# ITS2 AND trnH-psbA IN IDENTIFYING HIBISCUS SPECIES, MURRAYA PANICULATA AND MELALEUCA CITRINUS

ISSORY CHAVYN JORDAN

THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF BACHELOR OF BIOTECHNOLOGY (HONOURS)

FACULTY OF HEALTH AND LIFE SCIENCE INTI INTERNATIONAL UNIVERSITY PUTRA NILAI, MALAYSIA

## NON-PLAGIARISM DECLARATION

By this letter I declare that I have written this thesis completely by myself, and that I have used no other sources or resources than the ones mentioned.

I have indicated all quotes and citations that were literally taken from publications, or that were in close accordance with the meaning of those publications, as such. All sources and other resources used are stated in the references.

Moreover I have not handed in a thesis similar in contents elsewhere.

In case of proof that the thesis has not been constructed in accordance with this declaration, the Faculty of Health and Life Sciences has the right to consider the research dissertation as a deliberate act that has been aimed at making correct judgment of the candidate's expertise, insights and skills impossible.

I acknowledge that the assessor of this item may, for the purpose of assessing this item,

- reproduce this assessment item and provide a copy to another member of the University; and/or,
- communicate a copy of this assessment item to a plagiarism checking service (which may then retain a copy of the assessment item on its database for the purpose of future plagiarism checking).

In case of plagiarism the examiner has the right to fail me and take action as prescribed by the rules regarding Academic Misconduct practiced by INTI International University.

_Issory Chavyn Jordan Name	Signature	
	02   05   201 8 Date	

### **DECLARATION**

I hereby declare that the work in this thesis is my own except for quotations and summaries which have been duly acknowledged, and completed under the supervision of Dr. Geeta Selvarajah.

Issory Chavyn Jordan Student Name

Dr Geeta Selvarajah Supervisor

### **ACTKNOWLEDGEMENT**

I finished this dissertation by couple of month. Time was spent in a through reading all literature review available on the issue.

I must extend my wholehearted gratitude to Dr. Geeta Selvarajah, my project supervisor. Dr Geeta gave me her valuable comment and share me her previous experience on doing this dissertation. My coursemates also helping me a lot in the lab session. While when I had problem on writing, they will give me some of their opinion. We are helping each other to complete this dissertation.

Last but not least, I want to bestow my thanks to my parent for their support. They gave me support when I was stress during the period I prepared this dissertation and unflinching moral support.

#### **ABSTRACT**

The identification of plant species is still relevant today. It is used in ecology, medical field and for conservation efforts. DNA barcoding is a method able to provide a standardized and efficient tool for identification. It involves specific DNA sequences such as ITS2 and trnHpsbA known as DNA barcodes which are compared with known barcode sequences found on databases. Previous experiments to obtain ITS2 from Hibiscus resulted in poor quality sequence probably because of the heterogeneity of the barcode. One of the goals of this study was to see if the same problem exists in other members of the same clade, using Murraya panniculata, Hibiscus spp. and Callistemon citrinus as samples. Additionally, the effectiveness of trnH-psbA was assessed in the identification of the Roside clade. This study also attempted to identify a hibiscus species which had morphological characteristics which could not clearly identify it at species level. DNA was extracted using CTAB buffer followed by amplification of the markers by PCR after genomic DNA bands have been obtained on an agarose gel. It was hard to grind the leaves of the Hibiscus since foaming occurred. The amplification of the markers also required several attempts and modifications of sample concentrations. Nonetheless enough PCR products were successfully obtained for sequencing, and good quality sequences were obtained for all the samples. However, traces showed some multiple peaks at specific locations in the ITS2 sequences of Hibiscus rosa-sinensis and trnH-psbA sequences of M. panjculata; High mismatch ratios between the forward and reverse sequence were also encountered for these sequences and a consensus sequence could not be obtained using DNA BASER software. In these cases, only sequences from unidirectional reads were used. Using BLAST, the identification of the *Hibiscus* species and *M. panniculata* were ambiguous unlike C. citrinus. Maximum likelihood tree and neighbor joining tree using both markers could differentiate each species. Like the H. rosa-sinensis and M. panniculate, the unknown Hibiscus may be a hybrid; as the nuclear locus identified it as Hibiscus fragilis and the chloroplast loci identified it as *Hibiscus syriacus*.

