MOLECULAR DNA MARKERS IN PARENTAGE IDENTIFICATION AND CLONAL GENETIC STRUCTURE OF *Cryptocoryne* ×*purpurea* Ridl. nothovar. *purpurea* HYBRID POPULATIONS

ROSAZLINA BINTI RUSLY

UNIVERSITI SAINS MALAYSIA
2016

MOLECULAR DNA MARKERS IN PARENTAGE IDENTIFICATION AND CLONAL GENETIC STRUCTURE OF *Cryptocoryne* ×*purpurea* Ridl. nothovar. *purpurea* HYBRID POPULATIONS

by

ROSAZLINA BINTI RUSLY

Thesis submitted in fulfillment of the requirements
for the degree of
Doctor of Philosophy

July 2016

ACKNOWLEDGEMENTS

Alhamdulillah. First of all, I would like to mention how grateful I am to Allah S.W.T because with His entire blessing, I have finally completed my research dissertation. Undertaking this PhD has been a truly life-changing experience for me and it would not have been possible to do without the support and guidance that I received from many people. First and foremost, I would like to express my deepest appreciation to my supervisor, Professor Dr Ahmad Sofiman Bin Othman for his guidance, advices, comments and constructive suggestion throughout the period of my study. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. I also would like to address my deepest gratitude to my field supervisor, Associate Professor Dr Marian Ørgaard and Professor Dr Niels Jacobsen from University of Copenhagen for their valuable advice and friendly help. I would also like to grab this opportunity to express my utmost thanks to Professor Dr Amirul Al-Ashraf Abdullah who introduced me to this field and always giving me a good motivation, moral support and encouragement.

My deepest appreciation also goes to School of Biological Sciences' staff, En Bob, En Shukor, En Muthu and En Shanmugam for helping me during field sampling and preparation of herbarium specimens. I am indebted to them for their help. Also, thanks to Puan Sabariah who always helping me in the settlement of purchase order and provided lab apparatus. Completing this work would have been all the more difficult were it not for the support and friendship provided by the members of the Lab 409. I would like to thank all these people for their help and support throughout my study for the last couple of years; Fishah, Farah, Ila, Komala, Shafika, Shakina, Fasih, Jayaraj, Bad, Veera, Im Hin, Fahmi, Sughanti and also colleague friends from Lab 308; Nurul,

Mira, Daniel, Lim, Jam, Lia, Wani, Kak Adib, Fung and Kak Naz. Not forgetting my best friend Kak Fuzah, Ana, Fatimah, Zai and Fara for their care and joyousness they bring into my days.

To my family especially my father Hj Rusly Bin Yusoff and my mother Hjh Nina Bt Sahari, thank you very much for everything and I appreciate your love, endless pray and moral support. I owe my loving thanks to my family; Along, Angah, Kak Jia, Kak Tina and parent-in-law; Mak and Abah for their loving support. Not forgetting my adorable niece Adriana and Addin and my nephew Aariq those bringing colourful days in my life. To my beloved husband, Mohd Sallehudin Bin Mohd Hailani, thank you for your endless love, understanding and encouragements.

I also extend my gratitude to the Postgraduate Research Grant Scheme from Universiti Sains Malaysia (1001/PBIOLOGI/846007) for funding my research and MyBrain15 (Ministry of Higher Education, Malaysia) for providing scholarships for my financial support and university tuition fees.

Last but not least, a big thanks to everyone who has contributed, either directly or indirectly towards the success of this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF ABBREVIATIONS

AFLP	Amplified Fragment Length Polymorphism
AMOVA	Analysis of molecular variance
cpDNA	Chloroplast deoxyribonucleic acid
CTAB	hexadecyl-trimethylamonium bromide
DNA	Deoxyribonucleic acid
dNTP	Dinucleotide triphosphate
EDTA	Ethylenediamine tetra-acetic acid
EtBr	Ethidium bromide
FCA	Factorial correspondence analysis
HCl	Hydrochloric acid
HWE	Hardy-Weinberg Equilibrium
IPTG	Isopropyl *β-D*-1-thiogalactopyranoside
ITS	Internal Transcribe Spacer
LB	Luria-bertani
NCBI	National centre for biotechnology information
NJ	Neighbour joining
nrDNA	Nuclear ribosomal DNA
PCA	Principal component analysis
PCR	Polymerase chain reaction
RAPD	Random amplified polymorphic DNA
RNA	Ribonucleic acid
RNase	Ribonuclease enzyme
SSR	Simple sequence repeat
TFPGA	Tools for population genetic analysis

PENANDA DNA MOLEKUL DALAM PENGENALPASTIAN INDUK DAN STRUKTUR GENETIK KLONAL BAGI *Cryptocoryne* ×*purpurea* Ridl. nothovar. *purpurea* POPULASI HIBRID

ABSTRAK

Penghibridan semulajadi telah diyakini kerapkali berlaku pada *Cryptocoryne* Wydler dan dianggap sebagai sumber kepada kerumitan taksonomi pada genus ini. Penyelidikan ini melibatkan gabungan kajian daripada data jujukan DNA (kawasan tertranskripsi dalaman (ITS) DNA nuklear ribosom dan gen *mat*K daripada DNA kloroplas) untuk mengenalpasti induk kepada hibrid putatif *Cryptocoryne* daripada Semenanjung Malaysia. Berdasarkan kepada ciri-ciri morfologi tumbuhan ini dikenalpasti secara tentatif sebagai *Cryptocoryne* ×*purpurea* Ridl. nothovar. *purpurea*; tumbuhan steril yang telah lama dianggap sebagai hibrid, kemungkinan daripada dua spesies yang berkaitan; *Cryptocoryne cordata* Griff. var. *cordata* dan *Cryptocoryne griffithii* Schott. Status hibrid dan induk-induk ini dibuktikan secara bebas dengan kehadiran pada individu hibrid corak jujukan ITS daripada kedua-dua spesies induk ini. Tumbuhan hibrid ini berkongsi persamaan jujukan *mat*K daripada *C. cordata* var. *cordata* dan *C. griffithii*, menunjukkan kedua-dua spesies induk putatif ini telah menjadi induk betina. Penghibridan timbal balik di antara kedua-dua spesies ini dilihat sebagai simetri dan bukan satu arah. Kajian ini juga bertujuan untuk membangunkan penanda mikrosatelit menggunakan jujukan generasi (Roche 454 pyrosequencing) daripada DNA genomik *C. cordata* var. *cordata*. Sebelas lokus polimorfik baru telah berjaya dipencilkan dan kesemua lokus menyimpang daripada keseimbangan Hardy-Weinberg secara signifikan. Tiada alel nol dan tiada ketidakseimbangan untaian yang signifikan dikesan ke atas semua pasangan lokus. Amplifikasi silang spesies berbeza telah berjaya pada satu panel sebelas spesies *Cryptocoryne*. Kesamaan saiz alel yang tinggi di antara

