (Figure 3.3), supported by scale morphologies with *Sphaeropteris* as a sister to all of the other taxa. This finding seems to contradict Conant *et al.* (1994) in terms of which group is the most basal. Conant *et al.* (1994) in their publication did not agree with the hypothesis made by Tryon (1970) in which Tryon had determined *Sphaeropteris* as

the most primitive group of living tree ferns. However, Conant *et al.* (1994) then acknowledged that additional information was needed to resolve the conflict by adding more *Sphaeropteris* species as well as representatives from Dicksoniaceae to be an outgroup.

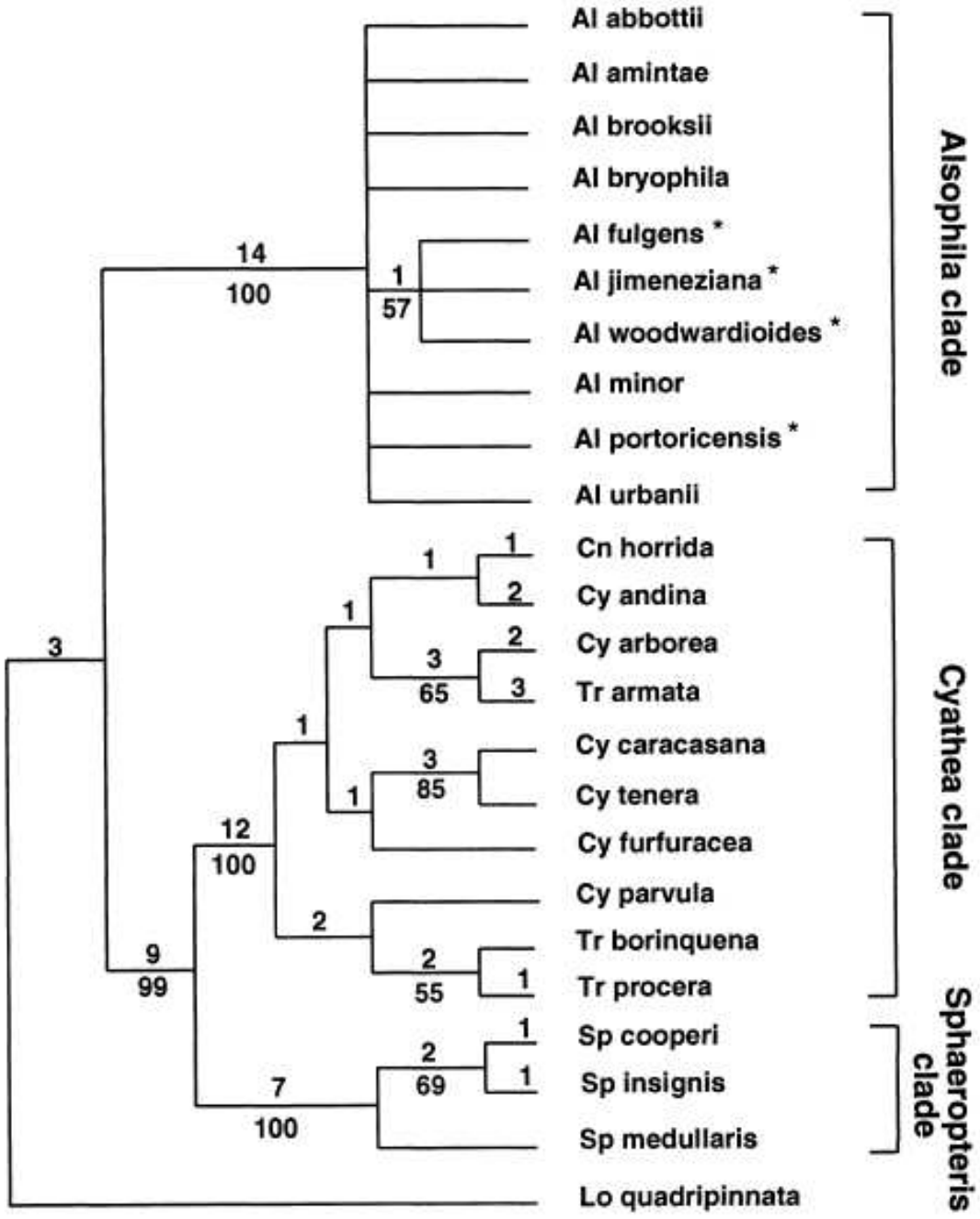
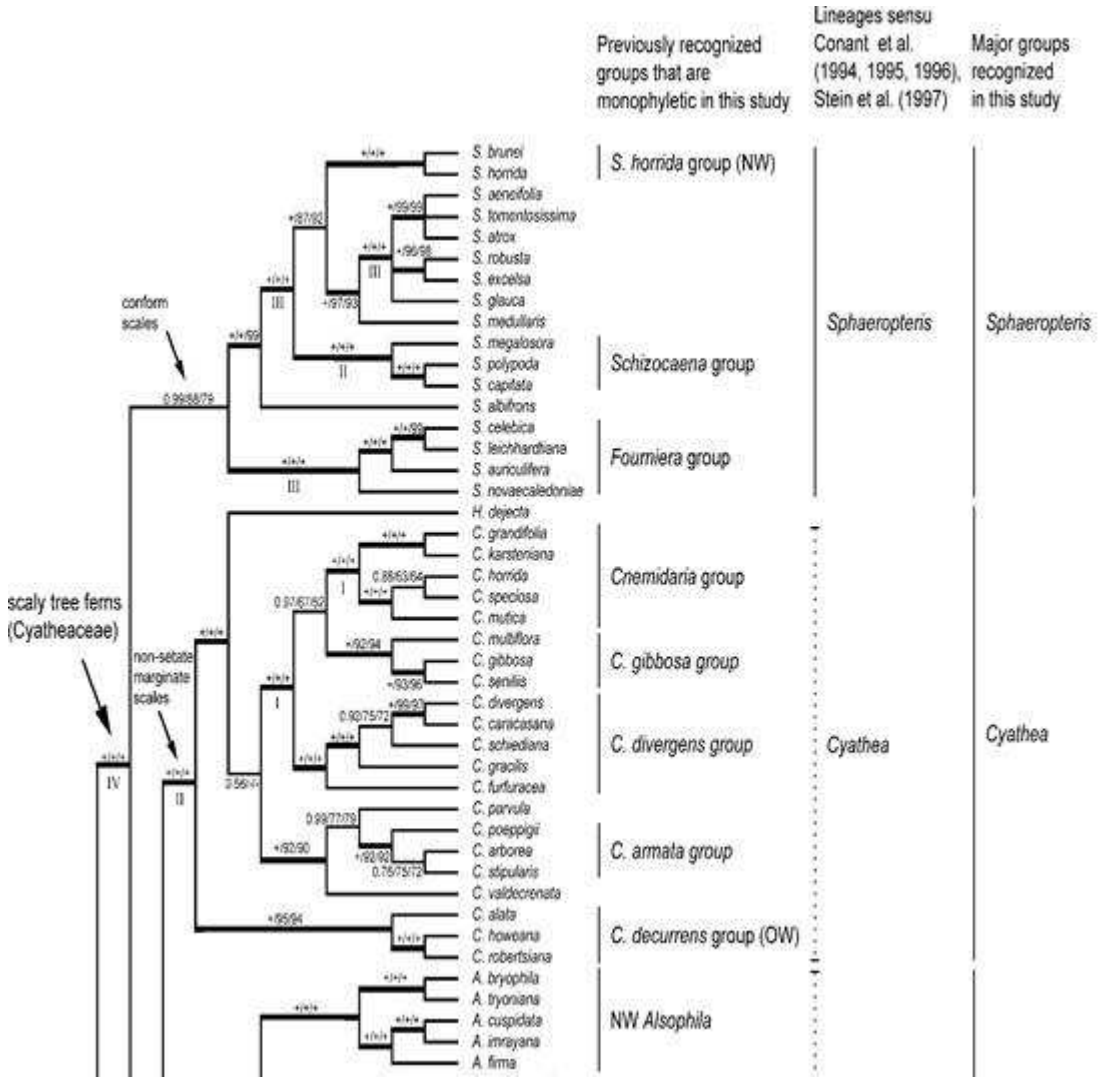


Figure 3.2 One of 60 equally most parsimonious Wagner trees of length 77 (excluding autapomorphies) and consistency index 0.75. Brackets at right indicate the tree major clades. Asterisks indicate species that Tryon and Tryon (1982) placed in the genus Nephelea. Numbers above and below the nodes indicate restriction site changes and number of times a monophyletic group appeared in 100 bootstrap replicates, respectively (Conant et al., 1994).


Previously recognized groups that are monophyletic in this study
Lineages sensu Conant et al. (1994, 1995, 1996), Stein et al. (1997)
Major groups recognized in this study
S. brunei
S. horrida
S. aenaifolia
S. tomentosissima
S. atrox
S. robusta
S. excelsa
S. glauca
S. medullaris
S. megalosora
S. polypoda
S. capitata
S. albifrons
S. celebica
S. leichhardtiana
S. auriculifera
S. novaecaledoniae
S. horrida group (NW)
Schizocaena group
Fourniera group
Sphaeropteris
Sphaeropteris
conform scales
0.99/88/79
H. dejecta
C. grandifolia
C. karsteniana
C. horrida
C. speciosa
C. mutica
C. multiflora
C. gibbosa
C. senilis
C. divergens
C. caracasana
C. schiediana
C. gracilis
C. furfuracea
C. parvula
C. poeppigii
C. arborea
C. stipularis
C. valdecrenata
C. alata
C. howeana
C. robertsiana
Cnemidaria group
C. gibbosa group
C. divergens group
C. armata group
C. decurrens group (OW)
Cyathea
Cyathea
scaly tree ferns (Cyatheaceae)
non-setate marginate scales
A. bryophila
A. tryoniana
A. cuspidata
A. imrayana
A. firma
NW Alsophila

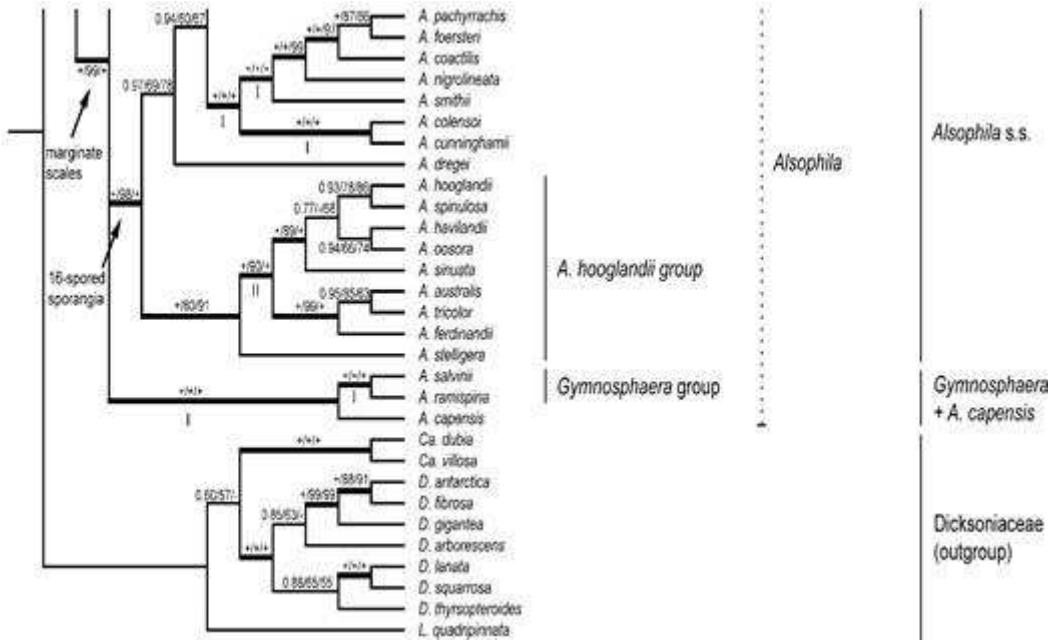


Figure 3.3 The 50% majority-rule consensus tree resulting from Bayesian (B/MCMC) analyses of the combined data set. Numbers above branches denote support values from Bayesian, maximum likelihood (ML), and maximum parsimony (MP) analyses, respectively: Posterior probabilities (PP)/ML bootstrap percentages (BP_{ML})/MP bootstrap percentages (BP_{MP}). A plus (+) represent PP =1.00, or BP_{ML}= 100, or BP_{MP} =100. A hyphen (-) represents bootstrap percentage <50%. Roman numerals below branches denote number of unambiguous indels (i.e., insertion or deletions events that are clearly delimited) in combined data set that support the node. Thickened branches are well supported (PP = 100, BP_{ML}, and BP_{MP} ≥90%). Previously recognized groups that are resolved as monophyletic in the study are indicated. The lineages of Conant et al. (1994, 1995) are indicated (dotted lines indicated non-monophyly). To the far right, the four major groups of scaly tree ferns that are recognized from the study are shown. A., Alsophila; C., Cyathea; Ca., Calochlaena; D., Dicksonia; H., Hymenophyllopsis; L., Lophosoria; S., Sphaeropteris. (Korall et al., 2007).

There were differences in basal position for the Cyatheaceae family between Conant *et al.* (1995) and Korall *et al.* (2007). These were probably due to higher species number and DNA regions used in the study made by Korall *et al.* (2007) thus resulting in a much more robust phylogeny. Korall *et al.* (2007); Korall and Pryer (2014) then proposed four genera for this family: *Sphaeropteris*, *Cyathea*, *Alsophila* and *Gymnosphaera.*

3.1.1 Aims

The aim of this study is to use DNA sequence from four plastid regions: *matK*, *rbcL*, *trnL-F* and *trnG-R* to investigate the phylogenetic relationships of the Peninsular Malaysian scaly tree ferns (Cyatheaceae) with the existing molecular phylogeny.

3.2 Materials and Methods

3.2.1 Sample Collection

Four hundred and nineteen samples described in Chapter 2 were used for DNA extraction.

3.2.2 DNA Extraction and Amplification

Total genomic DNA was isolated from approximately 0.03g of silica-gel dried plant material following the modified CTAB protocol from Doyle and Doyle (1987), which it had been altered initially for Daffodil extraction (Appendix 1) (Könyves, 2014). Protocol from Nunes *et al.* (2011) was also used (Appendix 2). However, due to the

high content of secondary metabolites present in the samples, the CTAB method had to be modified accordingly (Appendix 3). The extracted DNA was then stored in 100 µl of TE buffer at -20^{0}C for subsequent use.

Prior to amplification, a pilot run (Table 3.1) had been conducted on several primer pairs for each the four regions to find the most universal pairs which can be used to amplify the DNA. The samples tested were haphazardly selected from across the thirteen Cyatheaceae populations. The primers to use for each region were decided based from the pilot test outcomes. Using the finalized primer pairs (Table 3.2), all of the 419 Cyatheaceae DNA samples were amplified according to the recommended PCR conditions. Standard Polymerase Chain Reaction (PCR) was performed in 50 µl reaction mixtures containing 25µl of BioMix^{TM} Red from Bioline, 1µl of Bovine Serum Albumin (BSA), 1.75µl of 10uM of each primer: forward and reverse, 18.5µl of Millipore H_2O and 2µl of 50 to 100ng/µl template DNA. Detailed information on PCR profiles is given in Table 3.3. Amplifications of the templates were run on a Veriti® 96 well thermal cycler and the final PCR products were run on 1% agarose gel stained with ethidium bromide.

Table 3.1 Primers pairs tested prior amplification for all Cyatheaceae samples.

Loci	Primer Pairs	Reference	Sample Code											
			BL24	**LH14**	**PH24**	**LK23**	**MJ02**	**MP14**	**MB09**	**BK39**	**FH61**	**GH07**	**ML13**	**MA47**
rbcL	RBCL1187F/ ACCD816R	(Korall *et al.*, 2006)	✓	✓	✓	×	×	×	×	×	×	×	×	×
	FWrbcL392F/ FWrbcL874R	(Korall *et al.*, 2007)	×	×	×	×	×	×	×	×	×	×	×	×
	ESRBCL1F/ ESRBCL1361R	Schuettpelz and Pryer (2007)	✓	✓	✓	✓	✓	×	×	×	×	×	×	×
matK	Lb matK rYIY/ Tf matK rRLA	(Li *et al.*, 2011; Kuo *et al.*, 2011)	×	×	×	×	×	×	×	×	×	×	×	×
	matK390F/ matK1326R		×	×	×	×	×	×	×	×	×	×	×	×
	FWPtmatKF1/ FWPtmatKR522		×	×	×	×	×	×	×	×	×	×	×	×
	FWPtmatKF867/ FWPtmatKrAGK		×	×	×	×	×	×	×	×	×	×	×	×
	FWPtmatKrAGK/ FWPtmatKfEDR		✓	✓	✓	✓	×	×	×	×	×	×	×	×
trnL-F	trnLc/ trnFf	Taberlet *et al.* (1991)	×	×	×	×	×	×	×	×	×	×	×	×
	F/ FernLr1	Li *et al.* (2011)	✓	✓	✓	✓	✓	×	✓	×	×	×	×	×
trnG-R	TRNG1F/ TRNR22R	Nagalingum *et al.* (2007)	✓	✓	✓	×	✓	✓	×	×	×	×	×	×

Legend

✓	Successfully amplified
×	Failed to amplify

Table 3.2 List of primers used for amplification of all samples.

Region	Name	Primer sequence (5' to 3')	Reference
rbcL	ESRBCL1F	ATGTCACCACAAACGGAGACTAAAGC	Schuettpelz and Pryer (2007)
	ESRBCL1361R	TCAGGACTCCACTTACTAGCTTCACG	Schuettpelz and Pryer (2007)
matK	FWPtmatK rAGK	CGTATTGTACTYCTATGTTTRCCAGC	Kuo *et al.* (2011)
	FWPtmatK fEDR	ATTCATTCRATRTTTTTATTTHTGGAAGATAGATT	Kuo *et al.* (2011)
trnL-F	F	ATTTGAACTGGTGACACGAG	Taberlet *et al.* (1991)
	FernLr1	GGCAGCCCCCAGATTCAGGGGAACC	Li *et al.* (2011)
trnG-R	TRNG1F	GCGGGTATAGTTTAGTGGTAA	Nagalingum, Schneider and Pryer (2007)
	TRNR22R	CTATCCATTAGACGATGGACG	Nagalingum, Schneider and Pryer (2007)

Table 3.3 PCR profiles.

Region	Initial denaturing Temperature/ time	Denaturation temperature/ time	Annealing temperature/ time	Extension temperature/ time	Final extension temperature/ time	No. of cycles
rbcL	94^0C/5:00	94^0C/1:00	50^0C/1:00	72^0C/2:00	72^0C/10:00	35
matK	94^0C/5:00	94^0C/1:00	52.5^0C/1:00	72^0C/2:00	72^0C/10:00	35
trnL-F	95^0C/5:00	95^0C/0:50	57^0C/0:50	72^0C/1:00	72^0C/10:00	35
trnG-R	95^0C/2:00	95^0C/0:30	55^0C/0:30	72^0C/1:00	72^0C/5:00	35

3.2.3 Purification and Sequencing

Both purification and sequencing were done at the Source Bioscience in Biochemistry Department, University of Oxford, United Kingdom. All of the PCR products were sequenced using the Sanger method (Sanger and Coulson, 1975) using the same primers used for PCR amplification.

3.2.4 Genbank Data

All 64 species from Cyatheaceae and ten species from Dicksoniaceae sequences used in a previous study by Korall *et al.* (2007) to build a Cyatheaeceae phylogeny were downloaded from the Genbank (Table 3.4). The *Cyathea* lineage (including *Cnemidaria* and *Trichipteris*) is represented by 21 species, *Alsophila* (including *Nephelea*) by 25, and *Sphaeropteris* by 17. *Hymenophyllopsis* is represented by a single species. Only three DNA regions from the study made by Korall *et al.* (2007): *rbcL, trnL-F* and *trnG-R* were added to the analysis as the study previously conducted did not include the *matK* region.

Several other species from different families in Cyatheales were also added in the study to be used as an outgroup. The species included were from Metaxyaceae, Culcitaceae, Loxomataceae, and Thyrsopteridaceae. One species from Aspleniaceae was also added (Table 3.5).

Table 3.4 Genbank accession numbers for each species from Cyatheaceae used in (Korall et al., 2007).

Species	***rbcL***	***trnG-R***	***trnL-F***
Alsophila australis R. Br.	AM177319	AM410379	AM410314
Alsophila bryophila R. Tryon	AM177320	AM410364	NA
Alsophila capensis (L. f.) J. Sm.	AM177321	AM410381	AM410316
Alsophila coactilis (Holtt.) R. Tryon	AM410205	AM410404	AM410336
Alsophila colensoi Hook. F.	AM177322	AM410383	AM410318
Alsophila cunninghamii (Hook. F.) R. Tryon	AM410211	AM410410	AM410339
Alsophila cuspidata (Kunze) D. S. Conant	AM177323	AM410388	NA
Alsophila dregei (Kunze) R. Tryon	AM410194	AM410380	AM410315
Alsophila ferdinandii R. Tryon	AM410204	AM410403	AM410335
Alsophila firma (Baker) D. S. Conant	AM410207	AM410406	NA
Alsophila foersteri (Rosenst.) R Tryon	AM177324	AM410390	AM410324
Alsophila havilandii (Baker) R. Tryon	AM410189	AM410373	NA
Alsophila hooglandii (Holtt.) R. Tryon	AM177325	NA	AM410306
Alsophila imrayana (Hook.) D. S. Conant	AM410202	AM410395	AM410329
Alsophila nigrolineata (Holtt.) R. Tryon	AM410206	AM410405	AM410337
Alsophila oosora (Holtt.) R. Tryon	AM410209	AM410408	NA
Alsophila pachyrrachis (Copel.) R. Tryon	AM410186	AM410370	AM410305
Alsophila ramispina Hook.	AM177326	AM410389	AM410323
Alsophila salvinii Hook.	AM410184	AM410365	AM410300
Alsophila sinuata (Hook. & Grev.) R. Tryon	NA	AM410402	NA
Alsophila smithii (Hook. f.) R. Tryon	AM410210	AM410409	AM410338
Alsophila spinulosa (Hook.) R. Tryon	AM410212	AM410411	AM410340
Alsophila stelligera (Holtt.) Tryon	AM410198	AM410391	AM410325
Alsophila tricolor (Colenso) R. Tryon	AM410199	AM410392	AM410326
Alsophila tryoniana (Gastony) D. S. Conant	AM410208	AM410407	NA
Cyathea alata Copel.	AM177335	AM410363	NA
Cyathea arborea (L.) Sm.	AM177336	AM410396	NA
Cyathea caracasana (Klotzsch) Domin	AM410223	AM410422	AM410351
Cyathea divergens Kunze	AM177337	AM410386	AM410321
Cyathea furfuracea Baker	AM410224	AM410423	AM410352
Cyathea gibbosa (Klotzsch) Domin	AM177354	AM410397	AM410330
Cyathea gracilis Griseb.	AM410217	AM410416	AM410345
Cyathea grandifolia Willd.	AM177332	AM410367	AM410302
Cyathea horrida (L.) Sm.	AM410196	AM410385	AM410320

Cyathea howeana Domin	AM410188	AM410372	AM410308
Cyathea karsteniana (Klotzsch) Domin	AM410221	AM410420	AM410349
Cyathea multiflora Sm.	AM410197	AM410387	AM410322
Cyathea mutica (Christ) Domin	AM410220	AM410419	AM410348
Cyathea parvula (Jenman) Domin	AM177338	AM410384	AM410319
Cyathea poeppigii Domin	AM410201	AM410394	AM410328
Cyathea robertsiana (F. v. Muell.) Domin	AM410216	AM410415	AM410344
Cyathea schiediana (C. Presl) Domin	AM410218	AM410417	AM410346
Cyathea senilis (Klotzsch) Domin	AM410203	AM410399	AM410332
Cyathea speciosa H. & B. ex Willd.	AM177339	AM410398	AM410331
Cyathea stipularis (Christ) Domin	AM410219	AM410418	AM410347
Cyathea valdecrenata Dominc	AM410222	AM410421	AM410350
Hymenophyllopsis dejecta (Baker) Goebel	AF101301	AM410362	AM410299
Sphaeropteris aeneifolia (v. A. v. R.) R. Tryon	AM410185	AM410368	AM410303
Sphaeropteris albifrons (Fourn.) R. Tryon	AM410214	AM410413	AM410342
Sphaeropteris atrox (C. Chr.) R. Tryon	AM410225	AM410424	AM410353
Sphaeropteris auriculifera (Copel.) R. Tryon	AM177348	AM410401	AM410334
Sphaeropteris brunei (Christ) R. Tryon	AM177349	AM410366	AM410301
Sphaeropteris capitata (Copel.) R. Tryon	AM410192	AM410376	AM410311
Sphaeropteris celebica (Bl.) R. Tryon	AM410195	AM410382	AM410317
Sphaeropteris excelsa (Endl.) Tryon	AM410213	AM410412	AM410341
Sphaeropteris glauca (Bl.) R. Tryon	AM410193	AM410377	AM410312
Sphaeropteris horrida (Liebm.) R. Tryon	AM410200	AM410393	AM410327
Sphaeropteris leichhardtiana (F. v. Muell.) Copel.	AM410215	AM410414	AM410343
Sphaeropteris medullaris (G. Forst.) Bernh.	AM177350	AM410378	AM410313
Sphaeropteris megalosora (Copel.) R. Tryon	AM410190	AM410374	AM410309
Sphaeropteris novaecaledoniae (Mett.) R. Tryon	AM177351	AM410400	AM410333
Sphaeropteris polypoda (Baker) R. Tryon	AM410191	AM410375	AM410310
Sphaeropteris robusta (Watts) R. Tryon	AM410187	AM410371	AM410307
Sphaeropteris tomentosissima (Copel.) R. Tryon	AM177352	AM410369	AM410304
Calochlaena dubia	U05615	AM410425	NA

(R. Br.) M. D. Turner & R. A. White			
Calochlaena villosa (C. Chr.) M.D. Turner & R. A. White	AM177327	AM410426	AM410354
Dicksonia antarctica Labill.	U05919	AM410427	AM410355
Dicksonia arborescens L'He´r.	AM177340	AM410428	AM410356
Dicksonia fibrosa Col.	AM177341	AM410429	NA
Dicksonia gigantea H. Karst.	AM177342	AM410430	AM410357
Dicksonia lanata Col.	AM177343	AM410431	AM410358
Dicksonia squarrosa (G. Forst.) Sw.	AM177344	AM410432	AM410359
Dicksonia thyrsopteroides Mett	AM177345	AM410433	AM410360
Lophosoria quadripinnata (J. F. Gmel.) C. Chr.	AF101303	AM410434	AM410361

Table 3.5 Genbank accession numbers for outgroup species used in this study.

Species	***matK***	***rbcL***	***trnG-R***	***trnL-F***
Asplenium trichomanes L.	JF832256	EF463157	KP861389	JX475144
Culcita macrocarpa C. Presl	JF303913	AM177334	NA	NA
Dicksonia antratica Labill	HM021802		**See Table 3.4**	
Loxoma cunninghamii R. Br. & A. Cunn	JF303912	U30834	NA	NA
Metaxya lanosa A. R. Sm. & Tuomisto	JF303909	AF317701	KP244152	NA
Metaxya rostata (Kunth) C. Presl	KP244035	AF317700	KP244132	HQ157338
Thyrsopteris elegans Kunze	JF303910	AM177353	NA	HG422548

3.2.5 Sequence Assembly and Alignment

The raw sequence data for each of the four datasets (one dataset per DNA region) were assembled and edited using SeqMan® Pro version 13.0 (DNASTAR, 2016). The sequence data were then uploaded in BLAST search on GenBank to make sure none of the sequences acquired were contaminated. The resulting sequences together with the sequences from GenBank were then aligned using the multiple alignment Clustal

W algorithm as implemented in BioEdit version 7.2.5 (Hall, 1999) with further visual and manual adjustments, including misaligned regions. Sequences that could not be aligned were excluded and indels were treated as missing data.

A sequence alignment was prepared for each of the regions: *matK*, *rbcL*, *trnL-F* and *trnG-R* as well as the combination of all of the four regions. The combined matrix is the result from concatenating all of the four regions and treating the missing sequences from each region as missing data. Incomplete or partial sequences were also included and identical sequences were removed using Jalview version 2 software (Waterhouse *et al.*, 2009). The final list of *Cyathea* species as well as one sample from Ciboteaceae (*Cibotium barometz*) used in the analysis is presented in Table 3.6.

Table 3.6 List of species used in the phylogenetic analysis of Cyatheaceae in this study. Capital letters in sample codes represent sampling location followed by number in which order the sample was collected. NA is not applicable (no DNA sequence available).

Species	***matK***	***rbcL***	***trnG-R***	***trnL-F***
	Sample Codes			
Cyathea contaminans	LH01	LH01	LH01	LH01
Cyathea latebrosa	MJ09	MJ09	MJ09	MJ09
Cyathea borneensis	PH24	PH24	PH24	NA
Cyathea hymenodes	NA	MJ05	MJ05	NA
Cyathea obscura	BL15	BL15	NA	BL15
Cyathea trichodesma	BL07	BL07	NA	BL07
Cyathea polypoda	BF4	BF4	BF4	BF4
Cyathea assimilis	NA	MA20	NA	NA
Cyathea alternans	NA	MA16	NA	NA
Cyathea moluccana	HB1	HB1	HB1	HB1
Cyathea recommutata	NA	ML19	NA	ML19
Cyathea glabra	NA	BF8a	BF8a	BF8a
Cyathea lurida	NA	MB43	MB43	MB43
Cyathea gigantea	NA	ML24	ML24	NA
Cibotium barometz	MJ10	MJ10	MJ10	MJ10

3.2.6 Phylogenetic Analysis using Bayesian Inference

Bayesian inference was conducted in MrBayes version 3.1.2 (Ronquist *et al.*, 2012) by first determining the optimal substitution model using MrModelTest version 2.3 (Nylander, 2004). The best-fitting model of evolution for each region was selected with the Akaike Information Criterion (AIC) as a measure of optimality. The model determined for *matK* and *trnL-F* was GTR+G (nst=6 rates=gamma), while GTR+I+G (nst=6 rates=invgamma) for both *rbcL* and *trnG-R*. Two independent runs each with four Markov Chain Monte Carlo replicates (MCMC) (one cold and three heated) were run for 2,500,000 generations for all of the regions. Each tree was sampled every 10,000th generation. As for the combined matrix, two partitions with GTR+G and GTR+I+G were applied respectively for four character sets but the two independent runs with four MCMC (one cold and three heated) were run for 3,000,000 generations with each tree sampled every 10,000th generation. The sample frequency was set for every 10,000th generation to reduce convergence time as well as tree and parameter samples. Analysis was run until the convergence diagnostic and the average standard deviation of split frequencies reaches a value below 0.01.

A plot of negative log likelihoods (LnL) against tree likelihood (TL) was done using Markov chains to measure the burn-in. The output log files of the two independent runs for both individual regions and combined matrix were assessed using Tracer v1.6 (Rambaut *et al.*, 2014) to check for the convergence as well as the suitable burn-in. The 10% of the sampled trees were discarded as 'burn in' and the remaining trees were used to build a 50% majority rule consensus tree with posterior probability for

nodes. The consensus tree was exported and viewed using FigTree version v1.4.2 software (Rambaut, 2014).

3.3 Results

3.3.1 DNA Amplification, Sequencing and Assembling

Most of the regions amplified poorly. The most successful amplification was from the region *trnL-F*, but with very low sequence assembling success due to short sequences. Only 28 samples from three species (*Cyathea contaminans, C. polypoda* and *C. latebrosa)* were successfully amplified, sequenced, and assembled for all of the regions. Table 3.7 summarizes the percentage of success for amplification, sequencing and assembling success of the four regions used.

Table 3.7 Summary of amplification, sequencing and assembling success for the four regions used.

DNA Regions	N Individuals/ 419	% of Amplification Success	% of Sequencing Success	% of Sequence Assembling Success
rbcL	137	33	97	77
matK	130	31	89	93
trnL-F	313	75	98	59
trnG-R	214	50	99	75

3.3.2 The Individual DNA Region Phylogenies

The results from the analysis for the four regions show mostly well supported relationships with posterior probability (PP) greater than 0.70, unless otherwise stated. The trees generated generally conform to each other in topology except for *matK* due to the absence of taxa from Korall et al. (2007) for this region.

Analysis for the *matK* region includes 1380 characters with 642 conserved sites and 709 variable sites. The tree is well supported with most PP greater than 0.70, unless otherwise stated (Figure 3.4). All of the *Cyathea* species from Peninsular Malaysia are grouped together. Other Cyatheales species from *Dicksonia* and *Metaxya* are clustered together with *Cibotium* as a sister. Species from *Loxoma, Culcita,* and *Thyrsopteris* are positioned as a sister group to all others, in accordance with classification made by Smith *et al.* (2006). *Asplenium trichomanes* is positioned as an outgroup.

For the *rbcL* analysis, 1322 characters were included with 941 conserved sites and 371 variable sites. The tree (Figure 3.5) is a result from a combined analysis of the Peninsular Malaysian species and the species from Korall *et al.* (2007). The analysis showed *C. alternans, C. latebrosa, C. borneenis* and *C. assimilis* embedded among the *Alsophila* group. *C. contaminans* is well positioned in *Sphaeropteris* as well as *C. gigantea, C. moluccana, C. obscura, C. polypoda* and *C. trichodesma*. *C. recommutata, C. lurida* and *C. glabra* are grouped together with *Gymnosphaera*. *Cibotium barometz* is grouped together with Dicksoniaceae and other Cyatheales species with *Asplenium trichomanes* as outgroup.

The *trnG-R* region analysis comprised of 1328 characters with 580 conserved sites, 616 variable sites. The tree (Figure 3.6) is a result from a combined analysis of the Peninsular Malaysian species and the species from Korall *et al.* (2007). The analysis showed *C. latebrosa* and *C. borneensis* embedded among the *Alsophila* group while *C. lurida* and *C. glabra* are grouped together with *Gymnosphaera. C. contaminans* is well positioned in *Sphaeropteris* as well as *C. gigantea, C. obscura, C. moluccana, C. polypoda*. *Cibotium barometz* is grouped together with Dicksoniaceae and other Cyatheales species with *Asplenium trichomanes* as outgroup.

Analysis for the *trnL-F* region includes 1361 characters with 528 conserved sites, and 652 variable sites. The tree (Figure 3.7) is a result from a combined analysis of the Peninsular Malaysian species and the species from Korall *et al.* (2007) The analysis showed *C. latebrosa* and *C. borneensis* embedded among *Alsophila* group while *C. lurida* and *C. glabra* are grouped together with *Gymnosphaera. C. contaminans* is well positioned in *Sphaeropteris* as well as *C. gigantea, C. hymenodes, C. obscura, C.*

polypoda, *C. trichodesma* and *C. moluccana. Cibotium barometz* is grouped together with Dicksoniaceae and other Cyatheales species with *Asplenium trichomanes* as outgroup.

3.3.2.1 *matK*

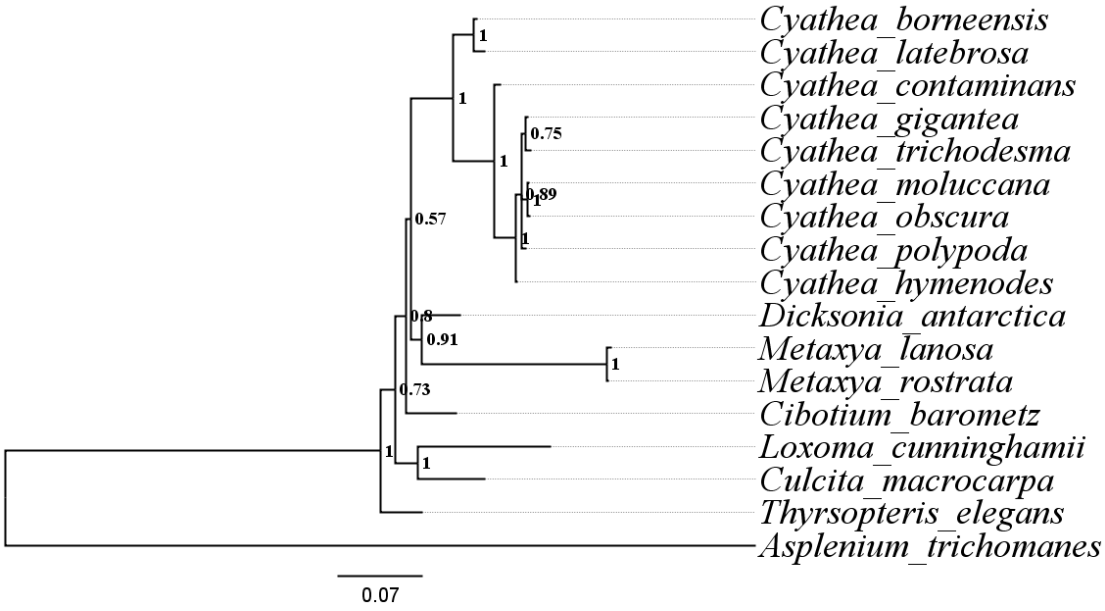


Figure 3.4 The 50% majority-rule consensus tree resulting from Bayesian (B/MCMC) analysis of matK region. Posterior probability (PP) is shown near branches.