# TABLE OF CONTENT

NON	- PLAGIARISM DECLARATION	PAGE II
DEC	LARATION .	Ш
ACK	NOLEDGEMENT	IV
ABS	ГКАСТ	V
TAB	LE OF CONTENT	VI
LIST	OF; TABLES	VIII
LIST	OF FIGURES	IX
LIST	OF ABBREVIATIONS	XI
СНАР	PTERS	
1.	INTRODUCTION	1
2.	LITERATURE REVIEW	3
	2.1 Hibiscus rošā-sinensis	3
	2.2 Murrdya pańiculata	4
	2.3 Melaleuca citrinus	4
	2.4 DNA barcoding	5
	2.5 Significance of DNA barcoding	5
	2.6 Comparison between dna barcoding with traditional method of identification	6
	2.7 ITS2 region and <i>trnH-psbA</i> as a plant DNA barcode	6
3.	METHODOLOGY	9
	3.1 Sample collection	9
	3.2 Preparation of reagents	10
	3.3 DNA extraction	10
	3.4 DNA quality assessment using agarose gel electrophoresis	11
	3.5 PCR amplification	12
	3.6 Analysis of PCR products by gel electrophoresis	14
	3.7 Sequencing and sequence analysis	14
	3.8 Sequence editing and BLAST analysis	14
	3.9 Multiple sequence alignment and tree topology	1.5

4.	RES	SULT	16
	4.1	Extraction of DNA and quality assessment using agarose gel electrophoresis	16
	4.2	Quality assessment of PCR products	17
	4.3	Sequence editing and BLAST analysis	19
	4.4	Multiple sequence alignment and phylogenetic tree construction	25
5.	DIS	CUSSION	27
	5.1	Sample collection and DNA extraction	27
	5.2	PCR amplification and product quality assessment by agarose gel	27
	•	electrophoresis	
	5.3	DNA sequencing and sequence editing	28
	5.4	BLAST analysis	29
	5.5	Tree topology	31
6	CO	NCLUSION	32
REFI	EREN	ICES	33
APPI	ENDI	CES	40
APPE	ENDIX	ζ A	40
APPE	ENDIX	X В	42
APPE			43
APPE		,	48
<b>APPE</b>	ENDIX	KE 😽 🗝	52

# LIST OF TABLES

Table	<b>t</b>	PAGE
3.1.1	The location of the plant from which the leaves sample were taken	9
3.5.1	Primer sequence of ITS2 and trnH-psbA used for amplification	12
3.5.2	Respective volume of each reagent in the PCR mix and the negative control	13
3.5.3	PCR conditions for ITS2 and trnH-psbA markers	13
4.3.1	Quality value (QV) for reverse and forward reaction for ITS2 sequences and <i>trnH-psbA</i>	20
4.3.2	BLAST analysis for ITS2 marker	22
4.3.3	BLAST analysis for <i>trnH-psbA</i> marker	24