C. ×*purpurea* nothovar. *purpurea*, *C. cordata* var. *cordata* dan *C. griffithii* telah menyokong idea bahawa *C. cordata* var. *cordata* dan *C. griffithii* merupakan induk kepada *C.* ×*purpurea* nothovar. *purpurea*. Kajian ini telah menyiasat enam populasi semulajadi *C.* ×*purpurea* nothovar. *purpurea* untuk memeriksa kepelbagaian klonal dan hubungan struktur genetik di antara populasi menggunakan analisis polimorfisme kepanjangan fragmen teramplifikasi (AFLP). Tahap kepelbagaian genetik klonal pada *C.* ×*purpurea* nothovar. *purpurea* adalah rendah kerana hadirnya kesterilan yang tinggi di dalam populasi disebabkan oleh asal usul hibrid. Walau bagaimanapun, kehadiran genotip yang berbeza pada sesetengah populasi memberi bukti terhadap kekerapan peristiwa pembentukan hibrid daripada populasi induk yang berbeza dan juga mutasi somatik. Analisis kluster mendedahkan dua kumpulan yang berbeza dengan majoriti variasi genetik tersebar di antara- berbanding di dalam populasi di antara wilayah dan menunjukkan kolerasi di antara jarak genetik dengan jarak geografi. Penemuan ini menunjukkan takson steril klonal ini boleh memelihara sejumlah variasi genetik.

MOLECULAR DNA MARKERS IN PARENTAGE IDENTIFICATION AND CLONAL GENETIC STRUCTURE OF *Cryptocoryne* ×*purpurea* Ridl. nothovar. *purpurea* HYBRID POPULATIONS

ABSTRACT

Natural hybridization has been confirmed to occur frequently in *Cryptocoryne* Wydler and considered a source of taxonomic complexity in this genus. This research involved a combined study of DNA sequencing data (internal transcribed spacer (ITS) of nuclear ribosomal DNA and *mat*K gene of chloroplast DNA) to identify the parentage of a putative *Cryptocoryne* hybrid from Peninsular Malaysia. Based on the morphological characters the plant was tentatively identified as *Cryptocoryne* ×*purpurea* Ridl. nothovar. *purpurea*; a sterile plant which has long been considered a hybrid, possibly from two related species; *Cryptocoryne cordata* Griff. var. *cordata* and *Cryptocoryne griffithii* Schott. The hybrid status and its putative parents was independently confirmed by the presence in hybrid individuals of an additive ITS sequence pattern from these two parental species. The hybrid plants shared the identical *mat*K sequences from *C. cordata* var. *cordata* and *C. griffithii*, which indicated that both putative parental species had functioned as the maternal parent. Reciprocal hybridization between the two species seems to be symmetrical rather than unidirectional. This study also aimed at developing microsatellite markers using next generation sequencing (Roche 454 pyrosequencing) from the genomic DNA of *C. cordata* var. *cordata*. Eleven new polymorphic loci were successfully isolated and all loci departed significantly from Hardy-Weinberg Equilibrium. No null alleles and no significant linkage disequilibrium were detected across any pairs of loci. Cross species amplification was successful across a panel of eleven *Cryptocoryne* species. The high similarities of allele sizes between *C.* ×*purpurea* nothovar. *purpurea, C. cordata* var. *cordata* and *C. griffithii* supported the idea that *C.*

cordata var. *cordata* and *C. griffithii* were the parents of *C.* ×*purpurea* nothovar. *purpurea*. This study investigated six natural populations of *C.* ×*purpurea* nothovar. *purpurea* to examine the clonal diversity and spatial genetic structure among populations using Amplified Fragment Length Polymorphism (AFLP) analysis. The level of clonal genetic diversity in *C.* ×*purpurea* nothovar. *purpurea* was low because of the apparent high sterility of the populations due to their hybrid origin. However, the occurrence of different genotypes in certain populations give an evidence of the frequency of hybrid formation events from different parental populations and also somatic mutations. Cluster analyses revealed two distinct groups with the majority of genetic variation distributed among- rather than within populations between regions and showed correlation between genetic distances with geographical distance. These findings demonstrate that this sterile clonal taxon can preserve substantial amounts of genetic variation.

CHAPTER ONE

GENERAL INTRODUCTION

Cryptocoryne Wydler is an aquatic plant genus belonging to the family Araceae commonly known as the Water Trumpet which refers to the spathe which is connate along its margin forming a water tight tube resembling a trumpet. *Cryptocoryne* is popular as ornamental plants for tropical aquaria and aquascaping in Europe since the 1950s (Jacobsen, 1976) due to the unique leaves and the flowers of various species that come in different attractive colours. The genus is native to South East Asia extending from Mainland India and Indo-China through Indonesia to Papua New Guinea. *Cryptocoryne* can be viewed as consisting of numerous populations in different river systems, and natural hybridization has been suggested to frequently occur (Jacobsen et al., 2002; Ipor et al., 2005; Othman et al., 2009; Ipor et al., 2015; Jacobsen et al., 2016). These events therefore would be a driving evolutionary force continuously producing new genotypes to be dispersed all over the ever changing river systems. To date, more than 25 *Cryptocoryne* natural hybrids have been discovered (Jacobsen et al., 2015).

This thesis will only concentrate on one natural hybrid which can be found in Peninsular Malaysia namely *Cryptocoryne* ×*purpurea* Ridl. nothovar. *purpurea*. The early identification of this hybrid was based on pollen fertility and morphological character analysis. Jacobsen (1977) observed that the pollen of *C.* ×*purpurea* nothovar. *purpurea* is completely sterile and has been suggested to have *C. cordata* Griff. var. *cordata* and *C. griffithii* Schott as parents owing to observable morphological characters (broad collar zone – *C. cordata,* and purple, rough limb of spathe – *C. griffithii*). Cytological analysis indicated that *C.* ×*purpurea* nothovar. *purpurea* shared the same

diploid chromosome numbers 2n = 34 with *C. cordata* var. *cordata* and *C. griffithii* (Jacobsen, 1977). The putative hybrid always shows intermediate morphological features of their parents. However, this character coherence is not always a reliable indicator of hybrid identity (Rieseberg and Ellstrand, 1993; Rieseberg et al., 1999) because morphological features are often under the influence of environmental conditions and thus can be unreliable and prone to misleading interpretation (Hegarty and Hiscock, 2005).

In recent years, use of molecular markers has been proven to be a good method in parentage identification and can provide considerable insight into plant hybridization (López-Caamal and Tovar-Sánchez, 2014). To date no DNA sequence and characterized molecular markers have been validated for hybrid identification in *Cryptocoryne*. Therefore, the first objective in this study is to determine the origin of *C.* ×*purpurea* nothovar. *purpurea* using a combination of DNA sequences namely from the internal transcribed spacer (ITS) of nuclear ribosomal DNA region and *mat*K gene of chloroplast DNA region. Next, microsatellite DNA markers were developed through Next generation Sequencing (NGS) origin and validate the markers through cross species amplification and then utilised to verify the hybrid origin.