3.3.2.2 *rbcL*

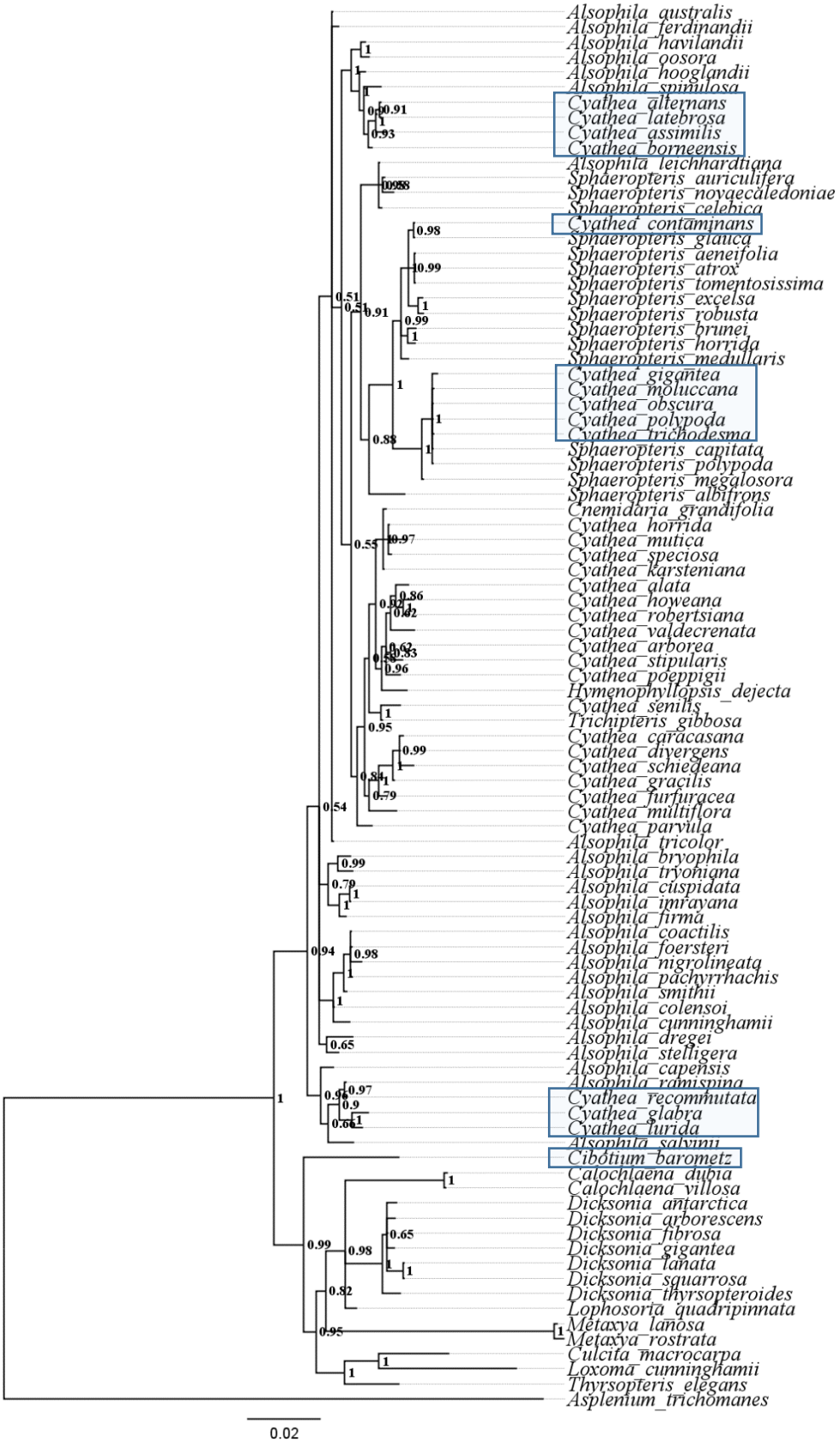


Figure 3.5 The 50% majority-rule consensus tree resulting from Bayesian (B/MCMC) analysis of rbcL region. Posterior probability (PP) is shown near branches. Highlighted boxes show species from this study.

3.3.2.3 *trnG-R*

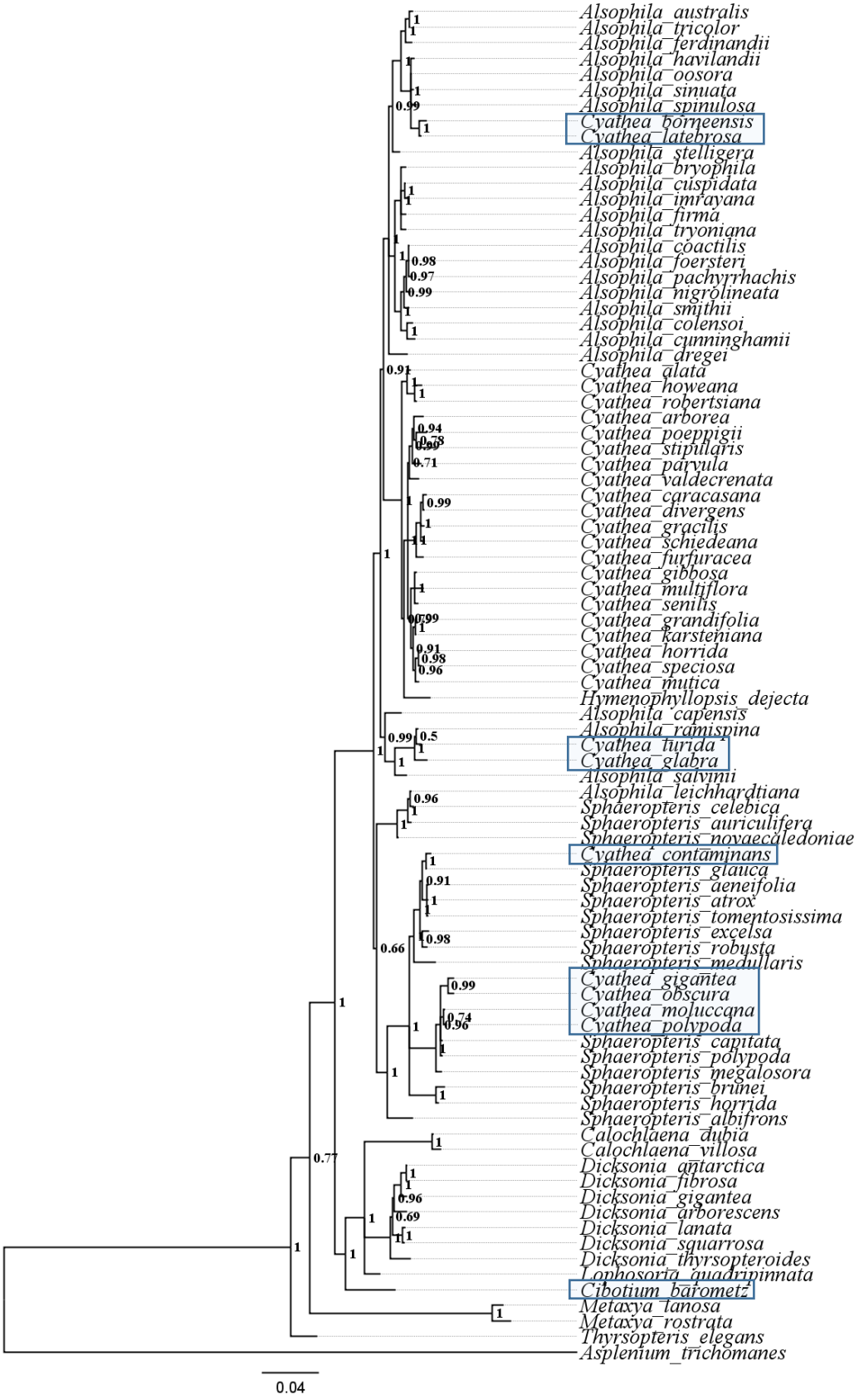


Figure 3.6 The 50% majority-rule consensus tree resulting from Bayesian (B/MCMC) analysis of trnG-R region. Posterior probability (PP) is shown near branches. Highlighted boxes show species from this study.

3.3.2.4 *trnL-F*

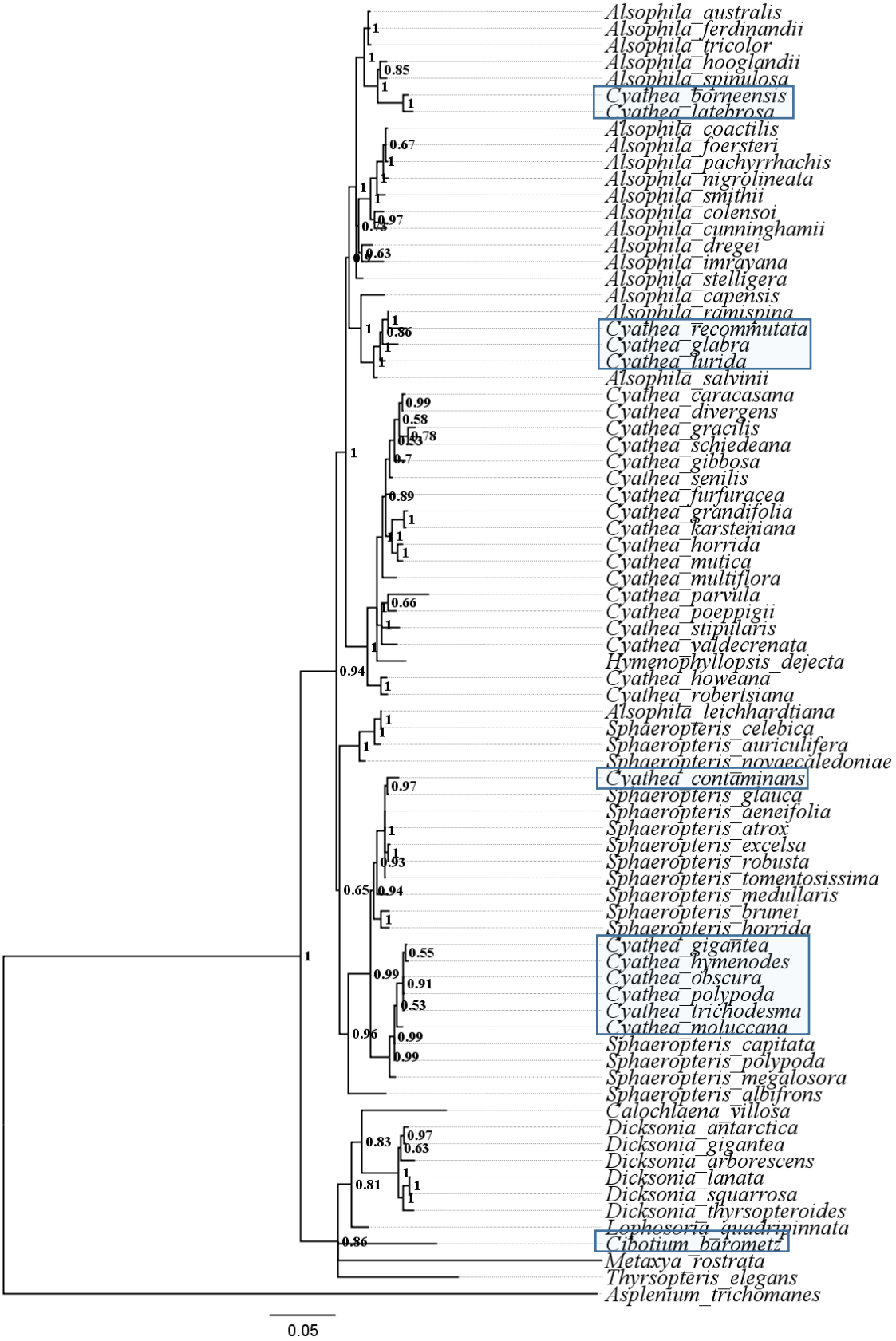


Figure 3.7 The 50% majority-rule consensus tree resulting from Bayesian (B/MCMC) analysis of trnL-F region. Posterior probability (PP) is shown near branches. Highlighted boxes show species from this study.

3.3.3 The Combined Matrix

The results from the combined matrix show mostly well-supported relationships with posterior probability greater than 0.70, unless otherwise stated. There is a basal dichotomy within Cyatheaceae (Figure 3.8), with a highly supported *Sphaeropteris* (PP of 1.00) sister to all other taxa. The sister group to *Sphaeropteris* is further separated to form a trichotomy: *Cyathea* (PP of 1.00), *Gymnosphaera* (PP of 1.00) and *Alsophila* (PP of 1.00). Dicksoniaceae (PP of 1.00) is a sister to all of these groups with other Cyatheales species as a sister group to Cyatheaceae and Dicksoniaceae.

C. borneensis, C. latebrosa, C. alternans and *C. assimilis* are embedded within the *Alsophila* group. *C. recommutata, C. lurida* and *C. glabra* are grouped together with *Gymnosphaera. C. contaminans* is well positioned in *Sphaeropteris* as well as *C. gigantea, C. trichodesma, C. obscura, C. polypoda, C. hymenodes* and *C. moluccana*. *Cibotium barometz* is grouped together with Dicksoniaceae and other Cyatheales species with *Asplenium trichomanes* as outgroup.

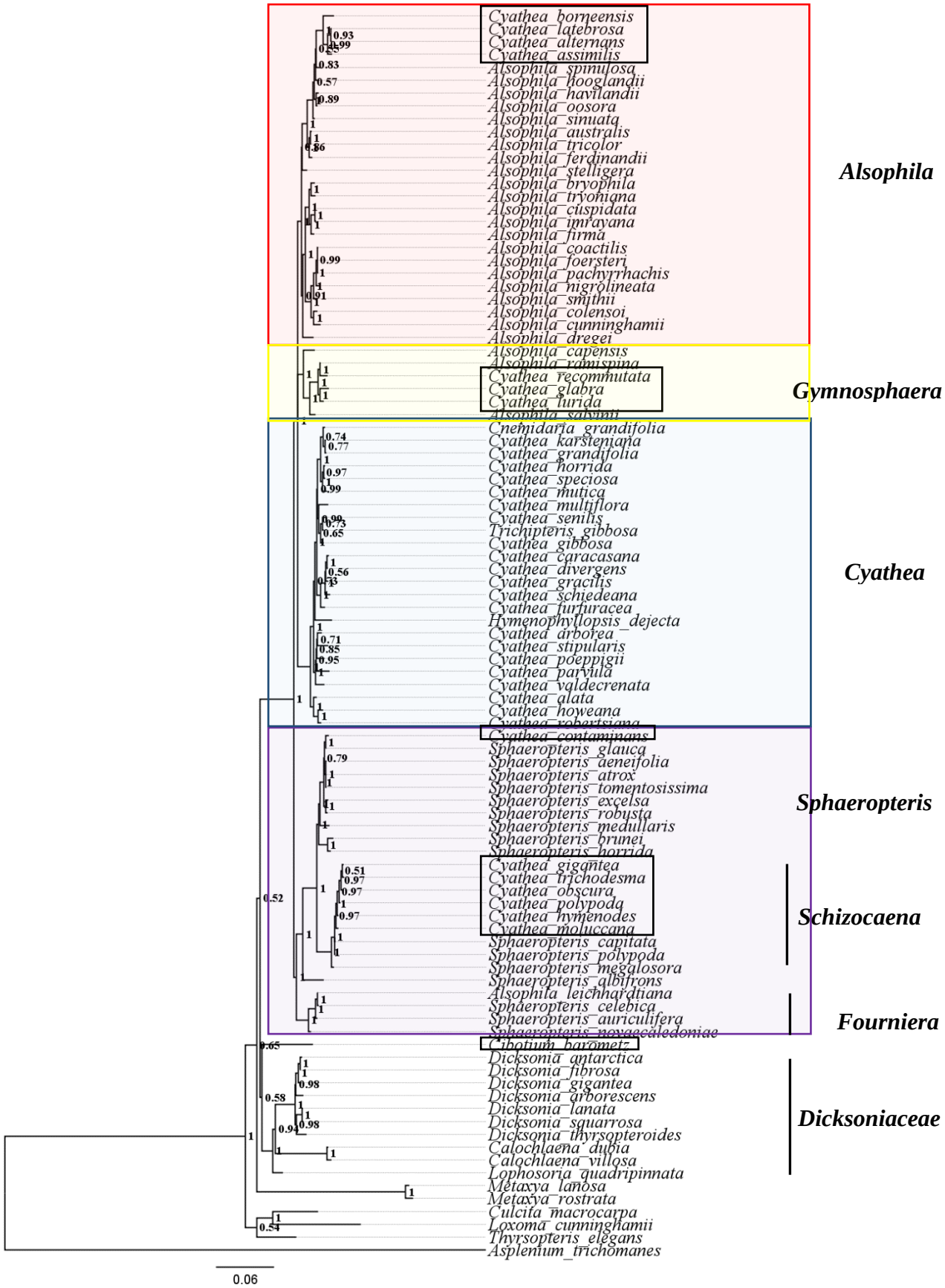


Figure 3.8 The 50% majority-rule consensus tree resulting from Bayesian (B/MCMC) analysis of the combined matrix. Posterior probability (PP) is shown near branches. Black boxes show species from this study.

3.4 Discussion

This work represents the first phylogenetic study on Cyatheaceae from Peninsular Malaysia. This study may contribute towards resolving and supporting the phylogeny of this family by adding 14 representative taxa from Peninsular Malaysia. The model-based estimate established from all four chloroplast DNA regions produced well-supported relationships. However, the Bayesian MCMC analysis using the four combined plastid regions and 78 in-group taxa was better supported compared with the tree generated from a single region analysis. The combined matrix had provided a posterior probability which is greater than 0.80 (unless otherwise stated) giving a better support compared with the single region tree. Species from two closely related genera *Cibotium* and *Dicksonia* were included in the analysis in order to determine their relationship to Cyatheaceae as well as several other species from different family in Cyatheales. The species included were from Metaxyaceae, Culcitaceae, Loxomataceae, and Thyrsopteridaceae. One species from Aspleniaceae was also added.

In Korall *et al.* (2007) four major groups were resolved: *Sphaeropteris, Cyathea, Alsophila* and *Gymnosphaera*, with *Sphaeropteris* being a sister group to a trichotomy containing the other three groups. In this study, it was found that all of the groups: *Sphaeropteris, Cyathea, Alsophila* and *Gymnosphaera* are well supported in the tree with *Sphaeropteris* as sister to the rest of the scaly tree ferns. This result agrees with study made by Korall *et al.* (2007).

This study also agrees with previous studies made by Korall *et al.* (2006, 2007); Korall and Pryer (2014) in which they concluded that this family is a monophyletic, based on a large-scale analysis of Cyatheaceae. In Korall *et al.* (2007) *Sphaeropteris* was shown to be moderately supported as sister to the other three well-supported groups. However in this study, *Sphaeropteris* is well supported (PP of 1.00) as a sister to the other three groups. Fourteen *Cyathea* species from Peninsular Malaysia can be found interspersed among the groups except in *Cyathea*. *C. latebrosa, C. borneensis, C. alternans* and *C. assimilis* are well positioned within *Alsophila,* specifically the *A. hooglandii* group. *C. lurida, C. recommutata,* and *C. glabra* can be found among the *Gymnosphaera,* adding the number of taxa in the group. A basal dichotomy had positioned the *Fourniera* group as sister to the rest of *Sphaeropteris,* in which the result agrees with Korall *et al.* (2007). The *Fourniera* group occurs from Malaysia to Australia and New Caledonia Korall *et al.* (2007). This group was previously recognized by Conant and Stein (2001) in which they concluded this group to be a distinct lineage and separated from the rest of *Sphaeropteris* species. *C. contaminans* is well positioned in *Sphaeropteris* as well as *C. gigantea, C. trichodesma, C. obscura, C. polypoda, C. hymenodes* and *C. moluccana.* The latter six species are specifically clustered in *Schizocaena* group in which Holttum (1963), Holttum and Edwards (1983) and Korall *et al.* (2007) stated this group to be confined to Malaysia and the Pacific. *S. albifrons* is positioned as a sister to the remaining *Sphaeropteris* (excluding the *Fourniera* group).

There are no *Cyathea* species from Peninsular Malaysia found in the *Cyathea* group. *Hymenophyllopsis* is positioned within the New World *Cyathea* group as sister to all other New World *Cyathea* species, in accordance to the study by Korall *et al.* (2007).

Note that there are errors with the species name in *Cyathea* group, *Cnemidaria grandifolia* which is supposed to be *Cyathea grandifolia* and *Trichipteris gibbosa* which is *Cyathea gibbosa.* These errors occurred as the analysis used *rbcL* sequences from the GenBank and retained the same name as per GenBank accession thus creating different taxa for the same species in the combined matrix tree.

Alsophila groupings recognized in previous phylogenetic studies by Conant *et al.* (1995); Conant and Stein (2001) and Korall *et al.* (2007) appear to be supported in this study. Most species of *Alsophila* used in Korall *et al.* (2007) study were a monophyletic group within the Old World taxa. *Cyathea* species from Peninsular Malaysia: *C. latebrosa, C. borneensis, C. alternans* and *C. assimilis* are clustered in the *A. hooglandii* group in which this group was previously recognized by Conant and Stein (2001).

Despite having incorporated 14 species from Peninsular Malaysia, there is still a need for a new, well-corroborated classification of Cyatheaceae. Further studies should be based on the current knowledge of phylogenetic relationships as well as morphological studies that will better support the groups within the family. More regions are needed to be added, for example *matK*, which will help to improve the tree topology and provide better view of the relationships. The problems in extracting, amplifying and sequencing the DNA for most of the species limited their use in the analysis, thus improved methods are needed in the future works.

3.5 Conclusion

The study previously done by Korall *et al.* (2007) resulted in the most extensive evaluation of Cyatheaeceae phylogeny, setting a platform for further research in large-scale evolutionary patterns of this family. This study was conducted with a primary motivation to use DNA sequences from four plastid regions: *matK, rbcL, trnL-F* and *trnG-R* to investigate the phylogenetic relationships of the Peninsular Malaysian scaly tree ferns (Cyatheaceae) with the existing molecular phylogeny. Phylogenetic analysis using Bayesian Inference based on the four plastid regions provided evolutionary information of 14 *Cyathea* species from Peninsular Malaysia. However, Peninsular Malaysian *Cyathea* species were found only interspersed among the groups of *Sphaeropteris, Alsophila* and *Gymnosphaera,* with none in *Cyathea* group. It is clear that much work remains to be done by continuing to include more taxa and additional data in order to be able to move even closer to a full understanding of Cyatheaceae evolution and diversification.

CHAPTER 4

THE DEVELOPMENT OF CHLOROPLAST DNA BARCODING MARKERS FOR PENINSULAR MALAYSIAN CYATHEACEAE SPECIES IDENTIFICATION

4.1 Introduction

First proposed by Hebert *et al.* (2003), DNA barcoding has a wide range of applications including revealing cryptic species (Hebert, Penton, *et al.*, 2004), linking biological samples to crime scenes (Sonet, 2013), and revealing misidentified species (Pryer *et al.*, 2010). Since its proposal, barcoding quickly gained popularity, leading to the formation of the Consortium for the Barcode of Life (CBOL) in 2004. The primary aim of CBOL is to promote the exploration and development of DNA barcoding as a global standard for species identification (CBOL Plant Working Group (2009). The consortium is composed of approximately 200 organizations, such as museums, herbaria and research institutes, from over 50 participating countries, with an obligation of making their barcoding sequences and voucher specimen data available through the Barcode of Life Database (BOLD) (CBOL Plant Working Group, 2009).

In their original work, Hebert *et al.* (2003) suggested the use of variations in short DNA sequences as labels for different taxa. In zoology, the usage of mitochondrial cytochrome c oxidase 1 (CO1) sequence proved very successful for taxon discrimination, making it a universal barcode for animals (Hebert *et al.*, 2004). In plants, however, CO1 presented much slower mutation rates than in animals, making it an inappropriate region for a universal plant barcode (Kress *et al.*, 2005). Even though extensive research has been conducted in the search for a universal barcode for plants, none of the tested loci were successful for all plant species during the time when barcodes was first introduced (Kane and Cronk, 2008; Chase and Fay, 2009).

In 2009, the Plant Working Group from CBOL proposed the two-locus combination of *matK+rbcL* as the core system for land-plant identification, accomplishing 70 to 75% successful discrimination at species level (Ferri *et al.*, 2015). Although a multi-locus approach has been proposed by different researchers (Kress and Erickson, 2008; Lahaye *et al.*, 2008) the idea has not been adopted formally (CBOL Plant Working Group, 2009; Hollingsworth, Graham and Little, 2011). The *matK* locus offers higher species resolution than *rbcL* but a universal primer set has not yet been found, and may not exist (Ferri *et al.*, 2015). However, *rbcL* seems to be more appropriate for barcoding in non-vascular plants than for seed plants (Dong *et al.*, 2014).

Several researchers have proposed the use of whole plastid genome sequence for plant discrimination but this idea has not yet been universally accepted (Erickson *et al.*, 2008; Sucher and Carles, 2008; Nock *et al.*, 2011; Yang *et al.*, 2013). Apart from concerns regarding high sequencing cost, there are obstacles involved in retrieving complete plastid genome sequences in comparison to the use of single-locus barcodes

(Li *et al.*, 2015). Hollingsworth *et al.* (2011) disputed that the full plastid haplotype is a good marker because it does not always track species boundaries. Despite extensive research, to date it remains unclear whether plastid regions can be regarded as suitable for barcoding (Li *et al.*, 2015).

Apart from plastid regions, the internal transcribed spacer region of the nuclear ribosomal DNA (nrITS) has been frequently used in molecular plant systematics research at the species level and is the most frequently sequenced locus (Álvarez and Wendel, 2003; Kress *et al.*, 2005; Hollingsworth, 2011). This region was proposed as a possible plant barcode locus as it posed greater discriminatory power than plastid regions and a large amount of sequence data for this region was available in GenBank (Kress *et al.*, 2005). There are, however, multiple limitations that prevent the use of nrITS as a primary element of the plant barcode (Hollingsworth *et al.*, 2011), such as reduced species-level variability in certain groups, divergent paralogues that require cloning of multiple copies, secondary structure problems resulting in poor quality sequence data (Álvarez and Wendel, 2003; Kress *et al.*, 2005), fungal contamination and difficulty in amplifying and sequencing the region from diverse sample sets (Hollingsworth *et al.*, 2011). Nevertheless, nrITS region can be successfully amplified in two smaller sections, a feature that is especially useful when degraded plant material is studied (Hollingsworth *et al.*, 2011).

4.2 Evaluation of Some of the Core Chloroplast Coding Regions Used for Ferns

Fern groups have been neglected in choosing the universal barcode for all land plants (Lahaye *et al.*, 2008; Hollingsworth *et al.*, 2011). Even though CBOL had announced the two-locus combination of *matK*+*rbcL* as the core system for land plant identification, it was found that most of the existing primer sets for *matK* and *rbcL* are not compatible for all lineages of land plants (Hollingsworth *et al.*, 2009). Difficulties in amplification of fern DNA, especially of the *matK* region were due to the strong rearrangement of the chloroplast genome during the evolution of the fern clade (Duffy *et al.*, 2009). As well as missing the flanking *trnK* exons, this region has been used for designing stable priming sites (Kuo *et al.*, 2011). While *rbcL* has frequently been used for fern phylogeny investigations, species discrimination has been proven to be insufficient and general identification below genus level remains uncertain (Schneider *et al.*, 2005; Schneider and Schuettpelz, 2006).

Ebihara *et al.* (2010) and de Groot *et al.* (2011) proposed adding *trnH-psbA* and *trnL-F* regions as alternatives to the land plant barcode markers. Studies made by Ma *et al.* (2010) on medicinal ferns had 90.2% successful identification rate, proving that the chloroplast *trnH-psbA* intergenic region has sufficient variation available for identification of ferns and can possibly be applied to wider taxa. Additionally, de Groot *et al.* (2011) reported successful amplification and sequencing of ferns using a restricted set of the universal and very reliable *trnL-F* primers, even with inadequate sample material (de Groot *et al.*, 2011).

Kuo *et al.* (2011) successfully designed and developed primers that are both universal and lineage-specific to overcome the *matK* amplification challenges for fern families. It was done by comparing the *matK* phylogenetic performance and sequence characteristics against *rbcL* and *atpA.* The studies found *matK* has the highest variability and substitution evenness but shows the least homoplasy, which can be used to gather representative sequences from all of the fern families (Kuo *et al.,* 2011).

4.2.1 *rbcL* and *matK* Markers for Cyatheaceae

Ebihara *et al.* (2010) used eight taxa from Cyatheaceae for DNA barcoding using *rbcL* and *matK* regions but none of the representatives were from Malesian region. Studies made by Li *et al.* (2011) and Kuo *et al.* (2011) only had one representative from Cyatheaceae for their barcode analysis but also demonstrated the value of the two DNA regions.

4.2.2 *trnL-F* and *trnG-R* as potential Barcoding Markers for Cyatheaceae

Taberlet *et al.* (2007) concluded that variation of the combined *trnL* and *trnL-F* spacer regions was unexpectedly high in ferns. Later Ebihara *et al.* (2010) reported a 100% resolving power for both genera and species when *rbcL* and *trnL-F* spacer regions were combined.

Even though *trnG-R* has never been evaluated as one of the plant DNA barcoding markers, a study using this region resulted in successful fern identification (Pryer *et*

al., 2010). There are also several phylogenetic studies on Cyatheaceae which used the *trnG-R* region in the analysis (Korall *et al.*, 2006, 2007, Korall and Pryer, 2014). These studies reported high DNA amplification and sequencing success (Korall *et al.*, 2006, 2007, Korall and Pryer, 2014).

4.3 Aims

To date there is no published survey of barcode markers for Peninsular Malaysian Cyatheaceae. This project aims to develop these.

4.4 Materials and Methods

All of the plant material collected (419 samples, 15 *Cyathea* species) for the analysis of Chapter 2 was used in this study. A detailed description of the sampling and sample preparation can be found in Chapter 2.1.1. Methods for DNA extraction, amplification, and alignment were discussed in details in section 3.3. Only samples with good quality DNA sequences and consist of species with more than 3 replicates were chosen.

4.4.1 DNA Barcoding Analyses

There are two widely used methods for barcoding analysis: distance based and tree based. These methods were used to search for a distinction and to test the resolving power of the regions. All reliable samples of the individual genes: *matK*, *rbcL*, *trnG-R* and *trnL-F* were used for initial analysis. However, only species which had been

sequenced for all of the four regions will be used in this study. Each species was represented by at least seven individuals and up to 12 individuals.

4.4.1.1 Distance based methods

Genetic distances were calculated using the Kimura two parameter (K2P) distance method as implemented in TaxonDNA (Meier *et al.*, 2006; Vaidya *et al.*, 2011). The intra- and interspecific congeneric pairwise (uncorrected) distances in each of the datasets were calculated using "pairwise summary" implemented in TaxonDNA (Meier *et al.*, 2006; Vaidya *et al.*, 2011). The minimum interspecific distance was plotted against the maximum intraspecific distance as recommended by Consortium for the Barcode of Life (CBOL) in order to assess the barcoding gap. The variability of each of the barcode regions used were assessed based on the number of variables sites using Mega 6.0 (Tamura *et al.*, 2013).

To assess the utility of DNA barcoding for accurate species discrimination, the "best match" and "best close match" functions in TaxonDNA (Meier *et al.*, 2006) were used. The "best match" is the least rigorous criterion as it finds the closest barcode match to each query sequences. This method was assessed on the 11 barcode datasets using uncorrected pairwise distance and a minimum 300bp sequences overlap.

4.4.1.2 Tree Based Method

Analysis using the tree based method was performed and two distance methods were used to evaluate the species monophyly clusters. Neighbor joining (NJ) and

Unweighted Pair Group Method with Arithmetic mean (UPGMA) trees were constructed in MEGA 6.0 (Tamura *et al.*, 2013) with the K2P model of nucleotide substitutions. Node support was obtained from heuristic searches of 1000 bootstrap replicates. The species were considered to be identified correctly when all of the individual species representatives clustered in a monophyletic clade.

4.5 Results

Of the 419 samples collected for the Malay Peninsula, sequencing and assembling were only successful for only part of the samples. There were 98 samples of the *matK* region (Table 4.1), 74 samples for *rbcL* region (Table 4.2), 114 samples for *trnG-R* region (Table 4.3) and 71 samples for *trnL-F* region (Table 4.4). Detailed information for each sample can be found in Appendix 4. Only 28 samples from three species: *Cyathea latebrosa, C. polypoda* and *C. contaminans* were sequenced for all of the four regions (Table 4.5). These 28 samples represent eight different populations throughout Peninsular Malaysia.

Table 4.1 List of 98 individuals belonging to Cyatheaceae family from populations all over Peninsular Malaysia from matK region.

Species	**Code**
C. borneensis	BK23, BL22, FH45, FH54, LH14, LH29, LK20, LK23, LK25, LK26, MB26, MB27, MB28, MB31, MB38, MP06, MP10, PH03, PH04, PH06, PH09, PH11, PH12, PH13, PH14, PH16, PH17, PH18, PH22, PH23, PH23, PH25, PH27
C. contaminans	BF7a, BF7b, BL01, BL09, BL12, BL16, BL18, BL35, GH17, LH01, LH06, LH09, LH18, LH25, LK09, LK14, LK15, MB40

C. latebrosa	BF1a, BF1b, BF1c, BF2, BF5a, BF5b, BF6b, BK03, BK06, BK07, BK08, BK18, BK33, BK34, BL02, BL24, BL33, FH05, FH12, FH24, FH28, FH41, HB3, LK01, LK07, LK18, LK19, LK24, LK29, MA09, MB41, MB44, MJ09, PH01, PH02, PH24
C. polypoda	BF3a, BF3b, BF3c, BL07, BL15, FH31, ML02, MA05

Table 4.2 List of 74 individuals belonging to Cyatheaceae family from populations all over Peninsular Malaysia from rbcL region.

Species	**Code**
C. borneensis	FH16, LH14, LH30, MA25, MB26, MB27, MB28, MB35, MB36, MP06, PH09, PH16, PH22, PH25
C. contaminans	BF7a, BF7b, BL09, FH57, LH01, LH06, LH09, LH25, LK15, MB40
C. glabra	BF8a, BF8b, FH43
C. latebrosa	BF1a, BF1b, BF1c, BF2, BF5a, BF5b, BF6a, BF6b, BK06, BL02, BL24, FH05, FH12, FH28, HB3, LK01, LK02, LK19, LK30, MA06, MA15, MA39, MB41, MB44, MJ09, MP12, PH24
C. obscura	BL30, FH06, FH52
C. polypoda	BF3a, BF3b, BF3c, BF4, FH31, FH56, MA05, ML02, ML03, ML06
C. recommutata	FH39, ML21, ML22, ML23, ML16, ML20, MA27

Table 4.3 List of 114 individuals belonging to Cyatheaceae family from populations all over Peninsular Malaysia from trnG-R region.

Species	**Code**
C. borneensis	BK07, BK09, BK13, BK18, BK23, BK24, BK26, BK30, BK33, BL02, FH14, FH24, FH38, FH40, FH45, FH46, FH51, FH54, FH58, LH29, LK25, MB28, MB29, MB30, MB36, MB37, MB38, MB39, MP05, MP06, MP07, MP08, MP10, MP11, MP14, PH02, PH03, PH04, PH06, PH09, PH11, PH12, PH13, PH14, PH15, PH16, PH17, PH18, PH21, PH22,

	PH25, PH27
C. contaminans	BF7a, BF7b, BL01, BL09, BL12, BL35, FH57, LH01, LH04, LH06, LH09, LK09, LK14, MB34, MB40
C. glabra	BF8a, BF8b, MB43
C. hymenodes	MJ04, MJ05, MJ06, MJ08
C. latebrosa	BF1a, BF1b, BF1c, BF2, BF5a, BF6a, BF6b, BK06, BK08, BK35, BK37, BK38, BL22, BL24, BL33, FH05, FH12, FH28, FH41, HB3, LH14, LK07, LK18, LK19, LK29, MA06, MB31, MB41, MB44, MJ03, MJ09, PH24
C. polypoda	BF3a, BF3b, BF3c, BF4, FH06, FH31, MA17

Table 4.4 List of 71 individuals belonging to Cyatheaceae family from populations all over Peninsular Malaysia from trnL-F region.

