DNA BARCODING FOR THE IDENTIFICATION OF DIFFERENT *HIBISCUS ROSA-SINENSIS* INTRASPECIFIC VARIANTS USING *rbcL* AND *ITS2* MARKERS

JOHN MARK WAMAHIU MUNENE

DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF BIOTECHNOLOGY (HONOURS)

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

2016

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.

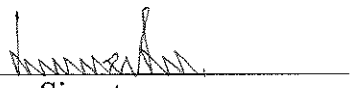
Moreover I have not handed in a thesis similar in contents elsewhere.

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 thesis 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.

John Mark Wamahiu Munene	
Name	Signature
I13004466	14th November 2016
I.D.Number	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.

John Mark Wamahiu Munene	Dr. Geeta Selvarajah
Student ID	(SUPERVISOR)

14/NOV/2016

ACKNOWLEDGEMENT

I take this opportunity to express deeply sincere gratitude to my supervisor, Dr. Geeta Selvarajah, for the guidance and support rendered to me in the midst of compiling this thesis. She inspired me to think critically and enabled me to obtain the understanding necessary to draw up this thesis by sharing her wealth of knowledge and experience. I would also like to make known my appreciation for her infinite patience, understanding and encouragement which served me well in conducting my experiments, gathering the required information and compiling this thesis.

I also take this opportunity to thank the lecturers and staff of INTI International University for providing an enabling learning environment that has spurred me towards achieving my academic potential. Not forgetting also my fellow student colleagues who provided helpful criticism and encouragement along the way.

ABSTRACT

DNA barcoding is a novel technique aimed at providing a means to rapidly and accurately identify species using short, standardized DNA sequences as tags in a fully automatable process. Its usefulness becomes apparent when need arises to support identifications of organisms made in accordance with the existing Linnaean classification system. It is universally applicable to all organisms and only a small set of markers are used. In this experiment, the genetic markers *rbcL* and *ITS2* were used with the aim of assessing their discriminatory powers in as far as differentiating between three varieties of *Hibiscus rosa-sinensis* namely, *Hibiscus rosa-sinenensis* el capitolio red, *Hibiscus rosa-sinensis* red – variegated leaves and *Hibiscus rosa-sinenesis* pink. Using the Edwards' method of DNA extraction, genomic DNA was first obtained from the plant samples before the *rbcL* and *ITS2* gene sequences were isolated from each of the *Hibiscus rosa-sinensis* variants and amplified by Polymerase Chain Reaction. The PCR products were then sequenced and the sequence data analyzed using various bioinformatics tools. The first step of analysis involved sequence editing and quality assessment, followed by BLAST analysis on the NCBI database for sequence comparison. The next step was to conduct Multiple Sequence Alignment in order to detect polymorphisms between the sequences and enable the construction of a phylogenetic tree. Finally, a DNA distance matrix was constructed which allowed for calculation of the barcoding gap and the resulting data was used to distinguish between the three variants of *Hibiscus rosa-sinensis*. In conclusion, the species discrimination powers of both markers were assessed and although both markers displayed the ability to discriminate between species (interspecific variation), the *ITS2* sequence was found to be more suitable for use in the identification of different variants within the species (intraspecific variation) as opposed to the *rbcL* sequence which was unable to identify differences between intraspecies variants.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFLP	Amplified fragment length polymorphism
BLAST	Basic Local Alignment Search Tool
BOLD	Barcode of Life Data System
COBL	Consortium for the Barcode of Life
COI	Cytochrome c oxidase 1 gene
DNA	Deoxyribonucleic acid
EDTA	Ethylene Diamine Tetraacetic Acid
ITS2	Internal transcribed spacer 2
MEGA 7.0	Molecular Evolutionary Genetics Analysis
NaCl	Sodium chloride
NCBI	National Center for Biotechnology Information
PCR	Polymerase chain reaction
PVP	Polyvinylpyrrolidone
rbcL	RuBisCo large subunit
RFLP	Restriction fragment length polymorphism
RAPD	Random amplification of polymorphic DNA
RNA	Ribonucleic acid
RNase	Ribonuclease
SDS	Sodium dodecyl sulfate
TBE buffer	Tris/ Borate/ EDTA buffer
TE buffer	Tris/EDTA buffer
TER	Tris/EDTA buffer with RNase A