Cryptocoryne ×*purpurea* nothovar. *purpurea* propagates only via vegetative propagation through rhizomes known as clonal reproduction. Clonal plants resulting from the replication of an individual by vegetative growth resulting in genetically identical individuals. Therefore all clonally reproducing organisms should have low amounts of genetic diversity. However, most studies have shown the opposite trends where sterile clonal plants tend to show higher genetic diversity than sexually reproducing plants (Ally et al., 2008; Gross et al., 2012; Bobiwash et al., 2013).

Therefore, another objective in this study is to define the possible clonal structure and genotypic diversity among *C.* ×*purpurea* nothovar. *purpurea* clones. The AFLP marker (Amplified Fragment Length Polymorphism) was used to achieve this objective since this marker can produce many polymorphic loci at a lower cost which is ideal for genet discrimination. The information generated in this study is expected give valuable genetic information for understanding the extent of hybridization in *Cryptocoryne*.

Thus the specific objectives in this study are:-

1. To verify the hybrid origin of *C.* ×*purpurea* nothovar. *purpurea* using a biparental inherited nuclear marker namely the internal transcribed spacer (ITS) region of nuclear ribosomal DNA (nrDNA).
2. To identify the maternal parent of *C.* ×*purpurea* nothovar. *purpurea* using the chloroplast *mat*K gene.
3. To develop polymorphic microsatellite DNA marker through Next Generation Sequencing (NGS) and utilize the markers for *C.* ×*purpurea* nothovar. *purpurea* parentage identification.
4. To define the possible clonal structure and genetic diversity in *C.* ×*purpurea* nothovar. *purpurea* clones using AFLP (Amplified Fragment Length Polymorphism) marker.

CHAPTER TWO

LITERATURE REVIEW

2.1 The Genus *Cryptocoryne*

Cryptocoryne Wydler is an aquatic plant genus, belonging to the aroid family (Araceae). *Cryptocoryne* are amphibious herbs and have proliferously dividing subterranean rhizomes, thereby enabling them to form large stands in streams and rivers (Jacobsen et al., 2015). The first *Cryptocoryne* species was described in 1779 as Arum spirale by Retzius. The genus was described and established by Wydler in 1830 (Othman et al., 2009). *Lagenandra* is another genus closely related to the genus *Cryptocoryne* (Cusimano et al., 2011). These two genus can be easily differentiated since the leaves of *Cryptocoryne* exhibit convolute vernation while *Lagenandra* exhibit involute vernation. The genus name *Cryptocoryne* is derived from *crypto* (Latin), meaning hidden, and *koryne* (Greek), meaning club or the spadix that is totally hidden inside the kettle (Othman et al., 2009).

The 'Water Trumpet' is another popular name of *Cryptocoryne* which refers to the spathe is connate along its margin forming a water tight tube resembling a trumpet. This genus is locally known as *Hati-hati Paya* or *Hati-hati Air* or *Keladi Paya* (Peninsular Malaysia), *Kiambang Batu* (Malays-Sarawak), *Kelatai* (Iban), and *Tropong Ajer* (Banjarmasin, Kalimantan) (Fung, 2008; Ipor et al., 2010). The genus is native to tropical regions of Asia extending from India in the west to the Philippines in the east, onwards to Malaysia, through Indonesia to Papua New Guinea (Othman et al., 2009). *Cryptocoryne* is widely used as an aquarium plant since 1910s (Jacobsen, 1982) due to appearance of their attractive colour and shapes of features including leaves, spathes and

limb of the spathes (Othman et al., 2009). They are also heavily exploited for aquarium plants and apparently fetch high prices in the international aquarium market (Mansor, 1991; Othman et al., 2009).

2.1.1 Natural hybridization in the Genus *Cryptocoryne*

To date, more than 25% of the about 91 named and unnamed *Cryptocoryne* have proven to be of hybrid origin (54 species, an additional 12 varieties and 25 natural hybrids) (Jacobsen et al., 2015). Natural hybridization has been considered to represent an important factor influencing the high diversity of the genus *Cryptocoryne*. This genus may frequently be observed in co-existence with two or more close related species inhabiting the same or adjacent streams and these phenomena may lead to hybridization and producing new hybrids within the same area (Jacobsen et al., 2002; Ipor et al., 2005; Othman et al., 2009; Ipor et al., 2015). The pollinating flies are a major factor in exchanging genes stochastically within operation distance of the flies (Jacobsen et al., 2015).

Most recently, the genus is made up of 57 species, 17 varieties and 7 named hybrids with several unnamed hybrids (Table 2.1) (Jacobsen et al., 2016). The uncertain status and tendency of *Cryptocoryne* to hybridise naturally may create more complexity in terms of taxonomic studies and classification. Recently a new hybrid was describe from Sarawak viz. *Cryptocoryne* ×*batangkayanensis* 2n = 85, postulated to be a hybrid between *C. cordata* Griff. var. *grabowskii* (Engl.) N. Jacobsen and *C. ferruginea* Engl. var. *ferruginea* (Ipor et al., 2015). Earlier, Ipor et al. (2008) assigned this hybrid with uncertain status under *C.* ×*purpurea* nothovar. *borneoensis* N. Jacobsen et al. Another hybrid, *C.* ×*timahensis* Bastmeijer (2n = 34) with completely sterile pollen can be found in Bukit Timah, Singapore and perhaps in the southern region of Malay Peninsula. It