Species	**Code**
C. contaminans	BF7a, BF7b, BL01, BL09, BL12, BL16, BL18, BL19, BL35, FH57, LH01, LH04, LH05, LH06, LH09, LH18, LH25, LK09, LK12, LK14, LK15, MB15, MB16, MB34, MB40
C. gigantea	ML24, ML25, ML27, ML28, ML29,
C. glabra	BF8a, BF8b, BL03, BL05, FH43, MA38
C. latebrosa	BF1a, BF5a, BF5b, BF6Bb, BL02, BL17, FH28, LK30, MB31, MB44
C. obscura	BL25, BL28, BL30, FH52
C. polypoda	BF3a, BF3b, BF3c, BF4, FH31, MA05, ML01, ML02, ML03, ML06
C. recommutata	ML15, ML16, ML18, ML19, ML20, ML22
C. trichodesma	MA34, MA43, MA44, MA49, MA50

Table 4.5 Samples with successful DNA assembly for all of the regions used in this study.

Species	Code	Location	Regions/Sequences Length (bp)			
			matK	***rbcL***	***trnL-F***	***trnG-R***
C. contaminans	BF7a	Fraser's Hill	812	1206	796	968
	BF7b	Fraser's Hill	811	1207	797	968
	BL09	Bukit Larut	794	1206	796	931
	LH01	Lojing Highlands	799	1206	796	968
	LH06	Lojing Highlands	802	1206	794	968
	LH09	Lojing Highlands	732	1205	796	875
	LH25	Lojing Highlands	811	1210	795	968
	LK15	Lake Kenyir	811	1206	796	821
	MB40	Mount Berinchang	790	1206	796	968
C. latebrosa	BF1b	Fraser's Hill	815	1206	697	940
	BF5a	Fraser's Hill	814	1206	875	940
	BF5b	Fraser's Hill	814	1216	876	941
	BF6b	Fraser's Hill	806	1206	875	940
	BL02	Bukit Larut	803	1206	822	835
	FH28	Fraser's Hill	799	1210	536	940
	HB3	Bangi Forest	814	1206	872	942
	LK30	Lake Kenyir	813	1206	868	942
	MA06	Mount Angsi	814	1206	853	938
	MB41	Mount Berinchang	814	1206	853	941
	MB44	Mount Berinchang	804	1211	850	942
	MJ09	Mount Jerai	814	1206	867	943
C. polypoda	BF3a	Fraser's Hill	809	1206	869	969
	BF3b	Fraser's Hill	813	1201	870	969
	BF3c	Fraser's Hill	811	1206	869	969
	BF4	Fraser's Hill	813	1207	869	969
	FH31	Fraser's Hill	796	1207	869	947
	MA05	Mount Angsi	811	1206	869	834
	ML02	Mount Ledang	811	1206	869	915

4.5.1 TaxonDNA Analysis

4.5.1.1 The Barcoding Gap

The DNA barcoding gap reflects the distributions of intra- and interspecific variability separated by a distance (Wiemers and Fiedler, 2007). Using pairwise analysis in the TaxonDNA software, the level of divergence between and within species was tested and calculated on all of the eleven barcode datasets, including the datasets for samples of the individual genes. In order to accurately analyze the sequence identification, the maximum intraspecific distance should be lower than the minimum interspecific distance (Wiemers and Fiedler, 2007). Table 4.6 summarizes the intra- and interspecific distances for samples of the individual genes, while Table 4.7 summarizes the intra- and interspecific distances for eleven barcode datasets of the 28 samples used. Figures 4.1 to 4.15 illustrate these relationships. All of the individual genes failed to create any barcoding gap. However, out of all the single region for eleven barcode datasets of the 28 samples, *trnL-F* gave the highest barcoding gap of 10.5% while *matK, rbcL* and their combination gave the lowest with 0.5%. The combined datasets gave 3.5% differences in intra- and interspecific distances. Even with the combination of all of the other regions, *trnL-F* still had the highest percentage of barcoding gap.

Table 4.6 Summary of sequence divergence for samples of the individual genes.

Barcode regions	Intraspecific distance (%)	Interspecific distance (%)
matK	<= 0.0 – 3.0	<= 0.0 – 10.5
rbcL	0.0 – 0.5	<= 0.0 – 6.5
trnL-F	<= 0.0 – 15.5	<= 0.0 – > 20.0
trnG-R	<= 0.0 – 2.0	<= 0.0 – 16.5

Table 4.7 Summary of sequence divergence for 11 barcode datasets of the 28 samples used.

Barcode regions	Intraspecific distance (%)	Interspecific distance (%)
matK	<= 0.0 – 1.5	2.0 – 7.5
rbcL	0.0 – 0.5	1.0 – 3.5
trnL-F	<= 0.0 – 0.5	11.0 – > 20.0
trnG-R	<= 0.0 – 0.5	3.0 – 15.0
matK+rbcL	0.0 – 1.0	1.5 – 5.5
matK+trnL-F	0.0 – 1.0	6.5 – > 20.0
matK+trnG-R	0.0 – 1.0	2.5 – 16.0
rbcL+trnL-F	0.0 – 0.5	5.5 – 20.0
rbcL+trnG-R	<= 0.0 – 0.5	2.0 – 11.5
trnL-F+trnG-R	0.0 – 0.5	7.0 – > 20.0
Combined	<= 0.0 – 0.5	4.0 – 18.5

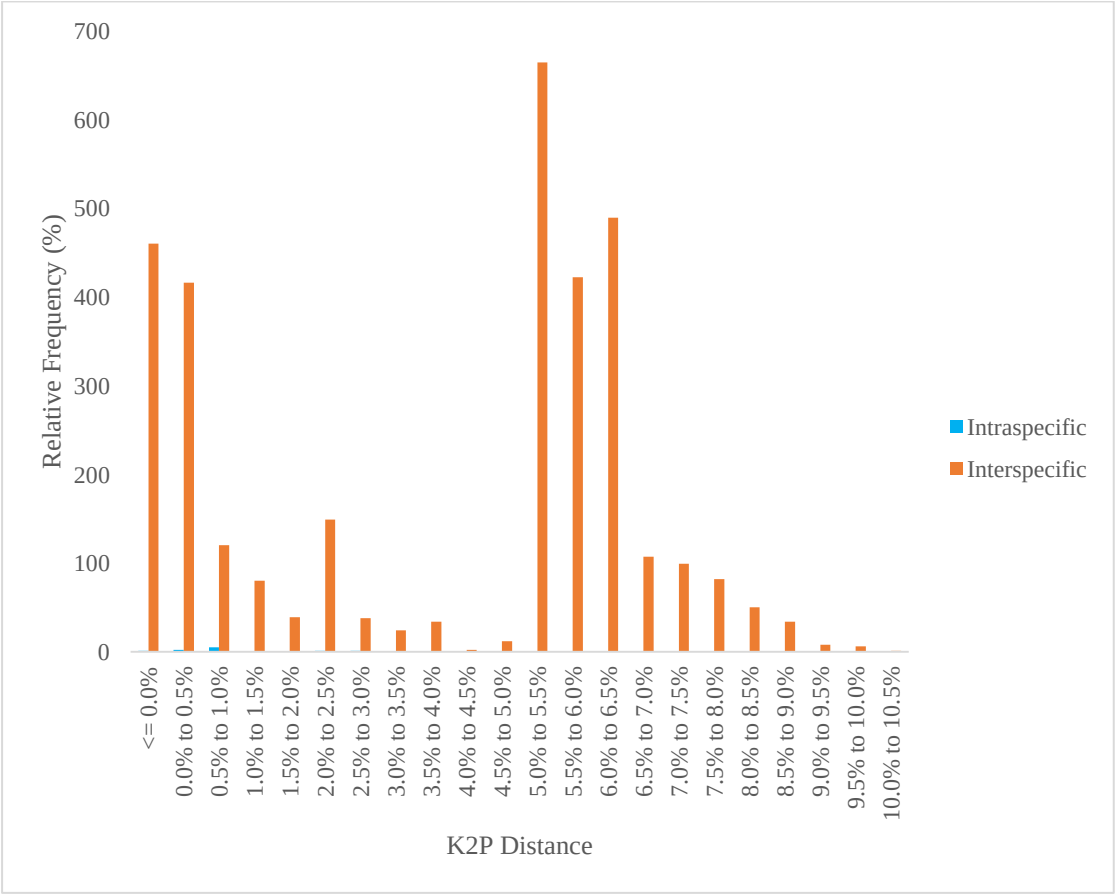


Figure 4.1 Intra- and interspecific distance of matK region for 98 samples.

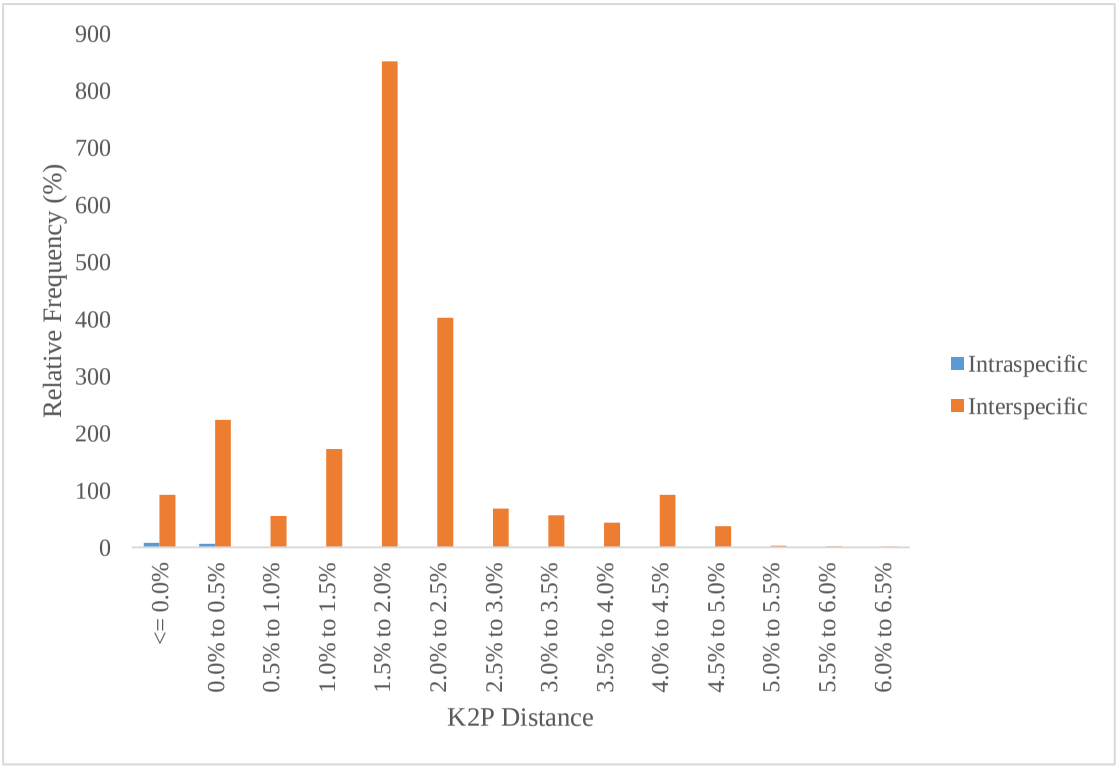


Figure 4.2 Intra- and interspecific distance of rbcL region for 74 samples.

.

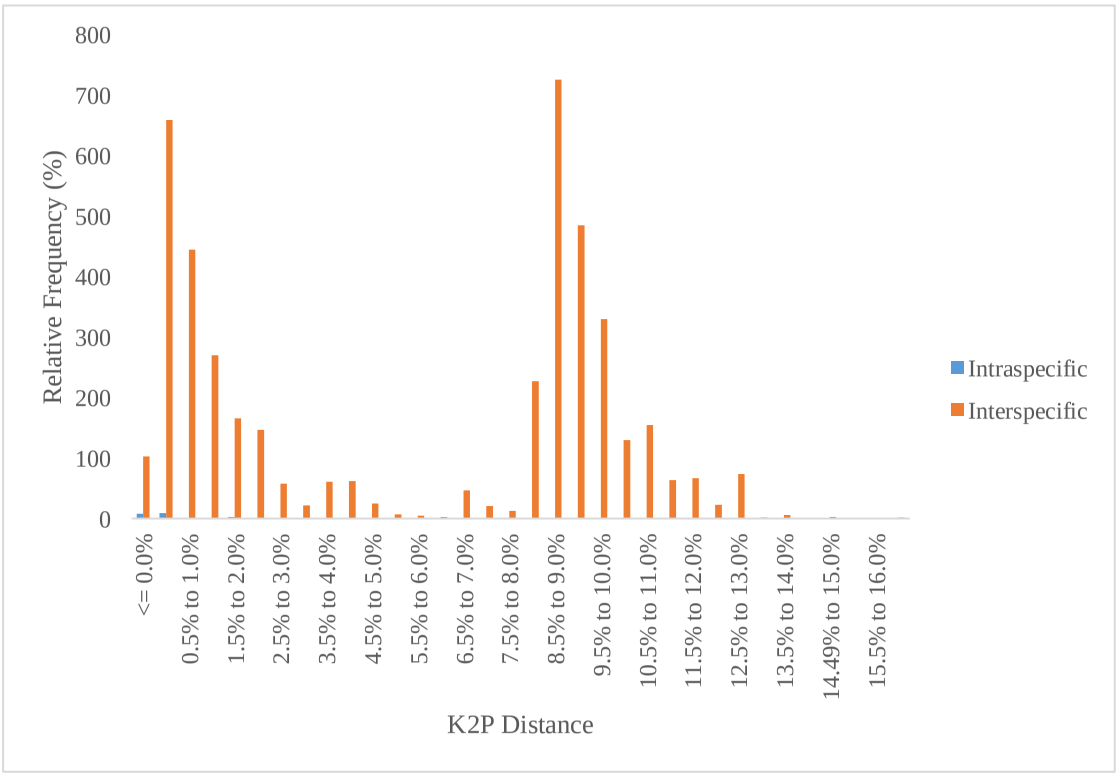


Figure 4.3 Intra- and interspecific distance of trnG-R region for 114 samples.

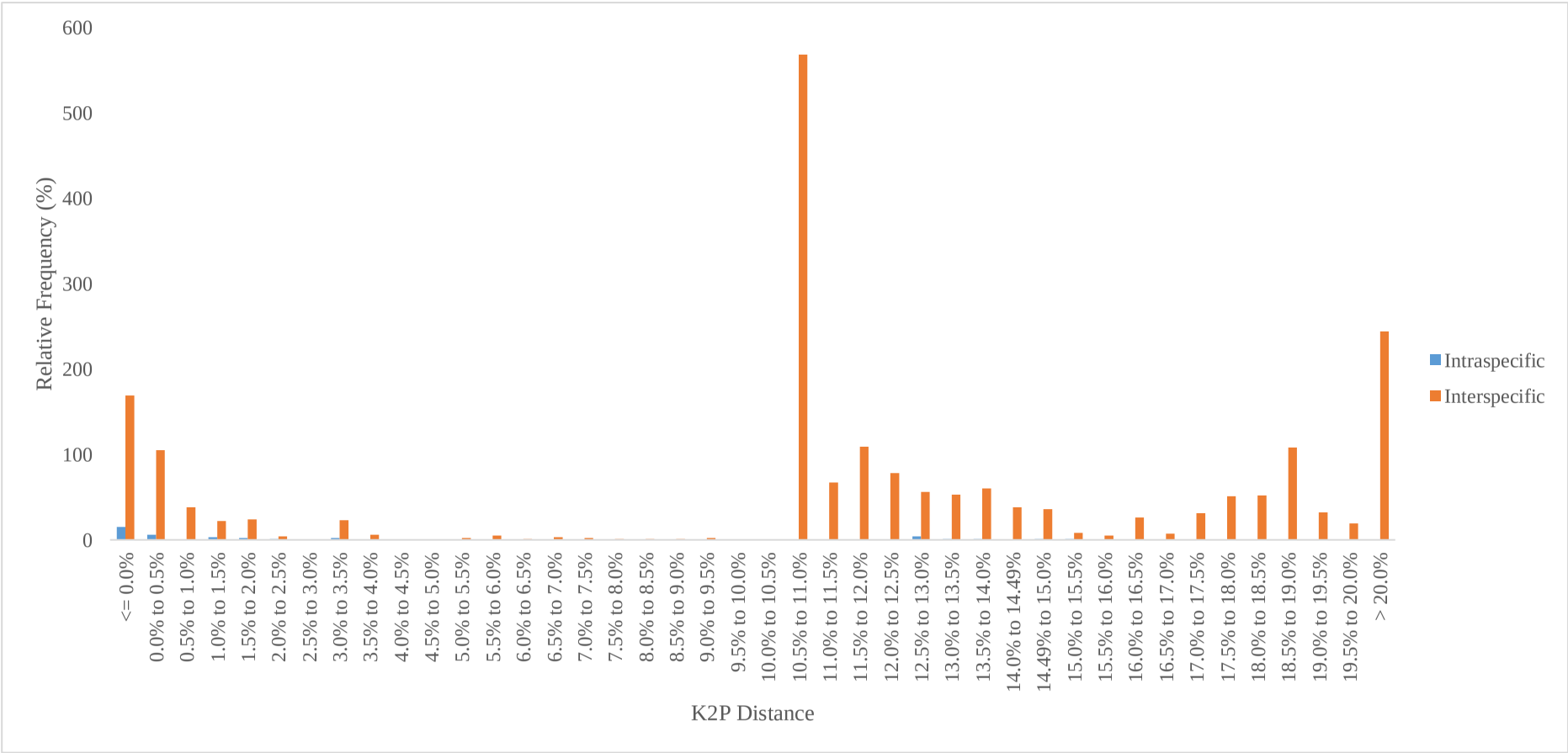


Figure 4.4 Intra- and interspecific distance of trnL-F region for 71 samples.

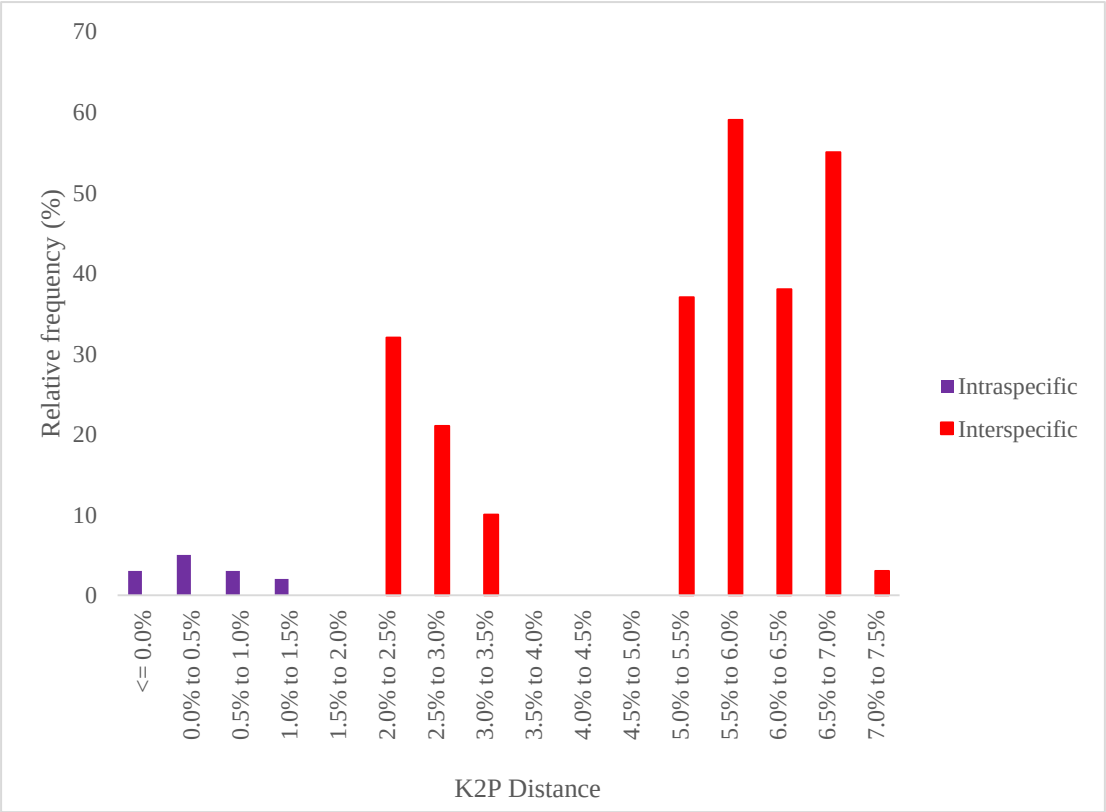


Figure 4.5 Intra- and interspecific distance of matK region for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

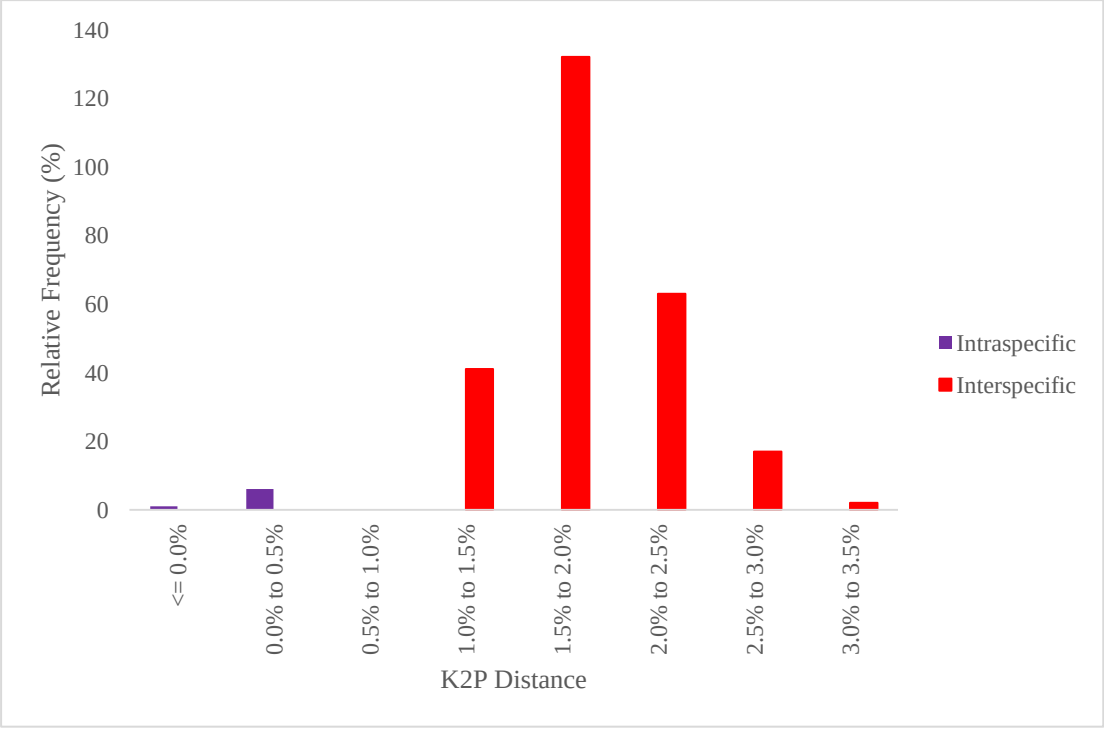


Figure 4.6 Intra- and interspecific distance of rbcL region for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

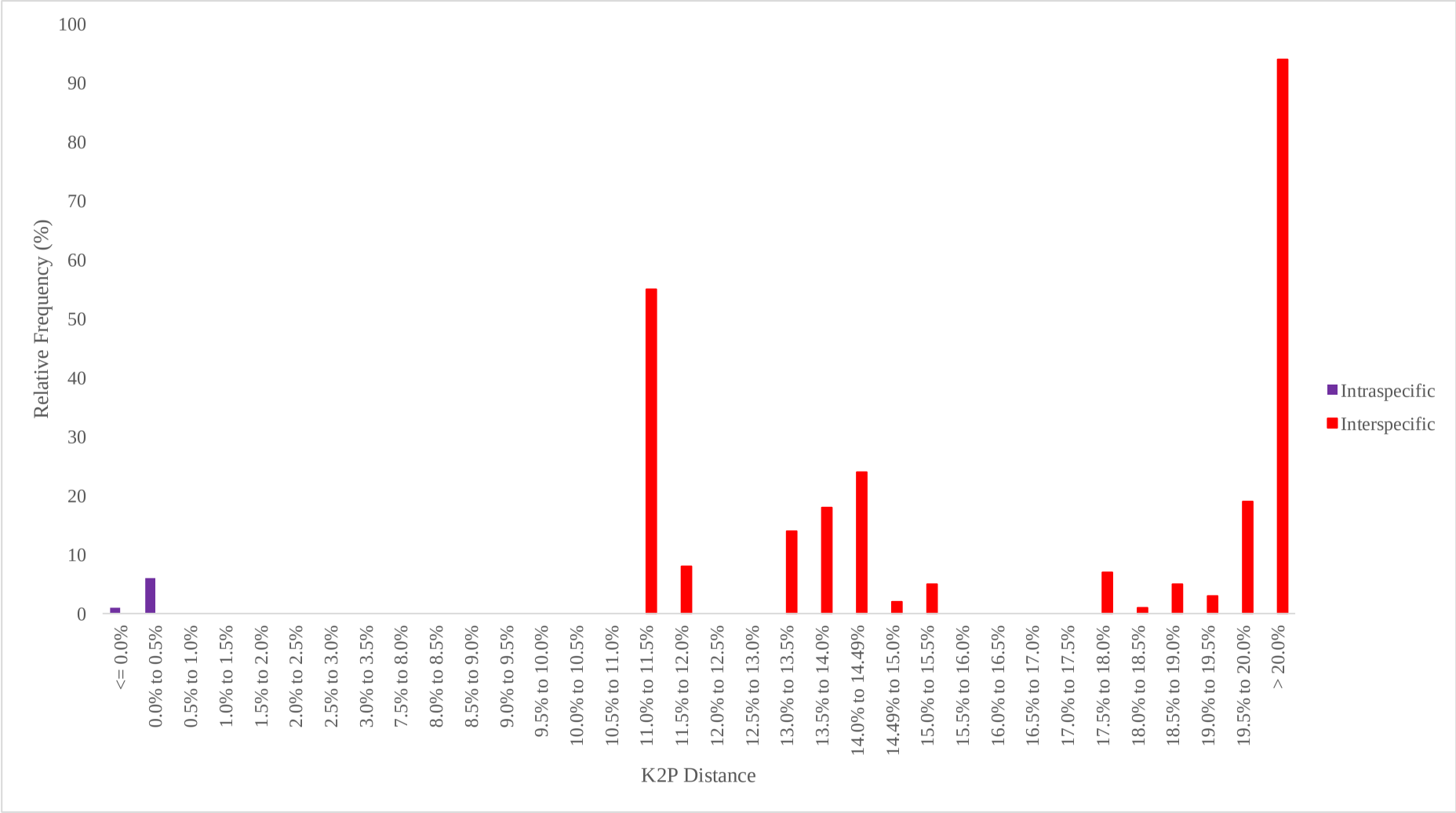


Figure 4.7 Intra- and interspecific distance of trnL-F region for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

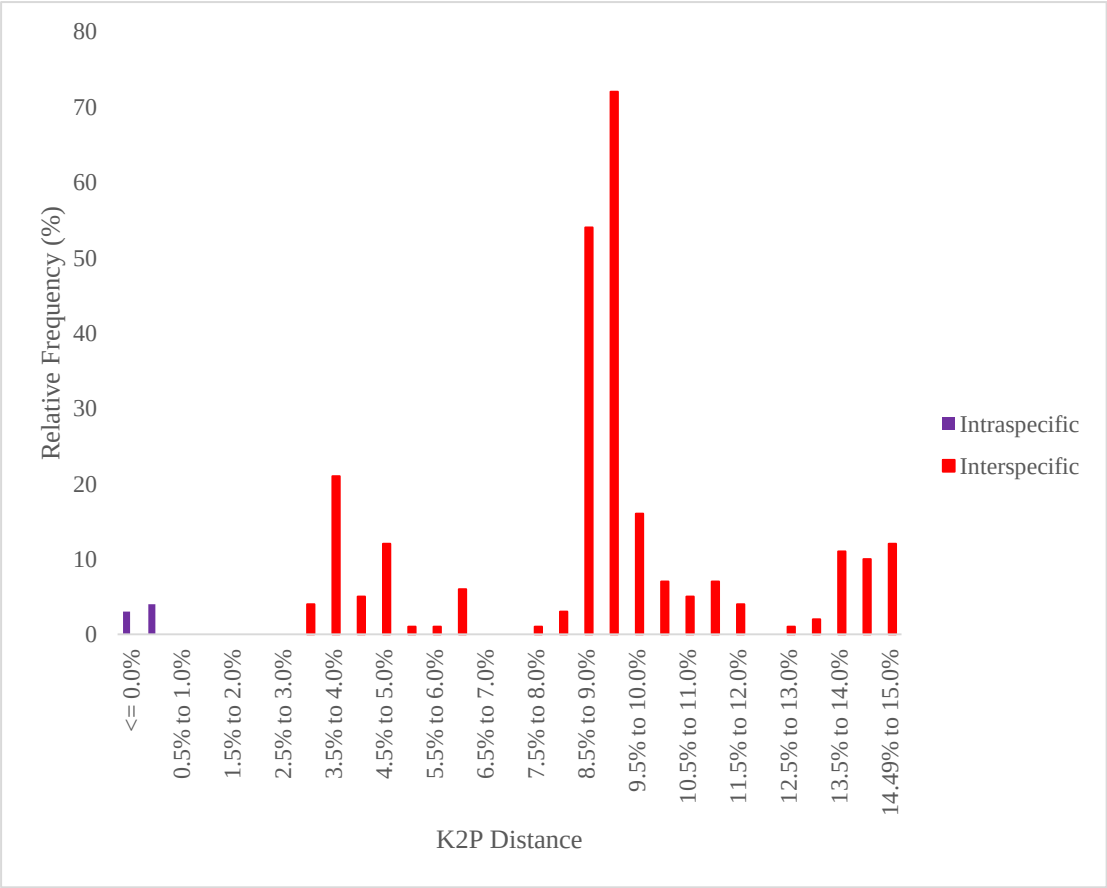


Figure 4.8 Intra- and interspecific distance of trnG-R region for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

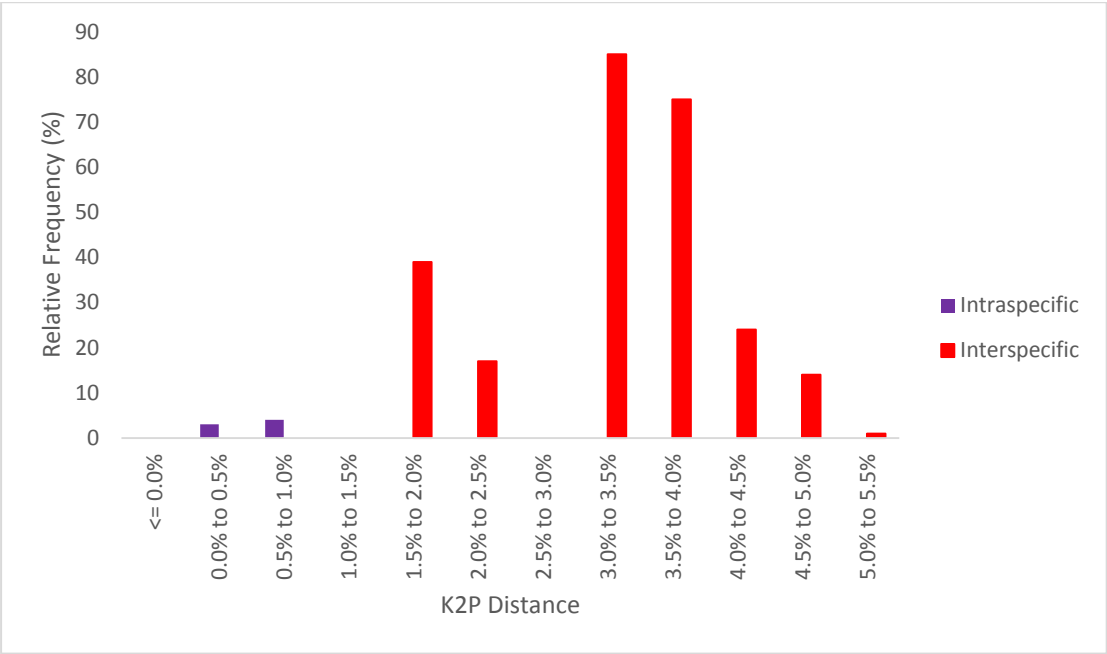


Figure 4.9 Intra- and interspecific distance of matK+ rbcL region for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

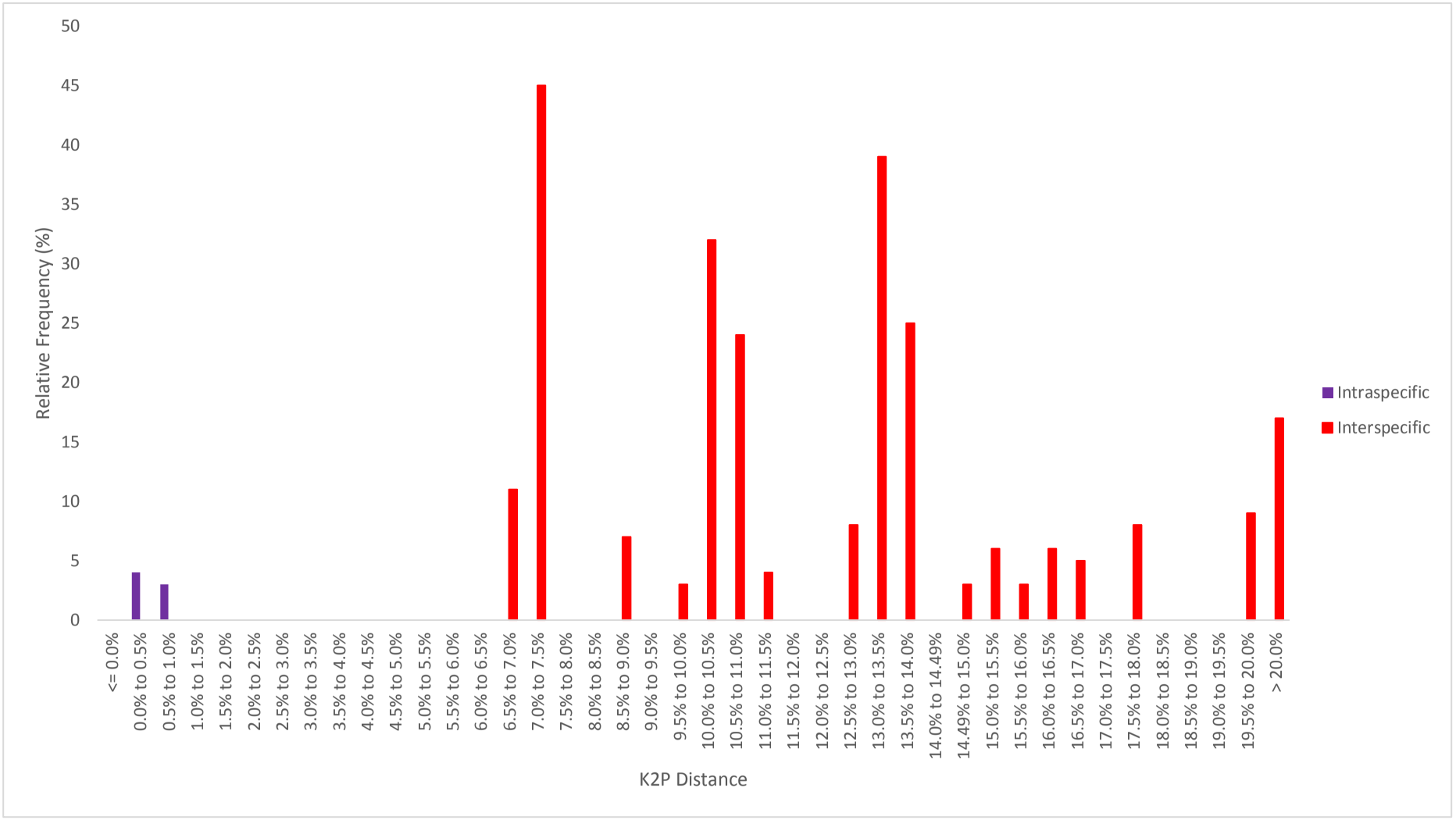


Figure 4.10 Intra- and interspecific distance of matK+trnL-F region for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