CHAPTER 1

INTRODUCTION

The identification of species is of great importance in scientific research, particularly in the fields of conservation, ecology, epidemiology and evolution (Kress, García-Robledo, Uriarte, & Erickson, 2015). It is in the pursuit of this goal, of accurately identifying the various species in existence, that the molecular technique known as DNA barcoding arose. Kress *et al.* (2015) elucidated on DNA barcoding as being a technique that utilizes particular sequences derived from a standard segment of the genome to identify organisms. DNA barcoding compares variation between different species for discrimination, but still allows for small amounts of variation that may be used to distinguish between individuals within the same species as well (Eby, Linson, Arockiasamy, Melchias & Prabhu, 2011). This is particularly useful when identifying variants within the same plant species such as *Hibiscus rosa-sinensis*, which occurs in several different varieties (Raja Rao, 2008; Ustinova, 1937).

The species under investigation in this study is *Hibiscus rosa-sinensis*. Ross (2003) described this particular species as a shrub belonging to the *Malvaceae* family – a plant with long, thin branches, spirally arranged along the stem, which can grow up to six metres. The plant produces single flowers, in the axils of the upper leaves, which exhibit cupule-shaped calyx of about 2.5 centimetres in length. The plant's origins are believed to lie in East Asia – China, and it can now be found, in wide distribution, within the tropics and subtropics (Lim, 2014). In regard to this proliferation, it is possible that the *Hibiscus* may have undergone hybridization with other plant species which would contribute to variation amongst affected *Hibiscus* species (Singh & Khoshoo, 1970). DNA barcoding may be carried out on such plants in order to identify and distinguish between hybrids and the original varieties. It has been noted that the flowers and leaves of *Hibiscus rosa-sinensis* are edible and some compounds isolated from the plant are known to exhibit medicinal properties such as cognitive-enhancing, antifertility,

cardioprotective, hypotensive, antimutagenic and wound-healing activity (Tanaka, 1976; Reddy *et al.*, 2007; Wongwattanasathien *et al.*, 2010; Lim, 2014). From this, the importance of accurately identifying the species is evident as it allows for use of the appropriate species for the desired purposes.

The two barcoding loci used to identify *Hibiscus rosa-sinensis* in this study are *rbcL* and *ITS2*. A number of sequences have been recommended for use as DNA barcodes and these include *psbA-trnH*, *matK*, *ITS*, *ycf5* and *rpoC1* (Chen *et al.*, 2010). Previous researchers on this project (Law, 2016) have amassed some information on *matK* and *trnH-psbA* and thus we would like to study the performance of two other recommended loci (*rbcL* and *ITS2*) in intraspecific identification of *Hibiscus*. A BLAST search on the NCBI database carried out on 14th June 2016 at 12:47 a.m. elicited 12 hits for *Hibiscus rosa-sinensis rbcL* and only 3 hits for *Hibiscus rosa-sinensis ITS2* – which did not specify the variants used. From this it may be realized that there is a necessity to carry out DNA barcoding for the *rbcL* and *ITS2* genes on *Hibiscus rosa-sinensis* in order to increase the amount of sequence data in the database which could be used to support the identification of individuals by differentiating them from other variants within the same species based on morphology. This information may prove particularly useful in conservation efforts and studies on evolution. Thus, the aims of this experiment were to determine the DNA sequences of *rbcL* and *ITS2* in *Hibiscus rosa-sinensis* and to determine whether these loci may be used to differentiate between the intra-species variants *Hibiscus rosa sinenensis* el capitolio red, *Hibiscus rosa sinensis* red – variegated leaves and *Hibiscus rosa sinenesis* pink.