Table 2.1 The list of known natural hybrids in *Cryptocoryne*

No	Name	Putative parent	Chromosome no	Origin	References
1	*Cryptocoryne* ×*purpurea* Ridl. nothovar. *purpurea*	*C. cordata* Griff. var. *cordata* × *C. griffithii* Schott	2n = 34	Peninsular Malaysia	(Jacobsen, 1982)
2	*C.* ×*decus-silvae* De Wit (*incl. C. jacobsenii* De Wit)	*C. cordata* var. *cordata* × *C. nurii* var. *nurii*	unconfirmed	Peninsular Malaysia	(Jacobsen et al., unpublished)
3	*C.* ×*zukalii* Rataj	*C. cordata* var. *cordata* × *C. minima* Ridl.	2n = 34	Peninsular Malaysia	(Jacobsen et al., unpublished)
4	*C.* ×*purpurea* nothovar. *borneoensis* N. Jacobsen et al.	*C. cordata* var. *grabowskii* N. Jacobsen (as *C. zonata* De Wit) × *C. griffithii*	2n = 51	Kalimantan, Indonesia	(Jacobsen et al., 2002)
5	*C.* ×*batangkayanensis* Ipor et al.	*C. cordata* var. *grabowskii* × *C. ferruginea* Engl. var. *ferruginea*	2n = 85	Sarawak, Malaysia	(Ipor et al., 2015)
6	*C.* ×*timahensis* Bastm.	*C. cordata* var. *cordata* × *C. nurii* Furt. var. *nurii* (today assumed to be a *C. nurii* Furt. var. *nurii* × *C. schulzei* De Wit)	2n = 34 (originally reported as 2n = 54)	Bukit Timah, Singapore	(Bastmeijer and Kiew, 2001)
7	*C.* ×*willisii* Reitz	*C. beckettii Trim/C. walkeri Schott* × *C. parva De Wit*	2n = 28	Sri Lanka	(Jacobsen, 1981;1987)
8	*C. beckettii hybrid - unnamed*	*C. beckettii* × *C. walkeri* (as *C. lutea* Alston)	2n = 28	Sri Lanka	(Jacobsen, 1981;1987)
9	*C. crispatula* Engl. hybrid - unnamed	*C. crispatula* var. *crispatula* × var. *balansae* (Gagnep,) N. Jacobsen	2n = 36	Phu Khieo, Thailand	(Jacobsen, 1980)
10	*C. ferruginea* var. *sekadauensis* Bastm. et al. hybrid - unnamed	*C. ferruginea* var. *sekadauensis* × *C. fusca* De Wit	2n = 34	Kalimantan, Indonesia	(Bastmeijer et al., 2013)
11	*C. crispatula* hybrids - unnamed	*C. crispatula* var. *crispatula* × other varieties	2n = 36	Cheng Khan, Thailand	(Idei, unpublished in Ipor et al., 2015)
12	*C. crispatula* hybrids - unnamed	*C. crispatula* var. *crispatula* × other varieties	2n = 36	Don Khon, Lao P. D. R. Thailand	(Idei, unpublished in Ipor et al., 2015; Jacobsen et al., 2016)
13	*C. crispatula* hybrids - unnamed	*C. crispatula* var. *crispatula* × *C. mekongensis* Idei et al.	2n = 36	Don Khon, Lao P. D. R. Thailand	(Idei unpublished in Ipor et al., 2015; Jacobsen et al., 2016)

was assumed that *C. cordata* Griff. var. *cordata* and *C. nurii* Furt. var. *nurii* are the putative parents (Bastmeijer and Kiew, 2001). However, Othman et al. (2009) later suggested to be *C. nurii* var. *nurii* and *C. schulzei* De Wit as the possible parents.

One of the natural *Cryptocoryne* hybrids that can be found in Peninsular Malaysia is *C.* ×*purpurea* Ridl. nothovar. *purpurea*. From early description by Ridley, much confusion arose because of wrong interpretation about *C.* ×*purpurea* nothovar. *purpurea* identity. In 1892, H.N. Ridley was the first to collect this plant at Kota Tinggi (Johor) (Othman et al., 2007; 2009). This plant was then cultivated at the Botanical Garden in Singapore and sent live to Kew Gardens in 1898 (Othman et al., 2009). Later on, it flowered in 1899 and plate, no. 7719, was published in Hooker's Icones Plantarum in 1900 under the name *C. griffithii* (Plate 2.1) (Bastmeijer, 2008). This was corrected shortly afterwards by Ridley (1904) who pointed out that the plate no. 7719 was actually a new species, namely *C.* ×*purpurea* nothovar. *purpurea*. Engler (1920) also mentioned this inconsistency. However, the name accompanying plate 7719 continued to mislead people and was thus up to the 1960s and 1970s incorrectly attached to *C.* ×*purpurea* nothovar. *purpurea* in cultivation in Europe (Othman et al., 2009).

The identification of this hybrid was based on pollen analysis by Jacobsen (1982) who found that the pollen of *C.* ×*purpurea* nothovar. *purpurea* is completely sterile and suggested *C. cordata* var. *cordata* and *C. griffithii* Schott as the parents owing to observable morphological characters (broad collar zone – *C. cordata* var. *cordata,* and purple, rough limb of spathe – *C. griffithii*). De Wit (1990) gave a comprehensive explanation of the differences between *C. griffithii*, *C. cordata* var. *cordata* and *C.* ×*purpurea* nothovar. *purpurea*. Cytological analysis indicated that *C.* ×*purpurea* nothovar. *purpurea* shared the same diploid chromosome numbers 2n = 34 with

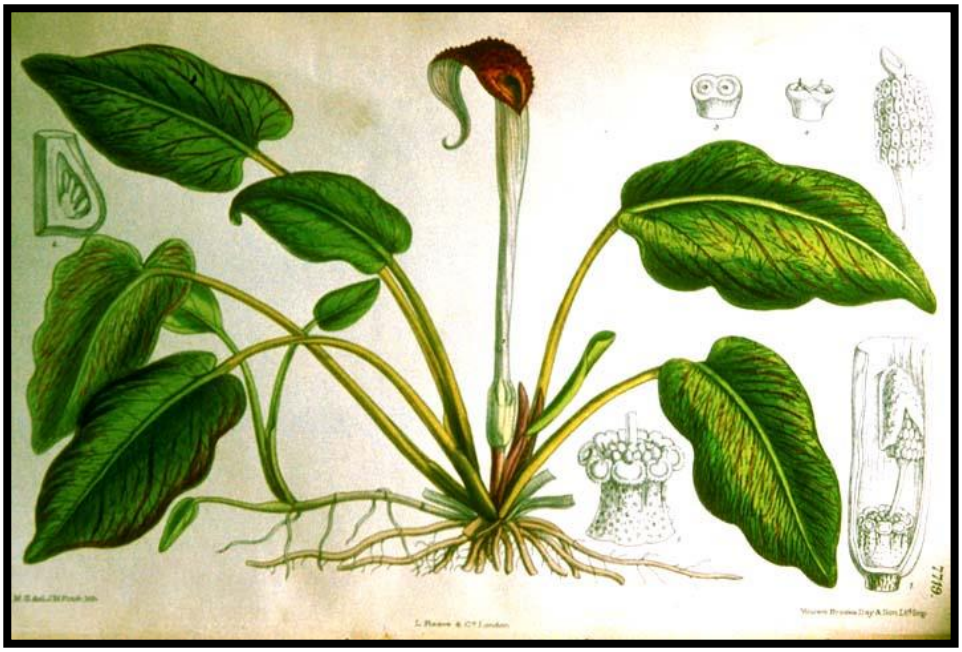

Plate 2.1 The images of *Cryptocoryne* ×*purpurea* Ridl. nothovar. *purpurea* was published in Hooker's Icones Plantarum in 1900 under the name *C. griffithii* (Hooker, 1900). (Image adapted from Bastmeijer, 2008).

C. cordata var. *cordata* and *C. griffithii* (Jacobsen, 1977; 1982). Moreover, *C. nurii* var. *nurii* and *C. schulzei* also became the suspected parents to *C.* ×*purpurea* nothovar. *purpurea* based on certain similarities in morphology characters.