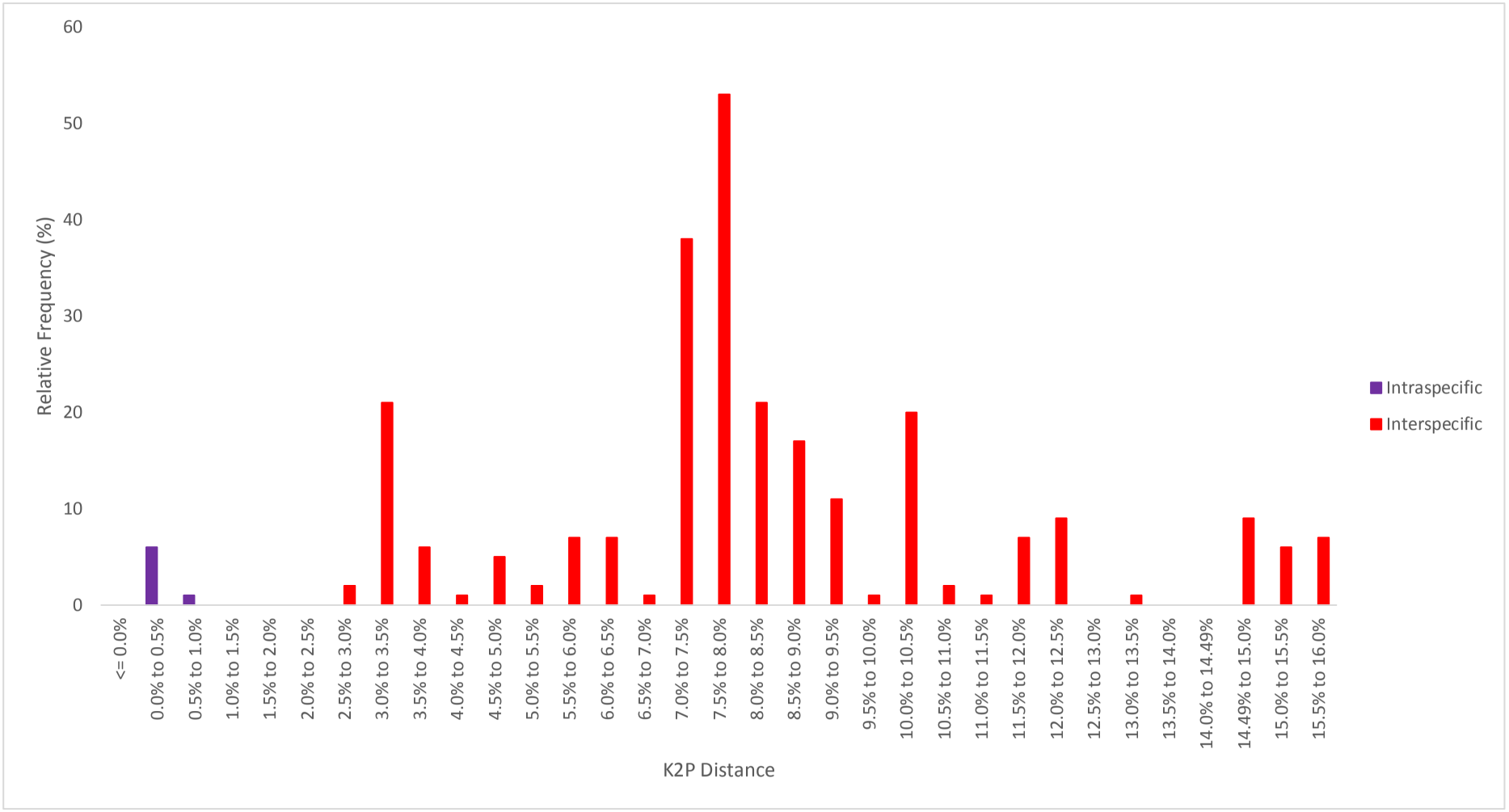


Figure 4.11 Intra- and interspecific distance of matK+trnG-R region for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

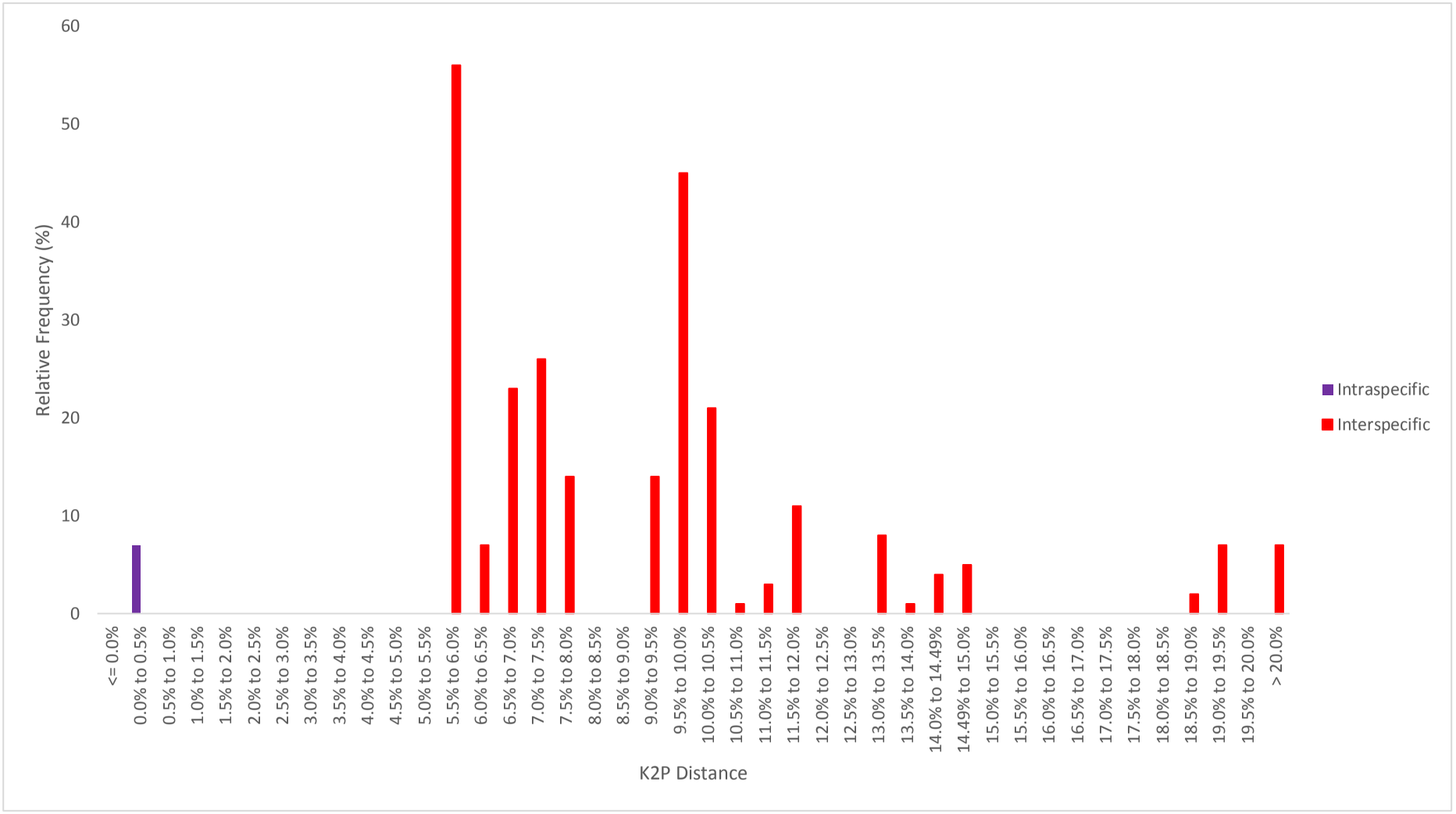


Figure 4.12 Intra- and interspecific distance of rbcL+trnL-F region for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

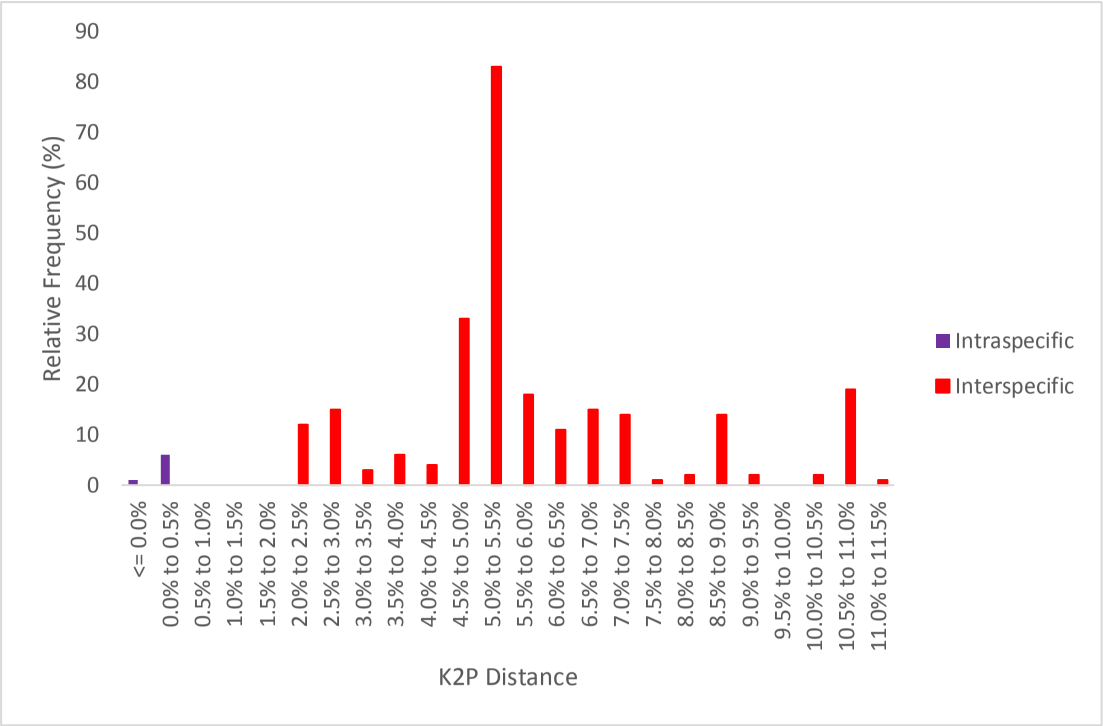


Figure 4.13 Intra- and interspecific distance of rbcL+trnG-R region for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

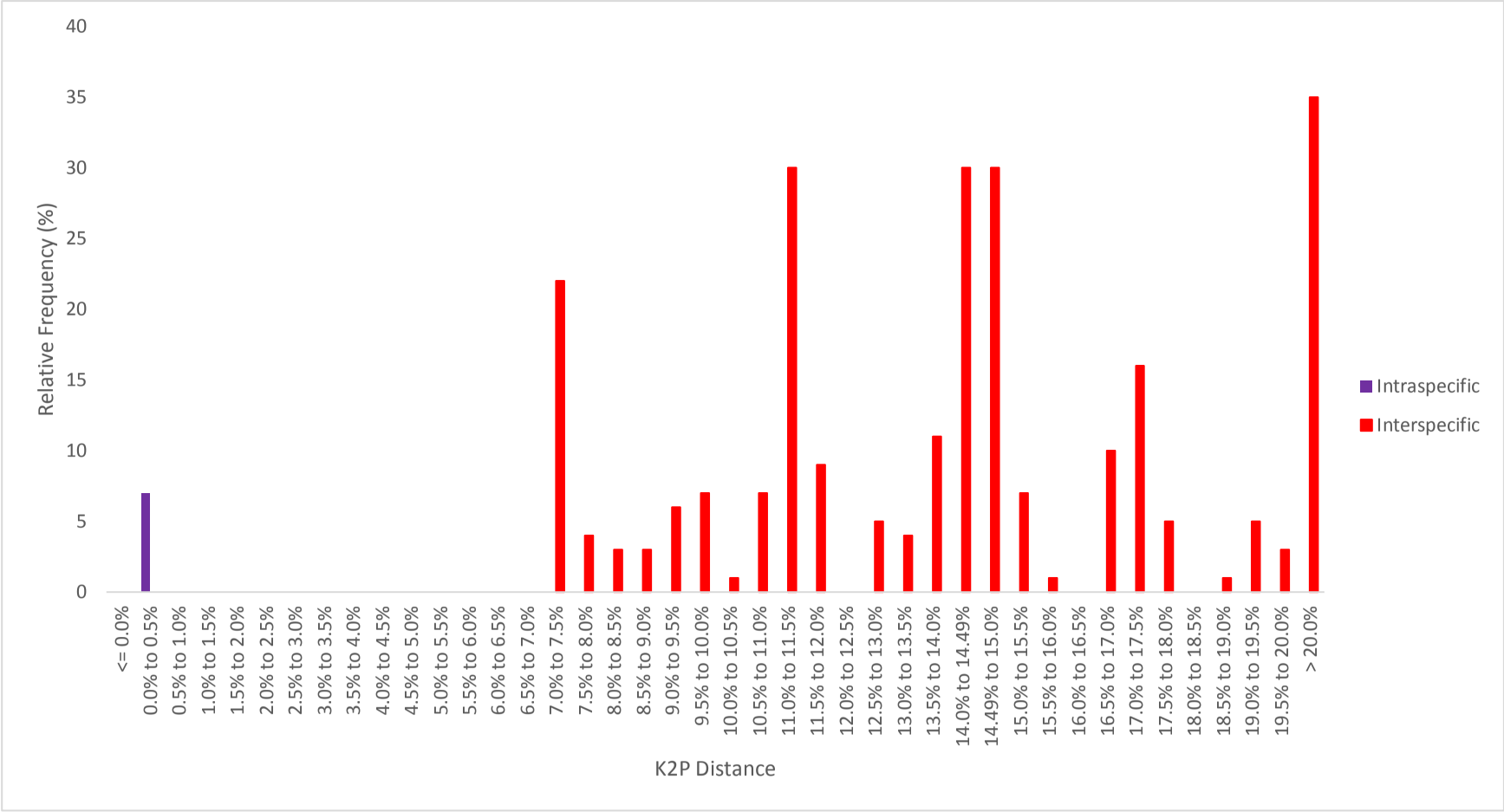


Figure 4.14 Intra- and interspecific distance of trnL-F+trnG-R region for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

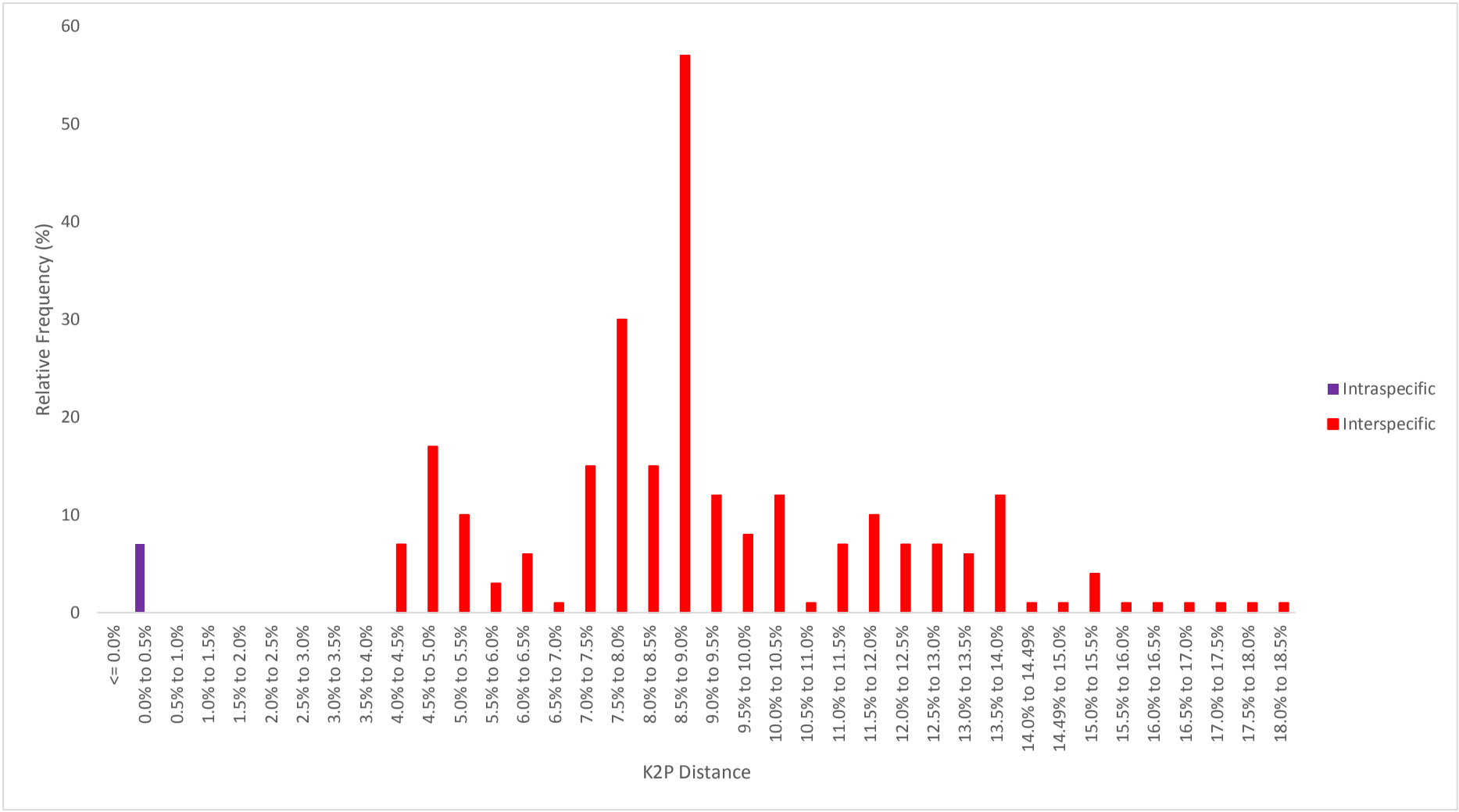


Figure 4.15 Intra- and interspecific distance of the combined region for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

4.5.1.2 Resolving Power

Table 4.7 summarizes the frequency of correct matches for samples of the individual genes. None of the genes showed incorrect matches, however the percentage of ambigous matches were more than 30%, with *trnL-F* having 9.85% "without any match closer than 3.0%" threshold.

Table 4.8 summarizes the frequency of correct matches for eleven barcode datasets of the 28 samples used. With the exception of *trnG-R* and the *rbcL+trnG-R* combination, with only 96.42% success, the others were 100% successful. Two single regions *matK* and *rbcL* had 100% correct match for both “best match” and “best close match” as well as the combination of both regions, *matK+rbcL*. In the “best close match” analysis, *trnL-F* and *trnG-R* had 89.28% and 92.85% correct match respectively and any combination with either of the regions resulted in less than 100% correct match. The combination for all of the regions gave 100% correct match in “best match” and 89.28% in “best close match”.

Table 4.7 Identification success for the single regions based from the "best match" and "best close match" analysis in the TaxonDNA.

Barcodes	**Best match (%)**			**Best close match (%)**			**Without any match closer than 3.0% (%)**
	Correct	**Ambiguous**	**Incorrect**	**Correct**	**Ambiguous**	**Incorrect**	
matK	43 (43.87)	55 (56.12)	0 (0.00)	43 (43.87)	55 (56.12)	0 (0.00)	0 (0.00)
rbcL	40 (54.05)	34 (45.94)	0 (0.00)	40 (54.05)	34 (45.94)	0 (0.00)	0 (0.00)
trnL-F	25 (35.21)	46 (64.78)	0 (0.00)	18 (25.35)	46 (64.78)	0 (0.00)	7 (9.85)
trnG-R	77 (67.54)	37 (32.45)	0 (0.00)	77 (67.54)	37 (32.45)	0 (0.00)	0 (0.00)

Table 4.8 Identification success for all of the barcodes regions based from the "best match" and "best close match" analysis in the TaxonDNA for three species (Cyathea latebrosa, C. polypoda and C. contaminans).

Barcodes	**Best match (%)**			**Best close match (%)**			**Without any match closer than 3.0% (%)**
	Correct	**Ambiguous**	**Incorrect**	**Correct**	**Ambiguous**	**Incorrect**	
matK	28 (100)	0 (0.0)	0 (0.0)	28 (100)	0 (0.0)	0 (0.0)	0 (0.0)
rbcL	28 (100)	0 (0.0)	0 (0.0)	28 (100)	0 (0.0)	0 (0.0)	0 (0.0)
trnL-F	28 (100)	0 (0.0)	0 (0.0)	25 (89.28)	0 (0.0)	0 (0.0)	3 (10.71)
trnG-R	27 (96.42)	0 (0.0)	1 (3.57)	26 (92.85)	0 (0.0)	0 (0.0)	2 (7.14)
matK+rbcL	28 (100)	0 (0.0)	0 (0.0)	28 (100)	0 (0.0)	0 (0.0)	0 (0.0)
matK+trnL-F	28 (100)	0 (0.0)	0 (0.0)	25 (89.28)	0 (0.0)	0 (0.0)	3 (10.71)
matK+trnG-R	28 (100)	0 (0.0)	0 (0.0)	27 (96.42)	0 (0.0)	0 (0.0)	1 (3.57)
rbcL+trnL-F	28 (100)	0 (0.0)	0 (0.0)	25 (89.28)	0 (0.0)	0 (0.0)	3 (10.71)
rbcL+trnG-R	27 (96.42)	0 (0.0)	1 (3.57)	27 (96.42)	0 (0.0)	0 (0.0)	1 (3.57)
trnL-F+trnG-R	28 (100)	0 (0.0)	0 (0.0)	22 (78.57)	0 (0.0)	0 (0.0)	6 (21.42)
Combined	28 (100)	0 (0.0)	0 (0.0)	25 (89.28)	0 (0.0)	0 (0.0)	3 (10.71)

4.5.2 Tree Based Analysis

Two different approaches were applied for the tree based analysis of the three species: *Cyathea latebrosa, C. polypoda* and *C. contaminans* (from the 28 samples used). Neighbor Joining (NJ) and Unweighted Pair Group Method with Arithmetic mean (UPGMA) analyses were used and the results showed clustering of conspecifics with mostly at more than 50% bootstrap support value. The results from the NJ analysis showed that most species were grouped at 99% to 100% bootstrap support value while other single regions were between 70% to 98% support (not shown here). The combined regions of *matK, rbcL, trnL-F,* and *trnG-R* gave a distinct species cluster at 100% bootstrap value (Figure 4.16). For the UPGMA analysis, *matK, rbcL* and *trnL-F* showed clustering of species at 100% support value, while *trnG-R* had species clustered at support values from 96% to 99% (not shown here). The combination of all the four regions, *matK, rbcL, trnL-F,* and *trnG-R* in the analysis gave support of 100% for all species cluster (Figure 4.17). From all of the tree based analysis, all individuals were found grouping with their conspecifics, even though there were differences in grouping orientation and support values.

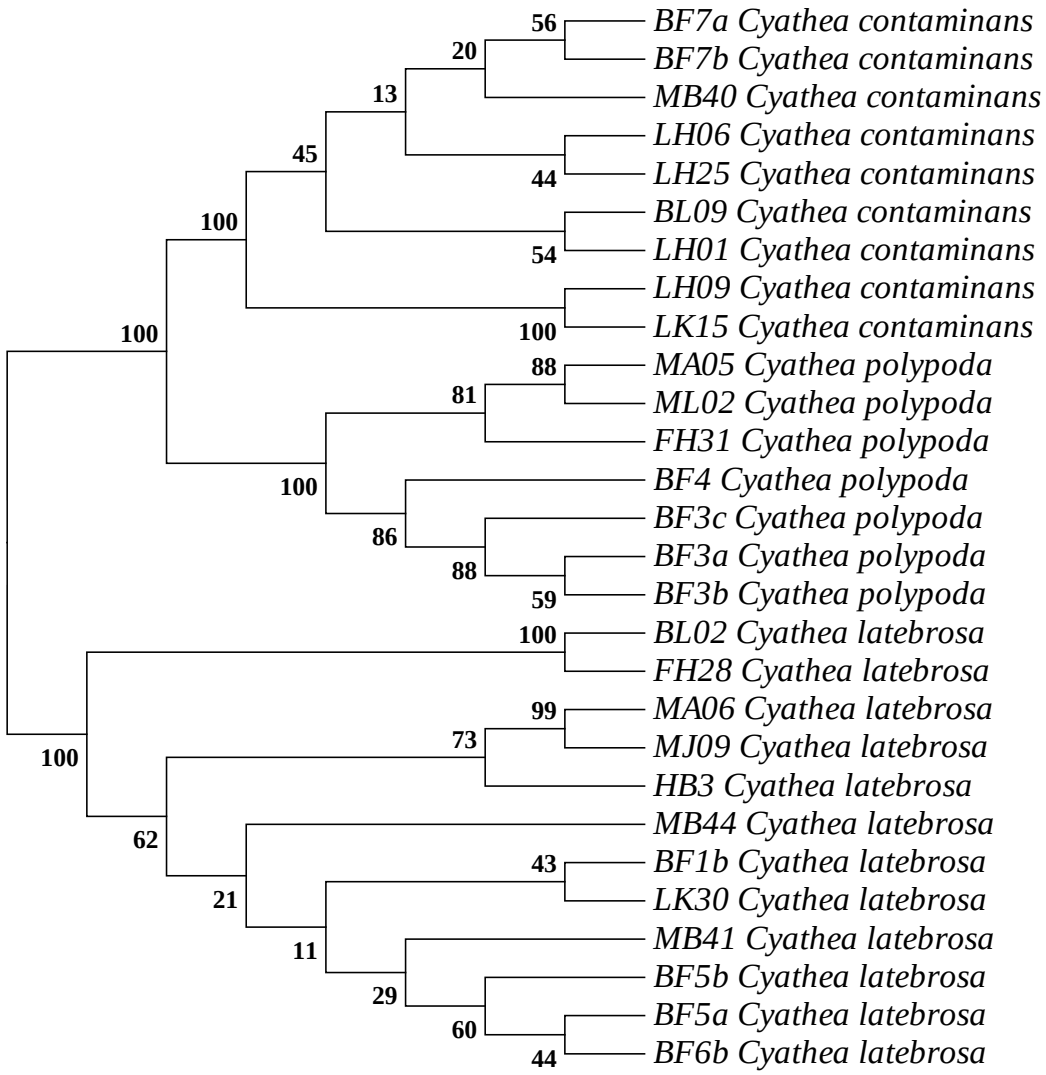


Figure 4.16 Consensus NJ tree based on K2P parameter model from combined (matK+rbcL+trnL-F+trnG-R) datasets. Numbers indicate bootstrap support values.

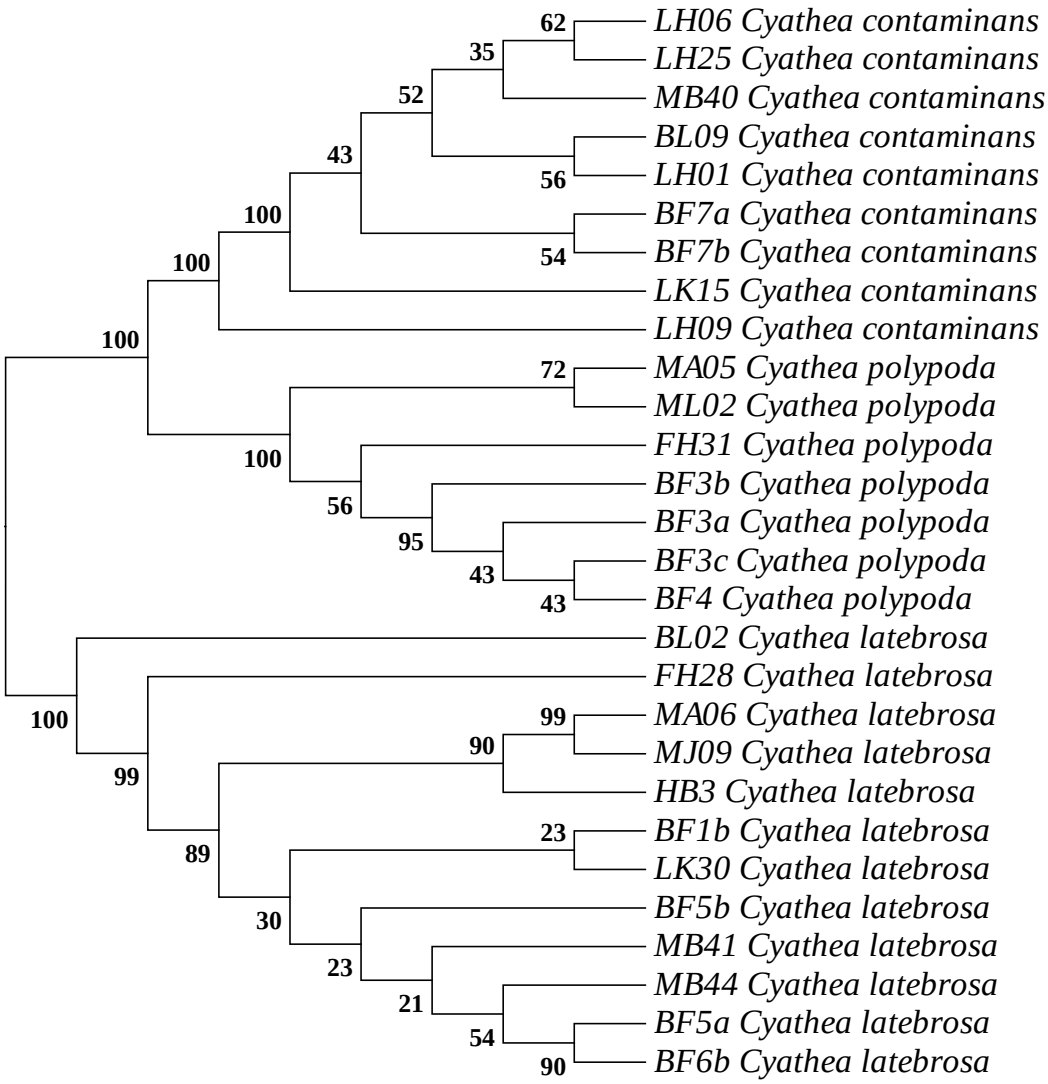


Figure 4.17 Consensus UPGMA tree based on K2P parameter model from combined (matK+rbcL+trnL-F+trnG-R) datasets. Numbers indicate bootstrap support values.

4.6 Discussion

Species identification of ferns by using DNA sequences has been applied in several studies most which focused either on single species identification (Schneider and Schuettpelz, 2006; Li *et al.*, 2009) or broad surveys (de Groot *et al.*, 2011). This study attempted to test DNA-based fern identification focusing on a specific taxonomic group in a defined geographical region which is Cyatheaceae from the Malaysian peninsula. Based on the suggestion from the Consortium for the Barcoding of Life (CBOL), a perfect DNA barcode should follow three criteria: primer universality, sequence quality and species discrimination (CBOL Plant Working Group, 2009). However, between these three criteria, primer universality must be the first and most important to consider (Chen *et al.*, 2013).

Even though the success rate of amplification and sequencing of regions from samples are important in the barcoding analysis in order to establish a suitable barcode marker, the extraction success of all the 419 samples had to be considered. The extraction from the samples was challenging, especially when searching for the most suitable method that would work for more than 90% of the total samples. Prior to amplification, CTAB methods were tested but failed to produce the desired DNA amount of at least 30 ng/µl needed in order to continue with amplification. Apart from CTAB, the extractions were also tested with DNeasy Plant Mini Kit from Qiagen, but the success rate was far lower, averaging between 5 to 10 ng/µl. The modified CTAB protocol used in

this study was the most successful method for the DNA extraction of the samples.

Among the four candidate regions tested in this study, the *trnG-R* region produced more amplified samples and higher sequencing rate than any of the other three regions: *matK, rbcL* and *trnL-F*. This outcome agreed with reports by Duffy *et al.* (2009) and de Groot *et al.* (2011) concerning the challenges regarding the universality of *matK* and *rbcL*. Based on the current outcomes, the use of *matK* and *rbcL* regions faces challenges.

The "best match/best close match" analysis using TaxonDNA (Meier *et al.*, 2006, Vaidya *et al.*, 2011) showed high percentange of ambiguous identification for the individual genes. The only explanation that can be hypothesized is that some of the samples collected were misidentified, thus showing conflicted species identification.

However, the "best match/best close match" analysis of the three positively identified species: *Cyathea latebrosa, C. polypoda* and *C. contaminans* from the 28 samples used showed that all of the regions and any of their combination resulted in 100% identification match in "best match" analysis, except for *trnG-R* and the *rbcL+trnG-R* combination. Nevertheless, the "best close match" resulted in various identification rate, ranging from 78.57% to 96.42%. This may due to the fact that this option is stricter because it depends on 95% pairwise distance threshold calculated by the "pairwise summary" function (Giudicelli *et al.*, 2015). The inter- and intraspecific divergence and

the success in identification from “best match” analysis were further supported by the tree-based analysis, as most of the species also formed their own well-supported monophyletic cluster in both NJ and UPGMA analyses.

4.7 Conclusion

This study was conducted to evaluate four potential DNA barcode regions: *matK, rbcL, trnL-F* and *trnG-R* for Peninsular Malaysian Cyatheaceae. Although *rbcL* and *matK* are two recommended markers for plant DNA barcoding by CBOL, in this study only *trnL-F* almost satisfied the three most important criteria: primer universality, sequence quality and species discrimination. However, with only 28 samples from three species (out of 15 species from 419 samples collected) working consistently, DNA barcoding in this study generally failed for the Cyatheaceae. The overall success is less than 10%. Nevertheless, the utility of four DNA barcoding markers in this study was tested and resulted with positive discrimination for the species of *Cyathea latebrosa, C. polypoda* and *C. contaminans.* More work needs to be done in order to include more taxa and to experiment with different primer pairs for better sequence quality. The plastid region *trnL-F* should be recommended as a DNA barcode for fern identification in future studies, at least for Cyatheaceae.

CHAPTER 5

THE DEVELOPMENT OF AN INTERACTIVE MULTI-ACCESS KEY FOR IDENTIFYING PENINSULAR MALAYSIAN CYATHEACEAE

5.1 Introduction

Over the years, only a handful of publications regarding Malaysian pteridopytes mentioning Cyatheaceae and its component Malaysian species have been written (Jaman and Latiff, 1998; Jaman and Latiff, 1999; Bidin and Jaman, 1999). Much of the information is published in traditional print format, such as *Malaysian Journal of Science* and *Sabah Park Nature Journal.* The information available needs to be synthesized and presented in a way that is accessible to the general public through the development of a computer-based, multi-access key. Few studies of Cyatheaceae in Malaysia have been made since Holttum (1963) in Flora Malesiana and none as extensive. Conant and Stein (2001) studied the phylogenetic and geographic relationships of Cyatheaceae on Mount Kinabalu and Latiff (2015) found a new species also from Mount Kinabalu.

Floristic research and all of its related fields depend on precise and usable species identification keys, which are usually in the form of dichotomous printed keys. Conventionally, these keys have been written by experts for someone with similar skills and have limited explanatory discussion (Lindsay and Middleton, 2009). However, there are several computer programmes developed that allow such experts to create a user-friendly multi-access key and other electronic identification tools for the use of a wider range of end-users with different levels of expertise (Lindsay and Middleton, 2009). Multi-access keys, as opposed to single access keys, do not require a sequential inclusion of features, giving the user the ability to only include features they can confidently observe. This is especially useful in cases of samples that are missing features such as sporangia. Prominent among these electronic identification tools is the LucID software (Norton, 2000).

5.1.1 The LucID Software

Multi-access keys are one of the many methods of overcoming the problem of the more traditional single-access keys. This will act as an instrument for conveying taxonomic expertise into a form that is easily accessed and utilized by the non-specialist. The development in Information Technology combined with the demand for accurate identification for conservation and management purposes led to the development of specialised software such as LucID (Norton, 2000). The LucID Builder version 3.3, which is the free version of this software, permits quick and easy development of multi-access identification keys. According to Norton (2000), LucID was exclusively developed for identification and analytical purposes, which permits

expert knowledge to be “duplicated” and distributed to the audience via CD or the Internet.