# LIST OF FIGURES

Figure	4	PAGE
2.1.1	ITS region	7
4.1.1	Genomic DNA bands on agarose gel under UV light	17
4.2.1	ITS2 marker bands on agarose gel under UV light	18
4.2.2	trnH-psbA marker bands on agarose gel under UV light	19 21
4.3.1	Log report obtained for poor quality ITS2 sequence of R1	22
4.3.2	Log report obtained for good quality ITS2 sequence of X3	22
4.4.1	Maximum likelihood tree computed using Jukes and Cantor parameter model and 1000 bootstrapping replications for ITS2 barcoding sequences from both <i>Hibiscus</i> species (R1 ITS2, R2 ITS2, W1 ITS2 and W2 ITS2, <i>Melaleuca citrinus</i> (T1 ITS2, T2 ITS2 and T3 ITS2) and Murraya panniculata (X1 ITS2, X2 ITS2 and X3 ITS2)	25
4.4.2	Maximum likelihood tree computed using Tamura parameter model and 1000 bootstrapping replications for <i>trnH-psbA</i> barcoding sequences from both <i>Hibiscus</i> species (R1 <i>trnH-psbA</i> , R2 <i>trnH-psbA</i> , W1 <i>trnH-psbA</i> and W2 <i>trnH-psbA</i> ), <i>Melaleuca citrinus</i> (T1 <i>trnH-psba</i> , T2 <i>trnH-psbA</i> and T3 <i>trnH-psbA</i> ) and Murraya panniculata (X1 <i>trnH-psbA</i> , X2 <i>trnH-psbA</i> and X3 <i>trnH-psbA</i> )	26
5.4.1	Screenshot of a phylogenetic tree that demonstrate <i>Glycosmis</i> and <i>Murraya</i> have a common ancestor	30
7.1	The unknown Hibiscus	42
7.2	Hibiscus rosa sinensis	42
7.3	Murraya paniculata	42
7.4	Melaleuca citrinus	42
7.5	Neighbor joining tree computed using Jukes and Cantor parameter model and 1000 bootstrapping replications for ITS2 barcoding sequences from both <i>Hibiscus</i> species (R1 ITS2, R2 ITS2, W1 ITS2 and W2 ITS2, <i>Melaleuca citrinus</i> (T1 ITS2, T2 ITS2 and T3 ITS2) and Murraya panniculata (X1 ITS2, X2 ITS2 and X3 ITS2)	52

Neighbor joining tree computed using Tamura parameter model and 1000 bootstrapping replications for *trnH-psbA* barcoding sequences from both *Hibiscus* species (R1 *trnH-psbA*, R2 *trnH-psbA*, W1 *trnH-psbA* and W2 *trnH-psbA*) *Melaleuca citrinus* (T1 *trnH-psbA*, T2 *trnH-psbA* and T3 *trnH-psbA*) and Murraya panniculata (X1 *trnH-psbA*, X2 *trnH-psAa* and X3 *trnH-psbA*)

### LIST OF ABBREVIATIONS

°C Degree Celsius

μL Microlitre

μM Micromolar

ug/mL Microgram per microlitre

% Percentage

BLAST Basic Local Alignment Search Tool

COI Cytochrome c oxidase 1 gene

CTAB Cetrimonium bromide

dATP Deoxyadenosine triphosphate

dCTP Deoxycytidine triphosphate

dGTP Deoxyguanosine triphosphate

dTTP Deoxythymidine triphosphate

dNTP Deoxynucleotide

DNA Deoxyribonucleic acid

EDTA Ethylene Diamine Tetraacetic Acid

g Gram

HCl Hydrochloric acid

hr hour

ITS Internal transcribed spacer

ITS2 Internal transcribed spacer 2

kb kilobases

M Moles

matK Megakaryocyte-Associated Tyrosine Kinase

MEGA 7.0 Molecular Evolutionary Genetics Analysis

MgCl<sub>2</sub> Magnesium Chloride

mg/mL Miligram per millilitre

min Minute

mL Millilitre

mM Millimolar

MW Molecular weight

NaCl Sodium Chloride

NaOH Sodium Hydroxide

NCBI National Centre for Biotechnology Information

nmoles nanomoles

PCR Polymerase chain reaction

pH Potential of hydrogen

rbcl Ribulose bisphosphate carboxylase

RNA Ribonucleic acid

rpm Revolutions per minute

sec Second

TBE buffer Tris/Borafe/EDTA buffer

TE buffer Tris/Borate buffer

TER Tris/EDTA buffer with RNase A

w/w Volume of solute in millilitres per 100 millilitres

w/v Weight in grams of a solute per 100 millilitres

### **CHAPTER 1**

### INTRODUCTION

Approximately 1.7 million species have been identified through the conventional taxonomic method that is using morphological traits (Mohamad et al., 2014). However, this figure is not totally reliable because this method is very subjective, thus many problems were encountered and it was a slow process (Mohamad et al., 2014). Accurate and quick identification of species is extremely crucial in scientific fields such as ecology, epidemiology, conservation, evolution (Kress, García-Robledo, Uriarte & Erickson, 2015) and recently even in crime detection for example illegal trading of timber (Hartvig, Czaka, Kjaer, Nelson & Theilade., 2015). Thus, DNA barcoding is used.