2.1.2 Morphological Characteristics and Habitat

The shape of the leaves and the shape and colours of the limb of the spathe are important diagnostic taxonomic characters in *Cryptocoryne* (Bastmeijer, 2015). The *C.* ×*purpurea* nothovar. *purpurea* leaf blades are ovate with cuneate to cordate base (Plate 2.2). The upper surface of the leaves is dark green to brownish and purplish mottled. The lower surface is often pale green, purplish mottled; upper and/or lower surface sometimes with a silvery luster. However, the morphological variation of the characters of *Cryptocoryne* leaves may to a large extent be due to the environment, especially to submerged and emergent habitat and also depending on the amount of light received (Othman et al., 2009). The *C.* ×*purpurea* nothovar. *purpurea* limb present at the upper most part of the spathe is ovate-acuminate, rugose, dull to bright red colour, absent collar, broad collar zone, red to reddish or more whitish to yellowish towards the opening (Othman et al., 2009). The lower parts of the spathe are tubular with the edges joined forming the kettle. A kettle contains the male and female flowers. Normally, a single whorl of female flowers of the *C.* ×*purpurea* nothovar. *purpurea* has 5 to 7 carpels with broadly rounded stigma and emarginated at the base of the spadix (Othman et al., 2009). In the middle of these female flowers, there is a single whorl of abortive, modified flowers known as the olfactory bodies. There is a long sterile zone above of the olfactory bodies, which is topped by a cluster of male flowers (Othman et al., 2009).

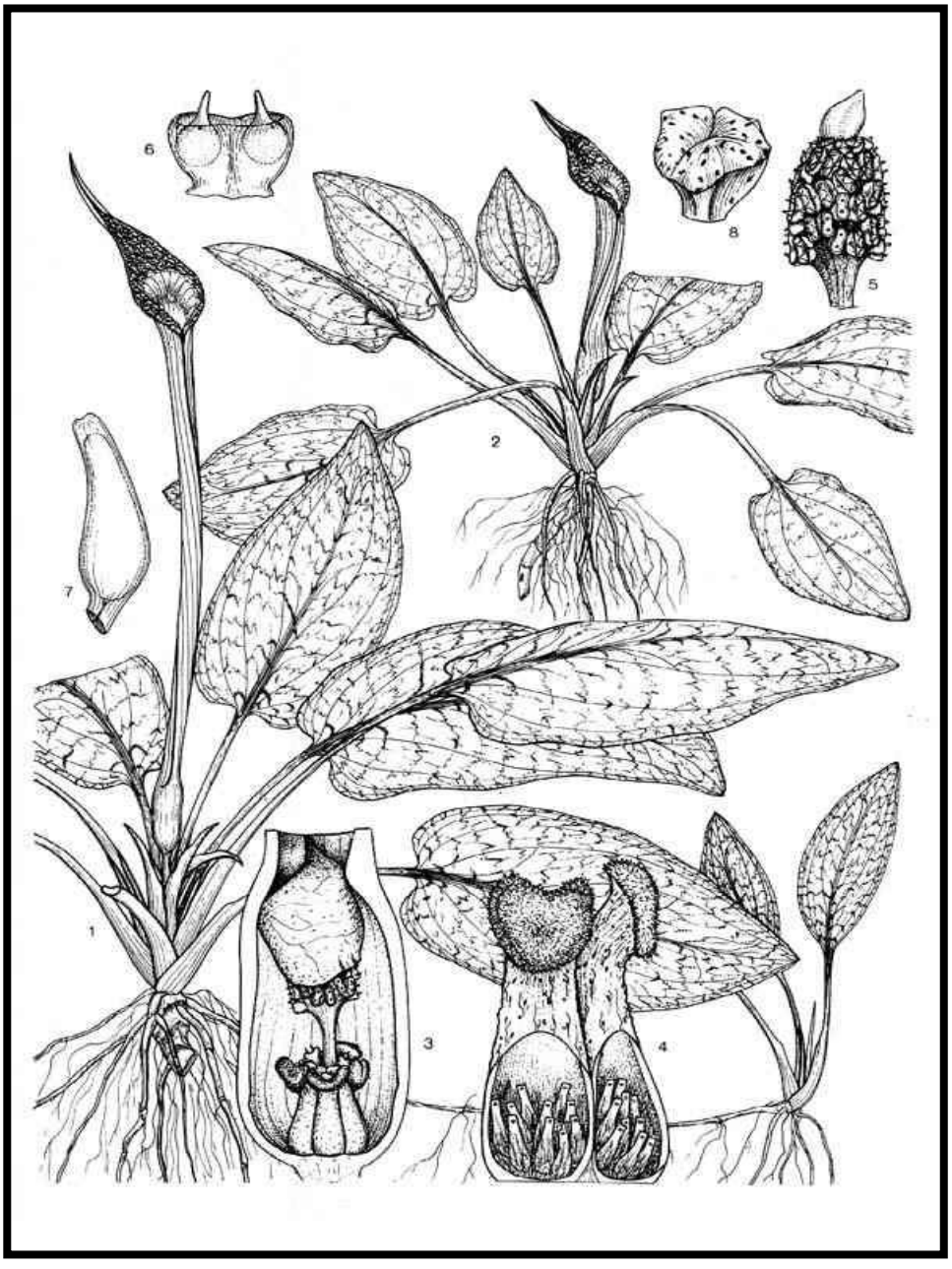

Plate 2.2 The morphological characteristics of *Cryptocoryne* ×*purpurea* nothovar. *purpurea*: 1 and 2. The whole plant with rhizomes; 3. The open kettle; 4. Female flower with olfactory bodies; 5, 6, 7, and 8. Male flower (De Wit, 1990).

Currently *C.* ×*purpurea* nothovar. *purpurea* have been found at eight documented locations in Peninsular Malaysia, three of those locations within Tasik Bera i.e., Pos Iskandar, Kg. Jelawat and Paya Kelantong (in the state of Pahang), two locations in the state of Johor (Kg. Sri Lukut, Sg. Sedili Kechil) and three locations in Melaka namely Kg. Pulau Semut, Padang Tembak and Sungai Udang. Other localities in southern Peninsular Malaysia no doubt also exist, but have not seen them. Interestingly, there are some variations in the colouration and also in the surface structure of the limb of the spathe in *C.* ×*purpurea* nothovar. *purpurea* from different localities. For example, the colour of the limb of the spathe of *C.* ×*purpurea* nothovar. *purpurea* found in Melaka is dull brownish yellow and this is slightly different to the hybrid found in Johor which have brighter red and *C.* ×*purpurea* nothovar. *purpurea* found in Pahang is dark red. Plate 2.3 shows the differences among the limbs of the spathe colouration in different locations. These differences may indicate that the hybrid has arisen several times independently from different parental populations and maybe from segregation in the F1 offspring. When comparing with the putative parents namely *C. cordata* var. *cordata*, *C. griffithii*, *C. nurii* var. *nurii* and *C. schulzei* (Plate 2.3) (Table 2.2), there are several similarities and differences in morphological characters in *C.* ×*purpurea* nothovar. *purpurea*.