Multi-access identification keys for the Malaysian flora are in development, with only four family keys currently available online for Flora Malesiana. The keys can be accessed from http://www.lucidcentral.com/en-us/keys173;/searchforakey.aspx (Figure 6.1).

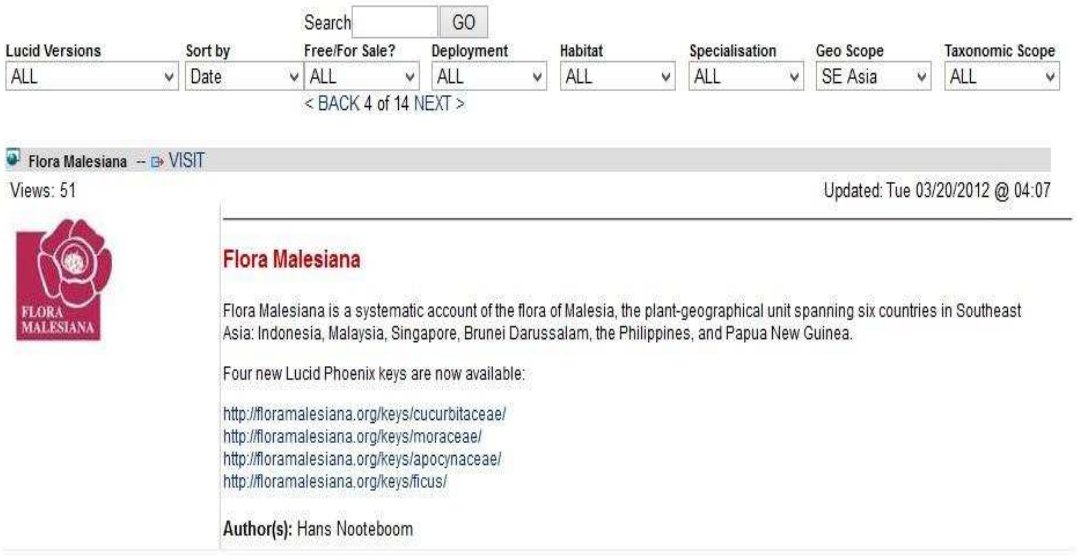


Figure 5.1 Online keys for plant families available for Flora Malesiana.

5.1.2 Aim

This study aims to develop a multi-access identification key for Peninsular Malaysian Cyatheaceae contributing to the Flora Malesiana online resources.

5.2 Materials and Methods

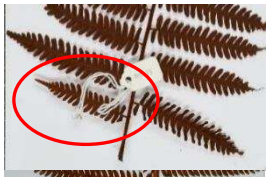
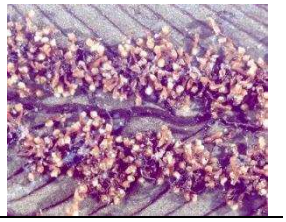
5.2.1 Plant Materials

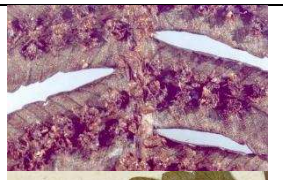
The plant collections reported in Chapter 2 were used in this study. A detailed description of the sampling and sample preparation can be found in Chapter 2.1.1. Herbarium materials from Kew Botanical Garden and Malaysia National University herbarium were also used (these specimens should be listed in an appendix).

5.2.2 Data Collection

Morphological traits for fronds, stipe, spines, scales, sori and indusia were observed and measured. These traits were chosen based from the existing dichotomous keys from Flora Malesiana Series II: Pteridophytes (Holttum, 1963) which should best reflect the variations among the species sampled. Based on the herbarium and 419 frond and stipe samples, numeric features as well as the descriptive characteristics of the species were recorded and used to build the multi-access LucID key for Peninsular Malaysia Cyatheaceae (Table 6.1). The information regarding trunk height and blade length were taken in the field. As certain species might reach up to 20 meter high, the trunk height was estimated as accurate as possible by placing known height (a person) next to the trunk. As for other features, all of them were observed and evaluated in the laboratory, with the help of a dissecting microscope (Leica DFC420) which had an attached digital camera with computer interface, making identification and recording information easier.

Table 5.1 The morphological information that was used in developing the multi-access key.

Part	Characters	Levels	
Trunk	Height	Continuous, measured in metres	
Fronds	Pinnation	Pinnate Bipinnate Tripinnate	
	Length	Continuous, measured in metres	
	Lower pinnae	Size is smaller from the rest	
		No differences in size	
Stipe	Colour (RHS Colour Chart)	Light to medium brown (N199) Medium to dark brown (200) Dark brown (N200) Purplish (N187) Dark or black (202/203)	
	Surface	Smooth Fine warts	
Spines	Present	Yes No	
Scales	Present	Yes No	
	Colour (RHS Colour Chart)	Light brown (199) Medium brown (N199) Brown (200) Dark brown (N200)	
	Finish	Glossy Not glossy	
Sori	Placement	Single row on the either side of the mid-vein	

		Near mid-vein	
		In groups of 1 to 3	
		Almost cover the lower surface of pinnule	
		In one to three rows on either side of the mid-vein	
		In three pairs of veins covering about half of the pinnae	
		In three or more pairs of veins on either side of pinnae	
Indusia	Present	Yes No	
	Sorus coverage	Complete cover Partial cover	
	Shape	Saucer-like Scale-like Bilobed	
	Colour	Light brown (199) Brown (200) Translucent	

5.2.3 Preparing the Key

A multi-access key for the Malaysian Peninsula Cyatheaceae was constructed using LucID software version 3.3, which can be downloaded from http://www.lucidcentral.com/en-us/software/lucid3.aspx following registration. The program comprises two elements, which are the LucID Builder (Figure 6.2) and Lucid Player (Figure 6.3). The builder was used to develop the identification key while the latter was used to assess and test the key. The key was assigned with the name 'Key for Cyatheaceae Species of Peninsular Malaysia' and consists of 15 quantitative and qualitative features, 44 states and 15 entities (species), as shown in Figure 6.4.

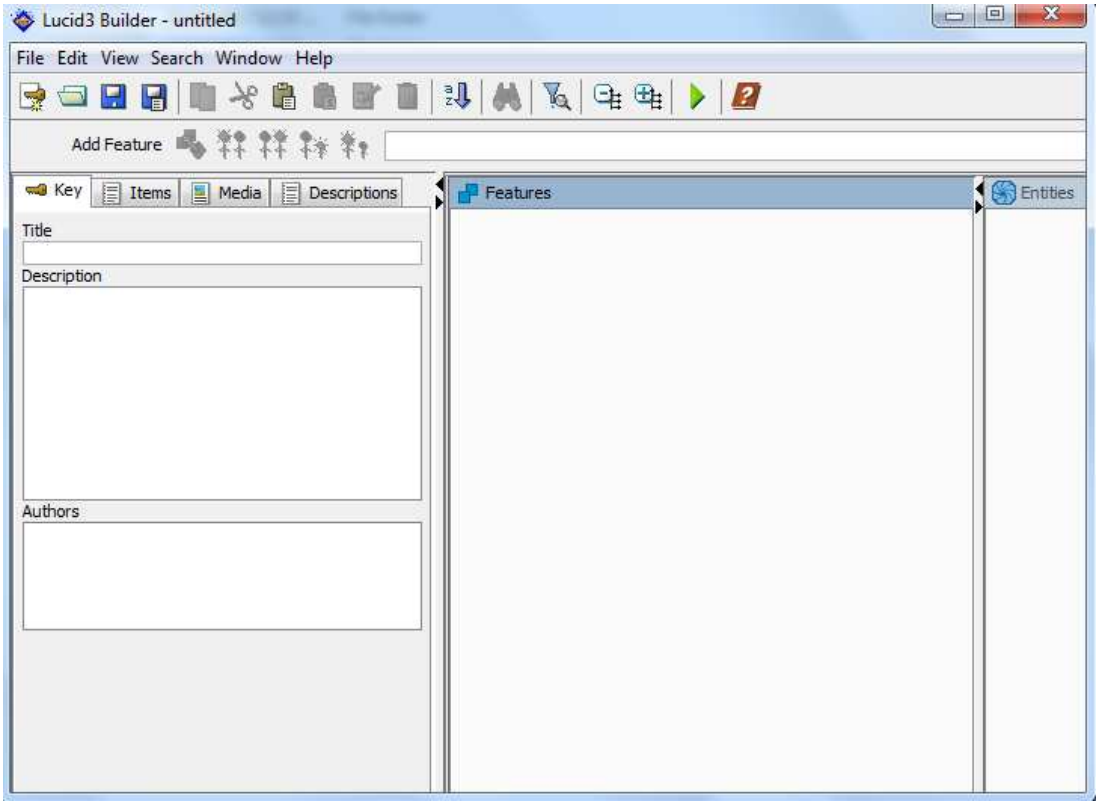


Figure 5.2 LucID Builder window before development of the multi-access key.

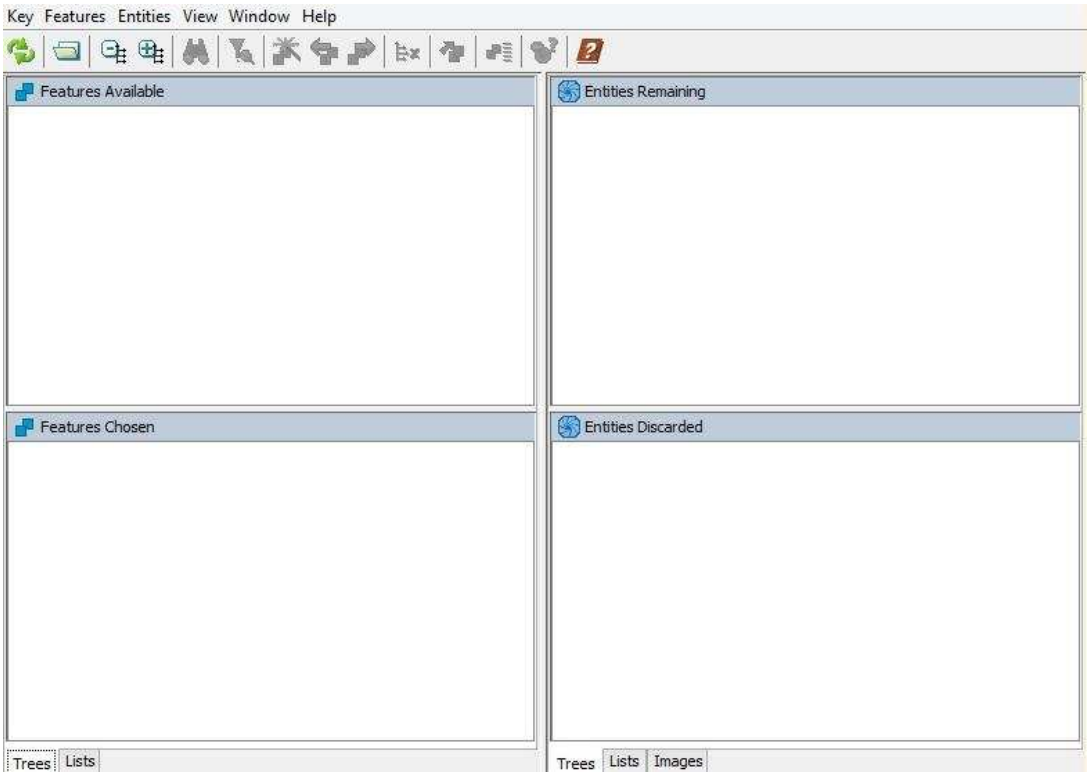


Figure 5.3 LucID Player window before any multi-access key available for assessing.

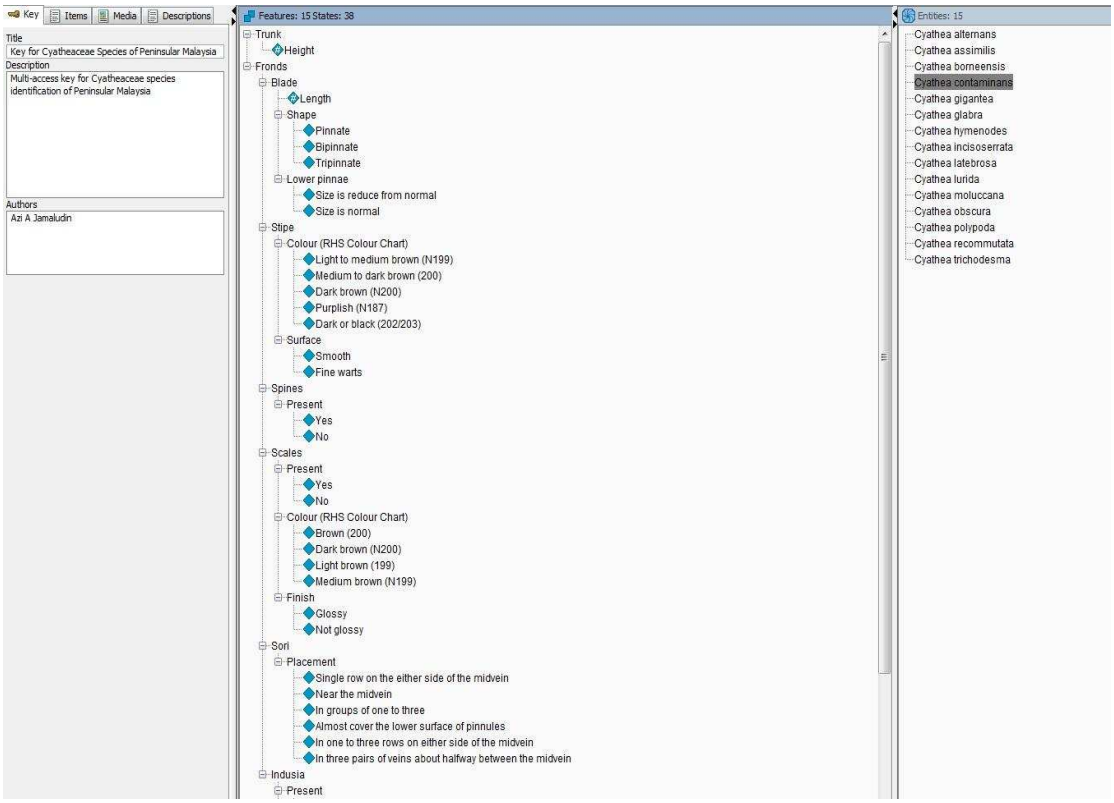


Figure 5.4 The key after entering the features, states and entities.

The different features and states were assigned individual images to improve identification success. Following this stage, all characters present in the key were scored using seven categories (common, rare, uncertain, common and misinterpreted, rare and misinterpreted, not scored and absent) (Figure 6.5). This completed the building stage of the key.

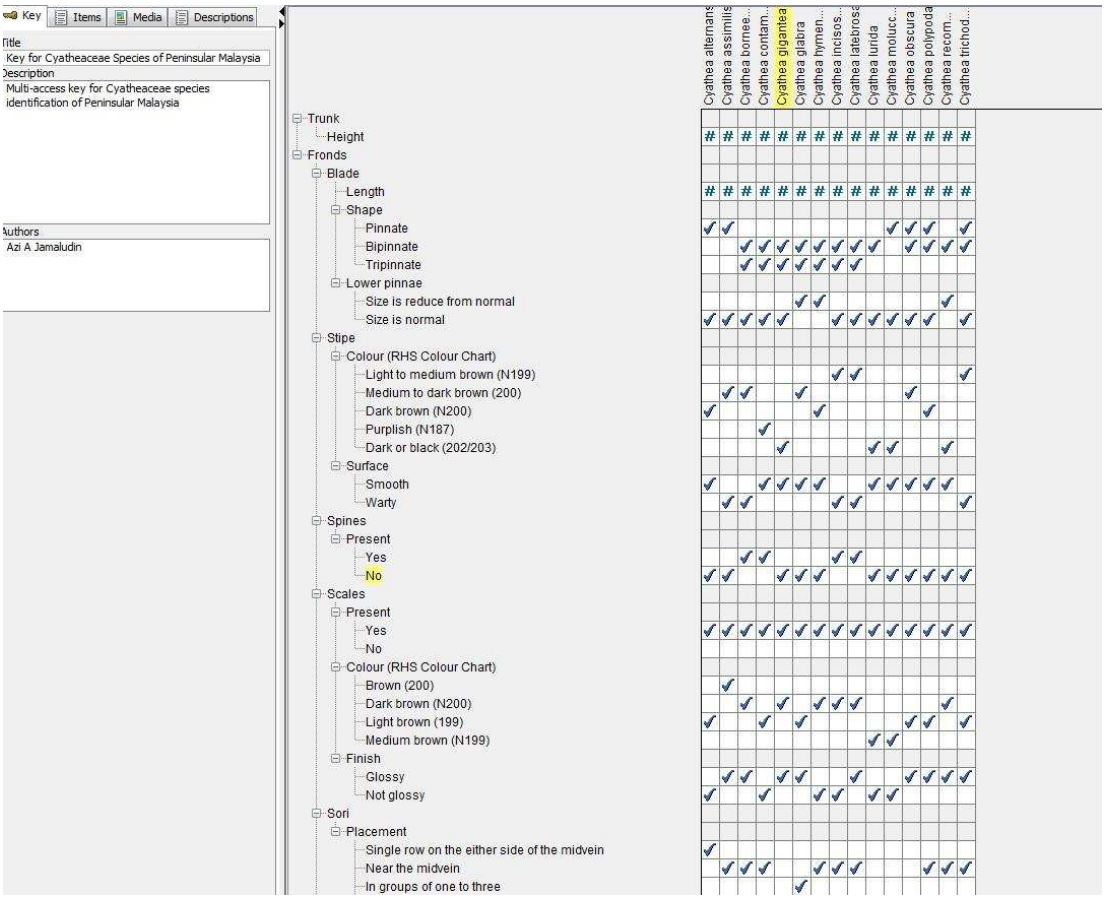


Figure 5.5 Scoring the species in the spreadsheet tab.

For features where a range is more appropriate than a categorical description, LucID provides the option to include numeric features, which come with four values: outside minimum, normal minimum, normal maximum and outside maximum. This option can be used to characterize the taxa where there is a range of natural variation between samples. ‘Normal minimum’ and ‘normal maximum’ values were determined once the

average of all of the measurements was acquired. It is also possible to use the same values of ‘normal minimum’ and maximum when scoring for outside minimum and maximum (Figure 6.6).

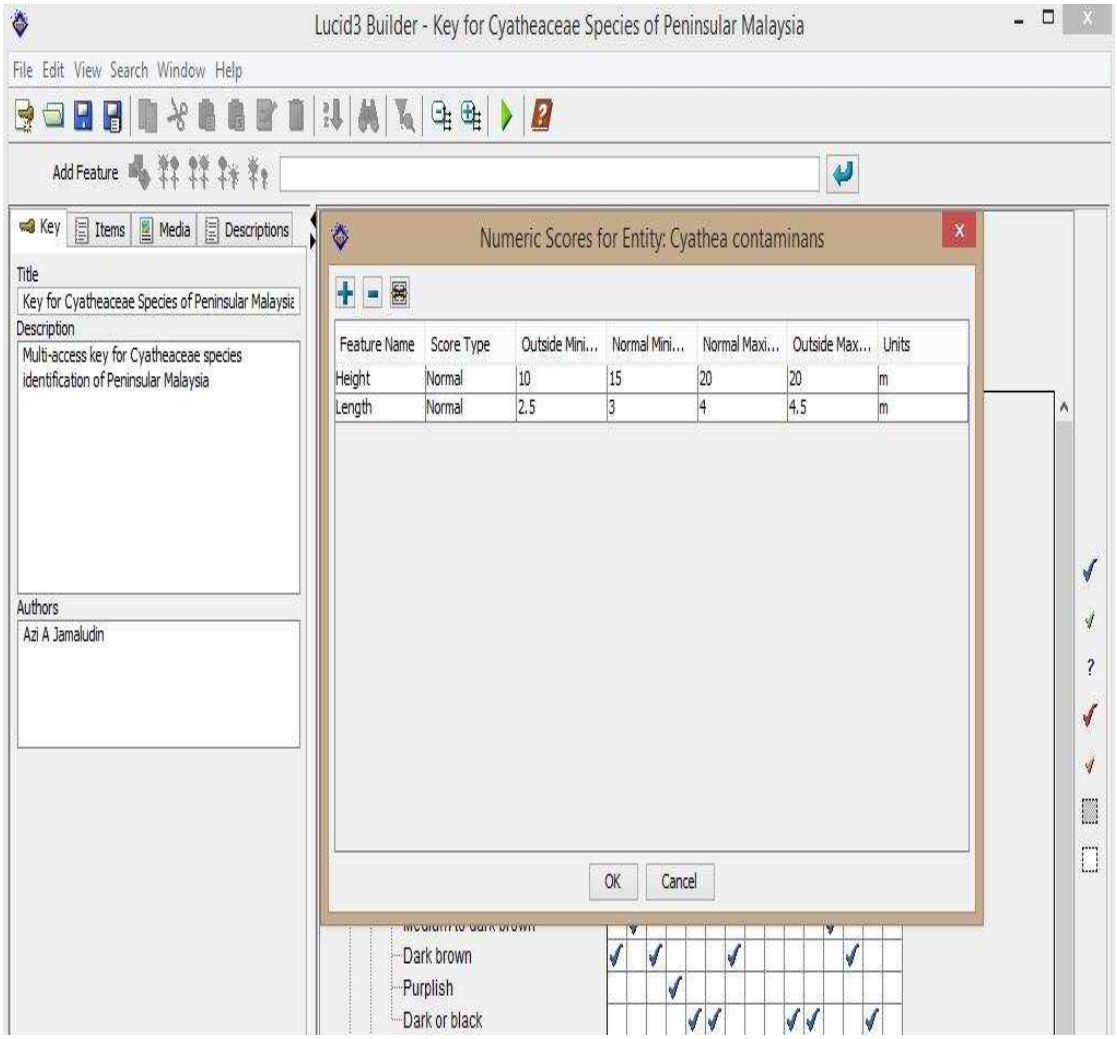


Figure 5.6 Scoring numeric features in the spreadsheet.

In the case where numeric features were used, images and comments regarding how to measure the samples were hugely beneficial. The comments can be added by choosing the items tab while the images can be added by choosing the media tab, in which both tabs can be found on the left side of the Builder window (Figure 6.7). The images then

can be viewed by clicking on the small image next to the features or states in LucID Player or easier comparison and understanding (Figure 6.8).

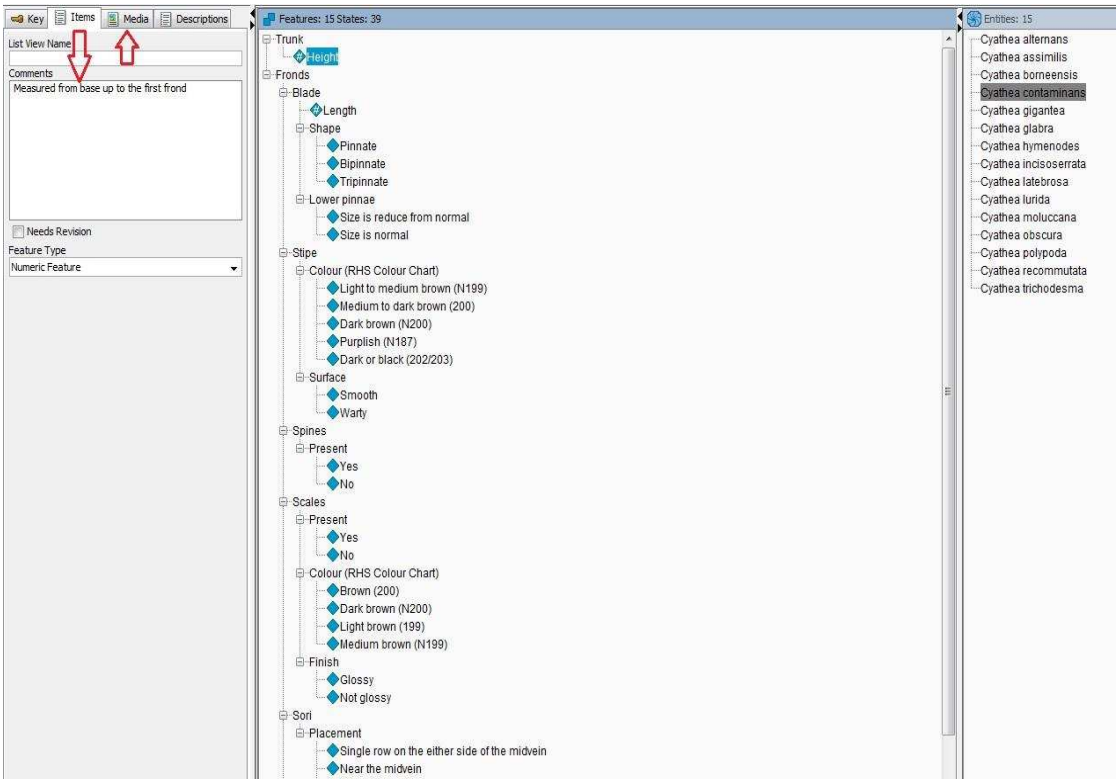


Figure 5.7 Red arrows shows the items and media tabs in LucID Builder. The description regarding the measurement can be added in the comments section in the Items tab.

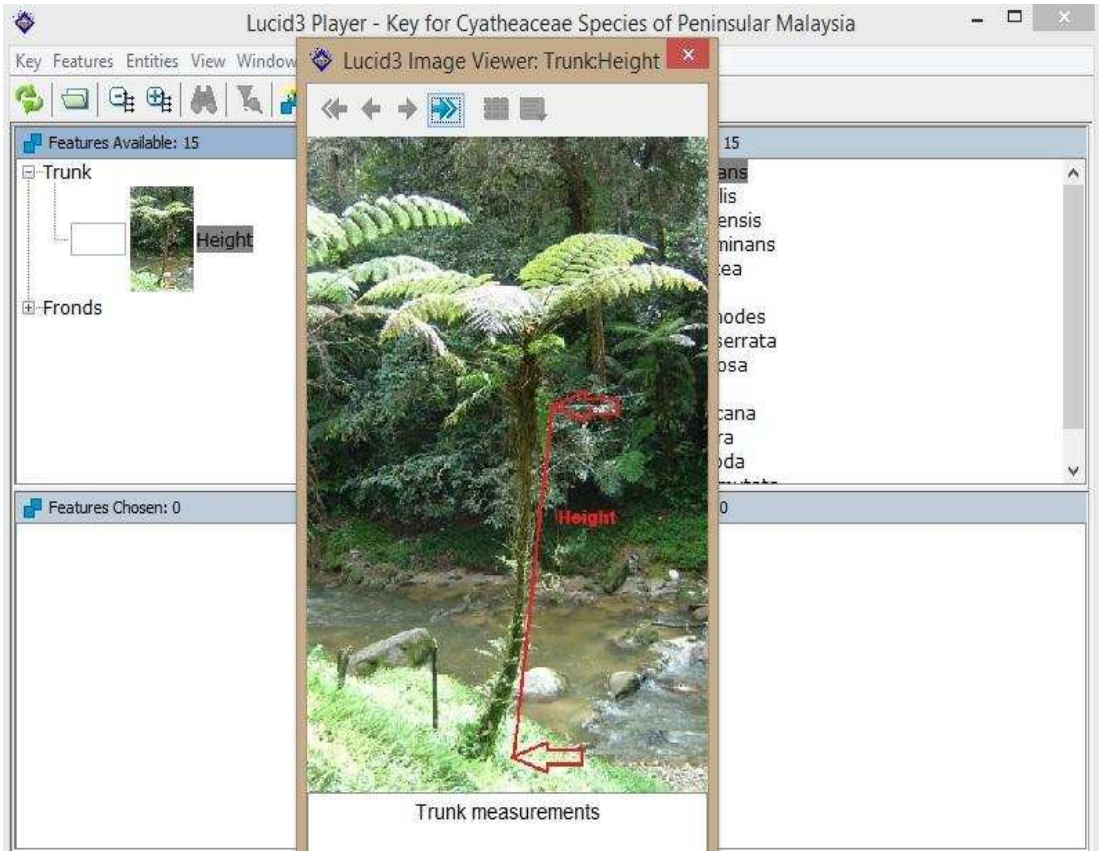


Figure 5.8 Image assigned to the feature describing the measuring of the height.

5.2.4 Assessing the Key

Accessing the end-user interface can be done through the LucID Player applications. Users can select the features they observe on their samples, and as they progress the key will start reducing the number of possible identifications based on their selections. The key will be assessed by the researcher and three other non-Cyatheaceae specialists.

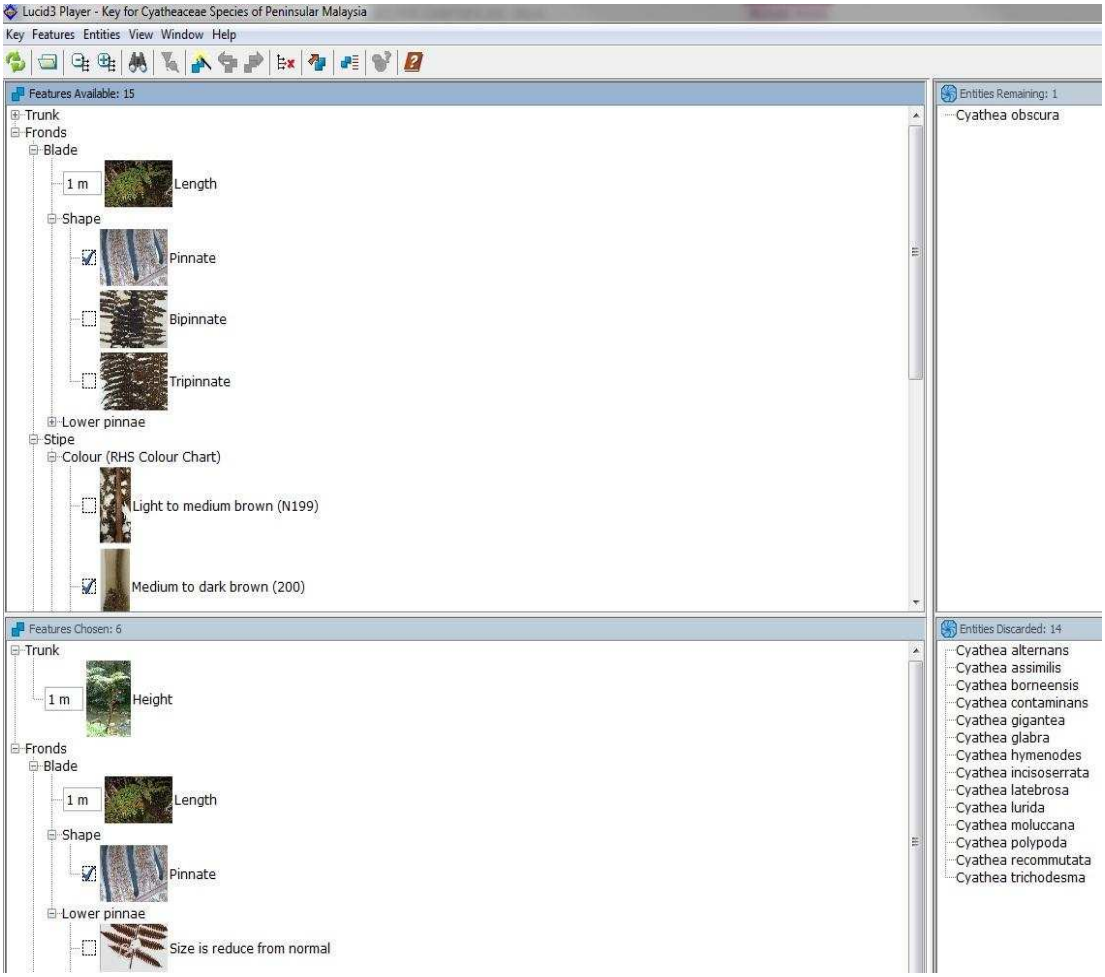


Figure 5.9 Features and states that had been selected managed to narrow down to one entity and excluded other entities of which did not match.

The LucID Player can also be used to compare different taxa by presenting the differences between them in order to assist identification (Figure 6.10).

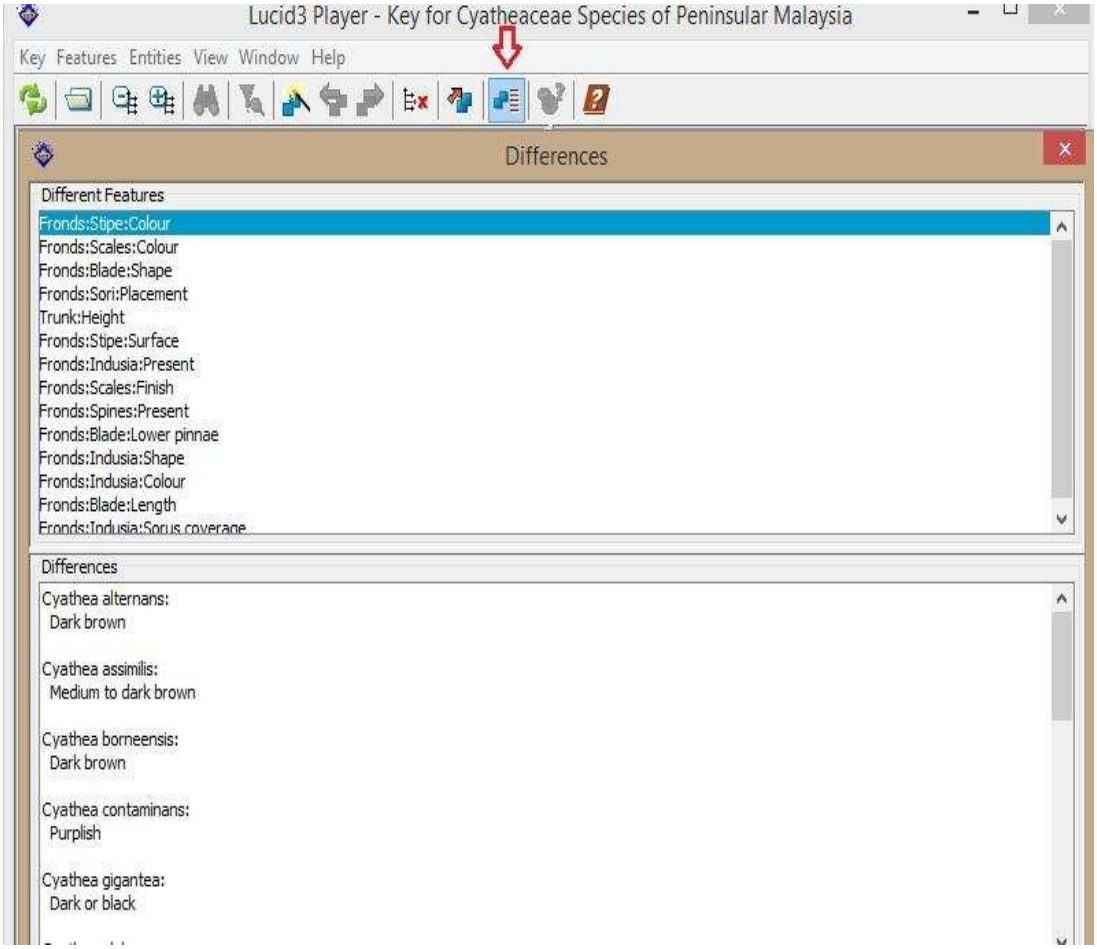


Figure 5.10 The differences shown between species for the feature of stipe colour.

5.3 Results

The 'Key for Cyatheaceae Species of Peninsular Malaysia' was assessed by the researcher and three other non-Cyatheaceae specialists (Table 6.2).

Table 5.2 Results from the assessment of the Cyatheaceae Key for 15 species.

Assessor	Correct Identification	Partial Identification	Wrong Identification
Researcher	15	0	0
PhD Student (Botany)	10	4	1
PhD Student (Bioinformatics)	6	7	2
PhD Student (Forensics)	5	8	2

The most valuable characters in distinguishing taxa were scales, spines, sori and indusia features (Figure 6.11).