CHAPTER 2

LITERATURE REVIEW

2.1 HIBISCUS ROSA-SINENSIS

Hibiscus rosa-sinensis is a plant belonging to the family of *Malvaceae*. As described by Lim (2014), the plant's origins lie in East Asia and it is characterised as having tropical or sub-tropical requirements, with cultivation as a garden ornamental plant that grows between altitudes of 0 and 500 m. It is sensitive to frost; freezes in mild winters and exhibits susceptibility to injury in severe winters. It flourishes under full sunlight in fertile, well-drained soils rich in organic matter. It is grown as a house-plant all over the world with most ornamental varieties being hybrids – many of which are as a result of crosses with the *Hibiscus schizopetalus* species from Africa (Ross, 2003).

Ross (2003) described this particular species as a shrub with long, thin branches, spirally arranged along the stem, which can grow up to six metres. The plant produces single flowers, in the axils of the upper leaves, which exhibit cupule-shaped calyx of about 2.5 centimetres in length and epicalyx of 5 to 7 bracteoles, each with a length of about 1 cm. The corolla occurs in various colours including rosy red, white, reddish, purplish or orange-yellow (Lim, 2014). The plant stands erect, is multi-branched and evergreen (Lim, 2014). The leaves are usually dark green and ovate with a serrated leaf margin (Lim, 2014).

Aside from being ornamental, the plant is edible and has several uses as a medicinal plant as well. In China, the bark or flowers of the plant are boiled in water and the solution taken orally as an emmenagogue (Burkill *et al.*, 1935; Lim, 2014). In the East Indies, flower extracts are administered orally to produce abortion (Burkill *et al.*, 1935; Lim, 2014). In the Cook Islands, dried leaves and flowers of the plant are boiled and the extract taken orally to treat gonorrhoea (Whistler, 1985; Lim, 2014). In Fiji, fresh

leaves of the plant are used to make a juice which is administered orally to treat diarrhoea and also to enhance childbirth (Singh, 1986; Lim, 2014). In Hawaii, the flowers are eaten to elicit lactation (Kobayashi, 1976; Lim, 2014). These are but some of the uses of the plant in various parts of the world. The plant's flower petals are also used in salads, to garnish dishes, eaten as pickles, used to make herbal teas and also used as food colouring (Lim, 2014). Some of the chemical components of this plant species that have been isolated include ergosterols (precursor to vitamin D2), flavonoids, quercetin (anti-inflammatory and antihistamine), stigmasterol (which has antimicrobial, antioxidant and anti snake venom activity), cyadinin (anti-diabetic) and taraxeryl acetate (anti-inflammatory agent) (Rehman et al., 2013; Lim, 2014; Acu-Cell, 2016; PubChem, 2016; University of Maryland Medical Center, 2016).

The variants to be investigated in this study are *Hibiscus rosa-sinenensis* el capitolio red, *Hibiscus rosa-sinensis* red – variegated leaves and *Hibiscus rosa-sinenesis* pink.

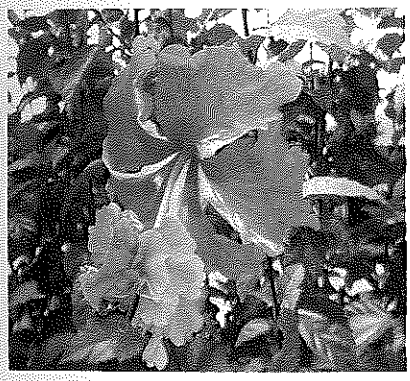


Figure 2.1.1 *Hibiscus rosa-sinenensis* el capitolio red (Grossman *et al.*, 2016)

Figure 2.1.2 *Hibiscus rosa-sinenensis* red – variegated leaves (Brugmansia-Québec)

Figure 2.1.3 *Hibiscus rosa-sinenensis* pink (Nilai, 2016)

2.2 DNA BARCODING

DNA barcoding is a novel technique that was developed by Canadian biologist Paul D. N. Hebert and his research group from the University of Guelph, aimed at providing a means to rapidly and accurately identify species using short, standardized DNA