The hybrid grows in different habitat types. The vast submerged stands of *C.* ×*purpurea* nothovar. *purpurea* were found in larger black water swamp in Tasik Bera, Pahang. The substratum of the black water swamp in the Tasik Bera is very acidic (pH= 4.2-5.2), consisting of decomposed leaves and branches from the swamp forest and sometimes grow together with *Barclaya motleyi* Hooker f. In Sg. Sedili Kechil, Johor, the plants were found in the freshwater tidal zone, while in Kg. Sri Lukut, Johor, the hybrid plants grow on muddy bottom in small forests streams. In Melaka, the hybrid

Plate 2.3 Upper; The differences among the limbs of the spathe of *Cryptocoryne* ×*purpurea* Ridl. nothovar. *purpurea* colouration in different locations. A: Padang Tembak, Melaka, B: Kg. Sri Lukut, Johor, C: Pos Iskandar, Pahang. Lower; The images of the putative parental species of the hybrid *C.* ×*purpurea* nothovar. *purpurea*. D: *C. cordata* Griff. var. *cordata*, E: *C. griffithii* Schott, F: *C. nurii* Furtado var. *nurii* and G: *C. schulzei* De Wit. Image A, B, C, E and G: Rosazlina Rusly. Image D and F: Niels Jacobsen.

Table 2.2 The summary of morphological characteristics of *C.* ×*purpurea* nothovar. *purpurea* with the putative parents (Othman et al., 2009)

Species	Morphological characteristics		
	Leaves	Spathe tube	Limb
C. ×*purpurea* Ridl. nothovar. *purpurea*	Blade ovate to cordate base. 3-11cm long. Upper surface dark green to brownish, purplish mottled. Lower surface often pale green.	8-17 cm long; whitish on the outside and inside.	2-5 cm long, ovate-acuminate, rugose; dull to bright red; collar absent; collar zone broad; red to reddish.
C. cordata Griff. var. *cordata*	Blade narrowly ovate to cordate; sometimes up to 20 cm long. Smooth upper surface with green or green-brownish or brownish markings. Lower surface paler with reddish veins.	10-30 cm long; whitish on the outside; sometimes brownish greenish towards the apex.	3-5 cm long; ovate with a shorter or longer point; yellow; collar zone broad.
C. griffithii Schott	Blade ovate to rounded; 3-9 cm long. Upper surface mostly purple green; lower surface paler or more clearly reddish.	5-10 cm long; upper part purplish on the outside; lower part whitish.	3-5 cm long; red to black purple; ovate; vertical to reflexed with short point, surface rough with rounded protuberances; collar prominent.
C. nurii Furtado var. *nurii*	Blade stiff, ovate to narrowly ovate to elliptic; dark olive green, distinctly darker, red lines, lighter mottled. Lower surface pale green.	5-20 cm long; the upper part brownish tinged on the outside.	3-5 cm long; cordate; usually deep red to dark purple, with conspicuous, large, irregular, branched protuberances; collar narrow.
C. schulzei De Wit	Blade lanceolate to ovate to obovate with a cuneate to cordate base. Leaves brownish to purplish; upper surface striped with prominent, purplish markings, lower surface paler with reddish.	4-12 cm long; whitish on the outside, sometimes purplish-brownish shaded, upper part greenish.	1-2 cm long, recurved; narrow to a tail somewhat irregular rugose, yellow, vertical opening; collar broad somewhat folded, black red-purplish.

grows in the muddy swamp area at small forests streams and there are also in stands in the pond exposed to the sun.

2.2 What is Plant Hybridization?

Hybridization is the process of interbreeding between individuals of different species (interspecific hybridization) or genetically divergent individuals from the same species (intraspecific hybridization) (Rieseberg and Wendel, 1993). The broader definition considers hybridization as the cross fertilization of individuals from populations that are distinguishable on the basis of one or more heritable characters (Harrison, 1990; Arnold, 1997). Hybridization may cause interactions involving a wide range of types and levels of genetic divergence between the parental forms (Abbott et al., 2013). Offspring produced by hybridization may be fertile, partially fertile, or sterile (Siegel, 2014). The hybridization may result in the duplication of a hybrid's chromosome complement (allopolyploid) or without a change in chromosome number by the stabilisation of a fertile hybrid segregant (homoploid hybrid) (Rieseberg, 1997; Soltis et al., 2010; Abbott et al., 2013).

Polyploidy is of major significance in plant evolution with the latest estimates indicating that all extant flowering plants have polyploidy in their ancestry (Wood et al., 2009; Jiao et al., 2011). Two types of polyploids are normally recognized: autopolyploids in which chromosome sets are derived from the same species and allopolyploids that contain chromosome sets from different species as a consequence of interspecific hybridization. Allopolyploidy is considered to be more common in nature than autopolyploidy (Soltis et al., 2007). Additionally, after polyploidy has occurred, species tend to become reduced in their chromosome number and become homoploid

over evolutionary time (Wisseman, 2007).

2.3 Effect of Plant Hybridization

Natural hybridization is a frequent evolutionary phenomenon in flowering plants (Rieseberg and Wendel, 1993; Whitney et al., 2010). Hybridization in plants has been found to be most common in species which have certain specific life-history characteristics, including perennial habit, outcrossing breeding systems and asexual reproduction (Wisseman, 2007). Hybridization plays an important evolutionary role since it may lead to a number of consequences that may affect either positively (formation of new species, increase of the intraspecific genetic diversity of the participating populations) as well as negatively (species extinction through genetic assimilation, increase generation of highly invasive genotypes).

The hybridization events depend on the genetic structure of the participating species, the environmental conditions (i.e., degree of disturbance) and the local abundance of the parental species (Levin and Francisco-Ortega, 1996; Arnold, 2006). The studies of hybrid zones are important in order to clarify the steps in speciation that yield to complete reproductive isolation between taxa (i.e., reinforcement). The first generation hybrids (F1) exhibit low pollen fertility and it has been proposed as a mechanism of reinforcement of the reproductive barriers between the participating species due to selection against hybrid genotypes (Marshall et al., 2002; Campbell et al., 2003). The fitness of hybrid individuals appears to be dependent on the environment - high degree of disturbance such as crops, floods, along roadsides and volcanic activity (Levin and Francisco-Ortega 1996; Lamont et al., 2003; Tucker and Behm, 2011). Although reinforcement is an important consequence of natural hybridization, it is not the only one; introgression and genetic assimilation may also occur. Introgression is the

movement of genes between species; once F1 individuals are formed, they may act as a bridge whereby alleles may cross from one species to another through repeated backcrossing with genetically distinguishable populations (Rieseberg and Carney, 1998). If the frequency of the parental species is similar, introgression may lead to an increase of the intraspecific genetic diversity of the parental species, which may enable them to colonize new areas (Caraway et al., 2001). However, when the frequency of the parental species differ, the introgression towards the less abundant species may lead to the loss of its genetic integrity, leading to its extinction through the process known as 'genetic assimilation' (Levin and Francisco-Ortega, 1996; Meyerson et al., 2010) and result in the formation of invasive species (Petit et al., 2004; Schierenbeck and Ellstrand, 2009).