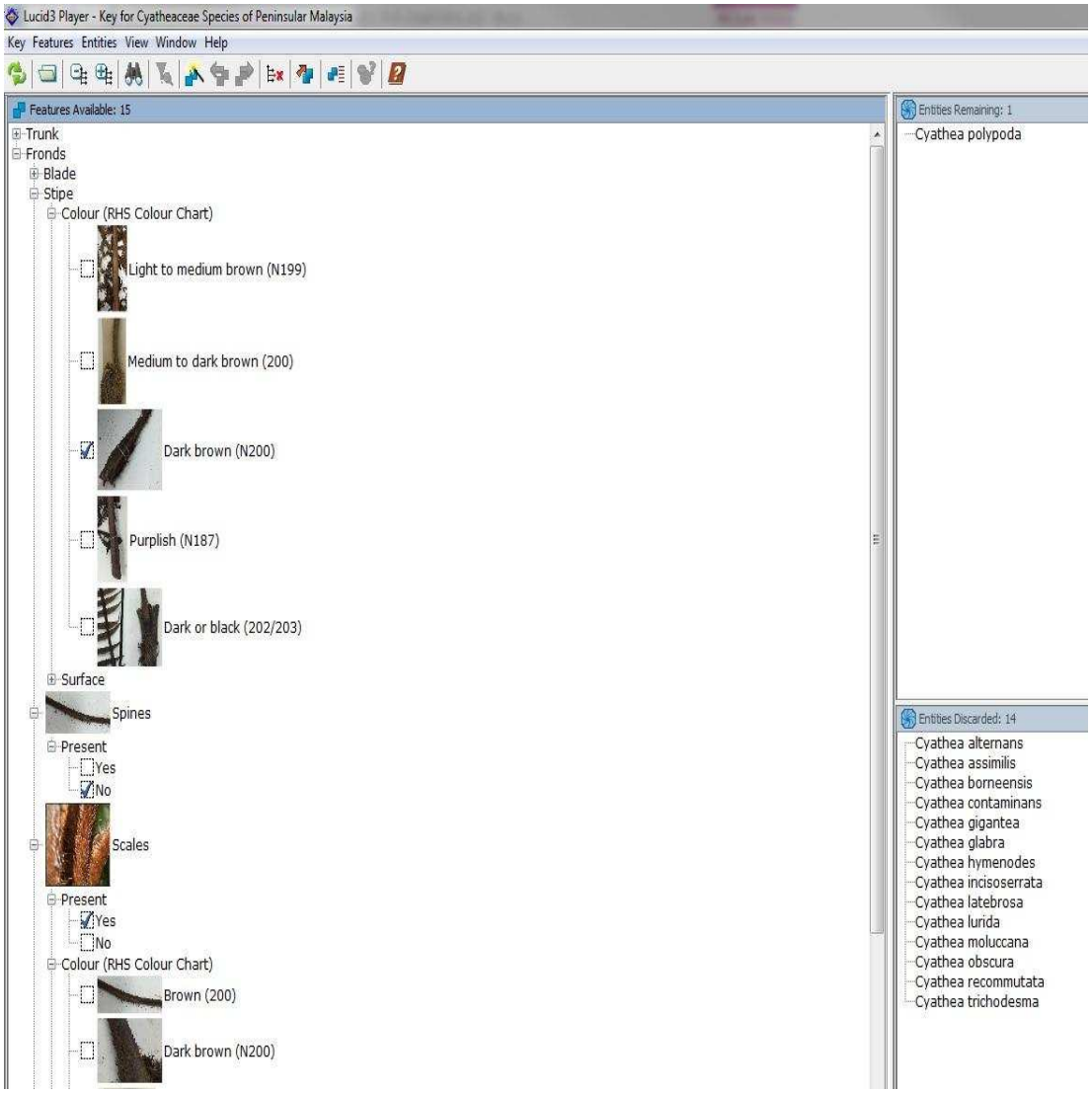


Figure 5.11 Assessing the key by choosing the features and states of the observed samples.

5.4 Discussion

The multi-access key for the Peninsular Malaysian Cyatheaceae was built with the idea of using it both in the field and in the herbarium, in order to provide faster and easier identification. The key is intended to be a professional identification tool that is also accessible to non-taxonomist or to the people who might be keen in Cyatheaceae identification, such as students and gardeners. In this study, the key was assessed by four users, including the researcher, with various identification success. Some improvement to the key needed to be done in terms of adding a glossary to the features used.

Identification of Cyatheaceae was difficult with a sterile frond, as important characters like indusia could not be observed. The interactive multi-access key in this study recorded all the available characters presented in the fertile frond from both the herbarium and own specimens, making identification possible even with incomplete material. This can be achieved as the researcher has a choice to record the features from the sterile frond, without a rigid sequence, in which species whose attributes do not match those of the specimen will be eliminated. The process then continues until one species remains, or at least no further character state values are left to be chosen, to have a partial Cyatheaceae identification.

As the key was largely developed based on the available characters of frond and stipe, the specimen identification is enhanced by assigning images of genuine reference specimens taken in the field and in the laboratory. This should give the end user a

visual comparison between the sample and the feature chosen. However, the identification should always be made from the sample rather than the image.

The multi-access key developed in this study will allow easier and faster changes to be made in the future as the key is not completed with only 15 species from Peninsular Malaysia. Further study should add more features and species from the peninsula as well as Sabah and Sarawak, complete with notes and photographs to produce a complete and comprehensive Cyatheaceae key in Malaysia.

5.5 Conclusion

This study was executed with the primary objective of developing a multi-access identification key for Peninsular Malaysian Cyatheaceae, contributing to the Flora Malesiana. The key was developed and assessed to identify the 15 species of Cyatheaceae, representing the 13 populations from all over the peninsula. However, the key developed is far from being complete as more features and species should be added, thus hampering the publication of the key to the online database for Flora Malesiana. For it to be eligible for Flora Malesiana, species from Cyatheaceae for the whole Malesian region will need to be evaluated. This work can be seen as a platform for further extension of developing the interactive multi-access key for Cyatheaceae in Flora Malesiana.

CHAPTER 6

UPDATING THE PENINSULAR MALAYSIA CYATHEACEAE SPECIES STATUS USING THE IUCN RED LIST CRITERIA

6.1 Introduction

Malaysia's national policy on conservation was set up in the Seventh Malaysia Plan (1996-2000) and the Eighth Malaysia Plan (2001-2005). The primary focus of the policy was to ensure biodiversity conservation while maintaining economic development (MNRE, 2006; MNRE, 2016). Even with the appropriate legislations in place, conservation in Malaysia is still challenging. The Malaysian government encourages both local and international experts to explore the country's unique tropical biodiversity in order to improve understanding and design better conservation policies. However, most of the botanical research in Malaysia focused on native woody plants compared to other plant groups. This can be seen from a number of publications such as *Endemic trees of the Malay Peninsula* and *Tree Flora of Sabah and Sarawak* (Ma, 2010). The country has also taken part in Flora Malesiana, an international collaboration that consist of six countries: Indonesia, Malaysia, Philippines, Papua New Guinea, Singapore and Brunei. This collaboration aims to produce family treatments, up to species level of the Malesian flora of approximately 41500 species of flowering plants and ferns with attention to the indigenous species (MNRE, 2006). The collaboration venture executed by a voluntary network of circa

130 taxonomists from all over the world (MNRE, 2006). Flora Malesiana currently consists of 18 volumes for seed plants and 5 volumes for ferns.

Malaysia is a developing country that underwent fast socio-economic growth. As a result, the country has lost most of its natural resources, such as forest, through ecosystem destruction and deterioration (Napis *et al.*, 2001). Malaysia lost 8.6% of its forest cover in 20 years (1990-2010) (FAO, 2010). Activities such as logging and hydroelectricity schemes led to endangerment of local biodiversity, raising concerns on the conservation status of species present (Napis *et al.*, 2001; MNRE, 2014).

The evaluation of the conservation status of a species is done through the International Union for Conservation of Nature (IUCN) Red List of Threatened Species™, which is broadly known as the most extensive, objective global approach for evaluating the conservation status of plant and animal species (IUCN, 2015). Since the introduction in 1994, the Red List has become a world standard when a strict method to decide risks of extinction was introduce that is applicable to all species (IUCN, 2015).

The status of only 15 Cyatheaceae species has been evaluated, none of which are found in Malaysia. (Figure 7.1).

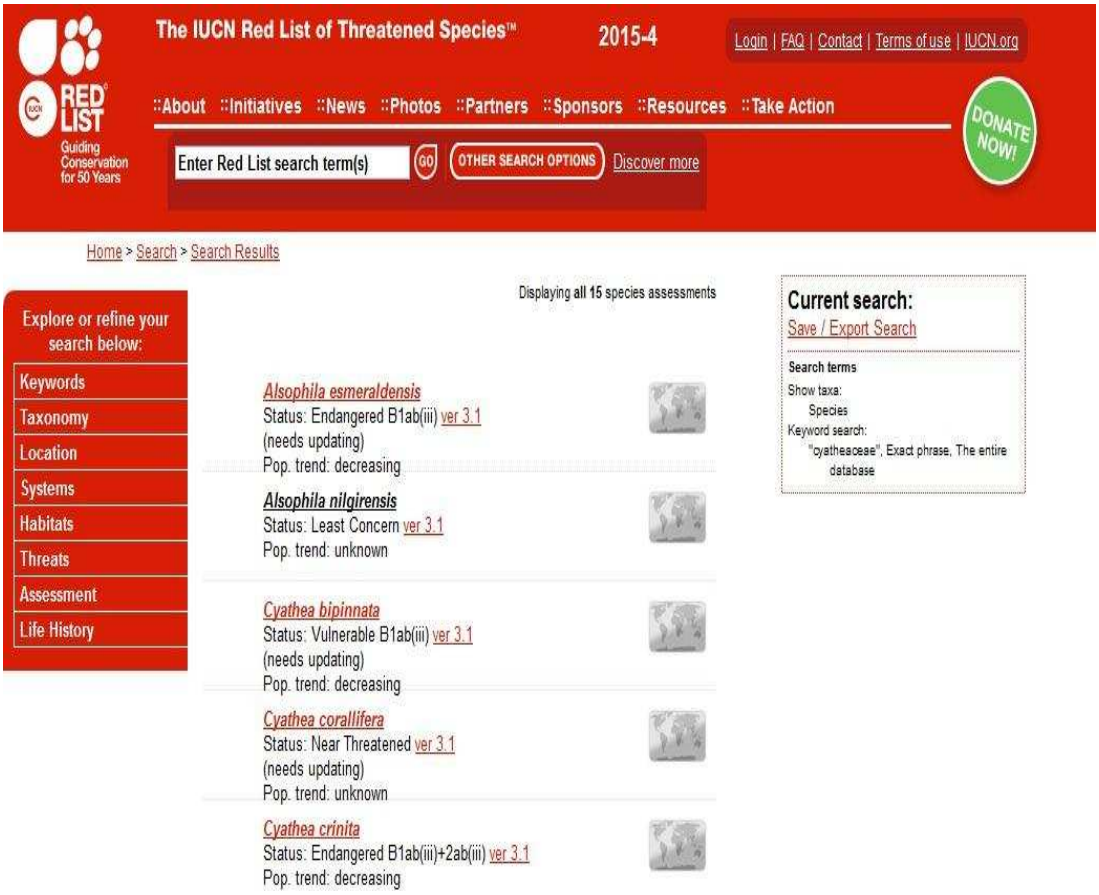


Figure 6.1 Fifteen species of Cyatheaceae evaluated in the IUCN Red List

6.1.1 Aim

This work aimed to assess and evaluate the conservation status of the species identified in Chapter 2. The outcomes will be submitted to the Red List database with the objective of increasing conservation efforts for this family within Malaysia.

6.2 Materials and Methods

6.2.1 Plant species selection

All fifteen species identified in Chapter 2 were used for evaluation as none of them had been previously assessed for the IUCN Red List (Table 7.1).

Table 6.1 Species of Peninsular Malaysian Cyatheaceae previously identified.

Genus	Species
Cyathea	*C. alternans*
	C. assimilis
	C. borneensis
	C. contaminans
	C. gigantea
	C. glabra
	C. hymenodes
	C. incisoserrata
	C. latebrosa
	C. lurida
	C. moluccana
	C. obscura
	C. polypoda
	C. recommutata
	C. trichodesma

6.2.2 Initial Screening

Initial screening was conducted by first evaluating and plotting the distribution range of the species. The screening is mainly based from previous studies and existing literatures, as well as from Malaysia Biodiversity Information System (MYBIS) database, and Global Biodiversity Information Facility (GBIF) database. The information is also extracted from the herbarium specimens, both from Malaysia National University (UKM) and Forest Research Institute Malaysia (FRIM) herbaria.

6.2.3 Applying the IUCN Red List Guidelines

The IUCN Red List Categories and Criteria guidelines were followed without modification. The species were evaluated for regional assessment. Risk of extinction of species was assessed following the threshold values listed under each of the criteria provided.

In this study, assessment was made to Cyatheaceae population in Peninsular Malaysia. However, larger scale assessment that includes other area in Malaysia as well as the neighboring countries is needed in order to provide the Red List with more information regarding the species. It is important for the assessment to be performed at population basis in each states. The outcome from the assessment can be used by the local government to plan for the level of conservation needed in that certain population.

6.2.4 Evaluation Process

Local taxonomist who have knowledge in this Cyatheaceae family and familiar with their distribution in Peninsular Malaysia will be appointed. This will be done to confirm the species evaluation done in this study before it can be submitted to IUCN for anonymous review. The species evaluation will be then amended and updated based on the feedback and review given by the appointed taxonomist. Once all of the evaluations are accepted, it will be then submitted to the Red List unit via the IUCN Species Information Service (SIS). The results of each of the species evaluation will be published online. A SIS account will be set up for each of the species with all the

necessary data, including bibliography and distribution map of the species which will be created with GeoCAT software (Bachman *et al.*, 2011).

6.2.4.1 Implementing the IUCN Criteria and Categories

6.2.4.1.1 IUCN Categories

The IUCN Red List consists of nine main categories (Figure 7.2) which are divided into two major groups; not evaluated (NE) and evaluated species. The evaluated species group is divided into two subgroups: data deficient species (DD) and adequate data species. The latter is then further divided into two main groups: 'non-threatened' (consisting of: Least Concern (LC), and Near Threatened (NT)), and 'threatened' (consisting of Vulnerable (VU), Endangered (EN), Critically Endangered (CR), Extinct in the Wild (EW) and Extinct (EX)).

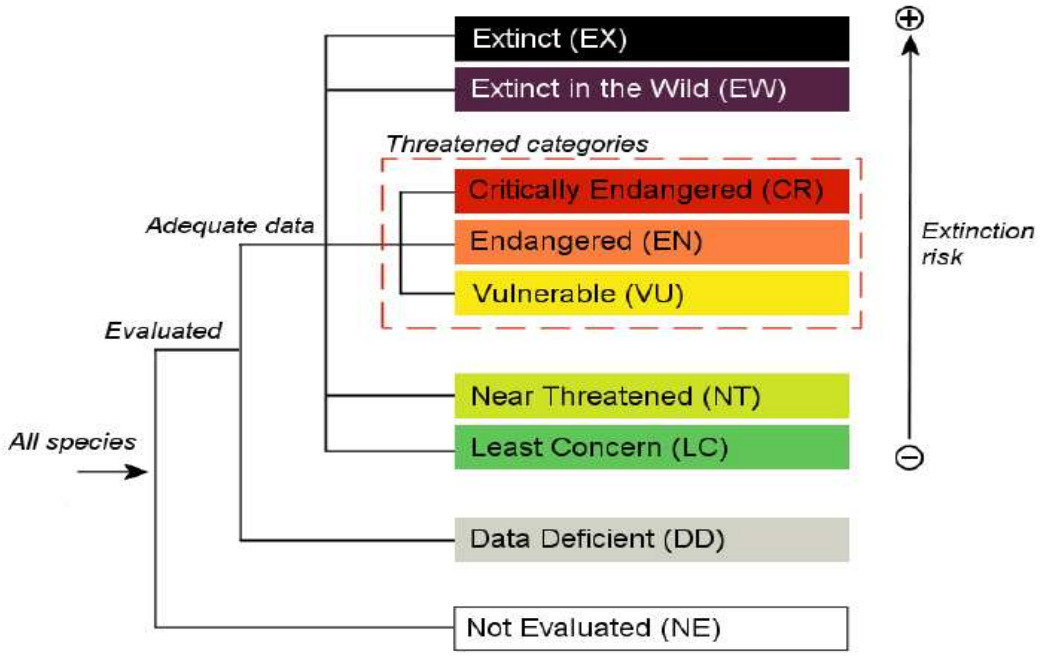


Figure 6.2 The International Union for Conservation Nature (IUCN) Red List Categories at the regional level (IUCN, 2015).

6.2.4.1.2 IUCN Criteria

'Threatened' is the most important category in the Red List according to the IUCN because it consist of 'Critically Endangered', 'Endangered', and 'Vulnerable' status. It contain five criteria; A, B, C, D and E. Each of these criteria is used to evaluate the risk of extinction of the species based on biological and ecological factors. The factors can be; A. Declining population (past, present and/or projected) (Table 7.2), B. Geographic range in the form of either B1 (extent of occurrence) and /or B2 (area of occupancy) (Table 7.3), C. Small population size and decline (Table 7.4), D. Very small or restricted population (Table 7.5) and E. Quantitative analysis of extinction risk (Table 7.6) (IUCN, 2015).

*Table 6.2 Summary of Criteria (**A**) in IUCN Red List criteria and categories used to assess the species and before deciding the status.*

A. Population reduction. Declines measured over the longer of 10 years or 3 generations based on any of A1 to A4.			
	Critically Endangered	Endangered	Vulnerable
A1	≥ 90%	≥ 70%	≥ 50%
A2, A3 & A4	≥ 80%	≥ 50%	≥ 30%
A1. Population reduction observed, estimated, inferred, or suspected in the past where the causes of the reduction are clearly reversible AND understood AND have ceased, based on and specifying any of the following: **(a)** Direct observation. **(b)** An index of abundance appropriate to the taxon. **(c)** A decline in area of occupancy (AOO), extent of occurrence (EOO) and/or habitat quality. **(d)** Actual or potential levels of exploitation **(e)** Effects of introduced taxa, hybridization, pathogens, pollutants, competitors or parasites. **A2**. Population reduction observed, estimated, inferred, or suspected in the past where the causes of reduction may not have ceased OR may not be understood OR may not be reversible, based on (a) to (e) under Al. **A3**. Population reduction projected or suspected to be met in the future (up to a maximum of 100 years) based on (b) to (e) under Al. **A4**. An observed, estimated, inferred, projected or suspected population reduction (up to a maximum of 100 years) where the time period must include both the past and the			

future, and where the causes of reduction may not have ceased OR may not be understood OR may not be reversible, based on (a) to (e) under Al.

*Table 6.3 Summary of Criteria (**B**) in IUCN Red List criteria and categories used to assess the species and before deciding the status.*

B. Geographic range in the form of either B1 (extent of occurrence) AND/OR B2 (area of occupancy)			
	Critically Endangered	Endangered	Vulnerable
B1. Extent of occurrence (EOO)	< 100 km²	< 5,000 km²	< 20,000 km²
B2. Area of occupancy (AOO)	< 10 km²	< 500 km²	< 2,000 km²
AND at least 2 of the following 3 conditions: (a), (b) and **(c)**			
(a) Severely fragmented, **OR** Number of locations	= 1	≤ 5	≤ 10
(b) Continuing decline in any of: **(i)** extent of occurrence; **(ii)** area of occupancy; **(iii)** area, extent and/or quality of habitat; **(iv)** number of locations or subpopulations; **(v)** number of mature individuals.			
(c) Extreme fluctuations in any of: **(i)** extent of occurrence; **(ii)** area of occupancy; **(iii)** number of locations or subpopulations; **(iv)** number of mature individuals.			

*Table 6.4 Summary of Criteria (**C**) in IUCN Red List criteria and categories used to assess the species and before deciding the status.*

C. Small population size and decline			
	Critically Endangered	Endangered	Vulnerable
Number of mature individuals	< 250	< 2,500	< 10,000
AND either C1 or C2:			
C1. An estimated continuing decline of at least:	25% in 3 years or 1 generation	20% in 5 years or 2 generations	10% in 10 years or 3 generations
C2. A continuing decline AND at least 1 of the following 3 conditions:			
(ai) Number of mature individuals in each subpopulation	≤ 50	≤ 250	≤ 1,000
(aii) % individuals in one subpopulation =	90–100%	95–100%	100%

(b) Extreme fluctuations in the number of mature individuals.			

*Table 6.5 Summary of Criteria (**D**) in IUCN Red List criteria and categories used to assess the species and before deciding the status.*

D. Very small or restricted population			
	Critically Endangered	Endangered	Vulnerable
D. Number of mature individuals	< 50	< 250	D1. < 1,000
D2. Only applies to the VU category. Restricted area of occupancy or number of locations with a plausible future threat that could drive the taxon to CR or EX in a very short time.	- -		10% in 10 years or 3 generations

Table 6.6 Summary of Criteria (D) in IUCN Red List criteria and categories used to assess the species and before deciding the status.

E. Quantitative Analysis			
	Critically Endangered	Endangered	Vulnerable
Number of mature individuals	> 50% in 10 years or 3 generations (100 years max.)	> 20% in 20 years or 5 generations (100 years max.)	> 10% in 100 years

The five main criteria are further divided into sub criteria or conditions in which a particular species is evaluated more specifically with a set of quantitative thresholds, under a particular category. If none of the thresholds are met, the species in question could be already Extinct (EX) or Extinct in the Wild (EW). If it nearly meets the conditions for a threatened category it is Near Threatened (NT) and if its current extinction risk is relatively low, it qualifies for Least Concern (LC). If the available data are insufficient to list the species under any category, the species qualifies as Data Deficient (DD) (IUCN, 2015).

Species of interest for evaluation need to be considered against all five criteria using all the available data. Even though a species may not meet all five criteria to qualify as threatened, it has to meet all of the conditions for at least one criteria in order for a conservation status to be made (IUCN, 2015).

The use of criteria A or E in this study is impossible as the information present at the moment could not meet the two requirements. This includes the generation length and the population reduction rate in the past, present or future due to the lack of quantitative data and population trend rates. Criteria C and D are also impossible to use as the information to meet all the requirements in both of these criteria were not sufficient. Relying on the current available information, criteria B was selected due to the availability of distribution range points, which were collected from herbarium labels and databases, and number of species found in the sampling location.

6.2.4.2 Red List Assessment Components

Each of the species in this study was allocated three main components to complete the evaluation. The first component was assigning a Red List category to the species based from the five main criteria. The second component was justifying each assessment with supporting information on the geographical range of the species or description of the habitats of the species or a description of the threats affecting the species populations and habitats. The final component was the distribution map for each of the species.

6.2.4.3 Data and Information Sources

The assessment was done using the updated version of IUCN Red List categories and criteria version 2015-4 (IUCN, 2015). The categories were then justified based from the **B** criteria factors: the estimated extent of occurrence (EOO), area of occupancy (AOO), number of locations, and number of mature individuals. In order to get a complete and correct assessment for Red List evaluation, wide range of data was required such as taxonomic information and distribution data, synonyms, habitat and ecology, uses and threats.

6.2.4.3.1 Taxonomic Information and Synonyms

Taxonomic information and synonyms for each of the species were verified from online databases including The Catalogue of Life, and The Plant List. The Flora Malesiana Series II: Pteridophytes reference book was also used (Holttum, 1963).

6.2.4.3.2 Distribution Information

Distribution information was derived from fieldwork collection and information gathered from the UKM and FRIM herbaria during the visit to Malaysia in 2013 as well as Kew herbarium. Another sources of information was from GBIF database and Flora Malesiana II: Pteridophytes.

The mapping for distribution was quantified by calculating the two main metrics, which were the extent of occurrence (EOO) and area of occupancy (AOO) using

GeoCAT software (Bachman *et al.*, 2011) while following the Red List regulations (Rankou *et al.*, 2015).

6.2.4.3.3 Habitat and Ecology

Information regarding the habitat and ecology of the species in Malaysia is limited. Most of the data available were from personal observation in the field or based on descriptions of the species habitats from scientific literature and herbarium labels.

6.2.4.3.4 Uses and Threats

The locals have long been using most of the species for many socio-economic purposes such as ornamental, construction, horticultural uses, food and medicine. Many of the locals used most of the species for the source of income, particularly the living specimens and this led to overharvesting the plants (Large and Braggins, 2004, Rout *et al.*, 2009). While various purposes of usage can lead to reduction in population sizes, the species are largely threatened by the ecosystem loss and deterioration due to land conversions such as new settlements, as well as illegal logging activities and clearing of forested area for agricultural activities (Napis *et al.*, 2001).

6.3 Results

6.3.1 *Cyathea alternans*

Red List status: **Least Concern (LC)**

Figure 6.3 The distribution of Cyathea alternans in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found in open forest and mostly lowlands and mountains up to 1430 meter in elevation. The Extent of Occurrence (EOO) of this species was 4659 km^2 and the Area of Occupancy (AOO) was 24000 km^2. Only one sample from this species was found while sampling in 2013, which was in Mount Angsi in Negeri Sembilan. However, according to the information from MYBIS and GBIF database as well as information extracted from the herbarium specimens both from UKM and FRIM herbaria, this species can be found in Selangor, Perak, Penang, Pahang, Terengganu and Negeri Sembilan forests.

6.3.2 *Cyathea assimilis*

Red List status: **Least Concern (LC)**

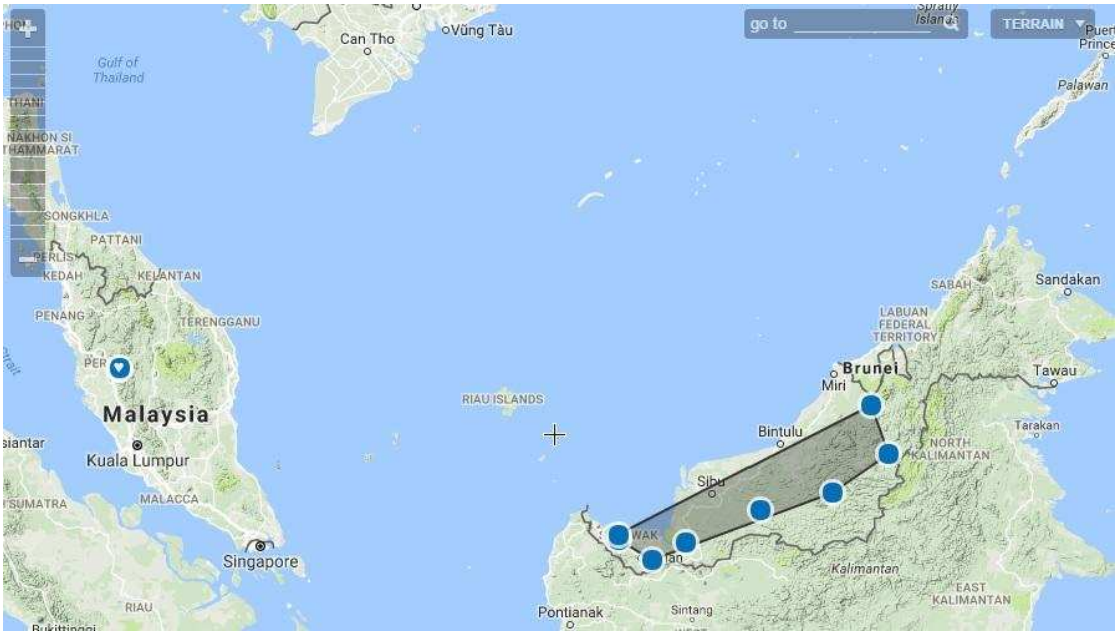


Figure 6.4 The distribution of Cyathea assimilis in Peninsula Malaysia spreading through to Sarawak showing the EOO using geoCAT software.

This species can be found in forest and mostly lowlands and mountains up to 1055 meter in elevation. The Extent of Occurrence (EOO) of this species was 53323 km^2 and the Area of Occupancy (AOO) was 48000 km^2. This species was found only in one place while sampling in 2013, which was in Mount Berinchang in Pahang. According to the information from MYBIS, GBIF database as well as information extracted from the herbarium specimens both from UKM and FRIM herbaria, this species can only be found in Sarawak forests making the species endemic. Thus, the samples identified in Peninsular Malaysia can be either considered as a new sightings or perhaps a misidentification.

6.3.3 *Cyathea borneensis*

Red List status: **Least Concern (LC)**

Figure 6.5 The distribution of Cyathea borneensis in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found in damp, shady forests and in lowlands and mountains from 100 up to 1200 meter in elevation. The Extent of Occurrence (EOO) of this species was 46863 km^2 and the Area of Occupancy (AOO) was 24000 km^2. This species was found in multiple places throughout Peninsular Malaysia while sampling in 2013. Supported by the information from MYBIS, GBIF database as well as information extracted from the herbarium specimens both from UKM and FRIM herbaria, this species can be found in Perak, Pahang, Selangor, Penang, Perlis and Kelantan.

6.3.4 *Cyathea contaminans*

Red List status: **Least Concern (LC)**

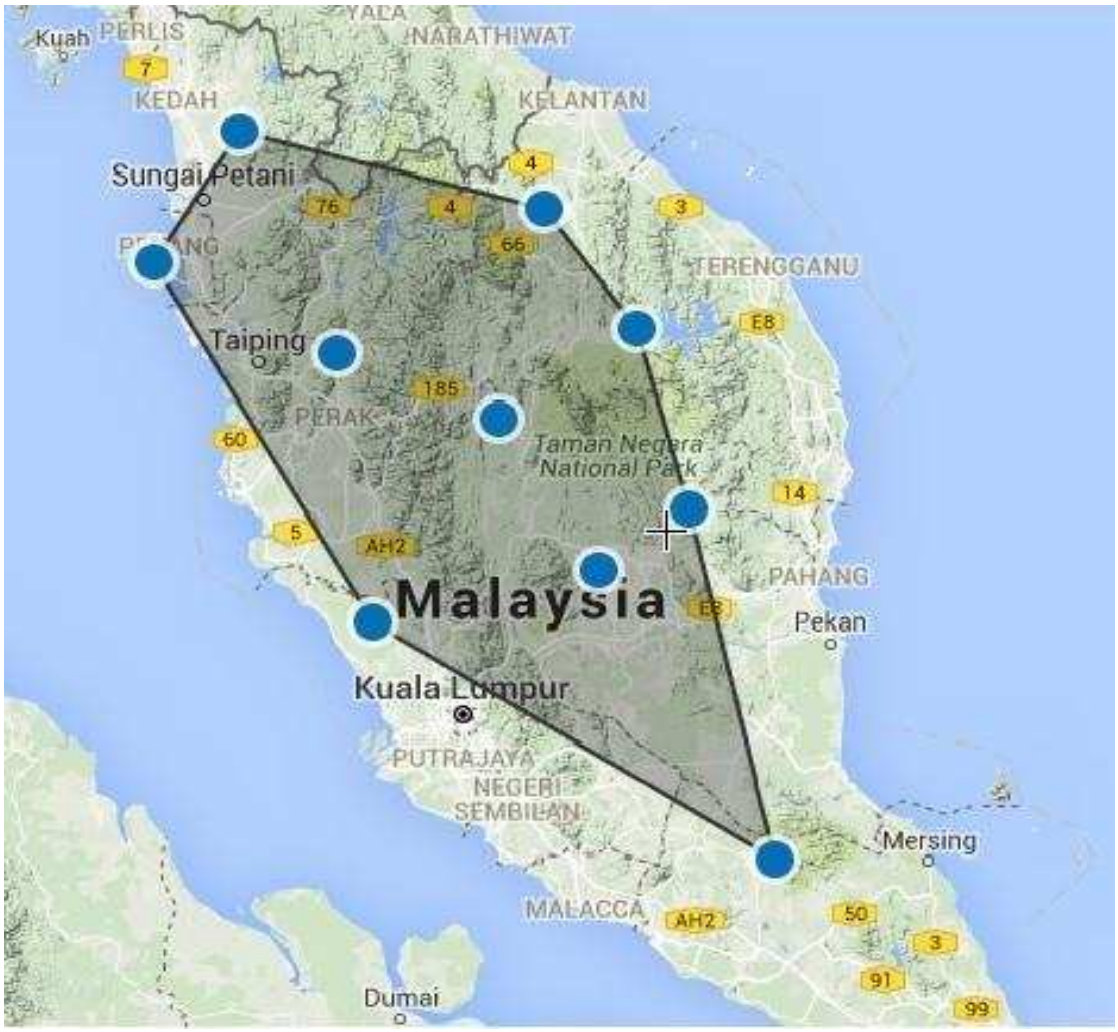


Figure 6.6 The distribution of Cyathea contaminans in Peninsula Malaysia showing the EOO using geoCAT software.

This species is common and can be found at edges and clearings of hill forest and mountains up to 2000 meter in elevation. The Extent of Occurrence (EOO) of this species was 60969 km^2 and the Area of Occupancy (AOO) was 40000 km^2. This species was found throughout Peninsular Malaysia while sampling in 2013. Supported with the information from MYBIS, GBIF database as well as information extracted from the herbarium specimens both from UKM and FRIM herbarium, this species can be found in Perak, Pahang, Selangor, Penang, Kedah, Johor, Terengganu and Kelantan.

6.3.5 *Cyathea gigantea*

Red List status: **Vulnerable (VU): B1ab (ii,iii,v)**

Figure 6.7 The distribution of Cyathea gigantea in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found in open forest up to 350 meter in elevation. The Extent of Occurrence (EOO) of this species was 10144 km^2 and the Area of Occupancy (AOO) was 16000 km^2. This species was found only in Mount Ledang in Johor while sampling in 2013. However, according to the information from MYBIS, GBIF database and information extracted from the herbarium specimens both from UKM and FRIM herbaria, the species can be found in Kedah, Penang, Perak and Johor.

6.3.6 *Cyathea glabra*

Red List status: **Least Concern (LC)**

Figure 6.8 The distribution of Cyathea glabra in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found in damp, shady forest and in lowlands and mountains up to 1700 meter in elevation. The Extent of Occurrence (EOO) of this species was 53978 km^2 and the Area of Occupancy (AOO) was 32000 km^2. This species was found in two places throughout Peninsular Malaysia while sampling in 2013, which were Bukit Larut in Perak and Fraser's Hill in Pahang. However, with the information collected from MYBIS, GBIF database and information extracted from the herbarium specimens both from UKM and FRIM herbaria, this species can be found in Perak, Kelantan, Terengganu, Pahang, Selangor, Kedah, Johor and Negeri Sembilan.

6.3.7 *Cyathea hymenodes*

Red List status: **Least Concern (LC)**

Figure 6.9 The distribution of Cyathea hymenodes in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found in montane forest from 700 to 2200 meter in elevation. The Extent of Occurrence (EOO) of this species was 45184 km^2 and the Area of Occupancy (AOO) was 24000 km^2. This species was found in multiple places throughout Peninsular Malaysia while sampling in 2013. Supported with the information collected from MYBIS, GBIF database and information extracted from the herbarium specimens both from UKM and FRIM herbaria, this species can be found in Pahang, Johor, Penang, Perlis and Kedah.

6.3.8 *Cyathea incisoserrata*

Red List status: **Near Threatened (NT)**

Figure 6.10 The distribution of Cyathea incisoserrata in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found in forest edges and clearings up to 1400 meter in elevation. The Extent of Occurrence (EOO) of this species was 41120 km^2 and the Area of Occupancy (AOO) was 24000 km^2. This species was found in two places throughout Peninsular Malaysia while sampling in 2013 which were Fraser's Hill in Pahang and Lake Kenyir in Terengganu. Supported by the information collected from MYBIS, GBIF database and information extracted from the herbarium specimens both from UKM and FRIM herbarium, this species can be found in Penang, Perak, Pahang, Terengganu and Johor.

6.3.9 *Cyathea latebrosa*

Red List status: **Least Concern (LC)**

Figure 6.11 The distribution of Cyathea latebrosa in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found in an open forest in lowlands and mountains up to 2000 meter in elevation. The Extent of Occurrence (EOO) of this species was 65710 km^2 and the Area of Occupancy (AOO) was 40000 km^2. This species was found in widespread throughout Peninsular Malaysia while sampling in 2013. Supported by the information collected from MYBIS, GBIF database and information extracted from the herbarium specimens both from UKM and FRIM herbaria, this species can be found in Kedah, Penang, Perak, Pahang, Kelantan, Terengganu, Selangor, Negeri Sembilan and Johor.