Several DNA barcodes have been proposed for land plants namely *trnH-psbA*, *matK*, *rbcL*, *rpiCl*, *ycf5*, ITS and ITS2 (Hui et al., 2010). The *matK* and *rbcL* combination is considered as a core two marker barcode because *rbcL* can be easily recovered and *matK* has good discriminatory power (Hollingsworth, Graham, & Little, 2011). However, this combination has drawbacks that makes its utility in identification questionable (Hollingsworth et al., 2011). *matK* does not always amplify with the current available primers and *rbcL* lacks sequence variation (Hollingsworth et al., 2011). Thus, DNA barcodes which exhibit higher levels of sequence variation and can be easily amplified may improve the ability to identify species. The barcodes having these characteristics are nuclear ITS and *trnH-psbA* (Hollingsworth et al., 2011; Huyen, Thi, Truong, Duy, & Hoang, 2017).

The success of using this technology, DNA barcoding, also relies on sequence homogeneity of barcode which ITS2 may not have. Being a nuclear gene it may possess biparental inheritance, meaning two nuclear ribosomal ITS2 sequences that differ exist together in an organism if it is a heterozygote (Hughes, Morris, & Reboredo-Segovia, 2015). Heterogeneity posed difficulties for Sanger sequencing as double or multiple nucleotide peaks were present in the sequence traces (Woo *et al.*, 2010).

The problem of poor sequence quality of ITS2 has been reported in Hibiscus, and it is suspected that it is due to the existence of heterogenous sequences since it is known that *Hibiscus rosa-sinensis* may have undergone hybridization with other plant species and thus result in many variations among the *Hibiscus* plant species (Singh & Khoshoo, 1970). Whether this is a problem that exists in other members of the same clade, rosids, is also a question.

Thus the aims of this study were to:

- Determine whether good quality sequences of ITS2 can be obtained from other plants of the same clade as Hibiscus, i.e., rosids, which are *Muraya paniculata* and *Melaleuca citrinus*.
- Determine whether ITS2 and *trnH-psbA* can effectively differentiate between *Hibiscus spp., Muraya paniculata* and *Melaleuca citrinus*
- Identify an unknown *Hibiscus* species using ITS2' and *trnH-psbA*

### **CHAPTER 2**

### LITERATURE REVIEW

### 2.1 Hisbiscus rosa-sinensis

The known *Hibiscus* species in this study is *Hibiscus rosa-sinensis*. It is a widely distributed plant which has many variants and used as a garden ornamental plant worldwide (Ross, 2003). It belongs to the family Malvaceae and it is believed to have originated from the East Asia, mostly China (Lim 2014).

H. rosa-sinensis is a shrub which can grow up to 20 feet (Ross, 2003). It is multi-branched plant where the branches are long and thin and spirally arranged along the stem (Ross, 2003). The H. rosa-sinensis's flowers are large, conspicuous and trumpet shaped single ones which can reach 18 cm breadth and located in the axils of the upper leaves (Ross, 2003). The corolla of the flower can be of different colours such as rosy red, white, reddish, purplish and orange-yellow (Lim, 2014). This plant normally remains upright and always retains its green leaves (evergreen) (Lim, 2014). The leaves are dark green in color and ovate with a serrated leaf margin (Lim, 2014).

Besides being used as an ornamental plant (Ross 2003), the plant possesses medical properties such as anti-depressant (Pallavi, Rupali & Yogesh, 2012), antibacterial (Ruban, & Gajalakshmi, 2012), cardio-protective (Gauthaman, et al., 2006), wound healing (Bhaskar, & Nithya, 2012), anti-dyslipidaemia and anti-oxidant activities (Kumar, *et a.l*, 2013). It can also be used to stimulate menstrual flow (Burkil, Birwistle, Foxworthy, Scrivenore, & Watson, 1935). It treats gonorrhoea (Whistler, 1985), diarrhoea and promotes childbirth as well as milk production in the mammary gland (Kobayashi, 1976).