2.4 Plant Hybrid Identification

Because of the importance of plant hybridization effects on the taxonomic and genetics, it is of great importance to make a correct identification of hybrid individuals. Some of the tools employed for hybrid recognition and their pattern of expression in hybrid individuals were morphological characters and secondary metabolite expression as well as chromosome number and DNA fingerprinting techniques. While morphological characters were thoroughly employed during the last century as the main marker for hybrid recognition, nowadays it is known that their pattern of inheritance is considered complex and usually unpredictable (Rieseberg et al., 1999; Hardig et al., 2000; Ritz and Wissemann, 2003). Although many hybrids have intermediate morphological features between their parents, character coherence is not always a reliable indicator of hybrid identity. It is because the morphological expression in hybrids is highly dependent on the environment (Kiær et al., 2007; Hegarthy et al., 2008). Also, morphological intermediacy may originate by processes other than

hybridization such as certain species retain plesiomorphic character states of their ancestral population, conducing to an erroneous interpretation of hybridization (Rieseberg, 1995; Judd et al., 2002; Arnold, 2006).

Plant secondary metabolites have a more reliable inheritance mechanism than morphological characters (Rieseberg and Ellstrand, 1993; Orians, 2000; Cheng et al., 2011). However, obtaining the chemical profile of hybrids is time consuming, expensive and technically difficult. Their low polymorphism and complex inheritance make them also unreliable tools for hybrid recognition in the absence of other markers. Also, it is a poor predictor of hybrid ancestry in later generation hybrids (Cheng et al., 2011).

Both chemical and morphological markers are phenotypic traits, their expression in hybrids is highly dependent on the environment, reducing their utility to detect hybridization under natural conditions (Mallet, 2005). The chromosome number of putative hybrids may provide information about the hybrid origin of individuals when these exhibit allopolyploidy (Lawton-Rauh, 2003; Strong and Ayres, 2013). However, sometimes hybrids exhibit a homoploid condition compared to its parental species (Gross and Rieseberg, 2005; Mallet, 2007; Abbott et al., 2010). Due to the complex pattern of expression of phenotypic data and the unreliable data provided by chromosome number counts in putative hybrids, DNA markers appear as a much better option for hybrid recognition due to their high availability in the genome, selectively neutral, strictly under Mendelian segregation ratios and the ease with which large amounts of data may be obtained (Rieseberg and Wendel, 1993; Travis et al., 2010; López-Caamal and Tovar- Sánchez, 2014).

2.4.1 Nuclear Ribosomal ITS Region

The plant nuclear genome (nDNA) is the largest genome in the plant cell. Plant nuclear genome size is constant in a species and can vary from 60 Mbp to 150 000 Mbp, a remarkable difference of 2300 times (Bennett and Leitch, 2011). The large genome size variation is because of multiplication of parts of, or complete, nuclear genomes (Heslop-Harrison and Schmidt, 2012; Schranz et al., 2012). The nuclear DNA contains coding and large number of regulatory sequences for genes and repetitive DNA (Kellogg and Bennetzen, 2004). Of all regions within nuclear DNA, the nuclear ribosomal DNA (nrDNA) is most widely used to infer plant phylogeny (Álvarez and Wendel, 2003).

The nrDNA in plants comprises three coding regions (18S, 5.8S and 26S regions), which are separated by two transcriptional regions - internal transcribed spacers (ITS1 and ITS2) (White et al., 1990) (Figure 2.1). The ITS region is phylogenetically informative at low taxonomic levels and is now extensively employed worldwide (Poczai and Hyvönen, 2010; Tripathi et al., 2013). ITS1 and ITS2 regions are inherently rich in G+C content and these core parts are evolutionary conserved within green plants (Hershkovitz and Lewis, 1996; Hershkovitz and Zimmer, 1996). In addition, the ITS2 region is a favourite marker in taxonomy because of the fast-evolving segment of the nuclear rRNA operon (Coleman and Mai, 1997; Joseph et al., 1999, Coleman, 2007) and 40% of ITS2 is found to be con served across all angiosperms studied (Hershkovitz and Zimmer, 1996). The nuclear ribosomal internal transcribed spacer regions (nrITS) is part of the ribosomal multigene family that includes hundreds to thousands of copies at one or more chromosomal loci and often used to obtain phylogenetic information due to the level of variation both within and among genera.

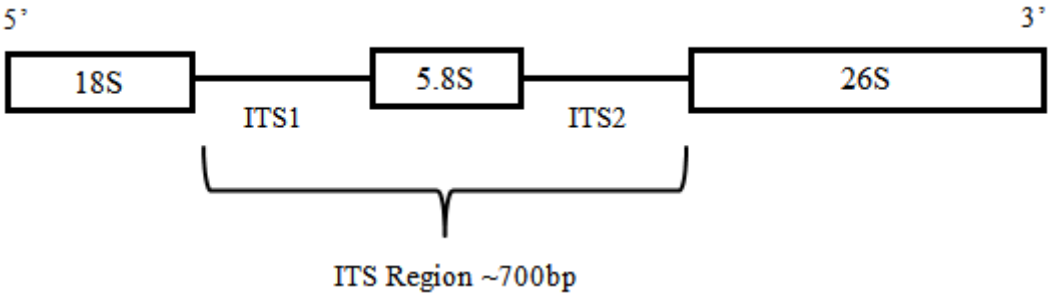

Figure 2.1 The three coding nrDNA repeat in plants. 18S, 5.8S and 26S are nrRNA genes. ITS1 and ITS2 are internal transcribed spacer regions. Modified from Soltis and Soltis (1998).