6.3.10 *Cyathea lurida*

Red List status: **Vulnerable (VU): B1ab (ii,iii,v)**

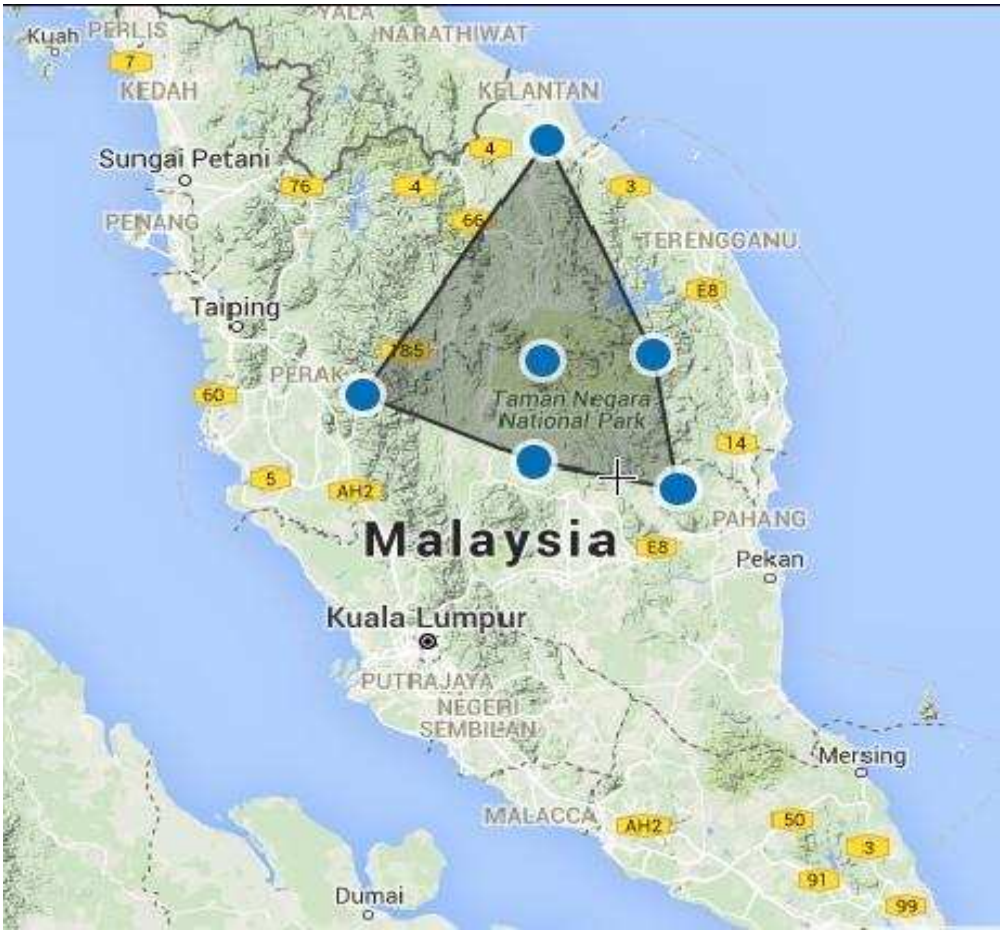


Figure 6.12 The distribution of Cyathea lurida in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found in montane forest between 1250 and 2220 meter in elevation. The Extent of Occurrence (EOO) of this species was 18464 km^2 and the Area of Occupancy (AOO) was 24000 km^2. This species was found only in Fraser's Hill in Pahang, sample courtesy of UKM herbarium. However, according to the information from MYBIS, GBIF database and information extracted from the herbarium specimens both from UKM and FRIM herbarium, the species can be found in Perak, Pahang and Kelantan.

6.3.11 *Cyathea moluccana*

Red List status: **Least Concern (LC)**

Figure 6.13 The distribution of Cyathea moluccana in Peninsula Malaysia showing the EOO using geoCAT software.

This species is common in secondary forest and can be found up to 1300 meter in elevation. The Extent of Occurrence (EOO) of this species was 64780 km^2 and the Area of Occupancy (AOO) was 32000 km^2. This species was found in one place which was Bangi Forest in Selangor, sample courtesy of UKM herbarium. According to the information collected from MYBIS, GBIF database and information extracted from the herbarium specimens both from UKM and FRIM herbaria, this species can be found in Kedah, Penang, Perak, Pahang, Selangor, Negeri Sembilan, Terengganu and Johor.

6.3.12 *Cyathea obscura*

Red List status: **Least Concern (LC)**

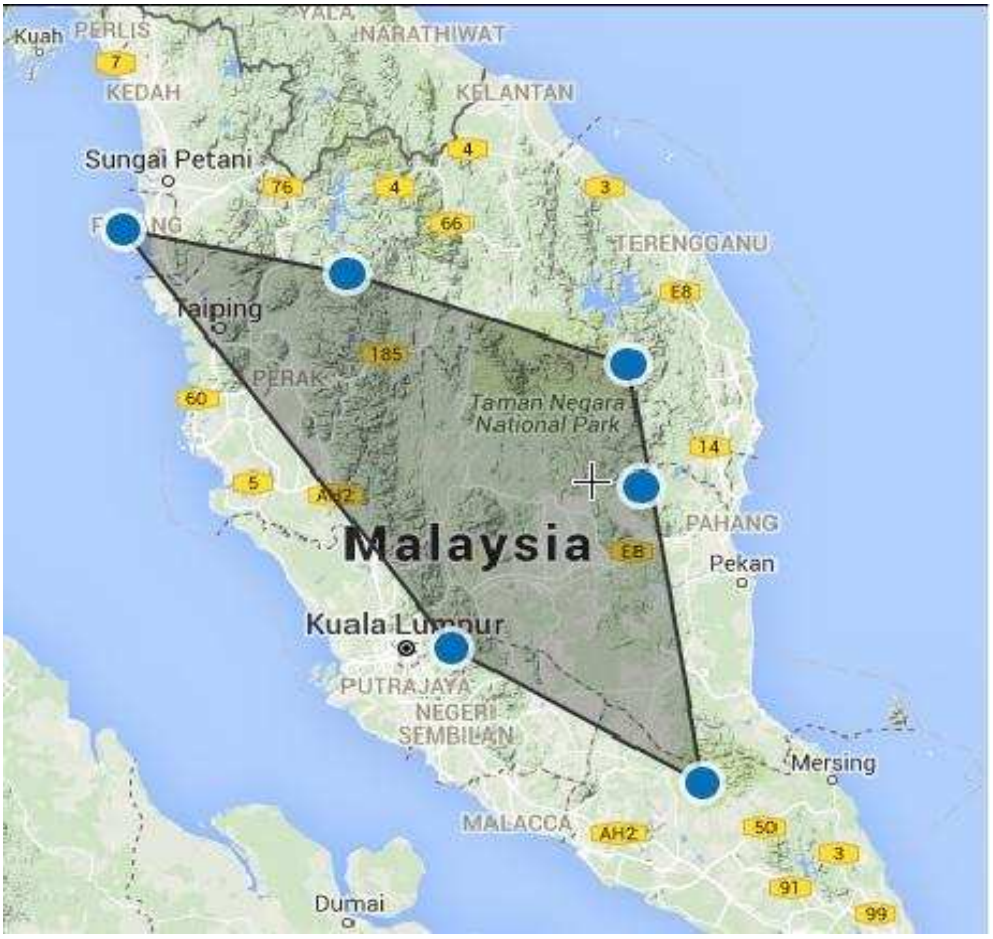


Figure 6.14 The distribution of Cyathea obscura in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found in hill forest at 900 to 2000 meter in elevation. The Extent of Occurrence (EOO) of this species was 45089 km^2 and the Area of Occupancy (AOO) was 24000 km^2. This species was found in three places which were Bukit Larut in Perak, Fraser's Hill and Genting Gighlands in Pahang while sampling in 2013. However, according to the information collected from MYBIS, GBIF database and information extracted from the herbarium specimens both from UKM and FRIM herbaria, this species can be found in Penang, Perak, Pahang, Selangor and Johor.

6.3.13 *Cyathea polypoda*

Red List status: **Near Threatened (NT)**

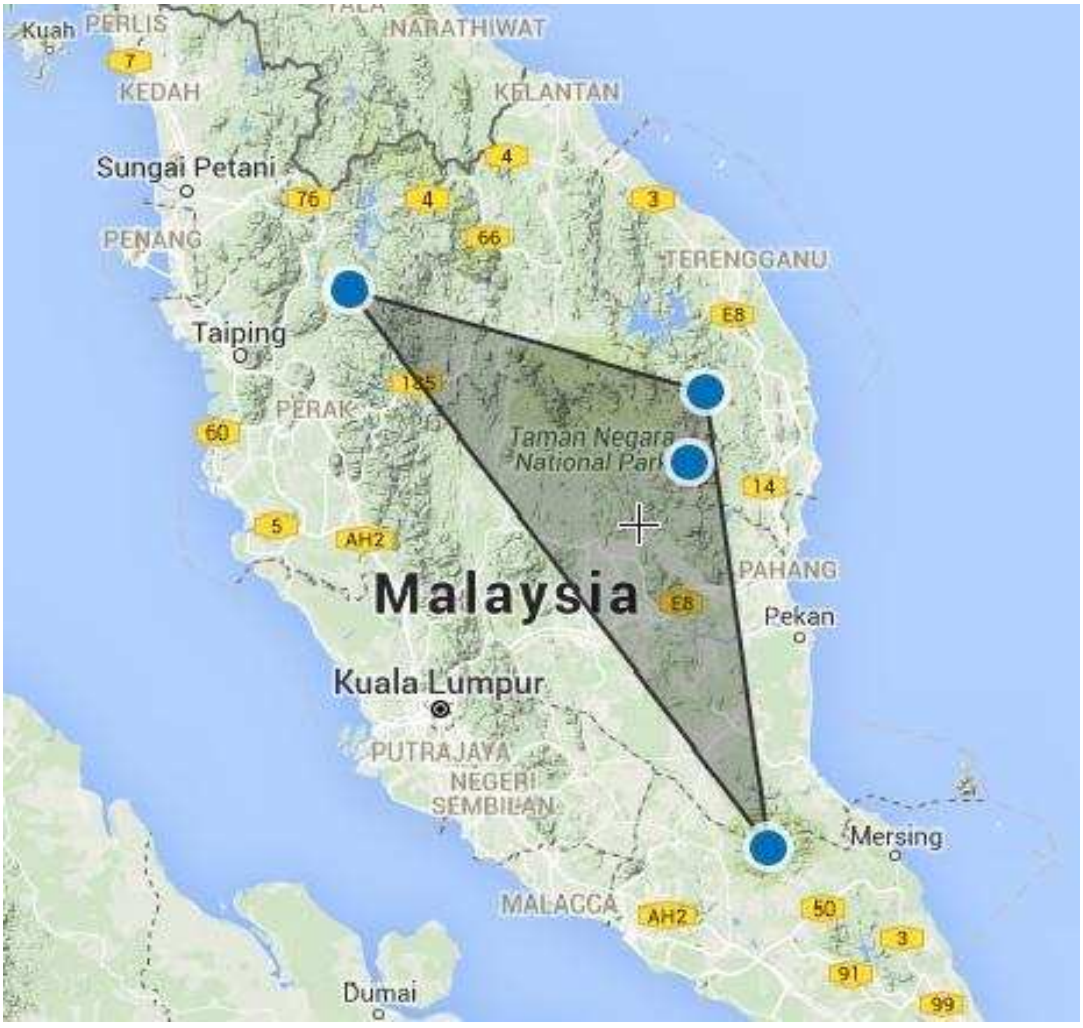


Figure 6.15 The distribution of Cyathea polypoda in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found on ridges in lower montane forest at 1100 to 1300 meter in elevation. The Extent of Occurrence (EOO) of this species was 21957 km^2 and the Area of Occupancy (AOO) was 16000 km^2. This species were found in only one place which was Fraser's Hill in Pahang while sampling in 2013. However, supported by the information collected from MYBIS, GBIF database and information extracted from the herbarium specimens both from UKM and FRIM herbaria, this species can be found in Perak, Pahang and Johor.

6.3.14 *Cyathea recommutata*

Red List status: **Near Threatened (NT)**

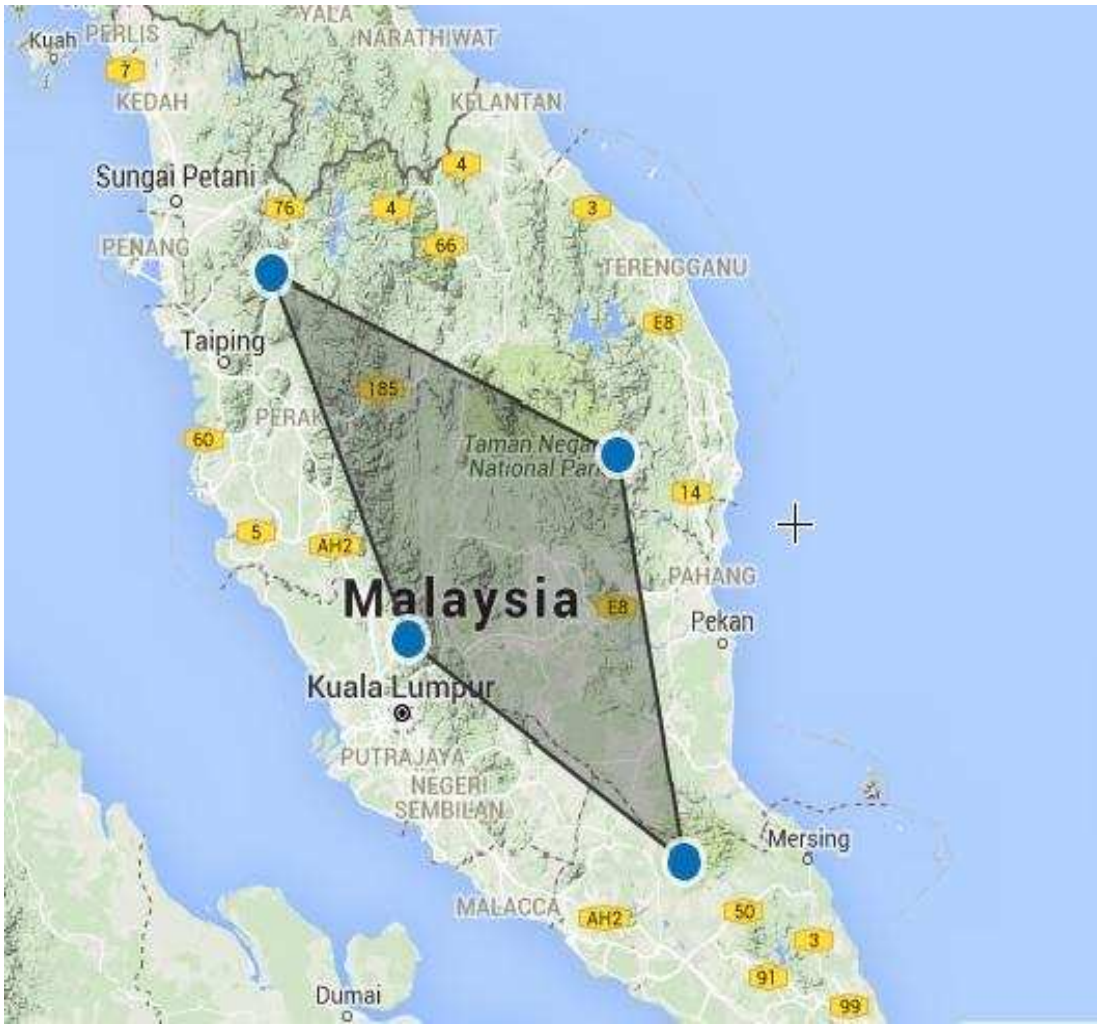


Figure 6.16 The distribution of Cyathea recommutata in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found in shaded montane forest at 600 to 1800 meter in elevation. The Extent of Occurrence (EOO) of this species was 31844 km^2 and the Area of Occupancy (AOO) was 16000 km^2. This species was found in only one place which was Mount Ledang in Johor while sampling in 2013. However, supported with the information collected from MYBIS, GBIF database and information extracted from the herbarium specimens both from UKM and FRIM herbaria, this species can be found in Perak, Pahang and Johor.

6.3.15 *Cyathea trichodesma*

Red List status: **Near Threatened (NT)**

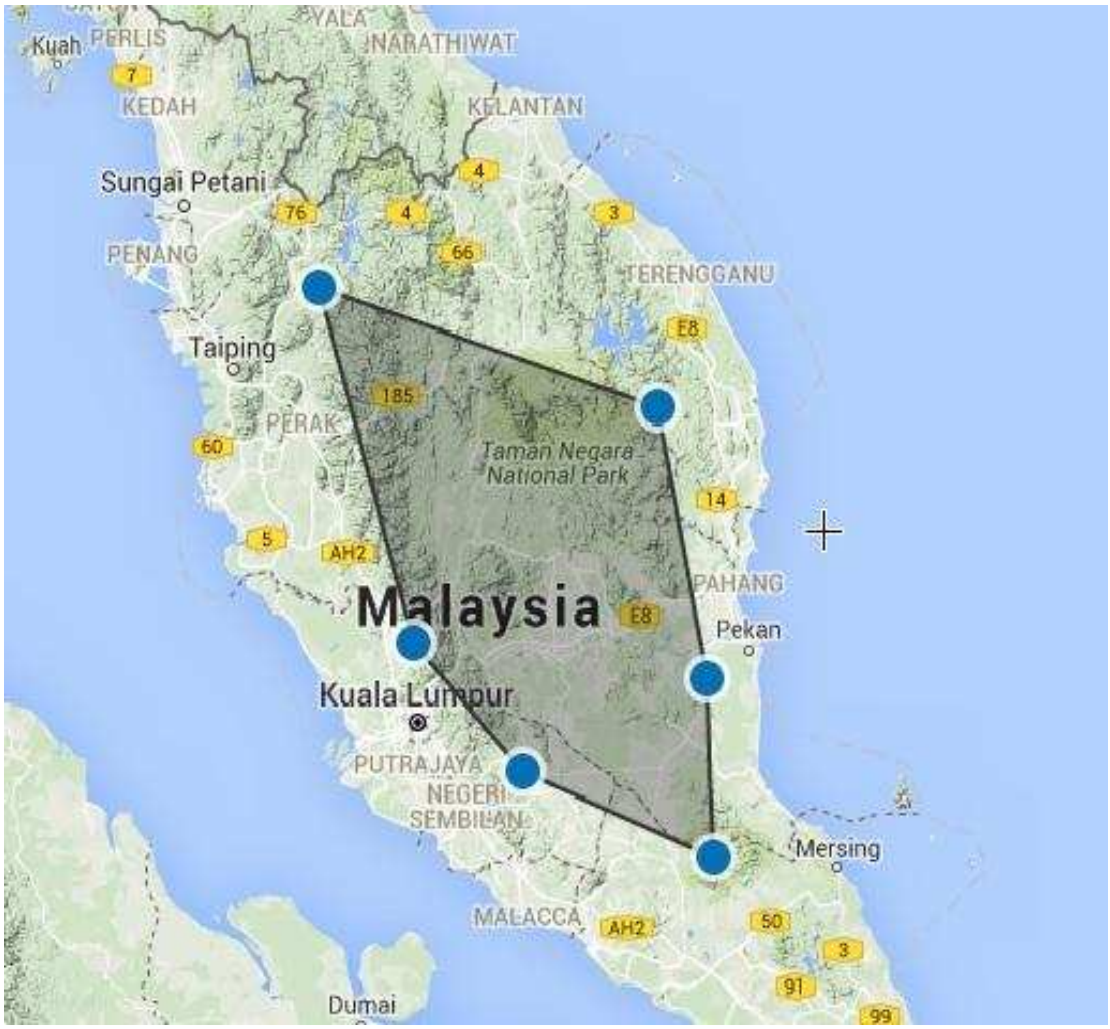


Figure 6.17 The distribution of Cyathea trichodesma in Peninsula Malaysia showing the EOO using geoCAT software.

This species can be found in swampy lowland forest up to 430 meter in elevation. The Extent of Occurrence (EOO) of this species was 39047 km^2 and the Area of Occupancy (AOO) was 24000 km^2. This species was found in two places which were Mount Angsi in Negeri Sembilan and Bukit Larut in Perak while sampling in 2013. However, supported with the information collected from MYBIS, GBIF database and information extracted from the herbarium specimens both from UKM and FRIM herbaria, this species can be found in Perak, Pahang, Selangor, Negeri Sembilan and Johor.

6.4 Discussion

The assessment in this study was done only for the Malaysian region. However, it is noted that national or regional assessments are not eligible for inclusion on the IUCN Red List, unless they are for endemic species (IUCN, 2015). Further assessment in the future will consider the rest of the species distribution. It was assumed that the habitat quality and number of mature individuals of these species will continue to decline as the species were sighted in areas for city development and tourist attractions. It was not known to date whether any conservation measures have been taken by the government as well as local authorities.

It was also found that four of the species, which are *C. incisoserrata, C. polypoda, C. recommutata* and *C. trichodesma* were near threatened **(NT)**. The assessment for these species should be taken into consideration for conservation measures as this status may change for more threatened status as many of the habitats now is undergoing a land conversions and more of the areas become fragmented each years.

The remaining nine species of the Peninsular Malaysia Cyatheaceae (*C. alternans, C. assimilis, C. borneensis, C. contaminans, C. glabra, C. hymenodes, C. latebrosa, C. moluccana*, and *C. obscura*) were all evaluated as least concern **(LC)**. Most of the species can be commonly found throughout the peninsula with multiple individuals in each population. Even though most of these species were common, the status may change due to habitat loss and land conversions, as well as natural disasters if conservation measurements are not implemented.

However, these evaluations have not yet been verified by the local experts who may change the species distributions information. This may lead to the change of the species conservation status. All of the samples in this study were personally identified by the researcher thus there was possibility of having a misidentification.

6.5 Conclusion

This study assessed the species identified from Chapter 2 and assigned the conservation status to each of them. Nine species fall under LC, followed by four species with NT and two species with VU status. The outcome can be used to propose the updated conservation status in the Red List database once the evaluation by local experts is complete and it has been further assessed in wider distribution. The IUCN Red Listing is able to inform and catalyse actions for more appropriate biodiversity conservation measures to be taken by the local government as well as creating reference points of which to observe any changes to the species.

CHAPTER 7

GENERAL DISCUSSSION

Cyatheaceae classification, especially the generic concepts, have been unstable throughout recent history. Different perceptions were made on the available evidence, leading to the suggestion of many evolutionary schemes. The incremental development of knowledge regarding Cyatheaceae structure over the time had provided more evidence in speculating possible relationships, and classification of this family has changed since (Christenhusz and Chase, 2014). The classification of Cyatheaceae in Malesia by Holttum (1963) includes all species and all genera, representing nested monophyletic groups. Here in this study, the nomenclature used follows Holttum (1963), i.e., a single genus, *Cyathea* with subgenera: *Cyathea* and *Sphaeropteris.* Even though Korall *et al.* (2007) has proposed four genera: *Cyathea, Alsophila, Sphaeropteris,* and *Gymnosphaera* and supported with morphological evidence, species from Peninsular Malaysia were not included in the study. Looking through the samples collected during the three month expedition to Peninsular Malaysia, identification was made mostly relying on the dichotomous keys from Flora Malesiana Series II: Pteridophyta (Holttum, 1963). The identification depends mostly on the available character of the frond and was difficult to make a clear-cut separation between the genera proposed by Korall *et al.* (2007). If the genera were to be defined based on the indusium, some of the genera are constant in indusial characters while others are not, thus underscoring the unnaturalness of the genera proposed.

In this study, *Cyathea* species from Peninsular Malaysia was incorporated into the existing phylogeny by using four plastid marker. The species were found interspersed within the three groups (*Alsophila*, *Sphaeropteris*, and *Gymnosphaera*) but interestingly, none of the species was found embedded inside *Cyathea* group, raising a question of whether species from the group exist in Peninsular Malaysia or separated from the rest of Malesian region. However, if using the classification of Holttum (1963), the family consist of two large groups: *Sphaeropteris* (which includes *Schizocaena* and *Fourniera*), and *Cyathea* (which includes *Alsophila* and *Gymnosphaera*). Thus, the Cyatheaceae phylogeny ought to be monophyletic.

Apart from studying the Cyatheaceae phylogeny, this research also use the molecular and identification tools on the local species to evaluate the status of Peninsular Malaysian Cyatheaceae. The work done in this study was designated toward this purpose, in which by using these tools, the information gathered will become the source for assessing the conservation status of Cyatheaceae in Peninsular Malaysia.

The sampling expedition that took place in most of the mountains and highlands in Peninsular Malaysia had witnessed major habitat conversion. Most of it was caused by anthropogenic effects that greatly affected the population of Cyatheaceae. This has a direct influence on the direction of current research in terms of preserving and protecting the local pteridophytes in general and the Cyatheaceae family in particular. The fieldwork conducted had gathered sample of the widest possible range of Cyatheaceae from Peninsular Malaysia, which have not been done previously. The sample collected had provided material to be used in morphological and molecular work in this study. Identification work on the sample found a wider range of

characters essential for Cyatheaceae field identification, especially when identifying species based on sterile individuals. In this study, species description of Peninsular Malaysia Cyatheaceae was updated, based on own observation and references from Holttum (1963) and Large and Braggins (2004), with detailed figures.

DNA barcoding was used in this study with the consequence of making the taxonomic system more accessible. It will benefit the conservation efforts as names and biological attributes of Cyatheaceae will be easily accessed. The idea was to assign specimens to known *Cyathea* species so that it will increase the species discoveries by allowing researchers to rapidly sort specimens, as well as recommending divergent taxa that may represent new species (Hebert and Gregory, 2005). However, this study only allows the proposal for the gene marker to be made, as the effort to develop the barcoding markers generally failed. Further efforts in developing the barcoding markers for this Cyatheaceae family in Malaysia could consider the proposed marker for Cyatheaceae identification in future studies.

The most vital step in conservation is correct identification and delimitation of the target species (Hartvig *et al.*, 2015). However, accurate identification in species-rich or taxonomically complex groups usually needs expert knowledge (Hartvig *et al.*, 2015). Using information gathered during fieldwork and morphological study, a multi-access key was successfully developed and tested for 15 Peninsular Malaysia *Cyathea* species. The key was one of many examples of making taxonomy approachable, as current local trends with newer generation in sciences are more towards biotechnology and taxonomy are dying out (Drew, 2011). Thus, more effort in attracting potential plant taxonomist is needed, especially in fern. Keeping updated with current

technologies will make the species identification key easily accessible from any platform, with easy access and user-friendly interface making plant species identification no longer seen as tedious works.

This research had a more personal note in terms of preserving and protecting this Cyatheaceae family. Having visited the field and witnessed the destruction majorly caused by human interference had triggered the realization regarding the measures needed to conserve the species. The 15 Peninsular Malaysian *Cyathea* species identified had been evaluated according to the IUCN Red List and established current conservation status. The evaluation process involved a detailed assessment of all available data, both online databases and conventional herbarium specimens. Apart from that, species identification and information in several Cyatheaceae specimen vouchers from Malaysia National University herbarium (HUKM) was successfully corrected and updated. The changes made in the herbarium specimens had a serious connotation on current plant taxonomy in Malaysia, especially in ferns. In conjunction with Malaysia's national policy on biodiversity conservation, this research will hopefully prompt other similar efforts, as conservation assessments of many species are need to be updated. The updated conservation information will hopefully help in develop appropriate conservation measures by the government as well as to be used by CITES for monitoring international trade of the *Cyathea* species (CITES, 2013; MNRE, 2016).

This work aimed to evaluate the status of Peninsular Malaysian Cyatheaceae and used the existing molecular and identification tools for the local species. All of the findings in this research show the correlation between the needs of taxonomic update and

conservation efforts. With further research and development needed, this research can be used as a platform or starting point for succeeding studies, taking into account both the negative and positive results.

REFERENCES

Abramoff, M.D., Magalhaes, P.J., Ram, S.J. (2004) Image processing with imageJ. *Biophotonics International*. (11), 36–42.

Álvarez, I., Wendel, J.F. (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*. **29**(3), 417–434.

Ashton, D.H. (2000) The Big Ash forest, Wallaby Creek, Victoria: changes during one lifetime. *Australian Journal of Botany*. **48**, 1–26.

Bachman, S., Moat, J., Hill, A.W., de Torre, J., Scott, B. (2011) Supporting Red List threat assessments with GeoCAT: geospatial conservation assessment tool. *Zookeys*. (150), 117–126.

Bidin, A. (1987) A preliminary survey of the fern flora of Langkawi Islands. *Garden Bulletin Singapore*. **40**(2), 77–102.

Bidin, A. (1983) Notes on the ferns and fern-allies of Carey Island. *Malayan Naturalist*. **37**(1), 13–19.

Bidin, A. (1985) The fern of the genus Brainea in Peninsular Malaysia. *Nature Malaysiana*. **10**(2), 26–28.

Bidin, A.A., Jaman, R. (1999) The Pteridophytes Of Tawau Hills Park, Sabah. *ASEAN Review of Biodiversity and Environmental Conservation (ARBEC)*. (2), 11.

Blows, M.W., Schwarz, M.P. (1991) Spatial-distribution of a primitively social bee: does genetic population structure facilitate altruism? . *Evolution*. **45**(3), 680–693.

CBOL, P.W.G. (2009) A DNA barcode for land plants. *Proceedings of the National Academy of Sciences*. **106**(31), 12794–12797.

Chase, M.W., Fay, M.F. (2009) Ecology. Barcoding of plants and fungi. *Science*. **325**(5941), 682–683.

Chen, C.W., Huang, Y.M., Kuo, L.Y., Nguyen, Q.D., Luu, H.T., Callado, J.R., Farrar, D.R., Chiou, W.L. (2013) trnL-F is a powerful marker for DNA identification of field vittarioid gametophytes (Pteridaceae). *Ann Bot*. **111**(4), 663–673.

Christenhusz, M.J., Chase, M.W. (2014) Trends and concepts in fern classification. *Annals of Botany*. **113**(4), 571–594.

Christenhusz, M.J.M., Zhang, X.-C., Schneider, H. (2011) A linear sequence of extant families and genera of lycophytes and ferns. *Phytotaxa*. **19**, 7–54.

Christensen, C. (1905) *Index Filicum*. Copenhagen: Hagerup.

Cicuzza, D. (2014) A rediscovery of Alfred Russel Wallace's fern collection from

Borneo at the Cambridge University Herbarium. *Notes and Records: the Royal Society journal of the history of science*. **68**(4), 403–412.

CITES (2013) The CITES Appendices. . (28/03/2016). [online]. Available from: https://cites.org/eng/app/index.php.

Conant, D.S., Raubeson, L.A., Attwood, D.K., Stein, D.B. (1995) The Relationships of Papuasian Cyatheaceae to New World Tree Ferns. *American Fern Journal*. **85**(4), 328–340.

Conant, D.S., Stein, D.B. (2001) Phylogenetic and geographic relationships of the tree ferns (Cyatheaceae) on Mount Kinabalu. *Sabah Parks Nature Journal*. **4**(25–43).

Conant, D.S., Stein, D.B., Valinski, A.E.C., Sudarsanam, P., Ahearn, M.E. (1994) Phylogenetic Implications of Chloroplast DNA Variation in the Cyatheaceae. I. *Systematic Botany*. **19**(1), 60–72.

DNASTAR (2016) SeqMan™ Pro.

Dong, W., Cheng, T., Li, C., Xu, C., Long, P., Chen, C., Zhou, S. (2014) Discriminating plants using the DNA barcode rbcLb: an appraisal based on a large data set. *Mol Ecol Resour*. **14**(2), 336–343.

Doyle, J.J., Doyle, J.L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phychem Bull*. **19**, 11–15.

Drew, L.W. (2011) Are We Losing the Science of Taxonomy? *BioScience*. **61**(12), 942–946.

Duffy, A.M., Kelchner, S.A., Wolf, P.G. (2009) Conservation of selection on matK following an ancient loss of its flanking intron. *Gene*. **438**(1–2), 17–25.

Ebihara, A., Nitta, J.H., Ito, M. (2010) Molecular species identification with rich floristic sampling: DNA barcoding the pteridophyte flora of Japan. *PLoS One*. **5**(12), e15136.

Erickson, D.L., Spouge, J., Resch, A., Weigt, L.A., Kress, W.J. (2008) DNA Barcoding in land plants: developing standards to quantify and maximize success. *Taxon*. (57), 1304–1316.

Ferri, G., Corradini, B., Ferrari, F., Santunione, A.L., Palazzoli, F., Alu, M. (2015) Forensic botany II, DNA barcode for land plants: Which markers after the international agreement? *Forensic Sci Int Genet*. **15**, 131–136.

Fountain-Jones, N.M., McQuillan, P.B., Grove, S. (2012) Beetle communities associated with the tree fern Dicksonia antarctica Labill. in Tasmania. *Australian Journal of Entomology*. **51**(3), 154–165.

Fraser-Jenkins, C.R. (2009) A brief comparison of modern pteridophyte classifications (families and genera in India). *Indian Fern Journal*. (26), 107–126.

FRIM, F.R.I.M. (2013) Biological Diversity Clearing House Mechanism: Flora Database.

Gastony, G.J., Yatskievych, G. (1992) Maternal inheritance of the chloroplast and mitochondrial genomes in cheilanthoid ferns. *American Journal of Botany*. (79), 716–722.

Giudicelli, G.C., M??der, G., de Freitas, L.B. (2015) Efficiency of ITS sequences for DNA barcoding in Passiflora (Passifloraceae). *International Journal of Molecular Sciences*. **16**(4), 7289–7303.

Grayer, S. (2009) The Royal Horticultural Society's Colour Chart: an everyday tool for use in the herbarium. Its past, present and future. *Natural Sciences Collections Association*. (18), 19–26.

de Groot, G.A., During, H.J., Maas, J.W., Schneider, H., Vogel, J.C., Erkens, R.H. (2011) Use of rbcL and trnL-F as a two-locus DNA barcode for identification of NW-European ferns: an ecological perspective. *PLoS One*. **6**(1), e16371.

Hall, R. (1998) The Plate Tectonics Of Cenozoic Se Asia And The Distribution Of Land And Sea. In R. Hall & J. D. Holloway, eds. *Biogeography and Geological Evolution of SE Asia*. Leiden, The Netherlands: Backhuys Publishers.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symposium*. (41), 95–98.

Hartvig, I., Czako, M., Kjaer, E.D., Nielsen, L.R., Theilade, I. (2015) The use of DNA barcoding in identification and conservation of rosewood (Dalbergia spp.). *PLoS One*. (10), e0138231.

Hasebe, M., Omori, T., Nakazawa, M., Sano, T., Kato, M., Iwatsuki, K. (1994) rbcL Gene Sequences Provide Evidence for the Evolutionary Lineages of Leptosporangiate Ferns. *Proceedings of the National Academy of Science of the United States of America*. **91**(12), 5730–5734.

Hasebe, M., Wolf, P.G., Pryer, K.M., Ueda, K., Ito, M., Sano, R., Gastony, G.J., Yokoyama, J., Manhart, J.R., Murakami, N., Crane, E.H., Haufler, C.H., Hauk, W.D. (1995) Fern Phylogeny Based on rbcL Nucleotide Sequences. *American Fern Journal*. **85**(4), 134–181.

Hebert, P.D., Cywinska, A., Ball, S.L., deWaard, J.R. (2003) Biological identifications through DNA barcodes. *Proc Biol Sci*. **270**(1512), 313–321.

Hebert, P.D., Gregory, T.R. (2005) The promise of DNA barcoding for taxonomy. . (54), 852–859.

Hebert, P.D., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W. (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. *Proc Natl Acad Sci U S A*. **101**(41), 14812–14817.

Hebert, P.D., Stoeckle, M.Y., Zemlak, T.S., Francis, C.M. (2004) Identification of Birds through DNA Barcodes. *PLoS Biol*. **2**(10), e312.

Hollingsworth, M.L., Andra Clark, A., Forrest, L.L., Richardson, J., Pennington, R.T., Long, D.G., Cowan, R., Chase, M.W., Gaudeul, M., Hollingsworth, P.M. (2009) Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. *Mol Ecol Resour*. **9**(2), 439–457.

Hollingsworth, P.M. (2011) Refining the DNA barcode for land plants. *Proceedings of the National Academy of Sciences*. **108**(49), 19451–19452.

Hollingsworth, P.M., Graham, S.W., Little, D.P. (2011) Choosing and using a plant DNA barcode. *PLoS ONE*. **6**(5).

Holttum, R.E. (1968) *A Revised Flora of Malaya. Fern of Malaya 2*. Singapore: Singapore Govt. Printing Office.

Holttum, R.E. (1963) Cyatheaceae. In C. G. G. J. Van Steenis & R. E. Holttum, eds. *Flora Malesiana Series II-Pteridophyta: Fern and Fern Allies. Volume 1 Part 2*. Netherlands: Martinus Nijhoff/Dr W. Junk Publishers.

Holttum, R.E. (1966) New Ferns from Malesia. *Kew Bulletin*. **20**(3), 6.

Holttum, R.E. (1964) The Genus Cyathea (sensu lato) in Malaysia. *American Fern Journal*. **54**(1), 9.