The nrITS sequence has also been proven as potentially effective in detecting the hybrid origin of plants or species, as well as in identifying reticulate evolution by

showing additive peaks on a sequencing electropherogram which contains evidence of hybridization when a species appears to have inherited repeat types from two parental species (Baldwin, 1992; Baldwin, 1993; Baldwin et al., 1995; Widmer and Baltisberger, 1999). Since the hybrids originate by joining of genomes from two different species, detection of parental genome in the putative hybrid taxa can be a direct evidence of a hybrid. The tandem repeats in nrITS are ideally suited for studying hybridization events because co-occurrence of parental nrITS types in a hybrid may be indicative of a recent hybrid origin (Koch et al., 2003) as concerted evolution usually leads to the rapid homogenization of divergent parental ribotypes (Wendel et al., 1995; Page and Holmes, 1998; Graur and Li, 2000). In the absence of sexual reproduction, concerted evolutionary homogenization of sequences by inter chromosomal crossing-over or gene conversion during chromosome pairing at meiosis would not be expected (Krieber and Rose, 1986; Elder and Turner, 1995; Li, 1997). Consequently, hybrids reproducing strictly vegetatively should retain copies of both divergent sequences for prolonged periods. In such instances, isolation of individual DNA sequences by molecular cloning can reveal the paternal origin of a hybridization event when sequences matching each parental species are recovered (Rossetto, 2005; Du et al., 2009; Zalewska-Gałosz et al., 2014). In recent years nrITS have been extensively used to investigate plant hybridization. Because of biparental inheritance of these markers, recent hybrids initially possess both divergent parental genotypes, as evidenced by DNA sequence polymorphisms (Zha et al., 2008; Les et al., 2009; Høibová et al., 2011; Kokubugata et al., 2011; Zalewska-Gałosz et al., 2014).

2.4.2 Chloroplast DNA Region; *mat*K

The angiosperm chloroplast genomes are double-stranded molecules, varying little in size, structure, and gene content, ranging from 120 to 200 kilobases (kb) (Soltis and Soltis, 1998). Chloroplast genomes contain a large 20 - 30 kb inverted repeat (IR_A and IR_B), which divides the remainder of the genome into two regions, one large single copy (LSC) and one small single copy (SSC) region (Figure 2.2) (Olmstead and Palmer, 1994).

Most of the genes within the chloroplast genome code for photosynthetic proteins, while the remainder are transfer RNA or ribosomal RNA genes and conserved ORFs (open reading frame) or potential protein-coding genes (Ravi et al., 2008). Wakasugi et al. (1998) constructed the updated gene map from tobacco (*Nicotiana tabacum* L.) which includes 105 different genes. There are many genes and intergenic spacers in the chloroplast genome that are widespread and sufficiently large (> 1 kb) to be generally useful in comparative sequencing studies and are highly conserved such as *mat*K, *rbc*L, *ndh*F, *atp*B, *rpH*6, *psa*B, *trn*L-*trn*F and *rbc*L-*acc*D (Oxelman et al., 1999; Chiang and Schaal, 2000; Soltis et al., 2000; Yuji et al., 2005; Heinze, 2007; Miz et al., 2008). These genes are suitable for a wide range of taxonomic levels and encompass a wide range of evolutionary rates (Olmstead and Palmer, 1994). The inheritance of the chloroplast DNA has historically been thought to be exclusively from the maternal parent in angiosperms (Corriveau and Coleman, 1988; Birky, 1995) and can infer a hybrid origin if a species appears to have inherited cpDNA from more than one maternal source (Clegg et al., 1993) due to chloroplast transfer from one species to another (Soltis and Soltis, 1998).

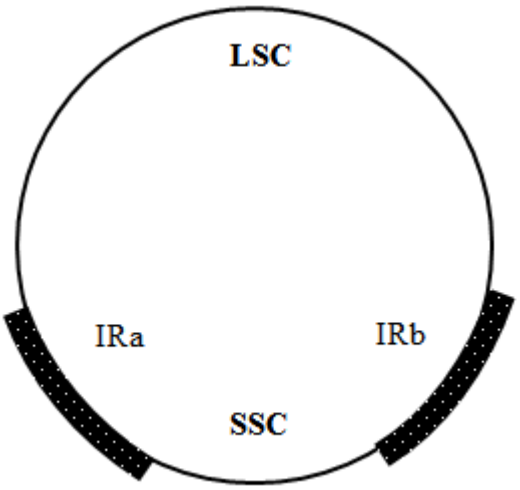

Figure 2.2 Chloroplast genome map showing the two inverted repeats (IRa and IRb) which separates the large single copy (LSC) from the small single copy (SSC). Modified from Soltis and Soltis (1998).

The Maturase K (*mat*K) gene was first identified by Sugita et al. (1985) from *N. tabacum* L. when they sequenced the *trn*K gene encoding the tRNA-lysine (UUU) of the chloroplast. The *mat*K gene, formerly known as *orf*K, is approximately 1500 base pairs long (bp) in most angiosperms and corresponding to around 500 amino acids for the translated protein product (Hilu et al., 1999). The *trn*K-*mat*K gene is located within an intron of approximately 2600 bp positioned between the 5' and 3' exons of *trn*K gene, in the LSC section adjacent to the inverted repeat (Figure 2.3) (Sugita et al., 1985; Hilu and Liang, 1997; Soltis and Soltis, 1998). This gene encodes a maturase-like polypeptide which might be involved in splicing Group II introns from RNA transcripts (Neuhaus and Link, 1987; Wolfe et al., 1992). The *mat*K gene has also been effective in addressing many systematic questions in various species that are important in molecular biology and evolution. Plant systematic studies have shown that the *mat*K gene are to be fast-evolving due to the fact that it has a high rate of nucleotide substitutions and more variable sites compared to other genes within cpDNA (Olmstead and Palmer, 1994; Johnson and Soltis, 1994; Soltis and Soltis, 1998). Olmstead and Palmer (1994) reported that out of 20 genes used in molecular systematics, the *mat*K had the highest nucleotide substitution rate. The rate of nucleotide substitution in *mat*K is three times faster than that of the large subunit of RubisCO (*rbc*L) in Saxifragaceae and six fold higher for the amino acid substitution rate (Olmstead and Palmer, 1994), denoting it as a fast- or rapidly-evolving gene. The resolution achieved with sequences of *mat*K is a relatively high rate of substitution in the conserved regions of the gene when comparing with eleven other genes combined representing multiple families and nine partial sequences representing monocot families from GenBank (Hilu and Liang, 1997; Hilu et al., 2003). Based upon the study of species representing different major plant groups, the

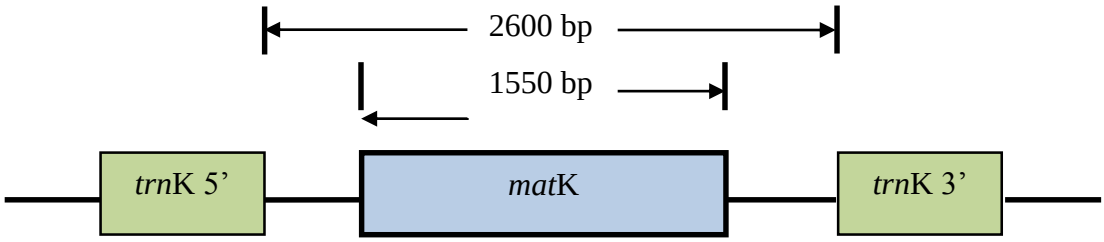

Figure 2.3 The *mat*K gene is an approximately 1.5 kb protein-coding region between two highly conserved 5’ and 3’ exons of *trn*K gene. Modified from Johnson and Soltis (1994).