Holttum, R.E. (1957) The Scales of Cyatheaceae. (With Special Reference to the Genus Schizocaena J. Sm.). *Kew Bulletin*. **12**(1), 5.

Holttum, R.E., Edwards, P.J. (1983) The Tree-Ferns of Mount Roraima and Neighbouring Areas of the Guayana Highlands with Comments on the Family Cyatheaceae. *Kew Bulletin*. **38**(2), 155–188.

IUCN (2015) The IUCN Red List of Threatened Species. *Version 2015-4*.

Jaman, R., Latiff, A. (1998) On Some Pteridophytes of Sayap-Kinabalu Park, Sabah. *ASEAN Review of Biodiversity and Environmental Conservation (ARBEC)*, 35.

Jaman, R., Latiff, A. (1999) The Pteridophyte Flora Of Pulau Tioman, Peninsular Malaysia. *The Raffles Bulletin Of Zoology*. **6**, 24.

Kane, N.C., Cronk, Q. (2008) Botany without borders: barcoding in focus. *Mol Ecol*. **17**(24), 5175–5176.

Könyves, K. (2014) *Taxonomy To Underpin Cultivar Identification In Hoop-Petticoat Daffodils.* University of Reading.

Korall, P., Conant, D.S., Metzgar, J.S., Schneider, H., Pryer, K.M. (2007) A molecular phylogeny of scaly tree ferns (Cyatheaceae). *American Journal of Botany*. **94**(5), 873–886.

Korall, P., Pryer, K.M. (2014) Global biogeography of scaly tree ferns (Cyatheaceae): evidence for Gondwanan vicariance and limited transoceanic dispersal. *J Biogeogr*. **41**(2), 402–413.

Korall, P., Pryer, K.M., Metzgar, J.S., Schneider, H., Conant, D.S. (2006) Tree ferns: monophyletic groups and their relationships as revealed by four protein-coding plastid loci. *Molecular Phylogenetics and Evolution*. **39**(3), 830–845.

Kramer, K.U. (1990) Cyatheaceae. In K. U. Kramer & P. S. Green, eds. *Pteridophytes and Gymnosperms: The Families and Genera of Vascular Plants Volume I*. Berlin: Springer-Verlag, pp. 69–74.

Kress, W.J., Erickson, D.L. (2008) DNA barcodes: genes, genomics, and bioinformatics. *Proc Natl Acad Sci U S A*. **105**(8), 2761–2762.

Kress, W.J., Wurdack, K.J., Zimmer, E.A., Weigt, L.A., Janzen, D.H. (2005) Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci U S A*. **102**(23), 8369–8374.

Kuo, L.Y., Li, F.W., Chiou, W.L., Wang, C.N. (2011) First insights into fern matK phylogeny. *Mol Phylogenet Evol*. **59**(3), 556–566.

Lahaye, R., van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., Maurin, O., Duthoit, S., Barraclough, T.G., Savolainen, V. (2008) DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci U S A*. **105**(8), 2923–2928.

Large, M.F., Braggins, J.E. (2004) *Tree Ferns*. Cambridge: Timber Press.

Latiff, A. (2015) Cyathea arjae Latiff (Cyatheaceae), A New Species of Dwarf Tree-Fern from Sayap, Mt. Kinabalu, Sabah, Malaysia. *Sains Malaysiana*. **44**(1), 4.

Lee, S.S., Swanson, E., Abuzeid, A., Akhlaghi, M., Amjadi, A., Batjargal, U., Cieslikowski, D., Dookhony, L., Fix, R., Hatakeyama, S., Hiraga, M., Horiuchi, K., Kalimili, B.B.N., Khaltarkhuu, B.E., Kwong, A., Lateef, K.S., Levent, I., Muhula, R., Sun, C., Vijayalakshmi, K.M., Zamora, E., Group, D.E.V.P.D.D., Bank, T.W. (2010) *World Development Indicators*. M. de Coquereaumont, B. Ross-Larson, & C. Trott, eds. Washington, USA: The World Bank.

Lehnert, M. (2006) New species and records of tree ferns (Cyatheaceae, Pteridophyta) from the northern Andes. *Organisms Diversity & Evolution*. **6**(4), 321–322.

Lehnert, M. (2009) Three new species of scaly tree ferns (Cyathea-Cyatheaceae) from the northern Andes. *Phytotaxa*. **1**, 43–56.

Lehtonen, S. (2011) Towards resolving the complete fern tree of life. *PLoS One*. **6**(10), e24851.

Lellinger, D.B. (1987) The Disposition of Trichopteris (Cyatheaceae). *American Fern Journal*. **77**(3), 90–94.

Li, F.-W., Tan, B.C., Buchbender, V., Moran, R.C., Rouhan, G., Wang, C.-N.,

Quandt, D. (2009) Identifying a mysterious aquatic fern gametophyte. *Plant Systematics and Evolution*. **281**(1–4), 77–86.

Li, F.W., Kuo, L.Y., Rothfels, C.J., Ebihara, A., Chiou, W.L., Windham, M.D., Pryer, K.M. (2011) rbcL and matK earn two thumbs up as the core DNA barcode for ferns. *PLoS One*. **6**(10), e26597.

Li, X., Yang, Y., Henry, R.J., Rossetto, M., Wang, Y., Chen, S. (2015) Plant DNA barcoding: from gene to genome. *Biol Rev Camb Philos Soc*. **90**(1), 157–166.

Liao, Y.Y., Yang, X.Y., Motley, T.J., Chen, J.M., Wang, Q.F. (2011) Phylogeographic analysis reveals two cryptic species of the endangered fern Ceratopteris thalictroides (L.) Brongn. (Parkeriaceae) in China. *Conservation Genetics*. **12**(5), 1357–1365.

Lindenmayer, D.B., Cunningham, R.B., Donnelly, C.F., Triggs, B.J., Belvedere, M. (1994) The conservation of arboreal marsupials in the montane ash forests of the central highlands of Victoria, South-Eastern Australia V. Patterns of use and the microhabitat requirements of the mountain brushtail possum Trichosurus caninus ogilby in retained l. *Biological Conservation*. **68**(1), 43–51.

Lindsay, S., Middleton, D.J. (2009) Development of a multi-access key to the ferns of Thailand. *THAI FOR. BULL. (BOT.)*, 4.

Ma, K. (2010) *The First Asian Plant Conservation Report*. Beijing, China.

Ma, X.-Y., Xie, C.-X., Liu, C., Song, J.-Y., Yao, H., Luo, K., Zhu, Y.-J., Gao, T., Pang, X.-H., Qian, J., Chen, S.-L. (2010) Species Identification of Medicinal Pteridophytes by a DNA Barcode Marker, the Chloroplast psbA-trnH Intergenic Region. *Pharmaceutical Society of Japan*. **33**(11), 6.

Meier, R., Shiyang, K., Vaidya, G., Ng, P.K.L. (2006) DNA Barcoding and Taxonomy in Diptera: A Tale of High Intraspecific Variability and Low Identification Success. *Syst Biol*. **55**(5), 715–728.

Metcalfe, I. (2002) Permian Tectonic Framework and Paleogeography of SE Asia. *Journal of Asian Earth Science*. **20**, 551–566.

MNRE (2014) *5th Report to Convention On Biological Diversity*. B. and F. M. Division, ed. Kuala Lumpur, Malaysia: The Government of Malaysia.

MNRE (2006) Biodiversity in Malaysia C. and E. M. Division, ed. *Multilateral Environmental Agreements: Capacity Building and Implementation Project*.

MNRE (2016) *National Policy On Biological Diversity 2016-2025*. Putrajaya, Malaysia: The Ministry of Natural Resources and Environment.

MSET (1998) Malaysia's National Policy on Biological Diversity E. and T. Ministry of Science, ed. , 27.

Nagalingum, N.S., Schneider, H., Pryer, K.M. (2007) Molecular Phylogenetic

Relationships and Morphological Evolution in the Heterosporous Fern Genus Marsilea. *Systematic Botany*. **32**(1), 16–25.

Napis, S., Mat Salleh, K., Itam, K., Latiff, A. (2001) Biodiversity databases for Malaysian Flora and Fauna: an update. In *Proceedings of Internet Workshop 2001*. National Institute of Informatics, Tokyo, Japan.

Nock, C.J., Waters, D.L., Edwards, M.A., Bowen, S.G., Rice, N., Cordeiro, G.M., Henry, R.J. (2011) Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnology Journal*. (9), 328–333.

Norton, G. (2000) Invasive Species: The Role Of Lucid Identification Keys. *Proceedings of International Seminar on Biological Invasions.*, 4.

Nunes, C.F., Ferreira, J.L., Nunes Fernandes, M.C., Breves, S.D.S., Generoso, a. L., Fontes Soares, B.D., Carvalho Dias, M.S., Pasqual, M., Borem, a., Almeida Cançado, G.M. (2011) Otimização de um método para extração de DNA genômico a partir de folhas de morangueiro. *Ciência Rural*. **41**(8), 1383–1389.

Nylander, J. (2004) MrModeltest, Version 2.

Oldfield, S. (1995) *International Trade of Tree-Ferns and Evaluation on the Application of CITES*. Cambridge: World Conservation Monitoring Centre.

Parris, B.S., Latiff, A. (1997) Towards a pteridophyte flora of Malaysia: a provisional checklist of taxa. *Malayan Nature Journal*. **50**, 235–280.

Pichi-Sermolli, R.E.G. (1973) Historical review of the higher classification of the Filicopsida. In: JERMY, A. C., CRABBE, J. A. & THOMAS, B. A. (eds.). *Botanical Journal Linnean Society.*

Piggott, A.G. (1988) *Ferns of Malaysia in Colour*. Kuala Lumpur: Tropical Press Sdn. Bhd.

Pryer, K.M., Schneider, H., Smith, A.R., Cranfill, R., Wolf, P.G., Hunt, J.S., Sipes, S.D. (2001) Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature*. **409**, 618–622.

Pryer, K.M., Schuettpelz, E., Huiet, L., Grusz, A.L., Rothfels, C.J., Avent, T., Schwartz, D., Windham, M.D. (2010) DNA barcoding exposes a case of mistaken identity in the fern horticultural trade. *Mol Ecol Resour*. **10**(6), 979–985.

Pryer, K.M., Schuettpelz, E., Wolf, P.G., Schneider, H., Smith, A.R., Cranfill, R. (2004) Phylogeny And Evolution Of Ferns (Monilophytes) With A Focus On The Early Leptosporangiate Divergences. *American Journal of Botany*. **91**(10), 17.

Pryer, K.M., Smith, A.R., Skog, J.E. (1995) Phylogenetic Relationships of Extant Ferns Based on Evidence from Morphology and rbcL Sequences. *American Fern Journal*. **85**(4), 205–282.

Rambaut, A. (2014) FigTree v1.4.2.

Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J. (2014) Tracer v1.6.

Rankou, H., Culham, A., Taleb, M.S., Ouhammou, A., Martin, G., Jury, S.L. (2015) Conservation Assessments and Red Listing of The Endemic Moroccan Flora (Monocotyledons). *Botanical Journal of the Linnean Society*. **177**, 504–575.

Richmond, S. (2010) *Malaysia, Singapore & Brunei*. C. Brash, A. Karlin, S. Low, & B. Presser, eds. Lonely Planet.

Ridley, H.N. (1912) A botanical survey of Pulau Adang. *Journal of the Straits Branch of the Royal Asiatic Society*. **61**, 45–65.

Ridley, H.N. (1908) A List of the Ferns of the Malay Peninsula. *Journal of the Straits Branch of the Royal Asiatic Society*. **50**, 1–59.

Ridley, H.N. (1926) *Ferns of the Malay Peninsula*. London: Reeves and Company Ltd.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic biology, 61, 539-542. *Systematic Biology*. (61), 539–542.

Rothfels, C.J., Larsson, A., Kuo, L.Y., Korall, P., Chiou, W.L., Pryer, K.M. (2012) Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. *Syst Biol*. **61**(3), 490–509.

Rothfels, C.J., Sunduem, M.A., Kuo, L.-Y., Larsson, A., Kato, M., Schuettpelz, E., Pryer, K.M. (2012) A revised family-level classification for eupolypod II ferns (Polypodiidae: Polypodiales). *Taxon*. **61**(3), 515–533.

Rout, S.D., Panda, T., Mishra, N. (2009) Ethnomedicinal studies on some pteridophytes of Similipal Biosphere Reserve, Orissa, India. *International Journal of Medicine and Medical Sciences*. **1**(5), 192–197.

Sanger, F., Coulson, A.R. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*. (94), 441–448.

Schneider, H., He, L., Hennequin, S., Zhang, X.-C. (2013) Towards a natural classification of Pteridaceae: inferring the relationships of enigmatic pteridoid fern species occuring in the Sino-Himalaya and Afro-Madagascar. *Phytotaxa*. (77), 49–60.

Schneider, H., Ranker, T.A., Russell, S.J., Cranfill, R., Geiger, J.M., Aguraiuja, R., Wood, K.R., Grundmann, M., Kloberdanz, K., Vogel, J.C. (2005) Origin of the endemic fern genus Diellia coincides with the renewal of Hawaiian terrestrial life in the Miocene. *Proc Biol Sci*. **272**(1561), 455–460.

Schneider, H., Schuettpelz, E. (2006) Identifying fern gametophytes using DNA sequences. *Molecular Ecology Notes*. **6**(4), 989–991.

Schuettpelz, E., Pryer, K.M. (2007) Fern Phylogeny Inferred from 400 Leptosporangiate Species and Three Plastid Genes. *Taxon*. **56**(4), 1037–1050.

Secretariat of the Convention on Biological Diversity (2010) Global Biodiversity Outlook 3. , 94.

Smith, A.R., Pryer, K.M., Korall, P., Schneider, H., Wolf, P.G. (2006) A classification for extant ferns. *Taxon*. **55**(3), 705–731.

Smith, A.R., Pryer, K.M., Schuettpelz, E., Korall, P. (2008) Fern Classification. In T. A. RANKER & C. H. HAUFLER, eds. *Biology and evolution of ferns and lycophytes.* Cambridge: Cambridge University Press.

Smith, S.Y., Rothwell, G.W., Stockey, R.A. (2003) Cyathea cranhamii sp. nov. (Cyatheaceae), anatomically preserved tree fern sori from the lower cretaceous of Vancouver island, British Columbia . *American Journal of Botany*. **90**(5), 6.

Sonet, G. (2013) DNA Barcoding Flies of Forensci Interest: Validation Test Using Local Entomofauna. *Barcode Bulletin*. **4**(2), 20.

Stein, D.B., Conant, D.S., Valinski, A.E.C. (1997) The implications of chloroplast DNA restriction site variation on the classification and phylogeny of the Cyatheaceae R. J. Johns, ed. *Holttum Memorial Volume*.

Sucher, N.J., Carles, M.C. (2008) Genome-based approaches to the authentication of medicinal plants. *Planta Medica*. (74), 603–623.

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., Willerslev, E. (2007) Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Res*. **35**(3), e14.

Taberlet, P., Gielly, L., Pautou, G., Bouvet, J. (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*. **17**, 1105–1109.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*. **30**(12), 2725–2729.

Tryon, R.M. (1952) A sketch of the history of fern classification. *Annals of the Missouri Botanical Garden*. (39), 255–262.

Tryon, R.M. (1970) The classification of the Cyatheaceae. *Contributions from the Gray Herbarium of Harvard University*. **200**, 3–53.

Tryon, R.M., Gastony, G.J. (1975) The biogeography of endemism in the Cyatheaceae. *Fern Gazette*. **11**, 73–79.

Tryon, R.M., Tryon, A.F. (1982) Ferns and Allied Plants, with Special Reference to Tropical America.

Vaidya, G., Lohman, D.J., Meier, R. (2011) SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*. **27**(2), 171–180.

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., Barton, G.J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. **25**(9), 1189–1191.

Wiemers, M., Fiedler, K. (2007) Does the DNA barcoding gap exist? - a case study in blue butterflies (Lepidoptera: Lycaenidae). *Front Zool*. (8), 8.

Wolf, P.G., Sipes, S.D., White, M.R., Martines, M.L., Pryer, K.M., Smith, A.R., Ueda, K. (1999) Phylogenetic relationships of the enigmatic fern families Hymenophyllopsidaceae and Lophosoriaceae: Evidence from rbcL nucleotide sequences. *Plant Systematics and Evolution*. (219), 263–270.

Yang, J.-B., Tang, M., Li, H.-T., Zhang, Z.-R., Li, D.-Z. (2013) Complete chloroplast genome of the genus cymbidium: Lights into the species identification, phylogenetic implications and population genetic analyses. *BMC Evolutionary Biology*. (13), 1–12.

APPENDIX 1

CTAB-V1 From Könyves (2014)

1. The day before extraction add polyvinylpyrrolidone (PVP) to the CTAB buffer and put it in a water bath at 65°C to dissolve PVP. If β-mercaptoethanol is needed add this at the same time. CTAB buffer with PVP needs to be used within 2-3 days, store capped.
2. Grind 0.03 g of silica-dried leaf in a 2 ml eppendorf tube with 2 beads and a small amount of sand using the BeadBeater at 30 Hz for 60 s. Turn the insert and grind again for another 60 s.
3. Remove the beads and add 800 µl of the pre-warmed (65°C) CTAB buffer onto the powder and grind a bit more. Incubate for at least 1 hour at 65°C. Mix by inverting every 5-10 minutes.
4. Centrifuge at 13000 rpm for 3 minutes. Following centrifugation, you should have the debris on the bottom.
5. Using a pipette carefully transfer the aqueous phase to a clean 1.5 ml eppendorf tube. Avoid removing any material from the debris.
6. Add an equal volume of chloroform/iso-amyl alcohol (24:1) and mix well to obtain an emulsion. Continue inverting for a further 1 minute.
7. Centrifuge at 13000 rpm for 5 minutes. Following centrifugation, you should have three layers: top = aqueous phase, middle = debris, bottom= chloroform. Go on to the next phase quickly so the phases do not remix.
8. Using a pipette carefully transfer the upper aqueous phase to a clean 1.5 ml eppendorf tube. Avoid removing any material from the interface.
9. Repeat the chloroform extraction. This time use 1.5 ml screw cap tube.

10. Add 0.08 volumes of cold 7.5 M ammonium acetate.
11. Add 0.54 volumes (using the combined volume of aqueous phase and added AmAc) of cold isopropanol.
12. Mix well and put in the freezer for 60 minutes or longer. Longer times tend to yield more DNA, but also more contaminats.
13. Centrifuge at 13000 rpm for 15 minutes to pellet.
14. Pour or pipette off the liquid, being careful not to lose the pellet with your DNA.
15. Add 700 µl of cold 70% Ethanol and mix. Leave it stand for a few minutes or until the pellet becomes free.
16. Centrifuge for 1 minute at 13000 rpm.
17. Pour or pipette off the liquid, being careful not to lose the pellet with your DNA.
18. Repeat 70% Ethanol wash.
19. Dry the pellets in the centrivap (35°C) or by inverting samples on a Kim-wipe and let stand until dry.
20. Resuspend samples with 100 µl of TE buffer. Put samples in the fridge overnight to resuspend the pellet. Before running the gel gently flick the tubes and pulse down.

APPENDIX 2

CTAB-V2 From Nunes *et al.* (2011)

1. The day before extraction add polyvinylpyrrolidone (PVP) to the CTAB buffer and put it in a water bath at 65°C to dissolve PVP. If β-mercaptoethanol is needed add this at the same time. CTAB buffer with PVP needs to be used within 2-3 days, store capped.
2. Grind 0.03 g of silica-dried leaf in a 2 ml eppendorf tube with 2 beads and a small amount of sand using the BeadBeater at 30 Hz for 60 s. Turn the insert and grind again for another 60 s.
3. Remove the beads and add 800 µl of the pre-warmed (65°C) CTAB buffer onto the powder and grind a bit more. Incubate for at least 1 hour at 65°C. Mix by inverting every 5-10 minutes.
4. Centrifuge at 13000 rpm for 3 minutes. Following centrifugation, you should have the debris on the bottom.
5. Using a pipette carefully transfer the aqueous phase to a clean 1.5 ml eppendorf tube. Avoid removing any material from the debris.
6. Add 450 µl of chloroform/iso-amyl alcohol (24:1) and mix well to obtain an emulsion. Continue inverting for a further 10 minute.
7. Centrifuge at 3000 rpm for 10 minutes. Following centrifugation, you should have three layers: top = aqueous phase, middle = debris, bottom= chloroform. Go on to the next phase quickly so the phases do not remix.
8. Using a pipette carefully transfer the upper aqueous phase to a clean 1.5 ml eppendorf tube. Avoid removing any material from the interface.

9. The aqueous phase was collected and transferred to a new tube containing 1 mL of of chloroform/iso-amyl alcohol (24:1). Mix well and then centrifuged at 13000 rpm for 15 minutes.
10. The aqueous phase was collected again and transferred to a new tube and then added 150 µl of mM ammonium acetate and 750 µl of chilled isopropanol.
11. Mix well and put in the freezer for incubation overnight at -20°C.
12. Centrifuge at 13000 rpm for 15 minutes to pellet.
13. Pour or pipette off the liquid, being careful not to lose the pellet with your DNA.
14. Add 700 µl of cold 70% Ethanol and mix. Leave it stand for a few minutes or until the pellet becomes free.
15. Centrifuge for 1 minute at 13000 rpm.
16. Repeat 70% Ethanol wash.
17. Dry the pellets in the centrivap (35°C) or by inverting samples on a Kim-wipe and let stand until dry.
18. Resuspend samples with 100 □l of TE buffer. Put samples in the fridge overnight to resuspend the pellet. Before running the gel gently flick the tubes and pulse down.

APPENDIX 3

CTAB-V1 Fern

1. The day before extraction add polyvinylpyrrolidone (PVP) to the CTAB buffer and put it in a water bath at 65°C to dissolve PVP. If β-mercaptoethanol is needed add this at the same time. CTAB buffer with PVP needs to be used within 2-3 days, store capped.
2. Grind 0.03 g of silica-dried leaf in a 2 ml eppendorf tube with 2 beads and a small amount of sand using the BeadBeater at 30 Hz for 60 s. Turn the insert and grind again for another 60 s.
3. Remove the beads and add 800 µl of the pre-warmed (65°C) CTAB buffer onto the powder and grind a bit more. Incubate for at least 1 hour at 65°C. Mix by inverting every 5-10 minutes.
4. Centrifuge at 13000 rpm for 3 minutes. Following centrifugation, you should have the debris on the bottom.
5. Using a pipette carefully transfer the aqueous phase to a clean 1.5 ml eppendorf tube. Avoid removing any material from the debris.
6. Add an equal volume of chloroform/iso-amyl alcohol (24:1) and mix well to obtain an emulsion. Continue inverting for a further 1 minute.
7. Centrifuge at 13000 rpm for 5 minutes. Following centrifugation, you should have three layers: top = aqueous phase, middle = debris, bottom= chloroform. Go on to the next phase quickly so the phases do not remix.
8. Using a pipette carefully transfer the upper aqueous phase to a clean 1.5 ml eppendorf tube. Avoid removing any material from the interface.
9. Repeat the chloroform extraction. This time use 1.5 ml screw cap tube.

10. Add 2/3 volume of ice cold isopropanol. Mix well and put it in the freezer for 60 minutes or longer. Longer times tend to yield more DNA, but also more contaminants.
11. Centrifuge at 3000 rpm for 5 minutes to pellet.
12. Pour or pipette off the liquid, being careful not to lose the pellet with your DNA.
13. Add 1000 µl of wash buffer. Leave it stand for a few minutes, then centrifuge at 3000 rpm for 5 minutes.
14. Pour or pipette off the liquid, being careful not to lose the pellet with your DNA.
15. Dry the pellets in the centrivap (35°C).
16. Resuspend the pellet in 90 µl of resuspension buffer. If a pellet does not dissolve, place in a 65°C water bath for up to 10 minutes. If a pellet remains, the DNA is contaminated with protein or polysaccharide in which case centrifuge at 3000 rpm for 1 minute and pipette the supernatant into a fresh screw cap tube.
17. Add 180 µl of RO water, 135 µl of 7.5 ammonium acetate and 1000 µl of ice-cold Ethanol. Gently invert and leave it in the freezer for 1 hour to precipitate.
18. Centrifuge at 13000 rpm for 10 minutes to pellet.
19. Pour or pipette off the liquid, being careful not to lose the pellet with your DNA.
20. Add 700 µl of cold 70% Ethanol and mix. Leave it stand for a few minutes or until the pellet becomes free.
21. Centrifuge at 13000 rpm for 1 minute.

22. Pour or pipette off the liquid, being careful not to lose the pellet with your DNA.

23. Dry the pellets in the centrivap (35°C) or by inverting samples on a Kim-wipe and let stand until dry.

24. Resuspend samples with 100 µl of TE buffer. Put samples in the fridge overnight to resuspend the pellet. Before running the gel gently flick the tubes and pulse down.

APPENDIX 4

List of samples with DNA sequences for at least one region

Species	Code	Location	Regions/Sequences Length (bp)			
			rbcL	***matK***	***trnL-F***	***trnG-R***
C. contaminans	BL01	Bukit Larut	-	926	903	1034
	BL08	Bukit Larut	-	-	853	-
	BL12	Bukit Larut	-	937	930	1009
	BL13	Bukit Larut	-	-	821	-
	BL16	Bukit Larut	-	1100	917	-
	BL18	Bukit Larut	-	1151	-	-
	BL35	Bukit Larut	-	929	904	997
	FH57	Fraser's Hill	1291	-	-	906
	MB34	Mount Berinchang	-	-	-	968
	LK09	Lake Kenyir	1247	925	-	-
	LK14	Lake Kenyir	-	944	-	971
	LH04	Lojing Highlands	-	1212	-	1088
	LH14	Lojing Highlands	1258	-	944	964
	LH18	Lojing Highlands	-	930	-	1004
C.borneensis	BL04	Bukit Larut	-	-	917	-
	BL14	Bukit Larut	-	-	913	-
	BL20	Bukit Larut	-	-	835	-
	BL22	Bukit Larut	-	953	902	923
	BL29	Bukit Larut	-	-	900	930
	PH03	Penang Hill	-	1027	953	937
	PH04	Penang Hill	-	927	-	959
	PH05	Penang Hill	-	-	-	968
	PH06	Penang Hill	-	1162	946	1010
	PH09	Penang Hill	1199	-	960	1021
	PH10	Penang Hill	-	-	-	953
	PH11	Penang Hill	-	935	956	928
	PH12	Penang Hill	-	1057	946	1005
	PH13	Penang Hill	-	1009	932	1018
	PH14	Penang Hill	-	970	959	1044
	PH15	Penang Hill	-	-	932	926
	PH16	Penang Hill	1274	925	952	1053
	PH17	Penang Hill	-	1080	958	962
	PH18	Penang Hill	-	933	-	976
	PH19	Penang Hill	-	-	-	930
	PH20	Penang Hill	-	-	948	939
	PH21	Penang Hill	-	-	935	931
	PH22	Penang Hill	1206	930	962	992

	PH23	Penang Hill	-	917	-	963
	PH24	Penang Hill	1205	958	-	970
	PH25	Penang Hill	1212	929	952	-
	PH27	Penang Hill	-	1009	-	945
	MP03	Mount Perlis	-	-	-	981
	MP04	Mount Perlis	-	-	-	932
	MP05	Mount Perlis	-	-	-	990
	MP06	Mount Perlis	1247	923	948	-
	MP07	Mount Perlis	-	-	936	994
	MP08	Mount Perlis	-	-	925	1017
	MP10	Mount Perlis	-	912	951	1080
	MP11	Mount Perlis	-	-	-	1008
	MP14	Mount Perlis	-	-	-	979
	FH07	Fraser’s Hill	-	-	851	-
	FH08	Fraser’s Hill	-	-	-	976
	FH23	Fraser’s Hill	-	-	926	922
	FH24	Fraser’s Hill	-	826	896	964
	FH27	Fraser’s Hill	-	-	873	-
	FH28	Fraser’s Hill	-	-	898	1004
	FH29	Fraser’s Hill	-	-	905	-
	FH38	Fraser’s Hill	-	-	909	919
	FH40	Fraser’s Hill	-	-	930	954
	FH41	Fraser’s Hill	-	802	902	963
	FH44	Fraser’s Hill	-	-	834	-
	FH45	Fraser’s Hill	-	966	912	919
	FH48	Fraser’s Hill	-	-	852	-
	FH50	Fraser’s Hill	-	-	863	944
	FH51	Fraser’s Hill	-	-	944	920
	FH54	Fraser’s Hill	-	890	962	955
	FH58	Fraser’s Hill	-	-	946	954
	GH17	Genting Highlands	-	879	-	956
	MB03	Mount Berinchang	-	-	513	-
	MB04	Mount Berinchang	-	-	921	-
	MB06	Mount Berinchang	-	-	772	-
	MB09	Mount Berinchang	-	-	899	-
	MB14	Mount Berinchang	-	-	955	939
	MB15	Mount Berinchang	-	-	-	1047
	MB17	Mount Berinchang	-	-	943	-
	MB20	Mount Berinchang	-	-	897	-
	MB23	Mount Berinchang	-	-	955	-
	MB26	Mount Berinchang	1217	992	955	-
	MB27	Mount Berinchang	1013	919	947	-

	MB28	Mount Berinchang	1264	-	948	982
	MB29	Mount Berinchang	-	-	953	916
	MB30	Mount Berinchang	-	-	-	971
	MB31	Mount Berinchang	-	936	-	966
	MB36	Mount Berinchang	1185	904	964	974
	MB37	Mount Berinchang	-	1162	-	972
	MB38	Mount Berinchang	1115	901	-	989
	MB39	Mount Berinchang	-	-	950	974
	LK18	Lake Kenyir	-	920	930	936
	LK20	Lake Kenyir	-	938	952	-
	LK21	Lake Kenyir	-	-	944	-
	LK22	Lake Kenyir	-	-	914	-
	LK23	Lake Kenyir	-	715	904	-
	LK24	Lake Kenyir	-	734	947	-
	LK25	Lake Kenyir	-	1000	958	869
	LK26	Lake Kenyir	-	-	-	934
	LK27	Lake Kenyir	-	-	946	-
	LK29	Lake Kenyir	-	908	949	977
	LH29	Lojing Highlands	-	934	950	987
	LH30	Lojing Highlands	1244	-	-	-
	BK03	Batang Kali	-	823	924	964
	BK23	Batang Kali	-	879	932	981
	BK24	Batang Kali	-	-	901	1020
	BK25	Batang Kali	-	-	-	863
	BK26	Batang Kali	-	-	930	966
	BK28	Batang Kali	-	-	879	-
	BK30	Batang Kali	-	-	897	899
	BK31	Batang Kali	-	-	-	768
	BK32	Batang Kali	-	-	899	982
	BK33	Batang Kali	-	757	883	991
	BK34	Batang Kali	-	787	873	983
	BK35	Batang Kali	-	-	-	915
	BK37	Batang Kali	-	-	908	1023
	BK38	Batang Kali	-	-	915	882
C. latebrosa	BL17	Bukit Larut	-	-	922	-
	BL19	Bukit Larut	-	1122	939	-
	BL33	Bukit Larut	-	931	923	952
	PH01	Penang Hill	-	836	959	962
	PH02	Penang Hill	-	1054	952	1017
	PH07	Penang Hill	-	-	597	-
	PH08	Penang Hill	-	-	950	923
	PH26	Penang Hill	-	-	949	-

	MJ04	Mount Jerai	-	-	-	968
	FH14	Fraser's Hill	-	-	867	973
	LK01	Lake Kenyir	1252	903	950	998
	LK07	Lake Kenyir	-	958	530	952
	LK19	Lake Kenyir	1266	-	955	977
	BK04	Batang Kali	-	-	907	-
	BK05	Batang Kali	-	-	891	-
	BK07	Batang Kali	-	931	903	959
	BK08	Batang Kali	-	795	903	976
	BK09	Batang Kali	-	-	900	921
	BK12	Batang Kali	-	-	882	-
	BK13	Batang Kali	-	-	901	911
	BK14	Batang Kali	-	-	902	-
	BK15	Batang Kali	-	-	882	-
	BK17	Batang Kali	-	-	850	-
	BK18	Batang Kali	-	789	902	976
	BK19	Batang Kali	-	-	879	-
	BK21	Batang Kali	-	-	762	-
C. glabra	BL03	Bukit Larut	-	-	915	-
	BL05	Bukit Larut	-	-	895	-
	FH06	Fraser's Hill	1236	-	866	949
	FH26	Fraser's Hill	-	-	859	-
	FH42	Fraser's Hill	1201	-	-	-
	FH43	Fraser's Hill	1234	-	920	952
	FH56	Fraser's Hill	1263	-	-	847
	BF8a	Fraser's Hill	1227	-	955	1002
	BF8b	Fraser's Hill	1242	-	986	1048
C. obscura	BL15	Bukit Larut	1275	911	898	-
	BL26	Bukit Larut	-	703	879	953
	BL27	Bukit Larut	-	-	904	-
	BL28	Bukit Larut	-	-	898	-
	BL30	Bukit Larut	1255	-	975	990
	FH39	Fraser's Hill	1203	-	939	956
	FH52	Fraser's Hill	1251	-	-	902
C. trichodesma	BL07	Bukit Larut	1264	953	-	972
	BL32	Bukit Larut	-	-	880	942
	MA08	Mount Angsi	-	-	849	-
	MA11	Mount Angsi	-	-	900	-
	MA14	Mount Angsi	-	-	896	-
	MA17	Mount Angsi	1266	-	-	1021
	MA18	Mount Angsi	-	-	884	-
	MA19	Mount Angsi	-	-	914	-

	MA22	Mount Angsi	-	-	892	-
	MA25	Mount Angsi	1183	-	906	-
	MA30	Mount Angsi	-	-	903	-
	MA45	Mount Angsi	-	-	934	-
	MA48	Mount Angsi	-	-	8710	-
C. hymenodes	MP12	Mount Perlis	-	-	-	995
	MJ05	Mount Jerai	-	-	921	966
	MJ06	Mount Jerai	-	976	930	1024
	MJ07	Mount Jerai	-	-	892	-
	MJ08	Mount Jerai	1262	-	929	1029
C. incisoserrata	FH15	Fraser's Hill	1275	-	-	-
	FH16	Fraser's Hill	1180	-	908	-
	LK02	Lake Kenyir	1253	-	903	-
	LK10	Lake Kenyir	1238	615	922	-
	LK11	Lake Kenyir	-	886	954	-
C. assimilis	MB22	Mount Berinchang	1277	-	-	-
	BF1a	Fraser's Hill	1277	914	-	1089
	BF1c	Fraser's Hill	1180	757	-	1020
	BF2	Fraser's Hill	1223	-	983	1081
	BF6a	Fraser's Hill	1225	-	951	1077
C. polypoda	ML02	Mount Ledang	1206	811	869	915
	ML06	Mount Ledang	1229	-	-	940
	ML20	Mount Ledang	1123	-	-	-
	ML25	Mount Ledang	1195	-	-	-
C. lurida	MB07	Mount Berinchang	-	-	-	957
	MB11	Mount Berinchang	-	-	951	-
	MB35	Mount Berinchang	1252	-	930	-
	MB43	Mount Berinchang	1213	-	930-	957
C. alternans	MA16	Mount Angsi	1284	-	938	-
C. moluccana	HB1	Hutan Bangi	1206	-	978	1090
	HB2	Hutan Bangi	-	879	981	1231
C. gigantea	ML23	Mount Ledang	1201	-	-	-
	ML24	Mount Ledang	1264	-	-	909
	ML27	Mount Ledang	-	-	-	911
	ML28	Mount Ledang	1248	-	-	976
C. recommutata	ML07	Mount Ledang	1238	-	-	-
	ML08	Mount Ledang	-	-	-	-
	ML15	Mount Ledang	1242	-	-	-
	ML17	Mount Ledang	-	-	614	-
Cyathea sp.	GH15	Genting Highlands	-	1195	-	-
	GH16	Genting Highlands	-	1264	-	-
	MA02	Mount Angsi	-	-	899	-

	MA15	Mount Angsi	-	1234	894	-
	MA20	Mount Angsi	-	1340	-	-
	MA27	Mount Angsi	-	1230	-	-
	MA38	Mount Angsi	-	1252	-	-
	MA39	Mount Angsi	-	1225	880	-
	MA40	Mount Angsi	-	1183	-	-
	MA46	Mount Angsi	-	-	869	-
	ML03	Mount Ledang	-	1231	-	-
	ML04	Mount Ledang	-	1211	-	-
	ML16	Mount Ledang	-	1256	-	-